



International Conference on

Stochastic Models in Ecology, Evolution and Genetics

9-13 December 2013, Angers, France



Program Booklet



Foreword

The organizing committee is pleased to welcome you to Angers for the international conference *Stochastic Models in Ecology, Evolution and Genetics*.

The conference is organized by a group of mathematicians and biologists from the University of Angers, Agrocampus Ouest and INRA, who are involved in MODEMAVE research project on mathematical modelling applied to vegetal biology.

The conference aims at bringing together both mathematicians and biologists who share interests in theoretical or applied aspects of ecology, evolution and genetic modelling, encouraging them to discuss the recent methodological and conceptual advances in these fields and to share their knowledge with young researchers. Throughout the week, plenary lectures, talks and poster sessions will highlight recent advances in various fields of research which are currently studied.

We also hope this conference will be an opportunity for you to visit Angers and its area. The cultural and historical heritage of the region Anjou is very rich and we encourage you to make the most of it during the week.

The organization of the conference is much indebted for the financial support of Université d'Angers, Région des Pays de la Loire, Laboratoire Angevin de Recherche en Mathématiques, Fédération Mathématiques des Pays de Loire, Centre de Mathématiques Henri Lebesgue, Angers Loire Métropole.

Finally, we are very grateful to all of you for attending and making the conference worthwhile by your contributions.

Angers, December 2013
The organizing committee

Schedule

Monday, December 9th

8:30 – 8:45 *Welcome*

8:45 – 9:40 **Hans Metz**

The interplay of infectivity that decreases with virulence and limited cross-immunity: (toy) models for respiratory disease evolution

9:45 – 10:25 **Frédéric Hamelin**

Allee effects and the evolution of polymorphism in cyclical parthenogens

10:30 – 11:00 *Coffee break*

11:00 – 11:40 **Nadav Shnerb**

Spatio-temporal heterogeneity and unstable coexistence

11:45 – 12:25 **Christian Tomasetti**

Stochastic modeling of the accumulation of passenger and driver mutations in cancer

12:30 – 14:00 *Lunch*

14:00 – 14:55 **Linda Allen**

Thresholds for disease extinction in stochastic and deterministic epidemic models

15:00 – 15:40 **Nicolas Champagnat**

Coalescent point process and applications to the size of large families in general branching processes

15:45 – 16:25 **Pierre Vallois**

A stochastic model of tumor growth based on branching processes

16:30 – 17:00 *Coffee break*

17:00 – 17:40 **Nathalie Krell**

Statistical estimation of a growth-fragmentation model observed on a genealogical tree

17:45 – 18:25 **Luca Dall'Asta**

Balancing selection in subdivided populations

Tuesday, December 10th

8:45 – 9:40 **Sylvie Méléard**

Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks

9:45 – 10:25 **Sylvain Billiard**

The interaction between genetics and demography causes population extinction

10:30 – 11:00 *Coffee break*

11:00 – 11:40 **Nicolas Bacaër**

Some population models in periodic or random environments

11:45 – 12:25 **Camille Coron**

Slow-fast stochastic dynamics and quasi-stationary behavior of a diploid population

12:30 – 14:00 *Lunch*

14:00 – 14:55 **Terry Speed**

Removing unwanted variation from high-throughput omic data

15:00 – 15:40 **Alain Hauser**

Causal inference from interventional data

15:45 – 16:25 **Solenn Stoeckel**

Integrating knowledge on genomic structure and ancestral diversity to detect outlying genomic regions

16:30 – 17:00 *Coffee break*

17:00 – 17:40 **Vincent Bansaye**

Letting Markov chains evolve along genealogies

17:45 – 18:25 **Małgorzata Bogdan**

Model selection approach for genome wide association studies in admixed populations

18:30 – 19:30 *Visit of Jean Lurçat Museum*

Wednesday, December 11th

8:45 – 9:40 **Amaury Lambert**

Coalescent trees of generalized birth-death processes

9:45 – 10:25 **Jean-François Delmas**

A population model with non-neutral mutations using branching processes with immigration

