CRiSM

Workshops

September 2015

Models and Inference in Population Genetics

14–16 september 2015

Warwick Statistics

EPSRC

Engineering and Physical Sciences Research Council

Models and Inference in Population Genetics
14th – 16th September 2015

Organising Committee: Bob Griffiths (Oxford), Paul Jenkins (Warwick),
Dario Spanò (Warwick)

Contents

1 Administrative Details

1.1 Workshop Webpage

- `http://warwick.ac.uk/populationgenetics2015`

1.2 Registration and Venue

- **Registration:** 8:00–8.55, Mon. 14th Sep, Maths & Stats Building, Lobby. (Map: 38, F4)
- **Talks:** Maths & Stats Building, MS.04. (Map: 38, F4)
- **Wine & Poster Session:** 17:00–18:00, Mon. 14th Sep, Maths & Stats Building, Lobby. (Map: 38, F4)
- **Breakfast (Participants with on-campus accommodation only):** 7:30–9:00, Arden or Rootes Social Building. (Map: 1, F2 or 55, C6)
- **Lunch:** Monday: 12:50–14:00, Maths and Stats Building, Lobby. (Map: 38, F4)
  Tuesday: 12:50–14:00, Maths and Stats Building, A0.05. (Map: 38, F4)
  Wednesday: 12:00–13:00, Maths and Stats Building, A0.05. (Map: 38, F4)
- **Dinner (Monday):** 19:00, Rootes Social Building (Rootes Restaurant). (Map: 55, C6)
- **Conference Dinner (Tuesday):** 19:00, Radcliffe House. (Map: 52, C4)
- **Conclusion:** 13:00, Wednesday 16th Sep.

1.3 Getting Here

- Information on getting to the University of Warwick from Coventry, as well as from other directions locally and further afield, can be found at

  `http://www.warwick.ac.uk/about/visiting/`

- See Appendix A & B for further details.

1.4 Accommodation

Accommodation is in en-suite rooms on campus either in Arden (map: 1, F2) or in the Arthur Vick residences (map: 3, F6).

- **Arden**: Check in time is from 15:00 and check out time is by 10:00 on your day of departure. Keys are available from the reception in Arden. Parking is available in Car Park 13 next to Arden; a parking permit can be obtained from reception on arrival. See Appendix A for further details on Arden.
- **Arthur Vick (part of the Warwick Conference Park)**: Check in time is from 15:00 and check out time is by 09:30 on your day of departure. Keys can be collected before 11pm from Conference Reception in the Students' Union (map: 63, D6). See Appendix B for further details on Warwick Conference Park.

Permit parking is available for conference delegates in any of the central campus shared use car parks (7, 8, 8a, 9, 10 and 15). If you would like to use any of these car parks please find in Appendix D of this booklet a copy of the parking permit ready to cut and use. Please email Olivia Garcia-Hernandez (Olivia.Garcia@warwick.ac.uk) to obtain disabled parking permit.

1.5 Internet Access

- **Campus:** Wireless access is most easily available via eduroam — `http://www.eduroam.org/` — which is supported across most of the Warwick campus. Speak to one of the organisers for details of other options.
- **Accommodation:** Wireless access is available; ask for log-in details when you check-in to your accommodation.

1.6 Start.Warwick

- The Start.Warwick app (available for iPads, iPhones and Android devices) provides useful information on travel and an interactive map of the campus amongst other things.

1.7 Facilities — See Map

- **Supermarket, Food and Drink Outlets:** `http://www.warwickretail.com`
  - See Appendix C for opening times.
- **Arts Centre:** `http://www.warwickartscentre.co.uk` (Map: 68, D5)
- **Sports Centre:** `http://www.warwick.ac.uk/sport/` (Map: 62, E5)
- **Health Centre:** `http://www.uwhc.org.uk` (Map: 21, D6)
- **Pharmacy:** Students Union Atrium (Map: 63, D6)

2 Help, Information & Telephone Numbers

2.1 Department

- **Address:** Department of Statistics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL
- **Telephone:** 024 7657 4812
- **Fax:** 024 7652 4532
- **Webpage:** `http://www.warwick.ac.uk/stats`

2.2 Emergency Numbers

- **Emergency:** Internal — 22222; External — 024 7652 2222
- **Security:** Internal — 22083; External — 024 7652 2083
- **Organisers:** Paul Jenkins: Internal — 74856, External — 024 765 74856;
  Dario Spanò: Internal — 75755, External — 024 765 75755

2.3 Transport

- **Swift Taxis (Coventry):** 024 7676 7676
- **Trinity Street Taxis:** 024 7699 9999
- **National Rail Enquiries:** 08457 484 950

3 Timetable

All activities will take place in the Mathematics & Statistics Building (map: 38, F4), with talks in room MS.04 (signposted from lobby), unless otherwise stated.

3.1 Monday 14th September

Time	Speaker	Title	Pg
7:30	**Breakfast (until 9:00)**	**Arden or Rootes Social Building**	**-**
8:00	**Registration**	**Maths & Stats building, lobby**	**-**
8:55	**Welcome**	**-**	**-**
9:00	Mark Beaumont	Computational approaches for detecting local selection	8
9:50	Christopher Quince	Inferring microbial species and strains directly from metagenome data	12
10:40	**Coffee**	**Lobby**	**-**
11:10	Maria De Iorio	Bayesian inference on population structure: from parametric to nonparametric modelling	8
12:00	Yee Whye Teh	Cancer subclonal reconstruction with Aldous' beta-splitting trees	12
12:50	**Lunch**	**Lobby**	**-**
14:00	Ellen Baake	Ancestral selection graph meets lookdown construction	8
14:50	**Poster summaries**	Eugenio Buzzoni Mathias Cronjäger & Alejandra Avalos-Pacheco Andreas Futschik Jere Koskela Sebastian Probst & Mareike Esser Simone Tiberi	14
15:40	**Coffee**	**Lobby**	**-**
16:10	Jochen Blath	Some new results for seed bank models	
17:00	**Posters & wine reception**	**Lobby**	**-**
18:00	**Finish**		**-**
19:00	**Dinner**	**Rootes Social Building (Rootes Restaurant)**	**-**