10:30 – 11:00 *Coffee break*

11:00 – 11:40 **Charline Smadi**

Genetic hitchhiking and eco-evolution

11:45 – 12:25 **Julien Berestycki**

The number of (selectively) accessible evolutionary paths in the House of cards model

12:30 – 14:00 *Lunch*

14:00 – 15:00 *Visit of Bon Pasteur convent*

16:00 – 18:00 *Visit of Angers medieval town*

19:30 – 20:30 *Conference dinner*

Thursday, December 12th

8:45 – 9:40 **Elizabeth Thompson**

Models for the inference and use of identity-by-descent in populations

9:45 – 10:25 **Simon Boitard**

Inferring the past dynamics of effective population size using genome wide molecular data

10:30 – 11:00 *Coffee break*

11:00 – 11:40 **Mikael Falconnet**

Modelling DNA sequence evolution with interacting particle systems

11:45 – 12:25 **David Kessler**

Family size statistics

12:30 – 14:00 *Lunch*

14:00 – 14:55 **Étienne Pardoux**

Lambda-coalescent and look-down model with selection

15:00 – 15:40 **Olivier Mazet**

Why does the effective size change in population history ?

15:45 – 16:25 **Dan Goreac**

Linearization methods in the control of PDMP associated to gene networks

16:30 – 17:00 *Coffee break*

17:00 – 17:40 **Yves Dumont**

On Impulse Fire events in a Tree-Grass interactions model in Savanna ecosystem

17:45 – 18:25 **Safia Slimani**

Stochastic stability for predator-prey model of Holling type II with term refuge

Friday, December 13th

8:45 – 9:40 **Bahram Houchmandzadeh**

Selection for altruism through random drift in variable size populations

9:45 – 10:25 **Beata Hat-Plewinska**

Cell fate decision control in the stochastic model of p53 regulatory pathway

10:30 – 11:00 *Coffee break*

11:00 – 11:40 **Carole Knibbe**

Genome size evolution: challenging intuition with modelling

11:45 – 12:25 **Richard Pymar**

Localisation in the Bouchaud-Anderson model

12:30 – 14:00 *Lunch*

Plenary lectures

Thresholds for disease extinction in stochastic and deterministic epidemic models

Linda Allen

(Texas Tech University)

Thresholds for disease extinction provide essential information for control, eradication or management of diseases. The next generation matrix approach yields a formula for the basic reproduction number $\mathcal{R}(0)$, a threshold for disease extinction in deterministic epidemic models. Similarly, a multitype branching process yields an estimate for probability of disease extinction in stochastic epidemic models. Relations between deterministic and stochastic models show that these thresholds are in agreement for discrete-time and continuous-time epidemic models with multiple infectious groups. These relations are illustrated, analytically and numerically, for several different epidemic models, including a vector-host model applied to West Nile virus in mosquitoes and birds, a stage-structured epidemic model and a spatially discrete epidemic model where risk of infection varies with location.

Coalescent trees of generalized birth-death processes

Amaury Lambert

(UPMC)

Birth-death processes produce random trees that can be represented as growing "vertically" as time goes by. Pruning the extinct lineages out of such trees leads to natural models for coalescent trees, or reconstructed trees (i.e., the tree spanned by points at the same distance from the root point).

Alternatively, reconstructed trees can be modelled by coalescent point processes (CPP), where trees grow "horizontally" by the sequential addition of vertical edges. Each new edge starts at some random branching time and ends at the present time; branching times are drawn from the same distribution independently.

CPP lead to extremely fast computation of tree likelihoods and simulation of reconstructed trees. Their topology always follows the uniform distribution on ranked tree shapes (URT).

We characterize which forward-time models lead to URT reconstructed trees and among these, which lead to CPP reconstructed trees. In particular, we show that for any "asymmetric" diversification model in which birth rates only depend on time and death rates only depend on time and on a non-heritable trait (e.g., age), the reconstructed tree is CPP, even under incomplete sampling of tips. If rates additionally depend on the number of species, the reconstructed tree is (only) URT (but not CPP).

Joint work with Tanja Stadler.

*Stochastic dynamics of adaptive trait and neutral marker driven
by eco-evolutionary feedbacks*

Sylvie Méléard

(École Polytechnique)

This talk presents a work in progress with Sylvain Billiard, Regis Ferriere and Chi Viet Tran. How the neutral diversity is affected by selection and adaptation is investigated in an eco-evolutionary framework. In our model, we study a finite population in continuous time, where each individual is characterized by a trait under selection and a completely linked neutral marker. The dynamics is ruled by births and deaths, mutations at birth and competition between individuals. The ecological phenomena depend only on the trait values but we expect that these effects influence the generation and maintenance of neutral variation. Considering a large population limit with rare mutations, but where the marker mutates faster than the trait, we prove the convergence of our stochastic individual-based process to a new measure-valued diffusive process with jumps that we call Substitution Fleming-Viot Process. This process restricted to the trait space is the Trait Substitution Sequence introduced by Metz et al. (1996). During the invasion of a favorable mutation, the marker associated with this favorable mutant is hitchhiked, creating a genetical bottleneck. The hitchhiking effect and how the neutral diversity is restored afterwards are studied. We show that the marker distribution is approximated by a Fleming-Viot distribution between two trait substitutions and that time-scale separation phenomena occur. The SFVP has important and relevant implications that are discussed and illustrated by simulations. We especially show that after a selective sweep, the neutral diversity restoration depend on mutations, ecological parameters and trait values.

The interplay of infectivity that decreases with virulence and limited cross-immunity: (toy) models for respiratory disease evolution

Hans Metz

(Leiden University)

Models for the evolution of virulence traditionally assume a trade-off between the inverse of disease-induced mortality rate and infectivity, resulting in intermediate virulence. The underlying intuition is that faster growing agent populations do both more damage and produce more infective particles. This intuition implicitly assumes a well-mixed host body. In reality both damage and infectivity depend mainly on the location in the body where the agents lodge. This is related i.a. to the surface proteins that allow agents to dock on and penetrate into different cell types. The typical example is respiratory diseases where more deeply seated ones are both less infective and more harmful. With the other standard assumption, full cross-immunity between disease strains, this would lead to evolution towards the tip of the nose. In reality cross-immunity depends on surface antigens and hence is at least in part connected to depth. In this talk I discuss a simple adaptive dynamics style model taking on board the aforementioned considerations. In doing so I will also shortly review salient aspects of the adaptive dynamics toolbox. Inferences are that in respiratory diseases (1) higher host population densities are conducive to a higher diversity, (2) diversity should be higher in the upper air passages than lower in the lungs, (3) emerging diseases will usually combine a high virulence with a low infectivity.

Lambda-coalescent and look-down model with selection

Étienne Pardoux

(Université d'Aix-Marseille)

We study the lookdown model with selection in the case of a population containing two types of individuals, with a reproduction model which is dual to the Λ -coalescent. In particular we formulate the infinite population " Λ -lookdown model with selection". When the measure Λ gives no mass to 0, we show that the proportion of one of the two types converges, as the population size N tends to infinity, towards the solution to a stochastic differential equation driven by a Poisson point process. We show that one of the two types fixates in finite time if and only if the Λ -coalescent comes down from infinity. In the case of no fixation, we discuss whether or not the non selective type goes extinct with probability one as time goes to infinity.

This is joint work with Boubacar Bah.

Removing unwanted variation from high-throughput omic data

Terry Speed

(University of California)