3.2 Tuesday 15th September

Time	Speaker	Title	Pg
7:30	**Breakfast**	**Arden or Rootes Social Building**	**-**
9:00	Steven Evans	Random trees in the large	9
9:50	Matthias Birkner	A conditional coalescent limit in fixed pedigrees	
10:40	**Coffee**	**Room A0.05**	**-**
11:10	Richard Everitt	A preliminary investigation of online Bayesian inference of clonal ancestry	9
12:00	Gerton Lunter	Using particle filters to infer past population sizes and migration rates from whole-genome sequences	11
12:50	**Lunch**	**Room A0.05**	**-**
14:00	Carolin Kosiol	Reversible polymorphism-aware phylogenetic models and their application to species tree inference	10
14:50	Amy Williams	Inferring local ancestry by jointly analyzing admixed samples	13
15:40	**Coffee**	**Room A0.05**	**-**
16:10	Simon Myers	A new approach to infer local ancestry in populations impacted by admixture	
17:00	**Finish**		
19:00	**Conference dinner**	**Radcliffe House**	**-**

3.3 Wednesday 16th September

Time	Speaker	Title	Pg
7:30	**Breakfast**	**Arden or Rootes Social Building**	**-**
9:00	Robert Griffiths	Neutral Wright-Fisher bridges	9
9:50	Asger Hobolth	Mathematical theory and statistical methods for the neutral Wright-Fisher model	10
10:40	**Coffee**	**Room A0.05**	**-**
11:10	Noemi Kurt	An individual-based model for the Lenski experiment, and the deceleration of the relative fitness	11
12:00	**Lunch**	**Room A0.05**	**-**
13:00	**Finish**	**-**	**-**

4 Talk Abstracts

Computational approaches for detecting local selection

Mark Beaumont
University of Bristol

There has been widespread interest in methods for detecting geographically localised selection in the genome. The basic aim is to find SNPs or sequences with allele or haplotype frequencies that vary among populations to a greater extent than expected under drift, given some demographic model. These approaches are controversial because they are generally confounded by unknown demography. This talk gives an overview of different modelling approaches for detecting local selection that have been taken, and then provides some recent results from the use of Expectation Propagation in conjunction with Monte-Carlo as a computational method for detecting local selection. Finally, the prospects of using Expectation Propagation as a general tool for combining Monte-Carlo-based likelihoods across the genome is discussed.

Ancestral selection graph meets lookdown construction

Ellen Baake
Universität Bielefeld

In a (two-type) Wright-Fisher diffusion with directional selection and two-way mutation, let x denote today's frequency of the beneficial type, and given x, let $h(x)$ be the probability that, among all individuals of today's population, the individual whose progeny will eventually take over in the population is of the beneficial type. Fearnhead (2002) and Taylor (2007) obtained a series representation for $h(x)$. We develop a construction that contains elements of both the ancestral selection graph and the lookdown construction and includes pruning of certain lines upon mutation. Besides being interesting in its own right, this construction allows a transparent derivation of the series coefficients of $h(x)$ and gives them a probabilistic meaning.

Bayesian inference on population structure: from parametric to nonparametric modelling

Maria De Iorio
University College London

Making inference on population structure from genotype data requires to identify the actual subpopulations and assign individuals to these populations. The source populations are assumed to be in Hardy-Weinberg equilibrium, but the allelic frequencies of these populations and even the number of populations present in a sample are unknown. In this talk we present a review of some Bayesian parametric and nonparametric models for making inference on population structure, with emphasis on model-based clustering methods. Moreover, we propose a Bayesian nonparametric model to infer population admixture, extending the Hierarchical Dirichlet Process (HDP, Teh et al. 2006) to allow for correlation between loci due to Linkage Disequilibrium. Our aim is to show how recent developments in Bayesian nonparametrics have been usefully exploited in order to introduce natural nonparametric counterparts of some of the most celebrated parametric approaches for inferring population structure.

Random trees in the large

Steven N. Evans
University of California at Berkeley

Trees are ubiquitous in population biology, phylogenetics and many other areas such as computer science, and there are numerous stochastic models that grow larger and larger trees sequentially. In work with various collaborators in the last few years I have examined the question of characterizing how several of these models "converge to infinity" using the notion of the Doob-Martin compactification of a countable state-space Markov chain. This talk will review some of those results and their connection with areas such as exchangeable arrays and dense graph limits.

A preliminary investigation of online Bayesian inference of clonal ancestry

Richard Everitt
University of Reading

Whole genome sequencing has had a big impact in studying the evolutionary history of pathogens. At the core of many studies is the need to infer the clonal ancestry of a sample from sequence data. This is usually performed using a Bayesian approach, with a coalescent prior on ancestries, and using Markov chain Monte Carlo (MCMC) for inference, as implemented in software such as BEAST. For most pathogens it is important to also account for recombination, and it may also be of interest to infer recombination events. ClonalFrame, again a Bayesian approach using MCMC, is a popular technique for tackling this problem, but it takes a prohibitively long time to run for a moderate number of sequences. This talk describes preliminary work on a sequential Monte Carlo approach to inference in the ClonalFrame model, with the aim of overcoming this computational intractability.

Neutral Wright-Fisher bridges

Robert Griffiths
University of Oxford

The path of the frequency of a gene which enters a population and is then lost at time T later is modelled as a bridge in a Wright-Fisher diffusion process beginning at 0 at time 0 when the gene enters and ending at 0 at time T when the gene is lost. Theoretical results that will be discussed in the talk are the infintesimable generator of the gene frequency in a bridge; the first passage time from 0 to a level y in a Wright-Fisher diffusion and the maximum frequency attained in a bridge. The coalescence structure in the population is useful in understanding the bridge. This talk is mostly from the paper Joshua Schraiber, Robert Griffiths, Steven Evans, (2013), Analysis and rejection sampling of Wright-Fisher diffusion bridges, *Theoretical Population Biology* **89** 64–74.

Mathematical theory and statistical methods for the neutral Wright-Fisher model

Asger Hobolth
Aarhus University

We consider the diffusion approximation of the Wright-Fisher process with mutation. Analytically tractable formulae for the first- and second-order moments of the allele frequency distribution are derived, and the moments are subsequently used to better understand key population genetics parameters and modelling frameworks. In particular we investigate the behaviour of the expected homozygosity (the probability that two randomly sampled genes are identical) in the transient and stationary phases, and how appropriate various distributions are for modelling the allele frequency distribution at different evolutionary time scales. In the univariate setting the Beta distribution is typically a fine approximation, and we find that the Dirichlet distribution is adequate for the multivariate pure drift model (no mutations allowed) and in most cases a suitable approximation for the multivariate parent independent mutation model. The Dirichlet distribution is, however, not sufficiently flexible for more general mutation models. We highlight recent suggestions to treat the multivariate Wright-Fisher process with a Kimura mutation model and with an infinite alleles mutation model. This is joint work with Jukka Siren from Helsinki University.