Over the last few years, many microarray-based gene expression studies involving a large number of samples have been carried out, with the hope of understanding, predicting or discovering factors of interest such as prognosis or the subtypes of a cancer. The same applies to proteomic and metabolomic data, and to other kinds of data. Such large studies are often carried out over several years, and involve several hospitals or research centers. Unwanted variation (UV) can arise from technical elements such as batches, different platforms or laboratories, or from biological signals such as heterogeneity in ages or different ethnic groups which are unrelated to the factor of interest in the study. They can easily lead to spurious conclusions. For example, when doing clustering to identify new subgroups of the disease, one might identify one of the UV factors if its effect on gene expression is stronger than the subgroup effect. Note that similar problems arise when the objective is to combine several smaller studies. A very important objective is therefore to remove these UV factors without losing the factors of interest. The problem can be more or less difficult depending on what is actually observed and what is not. For example, when doing differential expression studies or supervised learning when the factor of interest is known and all the UV factors (say technical batches or different studies) are also known, the problem essentially boils down to a regression, and methods such as *Combat* generally give good results. When the UV factors are modeled as unknown, the problem becomes more difficult because one has to estimate UV factors along with their effects on the genes, and several estimates may explain the data equally well while leading to very different conclusions. This is partially addressed by methods like *SVA*. When neither the factors of interest nor the UV are observed, the problem is even more difficult. It can occur if one is interested in any kind of unsupervised analysis like clustering, or if one simply wants to "clean" a large dataset from its UV without knowing in advance what factors of interest will be studied. Some authors use

SVD on the expression matrix to identify the UV factors. This approach may work well in some cases but relies on the strong assumption that all UV factors explain more variance than any factor of interest. Furthermore it will fail if the UV factors are too correlated with the factor of interest. Recently, we proposed a general framework to remove UV (called *RUV*) in microarray data using *control* genes. It showed very good behavior for differential expression analysis (i.e., with a known factor of interest) when applied to several datasets, in particular better performance than state of the art methods such as *Combat* or *SVA*. This suggests that controls can indeed be used to estimate and efficiently remove sources of unwanted variation. Our objective here is to describe ways of doing similar things when carrying our supervised and unsupervised learning. We propose methods exploiting the existence of replicate arrays. The methods are illustrated on a gene expression microarray dataset, an RNA-seq dataset, some metabolomic data and some methylation microarray data.

Joint work with Johann Gagnon-Bartsch and Laurent Jacob.

*Models for the inference and use of identity-by-descent in
populations*

Elizabeth Thompson

(University of Washington)

Identical by descent (IBD) is the term used to describe segments of DNA that descend from a single ancestral haploid genome to current individuals. Since DNA that is IBD has high probability of being of the same allelic type, IBD provides a basis for modeling similarities among relatives both at the population and pedigree level. Traditionally, IBD has been considered in the pedigree context, where the processes of meiosis and recombination, together with the defined pedigree structure, provide probabilities of IBD. Modern genetic data provide opportunities to infer and use IBD among individuals in a population in which pedigree relationships are unknown.

I will first develop population-based models and methods for inferring IBD among multiple haplotypes. I will then show how IBD can be used in the presence of allelic heterogeneity to detect causal genome regions in a case-control study. Finally, I will discuss current developments and future directions for the use of IBD in the analysis of trait data on members of a population.

This work is joint with Chaozhi Zheng, Chris Glazner, and Sharon Browning.

Talks

Some population models in periodic or random environments

Nicolas Bacaër
(IRD Bondy)

This presentation is about some recent work concerning multi-type birth-and-death processes in periodic or random environments.

Letting Markov chains evolve along genealogies

Vincent Bansaye
(École Polytechnique)

We consider Markov chains indexed by genealogies to model the evolution of a trait in a population whose size tends to infinity. We want to take into account non-homogeneity with respect to time, both in the genealogy and in the evolution of the trait. We also make the genealogy depend on the evolution of the traits of the individuals. We describe the asymptotic behavior of the number of individuals with a given trait. We pay a particular attention to branching genealogies and tackle the problem of local densities and extremal traits. The results rely on the study of the typical ancestral lineages, in the vein of spine decompositions.

The number of (selectively) accessible evolutionary paths in the House of cards model

Julien Berestycki
(UPMC)

Motivated by an evolutionary biology question, we study the following problem: we consider the hypercube $\{0, 1\}^L$ where each node carries an independent random variable uniformly distributed on $[0, 1]$, except $(1, 1, \dots, 1)$ which carries the value 1 and $(0, 0, \dots, 0)$ which carries the value $x \in [0, 1]$. We study the number θ of paths from the root $(0, 0, \dots, 0)$ to the opposite corner $(1, 1, \dots, 1)$ along

which the values on the nodes form an increasing sequence. We show that if the value on the root is set to $x = X/L$ then θ/L converges in law as L goes to infinity to $\exp(-X)$ times the product of two standard independent exponential variables.

As a first step in the analysis we study the same question when the graph is that of a tree where the root has arity L , each node at level 1 has arity $L - 1, \dots$, and the nodes at level $L - 1$ have only one offspring which are the leaves of the tree (all the leaves are assigned the value 1, the root the value $x \in [0, 1]$).

*The interaction between genetics and demography causes
population extinction*

Sylvain Billiard
(Université Lille 1)

The extinction of populations and species is a frequent event in Nature. The underlying processes leading to extinction are numerous and the interaction can be complex, but disentangling their relative effects is important, especially today, regarding the modern crisis of biodiversity. Here, we will discuss some recent developments and questions about the interaction between genetic and demographic processes and how it can affect the survival of population. We will especially show that even large population can go extinct because of this interaction.

*Model selection approach for genome wide association studies in
admixed populations*

Małgorzata Bogdan
(Wrocław University)

Genome Wide Association Studies (GWAS) are used to identify regions of the genome hosting genes influencing traits of interest. In such studies scientists test a large number of genetic markers for the association between their genotypes and a given trait. This creates a huge multiple testing problem and results in a relatively low power of detection of influential genes. In admixed

populations, which originate from a recent interbreeding between two previously isolated populations, one can locate influential genes by using admixture mapping, where the information on the genotypes of genetic markers is replaced with the information on the ancestry of a given region of the genome. Due to the strong correlation between ancestry states in the neighboring loci, the multiple testing correction for the admixture mapping can be substantially less stringent than in case of GWAS. This advantage is however counterbalanced by the non-perfect correlation between the genotype and the ancestry state. To utilize the strength of both approaches, some methods for genome wide association studies, which combine the information on the genotypes and admixture were recently proposed. These methods rely mainly on single marker tests. In this talk we will show that this idea can be applied in the context of model selection approach to GWAS. We will present an extension of the modified Bayesian Information Criterion, which works with the design matrix including the dummy variables both for the genotypes and the ancestry. Our simulation studies show that including the ancestry variables helps to detect influential genes in the regions of a low linkage disequilibrium without compromising the power of detection of other genes.

This is a joint work with F. Frommlet, P. Szulc and H. Tang.

*Inferring the past dynamics of effective population size using
genome wide molecular data*

Simon Boitard
(INRA Jouy-en-Josas)

Inferring the effective size of a given population, and its eventual expansions or reductions in the past, from genetic data, is a long standing question in population genetics. Due to the complexity and the high dimension of the mathematical models that are used in this context, exact inference is impossible and the most popular inference methods are based on numerical approaches as Markov Chain Monte Carlo, Importance Sampling or Approximate Bayesian Computation (ABC). Until recently, these methods were designed for data sets including a small number of independent markers or non recombining DNA sequences. However, the spectacular progress of genotyping and sequencing technologies during the last decade has enabled the production of high density genome wide data in many species, so new statistical methods are needed to take benefit of this new type of data. In this study we present an

ABC approach for inferring the past effective size of a single population. This approach is based on coalescent simulations and on the use of a large number of summary statistics related to allele frequencies and linkage disequilibrium. We illustrate the performance of this approach using cross validation. We compare different ABC strategies and discuss the influence of the different summary statistics. We finally apply this method to a set of 25 bovine sequences from the Holstein breed and compare our results with those obtained by the Pairwise Sequentially Markovian Coalescent approach of Li and Durbin (2011).