Reversible Polymorphism-Aware Phylogenetic Models and their Application to Species Tree Inference

Dominik Schrempf[1,2] and Carolin Kosiol[1]
[1]*Institut für Populationsgenetik, Vetmeduni, Vienna, Austria,* [2]*Vienna Graduate School of Population Genetics, Vetmeduni Vienna, Austria*

The availability of genome-scale inter- and intraspecies data leads to new opportunities in phylogenetics to improve tree accuracy and resolution as well as to take important steps towards understanding the process of speciation. We present a novel maximum likelihood implementation of a Polymorphism-Aware Phylogenetic Model (PoMo, De Maio et al., MBE 2013) that can do both, parameter estimation and species tree inference for genome-wide data of a moderate number of species while still allowing for many individuals per species. It extends any DNA substitution model and additionally accounts for polymorphisms in the present and in the ancestral population by expanding the state space to include polymorphic states. It is a selection-mutation model which separates the mutation process from the fixation process. Thereby, a Moran process is used to model genetic drift. Although a single phylogeny — the species tree — is considered, PoMo naturally accounts for incomplete lineage sorting because ancestral populations can be in a polymorphic state.
A large scale simulation study with four different scenarios for small and large trees (incomplete lineage sorting, anomaly zone, recent radiation and trichotomy) as well as applications to great ape data (12 populations in total, Prado-Martinez et al., 2013) show that PoMo is fast while being more accurate than other state-of-the-art methods (De Maio, Schrempf, and Kosiol, Syst. Biol. 2015).
Recently, we have derived a reversible version of PoMo and implemented it into IQ-Tree (Nguyen et al., 2015), an efficient and easy-to-use software package. We observe a reduction of runtime by a factor of 50 and no loss of accuracy. This demonstrates that PoMo is suitable to infer large scale phylogenies from population data.

An individual-based model for the Lenski experiment, and the deceleration of the relative fitness

Noemi Kurt
Technische Universität Berlin

The Lenski experiment investigates the long-term evolution of bacterial populations. Its design allows the direct comparison of the reproductive fitness of an evolved strain with its founder ancestor. It was observed by Wiser et al. (2013) that the relative fitness over time increases sublinearly, a behaviour which is commonly attributed to effects like clonal interference or epistasis. In this talk, we present an individual-based probabilistic model that captures essential features of the design of the Lenski experiment. We assume that each beneficial mutation increases the individual reproduction rate by a fixed amount, which corresponds to the absence of epistasis in the continuous-time (intraday) part of the model, but leads to an epistatic effect in the discrete-time (interday) part of the model. Using an approximation by near-critical Galton-Watson processes, we prove that under some assumptions on the model parameters which exclude clonal interference, the relative fitness process converges, after suitable rescaling, in the large population limit to a power law function. This ist joint work with Adrian Gonzalez Casanova, Anton Wakolbinger and Linglong Yuan.

Using particle filters to infer past population sizes and migration rates from whole-genome sequences

Gerton Lunter, Joe Zhu, Donna Henderson
Wellcome Trust Centre for Human Genetics, Oxford, UK

The complex interrelationships between genome sequences of individuals in a population are described by the Ancestral Recombination Graph (ARG). This object is of interest as it contains information about past population sizes and migration events. Inferring these parameters is difficult because the ARG is a very complex data structure, and it is only weakly determined by the observations, the genomic sequences in a population.
Inference methods for demographic parameters broadly fall into two categories: ones that approximate the ARG to explicitly marginalize over it, and ones that use summary statistics. Here we describe an inference method that stays very close to the true ARG, by using particle filters to sequentially construct an explicit sample from the posterior distribution of ARGs conditional on the observed sequence data.
We show that the method is able to infer past population sizes from a single diploid genome with a resolution and accuracy comparable to PSMC. Uniquely, it is able to do the same from two diploid genomes, improving the resolution for recent epochs. By taking the two diploid samples from two different populations, explicit migration rates between these populations can be inferred as well.
As a by-product, explicit ARGs are constructed, which can be used to detect loci with unusual genealogies due to e.g. balancing selection or recent selective sweeps, similar to the program ARGweaver. The simulation-based approach would also allow us to deal with more complex models, and in particular we are investigating whether explicitly modeling background selection would improve demographic inference.

Inferring microbial species and strains directly from metagenome data

Christopher Quince
Warwick Medical School, University of Warwick

In metagenome sequencing DNA from an entire microbial community is sequenced typically with short reads. The assemblies produced from these studies are usually highly fragmented comprising hundreds of thousands of partial assemblies or contigs. This is an inevitable consequence of intra- and inter-genomic repeats. Only from very simple communities can complete genomes be assembled. However, determining which contigs derive from which species or strain is almost as useful as a complete genome revealing gene complement. Metagenome analyses often comprise multiple samples from longitudinal analysis of the same community over time or horizontal sampling of multiple similar communities. We exploit this in a method, CONCOCT: Clustering cONtigs on COverage and ComposiTion, that combines sequence composition and coverage across multiple samples to automatically cluster contigs into species genomes. CONCOCT uses a dimensionality reduction coupled to a Gaussian mixture model, fit using a variational Bayesian algorithm, which automatically identifies the optimal number of clusters. We demonstrate high recall and precision rates on artificial as well as real human gut metagenome datasets. We then extend this principle, developing a probabilistic model of variant frequencies across samples on core genes within species clusters. These frequencies depend on the relative abundances of strains in each sample and their haplotype. Using a Gibbs sampling algorithm we can use this model to reconstruct the abundances of the strains and their genotypes on the core genes. These genotypes can then be used to determine the phylogenetic relationships between the strains present. Finally, we can apply this information to all the contigs associated with the species to reconstruct the accessory genomes of the different strains. This provides a methodology for *de novo* extraction of strain genome composition from metagenome analyses that does not rely on long read sequencing.

Cancer Subclonal Reconstruction with Aldous' Beta-Splitting Trees

Yee Whye Teh
University of Oxford

We describe a problem of reconstructing the history and structure of subclonal populations of cancer cells in tumours. We take a Bayesian nonparametric modelling approach to infer both the number of populations as well as their phylogenetic relationships from data obtained from deep read sequencing of tumour biopsies. As prior over the phylogenies, we use Aldous' beta-splitting trees, random binary trees constructed by recursively partitioning a unit interval into two subintervals. Due to the complex discrete-continuous nature of the model, posterior simulation is quite involved and we develop a Markov chain Monte Carlo methodology based on Hamiltonian Monte Carlo, reversible jump, and the Wang-Landau algorithm. Empirically, we demonstrate an improvement over a recent state-of-the-art model.
Joint work with Levi Boyles, Amit Deshwar, and Quaid Morris.