Coalescent point process and applications to the size of large families in general branching processes

Nicolas Champagnat
(INRIA Nancy)

We consider general branching processes with i.i.d. life lengths with arbitrary distribution and birth occurring in a Poissonian manner. The corresponding genealogical trees are called splitting trees and can be characterized by a contour process with jumps which is a Lévy process without negative jumps. This contour process allows to describe the genealogy of a population at a given time with a so-called coalescent point process (A. Lambert, Ann. Probab. 2010). In this talk, we consider a supercritical population with Poissonian mutations within the infinite allele framework. The population at time t is partitioned into several "families" with different alleles. We study the size of the largest families when t goes to infinity, depending whether the clonal process is subcritical, critical or supercritical. In particular, we are able to prove the convergence of the conveniently scaled point process of largest families.

This is joint work with Amaury Lambert.

Slow-fast stochastic dynamics and quasi-stationary behavior of a diploid population

Camille Coron
(Université Paris-Sud)

We consider a diploid population of hermaphroditic individuals characterized by their genotype at one bi-allelic locus. The population is modeled by a 3-types birth-and-death process whose birth rates model Mendelian reproduction. Under a large population approximation we obtain the convergence of the population toward a slow-fast stochastic dynamics. In particular, the population converges rapidly toward Hardy-Weinberg equilibrium, while the population size and proportion of a given allele converge toward a 2-dimensional diffusion. After an appropriate change of variable, this diffusion can be written as

$$dS_t = dW_t - \nabla Q(S_t) dt,$$

where W is a 2-dimensional Brownian motion and Q is a known real-valued function of two variables. We finally study the quasi-stationary behavior of this latter diffusion $(S_t, t \geq 0)$, and in particular we consider the long-time coexistence of the two alleles in the population conditioned on non-extinction.

Balancing selection in subdivided populations

Luca Dall'Asta
(Politecnico di Torino)

Balancing selection is a major force for the maintenance of polymorphism and biodiversity. Focusing on simple di-allelic models, and using non-rigorous methods (diffusion approximation) and numerical simulations, I will discuss how fixation/coexistence properties are affected by balancing selection in subdivided populations. In particular I will show that in mean-field structures, such as finite island models, the mean fixation time is a non-monotonic function of migration rate, provided that selection is sufficiently strong. In spatial meta-populations (e.g. 1-d stepping-stone models) I will show the existence of

fixation/coexistence transitions and discuss the non-equilibrium properties of the two regimes.

A population model with non-neutral mutations using branching processes with immigration

Jean-François Delmas
(École des Ponts ParisTech)

We consider a stationary continuous model of random size population with non-neutral mutations using a continuous state branching process with non-homogeneous immigration. We assume the type (or mutation) of the immigrants is random given by a constant mutation rate measure. We determine some genealogical properties of this process such as: distribution of the time to the most recent common ancestor (MRCA), bottleneck effect at the time to the MRCA (which might be drastic for some mutation rate measures), favorable type for the MRCA, asymptotics of the number of ancestors. This is a joint work with H. Bi.

On impulse fire events in a tree-grass interactions model in savanna ecosystem

Yves Dumont
(CIRAD)

Savanna is a grassland ecosystem characterized by various trees density. Since decades, this ecosystem has been extensively studied to understand its longterm and spatial evolution under several environmental factors, like climate changes, rainfall, browsers... and big fire events. To this end various theoretical models have been developed and discussed the last forty years. In this talk, we mainly focus on the modelling of fire events. Until now, in most of the models, fire events have been taken into account continuously. This is not really realistic. We propose to consider fire as discrete events that lead to a system of impulse differential equations. The model is mathematically well posed and a qualitative analysis shows that it derives richer longterm dynamics than the

related continuous fires model. In addition, we will introduce some randomness in the fire events. Finally, using appropriate numerical methods, we will illustrate the talk with various numerical simulations.