Inferring local ancestry by jointly analyzing admixed samples

Amy Williams
Cornell University

Local ancestry inference identifies the population of origin of DNA at each position in an admixed individual's genome. This information is necessary to perform admixture mapping, an approach to disease and trait association mapping that identifies loci with significant deviations in ancestry proportions among a set of admixed individuals that have the trait of interest. Local ancestry information has also been widely used to study human demographic history, and because rare variants correlate with local ancestry, local ancestry may be useful in addressing fine-scale population stratification in rare variant association studies.

Numerous methods exist for inferring local ancestry, but the accuracy of their inference in Latino and other groups is not ideal for case-only admixture mapping and has potential to be improved. The most accurate local ancestry methods require panels of unadmixed individuals to help build models of the ancestral populations that contribute ancestry to the admixed individuals. A challenge for such methods that rely heavily on panels is that ancestral populations may only exist in admixed form, with related unadmixed populations being drifted from those ancestors. An alternative to requiring panels is to attempt to leverage information from within the admixed samples themselves. The method RFMix performs an iterative approach that, beginning with unadmixed panels, also utilizes admixed individuals to infer local ancestry. RFMix has improved results compared to other approaches, but may not fully capture the latent information within admixed samples because of its initial dependence upon panels for inference in the first iteration.

We introduce MIX-HAPI, a method that extends the haplotype inference framework HAPI-UR to infer local ancestry. The algorithm uses global ancestry estimates as a prior probability for local ancestry and leverages all input samples with any combination of admixed or unadmixed individuals. Our initial evaluation using simulated data shows that MIX-HAPI is extremely accurate in inferring local ancestry in the presence of unadmixed individuals that derive from a different population than the true ancestral group. The approach can be iterated via expectation-maximization in order to better infer local ancestry in regions that are initially uncertain. We are currently evaluating the method on simulated data, including Latinos, with variable numbers and population origins of unadmixed individuals included in the analysis. Because unadmixed Native American groups may be differentially related to the ancestors of Latinos, this method holds promise to improve inference accuracy in real data and has application to other admixed population groups.

5 Poster Abstracts

Relation between the Wright-Fisher diffusion with seed bank and the two island diffusion model

Eugenio Buzzoni
TU Berlin

We investigate the relation between the Wright-Fisher diffusion with two islands investigated by Nath and Griffiths (1993) and later by Kermany et al. (2008) and the Wright-Fisher model with seed bank component introduced by Blath, Eldon, Gonzalez Casanova, Kurt and Wilke Berenguer (2014) by means of duality arguments. In the former, we consider a haploid population which is divided into two islands. Moreover, we assume both mutation and migration between the two islands (Kermany et al., 2008). In the latter model, there is only one population; however, in each generation, while most of the population stems from the previous generation, a few individuals obtain their type from a parent having lived in the (possibly far) past. This allows old genes to be reactivated after some time in a so-called seed bank effect (Blath, Eldon, Gonzalez Casanova and Kurt, 2014). These two models both exhibit striking similarities and qualitative differences. The similarities stem from the fact that the sdes linked to the models are almost the same. The main difference is that, as we can show, the seed bank diffusion can be reformulated as a one-dimensional stochastic delay differential equation. Based on joint work with Prof. Dr. Jochen Blath and Adrian Gonzalez Casanova (both TU Berlin).

Efficient maximum-likelihood inference for the "isolation with initial migration" (IIM) model.

Rui J. Costa
University College London

The "isolation with migration" (IM) model is a common tool to detect gene flow during speciation and model the speciation process in general. Recent papers have questioned the reliability of IM model estimates, especially due to its assumption of constant gene flow until the present. In this paper, we deal with an extension to the IM model, the isolation with initial migration (IIM) model. Our IIM model allows for one parent species, two descendant species, an initial period of (potentially asymmetric) gene flow between the descendant species, and a more recent period of complete isolation. We derive and describe a fast method of fitting this IIM model, or any of its nested models (including the IM model and the complete isolation model) to real data. This is a maximum likelihood method, applicable to observations on the number of segregating sites between pairs of DNA sequences from a large number of independent loci. To derive the likelihood, we define the coalescent process of a pair of sequences and solve it using an eigendecomposition of the generator matrix. In addition to obtaining parameter estimates, our method can also be used to distinguish between alternative models representing different evolutionary scenarios, by means of likelihood ratio tests or AIC scores. We illustrate the procedure on pairs of Drosophila sequences from approximately 30,000 loci. The fitting time for the most complex model version, using this data set, does not exceed a couple of minutes.

An almost infinite sites model

Alejandra Avalos-Pacheco and Mathias C. Cronjäger
University of Oxford

Two of the most used mutation models in population genetics are the infinite sites and finite sites models. In the first model, mutations occur at most once at each site of the sequence, being just an approximation to reality, where one site can mutate more than once. The finite sites model on the other hand imagines mutations as hitting 1 out of a finite number of possible sites. However, for real data with many polymorphic sites this model becomes overly complex and infeasible to implement. We introduce an "Almost infinite sites model", which incorporates ideas from both models. Along with the model itself, we present a generalisation of classic recursions from Griffiths (1989) along with numerical results.

Improving popular population genetic estimates by using shrinkage

Andreas Futschik
JK University Linz, Austria

Two key parameters in population genetic inference are the scaled mutation parameter theta and the scaled recombination rate rho. Estimates of these parameters have been obtained either from suitable summary statistics, or via more sophisticated methods such as (composite) likelihood. We explore how shrinkage can be used to improve the above mentioned population genetic estimates. Based on work by Futschik and Gach (2008), we explain why the popular Ewens-Watterson estimate of theta is inadmissible in terms of the mean squared error under the classical Wright-Fisher model and can be improved uniformly by shrinkage. Also other estimates of  such as the MLE can also be improved, although to a smaller extent. Then we look at estimating the scaled recombination rate rho, and explore possible gains that can be obtained by optimizing the bias/ variance trade-off for the two popular estimates LDhat and LDhelmet. It turns out that there is room for improvement also with these estimates, although the improvement is often achieved by an appropriate bias reduction.
Both for Wattersons estimate and when estimating recombination, the improvement can be achieved without a lot of computational effort. However, as no explicit formulas for bias and variance are available when estimating rho, the coefficients turning up in the resulting formula for the improvement need to be estimated from simulations.

Robust nonparametric Bayesian inference of Lambda-measures from molecular data

Jere Koskela
University of Warwick

Skewed offspring distributions, natural selection and range expansion of populations have all been shown to give rise to genealogies described by a family of coalescent processes known as Lambda-coalescents. The family provides a rich modelling framework, which is able to capture skewed branching of ancestral trees, simultaneous mergers of more than two lineages and associated insensitivity of nucleotide diversity to population size. Such features have been observed recently in genealogies of Atlantic cod and Pacific oysters populations, whale louse mitochondrial DNA as well as influenza viruses under strong selective immune pressure.
In contrast to the universality of the Kingman coalescent, the statistics of Lambda-coalescent trees depend on the details of the skewness in offspring distribution, the strength and mechanism of natural selection or population expansion etc. Hence it is of great interest to ask whether they can be correctly inferred from molecular data. This poster demonstrates that such identifiability fails in general when observations are taken over relatively few generations, such as a single time point, even when many unliked loci are available. A long time-series of molecular data is shown to result in much stronger identifiability under verifiable modelling assumptions. Recent advances in sequencing technology have resulted in time series data becoming available and these results provide a strong motivation for its use, particularly whenever the usual Kingman assumptions are violated.

Partitioning, duality, and linkage disequilibria in the Moran model with recombination

Sebastian Probst, Mareike Esser
University of Bielefeld

We consider a Moran model with recombination, more precisely, a finite haploid population of finite size N evolving under recombination and resampling in continuous time. Forward in time its evolution is described by a continuous time Markov chain, using so called recombinators. Backwards in time, the process is described in terms of a partition valued Markov process. Instead of considering the full (and complicated) multilocus ancestral recombination process, we use a marginalised version, where each locus is observed in only one of the sampled individuals. The transitions of this process are given in terms of splitting and coalescence events, leading to an explicit representation of its generator, which allows to take limits under different scaling regimes.
We prove, that the forward and the backward Markov processes are dual with respect to a sampling function given by the Möbius inverse of a certain generalised recombinator. This also leads to a closed system of ordinary differential equations for the expectations of the sampling functions and gives rise to explicit expressions for the expected decay of certain correlation functions, called linkage disequilibria.

Hierarchical Bayesian stochastic analysis (via a particle filtering) of oscillatory multiple single cell Nrf2 protein levels

Simone Tiberi
University of Warwick

Our work focuses on the hierarchical analysis of multiple single cell Nrf2 reporter levels in nucleus and cytoplasm. Nrf2 is a transcription factor that regulates the expression of several defensive genes protecting against various cellular stresses. Our analysis aims to gain an insight into this essential cellular protective mechanism.
We propose a reaction network based on five reactions, which include a distributed delay and a non-linear term, for the amount of Nrf2 in nucleus and cytoplasm. The diffusion approximation (DA) is used to approximate this Markov jump process with a stochastic differential equation (SDE). Being this continuous process only observed at discrete time points, a second approximation, the Euler-Maruyama approximation (EMA) of the DA, is needed to obtain an approximated likelihood of this bivariate process.
Furthermore, to make use of multiple single cell data, we embed the model in a hierarchical framework. A measurement equation, which involves a proportionality constant and a bivariate error, for the nuclear and cytoplasmic measurements, is necessary to relate the original unobserved population levels, X, to the observations, Y. Unlike most applications, the error term is assumed not to be i.i.d. in time and to follow instead an auto-regressive (AR) model of order 1. A Bayesian analysis is performed, via a particle marginal Metropolis-Hasting (PMMH) sampler, to explore the highly dimensional posterior space which includes a bivariate latent process X for every cell.
We will show the results obtained on simulation studies, proving the validity of the methodology, and on a real data application, composed of 35 single cell fluorescent levels, observed every two minutes, for 3–12 hours.

6 Participant List

Name	Affiliation
Prof Mark Achtman	University of Warwick
Dr Nabil-Fareed Alikhan	University of Warwick
Miss Alejandra Avalos-Pacheco	University of Warwick
Prof Ellen Baake	Faculty of Technology, Bielefeld University
Prof Mark Beaumont	University of Bristol
Prof Matthias Birkner	Johannes-Gutenberg-Universität Mainz
Prof Jochen Blath	TU Berlin
Mr Eugenio Buzzoni	TU Berlin
Mr Rui Costa	University College London
Mr Mathias Christensen Cronjager	University Of Oxford
Dr Maria De Iorio	Dept of Statistical Science, UCL
Mrs Mareike Esser	Faculty of Technology, Bielefeld University
Prof Steven Evans	University of California at Berkeley
Dr Richard Everitt	University of Reading
Dr Andreas Futschik	Dept. of Applied Statistics, JKU Linz
Mr Habib Ganjgahi	University of Warwick
Prof Robert Griffiths	University of Oxford
Dr Kari Heine	UCL
Mr Asger Hobolth	Aarhus University
Miss Watthanan Jatuviriyapornchai	University of Warwick
Dr Paul Jenkins	University of Warwick
Dr Carolin Kosiol	Vetmeduni Vienna
Mr Jere Koskela	University of Warwick
Prof Dr Noemi Kurt	TU Berlin
Dr Krys Latuszynski	University of Warwick
Dr Anthony Lee	University of Warwick
Dr Gerton Lunter	University of Oxford
Prof Simon Myers	University of Oxford
Prof Thomas Nichols	University of Warwick
Mr Valerio Perrone	University of Warwick
Dr Murray Pollock	University of Warwick
Mr Sebastian Probst	Faculty of Technology, Bielefeld University
Dr Christopher Quince	University of Warwick
Dr Mathias Rafler	TU Berlin
Prof Christian Robert	University of Warwick
Dr Franziska Schade	Fraunhofer EMB Lubeck
Ms Maria Simonsen	Bioinformatics Research Centre, Aarhus University
Dr Dario Spano	University of Warwick
Mr Yee Whye Teh	University of Oxford
Mr Simone Tiberi	University of Warwick
Dr Sara Wade	University of Warwick
Prof Amy Williams	Cornell University
Mr Giacomo Zanella	University of Warwick
Dr Zhemin Zhou	University of Warwick
Miss Anyi Zou	University of Warwick

Arden
Delegate Information

warwick conferences

Checking in/out
Rooms will be available after 3.00pm for check in Monday – Friday and after 6.00pm at weekends. Please vacate your bedroom by 10.00am on your day of departure. You should return your key to Reception as soon as possible after vacating your bedroom. Accounts may be settled by personal cheque supported by a banker's card, cash, Maestro, Delta, Visa, Mastercard or American Express.

Car parking
Free car parking is available to the rear of Arden whilst disabled parking spaces are available at the front of the main building. Due to its location, parking at Arden is subject to availability – please contact Reception to discuss alternative car parking arrangements.

Reception
Our Reception staff will be pleased to answer your queries between 7.00am – 11.00pm, on extension 0. A Night Porter is available outside of these hours on the same extension. All internal numbers are free. Should you require a private room to pray, please contact Reception.

Luggage
If you require assistance with your luggage, please ask at Reception. For your convenience we provide a secure area to store luggage near Reception.

Meals
A wide breakfast selection is followed by a delicious carvery lunch and an imaginative three course menu served at dinner. We also cater for special dietary requirements.
Meals are usually served between the following times:
Breakfast 7.00am – 9.30am
Lunch 12.00pm – 1.45pm
Dinner 7.00pm – 9.00pm
Please check with your Course Tutor or Event Organiser for the specific dining times arranged for your group. You are asked to respect our informal but smart dress code and not to use mobile telephones in the Restaurant. If you're planning a special dinner, choose our stylish private room and we'll create a menu exclusively for you. Please contact Reception for further information.

Your room
Our 121 en-suite bedrooms provide a comfortable and relaxing stay. All rooms provide a study desk area, free internet access via a network point, telephone, television, radio alarm clock, tea and coffee making facilities, hairdryer, towels and toiletries and an iron and ironing board. A small number of double bedrooms are available – please ask at Reception for more information.

Coffee lounge / Bar
Arden has two lounges offering unlimited refreshments during the day. With comfortable seating and direct access to outdoor spaces, our lounges are the perfect place to help you to unwind. Our Bar is a welcoming retreat in which to socialise and is open from 12.00pm – 2.00pm and 5.30pm until late.

Business Centre
The PC's in our Business Centre have the latest software and IT applications. You can access email and internet free of charge. There is also a PC enabled database with all the information you may require whilst staying at Arden. You can access the internet free of charge. Print, fax and photocopy services are also available for a small fee.

Arden - Warwick Conferences
The University of Warwick
Kirby Corner Road
Coventry, CV4 8UW

Tel - 024 7652 3904
Fax - 024 7652 2848
Email - arden@warwick.ac.uk
Web - www.warwickconferences.com

Conference
with confidence

www.warwickconferences.com

warwick conferences

Sports facilities
An Exercise Room is available in Arden and guests can also make use of the free sports and swimming facilities in the Sports Centre on the University campus. Details and opening times are available at Reception.

Emergency
In the event of an emergency or accident, please dial 0 on any internal phone for Reception who will summon assistance. Please inform Reception on arrival of any special requirements you may have in the unlikely event of an evacuation (e.g. hearing or walking difficulties).

Getting here
Arden is located at the University of Warwick on the outskirts of Coventry and is easily accessible by road, rail and air.

By road
The University of Warwick is located at the centre of the Midlands Motorway Network. A Location Map with detailed instructions can be obtained from your Event Organiser. Alternatively, you can download the map as a pdf from our website; www.warwickconferences.com.

By rail
Coventry Train Station is on the Main West Coast Line with regular trains connecting with other major towns and cities. Trains run every half hour to London with an approximate journey time of 59 minutes. Services to Birmingham International and Birmingham New Street are frequent throughout the day. There is a large taxi rank at Coventry Station, just four miles from Arden. Alternatively, a regular bus service is available.

By air
Birmingham International Airport is only 20 minutes by taxi. A frequent train service operates to Coventry from Birmingham International Station, adjacent to the airport.

Around and about
From the home of rugby to the home of the greatest English playwright, a visit to Coventry and Warwickshire has something for everyone. With cathedrals, castles, stately homes and gardens, museums and galleries, to the largest Arts Centre outside London, Coventry and the surrounding Warwickshire region offer the best of town and country. If you'd like to know more about what you can do to make your stay even more enjoyable, ask a member of our team, or visit www.visitcoventryandwarwickshire.co.uk.

We strive to continually improve your meeting experience – whether it's the quality of our product or the standard of service. We want to hear your thoughts, opinions and suggestions. Tell us what we do well and what we could do better - please take a moment to complete a feedback form or speak to a member of Arden's team.

We hope you enjoy your stay at Arden and look forward to welcoming you back again soon.

Arden - Warwick Conferences
The University of Warwick
Kirby Corner Road
Coventry, CV4 8UW

Tel - 024 7652 3904
Fax - 024 7652 2848
Email - arden@warwick.ac.uk
Web - www.warwickconferences.com

Appendix B: For guests staying in Arthur Vick

Delegate Joining Instructions
Warwick Conferences' Conference Park

We are delighted that you will be joining us at the University of Warwick. We hope that the information provided in this document will help you get the most from your event. Please bring these instructions with you as you will find them useful whilst you are on campus.

The Conference Park is on the **Central Campus** of the University of Warwick located on the outskirts of Coventry, which is accessible by road, rail and air. You can download further information from the website at www.warwickconferences.com following the link 'how to find us'. A further link can be found for any relevant traffic information at http://www.warwickconferences.com/delegates/delegates-conference-park. The Conference Park is the name given to the facilities provided by Warwick Conferences on the **Central** University campus.

Getting to campus and car parking:

Complimentary car parking is available for conference delegates in any of the **Central** campus multiple shared use car parks (7, 8, 8a, 9, 10 and 15). Please ensure you have printed your parking permit before arriving on **Central** campus and display this in your windscreen before leaving your vehicle (the permit will be sent by your Event Organiser). Disabled parking spaces are available within these allocated car parks and close to some accommodation areas, please inform your Event Organiser if you require a disabled car parking space.
As a University campus, from time to time these car parks become full and when this happens alternative parking will be available, which you will be directed to on the day. We advise that you allow sufficient time, for up to a fifteen minute walk to get to your destination on the Conference Park as car parks are not adjacent to the registration and meeting rooms. The University encourages conference attendees to consider car sharing where possible.

Accommodation:

Please check with your Event Organiser as to which type of accommodation has been reserved for your event and what facilities are available.

Conference Reception:

Located within the Students Union Building. The Reception team are available to answer your queries between 07:00 – 23:00. Here you can also collect information on:

- How to connect to the Wi-Fi around campus
- Ask about any lost property
- Request additional bedroom supplies such as pillows, blankets, clock radio or a bath mat
- Arrange for secure luggage storage

Keys:

You will be provided with a key or key card which will access your room and entry door to the residence. Keys can be left at Conference Reception, Rootes Restaurant (in Rootes Building) or one of the boxes situated in the entrance halls of each residence on the day of your departure.

Bedroom check in/out:

Bedroom keys will be available from 15:00 to 23:00 at Conference Reception. If you plan to arrive after 22.45, please contact Conference Reception to arrange late key collection (wcpreception@warwick.ac.uk). Rooms need to be vacated by 09:30 on your day of departure and all luggage and belongings should be removed by that time. Please inform Conference Reception on arrival, of any difficulties you may have in the unlikely event of an evacuation from your accommodation (e.g. hearing or mobility difficulties).

Disability services:
The University of Warwick aims to be accessible and welcoming to everyone and we are committed to making your visit as easy and enjoyable as possible. If you have any particular requirements that we should be aware of, then please discuss these with your Event Organiser.

Launderette facilities:
The launderette is situated between Rootes Building and Rootes residences, opening times are available from Conference Reception for self-service washing and drying. All machines and dryers are operated using a card payment system, the cards can be purchased from the launderette at a cost of £10 (£2 for the card and £8 of available spend for use on machines or dryers). You will need to provide your own washing powder and fabric softener.

Wi-Fi access across campus:
To connect your device to the *Warwick Guest* Wi-Fi network.

- If the Warwick Guest gateway page does not open automatically then open a browser such as Explorer, Chrome or Safari and you should be redirected
- If you are not yet registered use *Click Here* to create a new account and then *Attending a Conference*
- Fill in your details and you will be sent a password by text
- If you are unable to use your mobile phone then at the bottom of the page use *Click Here* to register and the details will be sent to your email address. Please ask a member of staff if you need access to a computer to retrieve your login details
- Return to the main login page and enter your details. You may use up to three devices simultaneously and your account will be valid for one week
- If you still cannot get to the login page try typing the following address into your browser: *cppm.csv.warwick.ac.uk/guest/guest_login_uow.php* (Note that if you use this method you will have to manually login this way each time)
- If you lose your password you can re-register at any time with the same phone number and email address. Your previous password will be overwritten
- Your devices will be logged out after 24 hours of inactivity and you will be required to login again

Food and Drink:
All meals are provided in Rootes Restaurant located on the first floor of Rootes Building for all delegates (unless your programme indicates otherwise). The restaurant offers an assisted style service of breakfast, lunch and dinner including a range of hot and cold drinks. Your Event Organiser will be able to advise you regarding the specific arrangements for your event. Please have with you your conference badge or room key to gain access to the restaurant. If you have any special dietary requirements then please inform your Event Organiser.

The bar is located on the first floor of Rootes Building and is the ideal place to network and relax after a day's session. There are also alternative bars in Warwick Arts Centre and Students Union building (check opening times locally)
Payment for all sundry items is by cash or credit card payment only.

Shops, Banks, Cafés and Bars on Central campus:

The campus has many facilities available to all delegates, for all information and opening times please see the website: http://www.warwickretail.com . Warwick Arts Centre cinema offers discounted cinema ticket prices, these can be purchased from the box office and proof of delegate status is required (not applicable for Met Opera Live or NT Live screening).

Sports facilities:

Your room key gives you free unlimited access to Warwick Sport's premium leisure facilities on campus. The opening hours are 07:00 to 21:30 (Monday to Friday) and 08:15 to 19:30 (Saturday and Sunday) with free use of the following with no booking required:

- 25 metre swimming pool & sauna
- State of the art gym
- Running track

The following facilities are also free of charge, but need to be booked in advance:

- Badminton courts
- Squash courts

To make a booking please call the Sports Centre Reception on **23011** on any internal phone or **024 76523011** externally or email warwicksport@warwick.ac.uk

Other sports facilities that can be used but at **additional charge** are:

- Tennis centre with 4 indoor and 4 outdoor courts
- Climbing centre with 14 metre high walls and a bouldering room
- Sports halls for basketball, 5-a-side football, netball or volleyball
- Outdoor Astro or grass pitches for football, rugby or cricket

You can take the towel from your bedroom or alternatively you can hire a towel from the Sports Centre at a cost of £2. You need to bring your room key with you to access any of the facilities and also a £1 coin if you require a locker, which is refunded after use.

For further information such as equipment hire fees and gym induction video, please visit: www.warwick.ac.uk/sport

For more information:

You can also refer to our Frequently Asked Questions document (FAQ's) which can be obtained from our website: http://www.warwickconferences.com/delegates/delegates-conference-park

Appendix C

Opening times

Summer vacation: Saturday 27 June – Saturday 3 October 2015

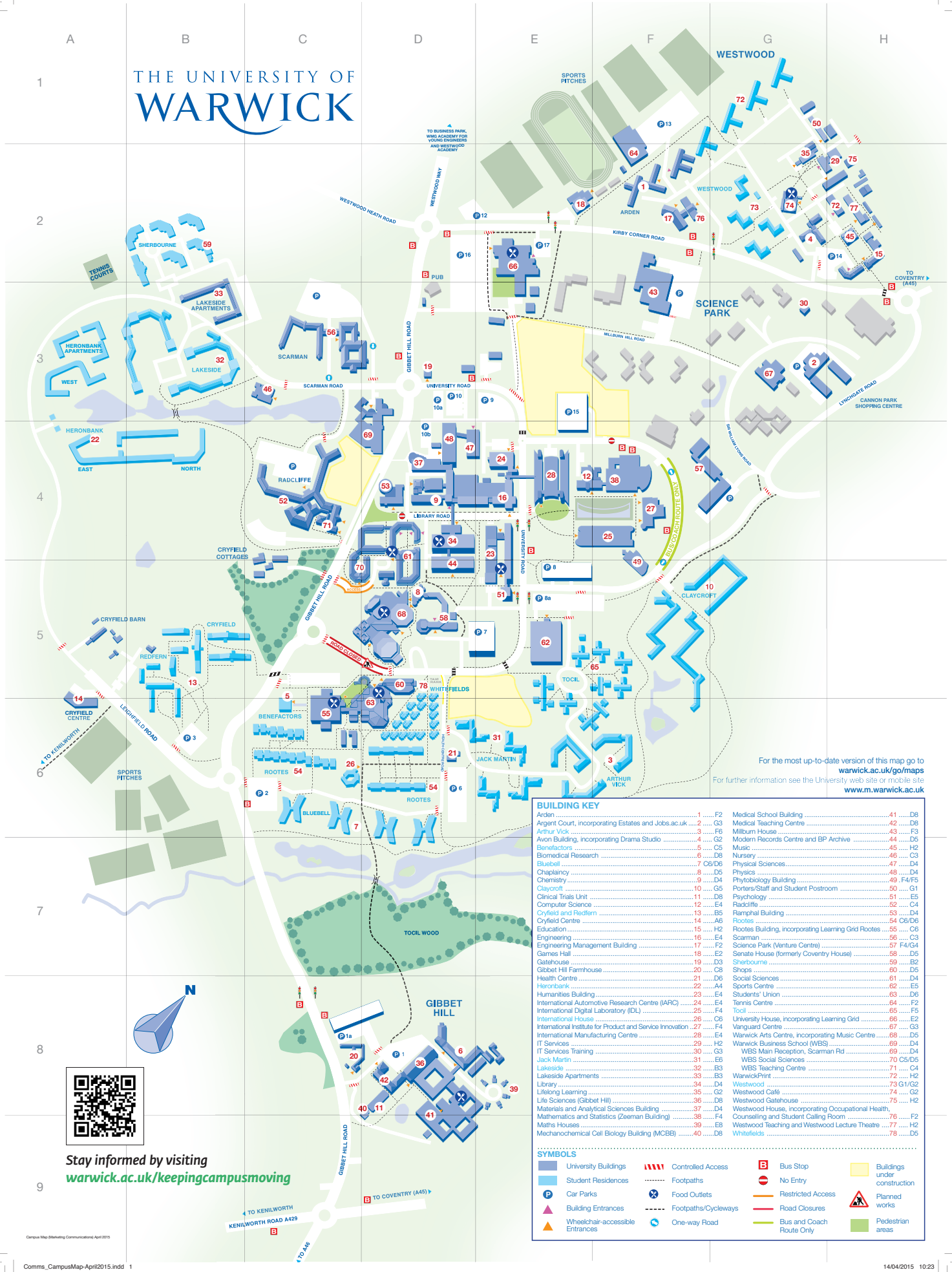
Café Library	Mon – Fri	8am – 3pm		
Café Library extension	Mon – Fri	8am – 6pm	Sat – Sun	10am – 4pm
Library Coffee Bar	Mon – Fri	8.30am – 3.30pm		
Café Social (closed August)	Mon – Fri	9am – 2pm		
Café Humanities	Mon – Fri	8.30am – 3pm		
Café Gibbet Hill	Mon – Fri	8am – 4.30pm		
Café Bar	Mon – Sat	8am – 9pm	Sun	3pm – 8.30pm
Theatre Bar	*Dependent upon performances*			
University House restaurant	Mon – Fri	8am – 2.30pm		
University House atrium coffee bar	Mon – Fri	8am – 5pm		
Fusion	Mon – Sun	12noon – 3pm		
The bar	Mon – Sun	12noon – 10pm		
Le Gusta Oven & Bar	Mon – Fri	12noon – 3pm	Sat	4pm – 10pm*
		6pm – 10pm*	*Sun 19 July only*	4pm – 10pm*
H-van (behind WMG building)	Mon – Fri	9am – 3pm		
Costa	Mon – Fri	8am – 5pm	Sat – Sun	9am – 4pm
Café Westwood	Mon – Fri	8am – 3pm		
Warwick Business School	*New café opening September 2015*			
The Bookshop	Mon – Fri	9am – 5.30pm		
Rootes grocery store	Mon – Fri	8am – 8pm	Sat	9am – 8pm
			Sun	11am – 5pm

* Last food orders 9pm. In exceptional circumstances management reserve the right to change opening times without prior notice.

Complimentary parking permit for for Conference Park residents.

Cut and use the following permits in car parks (7, 8, 8a, 9, 10 and 15). Display the permit in your windscreen before leaving your vehicle.


2015 Conference Parking Permit
Keeping Campus Moving
Expires: Sunday
20 Sep
THE UNIVERSITY OF
WARWICK



For the most up-to-date version of this map go to
warwick.ac.uk/go/maps
For further information see the University web site or mobile site
www.m.warwick.ac.uk

BUILDING KEY

Arden	1 F2	Medical School Building	41 D8
Argent Court, incorporating Estates and Jobs.ac.uk	2 G3	Medical Teaching Centre	42 D8
Arthur Vick	3 F6	Millburn House	43 F3
Avon Building, incorporating Drama Studio	4 G2	Modern Records Centre and BP Archive	44 D5
Benefactors	5 C5	Music	45 H2
Biomedical Research	6 D8	Nursery	46 C3
Bluebell	7 C6/D6	Physical Sciences	47 D4
Chaplaincy	8 D5	Physics	48 D4
Chemistry	9 D4	Phytobiology Building	49 F4/F5
Claycroft	10 G5	Porters/Staff and Student Postroom	50 G1
Clinical Trials Unit	11 D8	Psychology	51 E5
Computer Science	12 E4	Radcliffe	52 C4
Cryfield and Redfern	13 B5	Ramphal Building	53 D4
Cryfield Centre	14 A6	Rootes	54 C6/D6
Education	15 H2	Rootes Building, incorporating Learning Grid Rootes	55 C6
Engineering	16 E4	Scarman	56 C3
Engineering Management Building	17 F2	Science Park (Venture Centre)	57 F4/G4
Games Hall	18 E2	Senate House (formerly Coventry House)	58 D5
Gatehouse	19 D3	Sherbourne	59 B2
Gibbet Hill Farmhouse	20 C8	Shops	60 D5
Health Centre	21 D6	Social Sciences	61 D4
Heronbank	22 A4	Sports Centre	62 E5
Humanities Building	23 E4	Students' Union	63 D6
International Automotive Research Centre (IARC)	24 E4	Tennis Centre	64 F2
International Digital Laboratory (IDL)	25 F4	Tocil	65 F5
International House	26 C6	University House, incorporating Learning Grid	66 E2
International Institute for Product and Service Innovation	27 F4	Vanguard Centre	67 G3
International Manufacturing Centre	28 E4	Warwick Arts Centre, incorporating Music Centre	68 D5
IT Services	29 H2	Warwick Business School (WBS)	69 D4
IT Services Training	30 G3	WBS Main Reception, Scarman Rd	69 D4
Jack Martin	31 E6	WBS Social Sciences	70 C5/D5
Lakeside	32 B3	WBS Teaching Centre	71 C4
Lakeside Apartments	33 B3	WarwickPrint	72 H2
Library	34 D4	Westwood	73 G1/G2
Lifelong Learning	35 G2	Westwood Café	74 G2
Life Sciences (Gibbet Hill)	36 D8	Westwood Gatehouse	75 H2
Materials and Analytical Sciences Building	37 D4	Westwood House, incorporating Occupational Health, Counselling and Student Calling Room	76 F2
Mathematics and Statistics (Zeeman Building)	38 F4	Westwood Teaching and Westwood Lecture Theatre	77 H2
Maths Houses	39 E8	Whitefields	78 D5
Mechanochemical Cell Biology Building (MCBB)	40 D8		

SYMBOLS

- University Buildings
- Student Residences
- Car Parks
- Building Entrances
- Wheelchair-accessible Entrances
- Controlled Access
- Footpaths
- Food Outlets
- Footpaths/Cycleways
- One-way Road
- Bus Stop
- No Entry
- Restricted Access
- Road Closures
- Bus and Coach Route Only
- Buildings under construction
- Planned works
- Pedestrian areas