Modelling DNA sequence evolution with interacting particle systems

Mikael Falconnet
(Université d'Évry)

To study the changes in DNA sequences, one usually deals with the finite product of i.i.d. continuous-time Markov chains modelling single site nucleotide substitutions. A consequence of the independence is that in a long DNA sequence at equilibrium the frequency of a dinucleotide xy should be the product of the x and y frequencies. But this is actually not the case in some biological contexts. Indeed, it is well known that the dinucleotide CpG is less frequently present in many mammals DNA than it would be expected from base composition. This phenomenon is related to DNA methylation: the substitution rate of cytosine by thymine is higher in methylated CpG's than in other dinucleotides. Therefore, more realistic substitution models incorporating such neighboring effects have been introduced. But substitutions are not the only way to alter DNA sequences. For example, one may add several extra nucleotides to a DNA sequence by insertions, or remove them by deletions. We consider Markov processes defined on the integer lattice, that allow for substitution according to a Markovian kernel depending of the neighborhood and also for a single "cut-and-paste" mechanism, and we provide sufficient conditions for the process to be ergodic. Joint work with N. Gantert and E. Saada

Linearization methods in the control of PDMP associated to gene networks

Dan Goreac
(Université Paris-Est Marne-la-Vallée)

The aim of this talk is to present some applications of linear programming methods in the control of piecewise deterministic Markov processes associated

to stochastic gene networks. Using viscosity solutions, we will show how classical (nonlinear) control problems can be interpreted in connection to a linear programming problem. This is very useful in the study of asymptotic properties: Zubov's method for stability and control problems with long-run average costs. The theoretical aspects are applied to Cook's model for haploinsufficiency and Hasty's model for bacteriophage lambda.

Allee effects and the evolution of polymorphism in cyclical parthenogens

Frédéric Hamelin
(Agrocampus Ouest)

Cyclic parthenogens alternate asexual reproduction with periodic episodes of sexual reproduction. Sexually produced free-living forms are often their only way to survive unfavorable periods. When sexual reproduction requires the mating of two self-incompatible individuals, mating limitation may generate an Allee effect, which makes small populations particularly vulnerable to extinction; parthenogenetic reproduction can attenuate this effect. However, asexual reproduction likely trades off with sexual reproduction. To explore the evolutionary implications of such a trade-off, we included recurrent mating events associated with seasonal interruptions in a simple population dynamics model. Following an adaptive dynamics approach, we showed that positive density dependence associated with Allee effects in cyclic parthenogens promotes evolutionary divergence in the level of investment in asexual reproduction. Although polymorphism may be transient, morphs mostly investing into sexual reproduction may eventually exclude those predominantly reproducing in an asexual manner. Asexual morphs can be seen as making cooperative investments into the common pool of mates, while sexual morphs defect, survive better, and may eventually fix in the population. Our findings provide a novel hypothesis for the frequent coexistence of sexual and asexual lineages, notably in plant parasitic fungi.

Cell fate decision control in the stochastic model of p53 regulatory pathway

Beata Hat-Plewinska

(Institute of Fundamental Technological Research, Poland)

The p53 protein known as the "guardian of the genom" is responsible for the detection of DNA damage, the cell cycle arrest, initiation of DNA repair process, which can lead to cell recovery or apoptosis in the case of the irreparable damage. The cell fate is decided in the interplay of the pro-apoptotic and anti-apoptotic p53 network components.

We propose a mathematical model of the p53 network, which successfully reproduces apoptotic-survival decisions in response to irradiation as a function of particular proteins that are known to be varied among cancer cell lines. The model is described by Markov process and is analyzed either by Gillespie algorithm simulation, or is approximated by the system of ODEs solved by standard MATLAB algorithms.

The dynamical structure of the regulatory system is controlled by 4 negative feedback loops responsible for oscillatory responses, and 2 positive mediated by the phosphatases Wip1 and PTEN. The analysis of the model shows that the critical dose of radiation initializing apoptosis is dependent on expression levels of these two phosphatases. In conclusion, the model predicts different susceptibility of cancer cell lines to the irradiation induced apoptosis, which can be helpful in selecting the most appropriate therapy based on expression signature.

Causal inference from interventional data

Alain Hauser

(University of Bern)

Causal models represented by DAGs, directed acyclic graphs, are used in many fields of biology, ranging from epidemiology to systems biology. Beyond statistical properties such as correlation and (conditional) independence of variables, causal models also provide predictions about effects of interventions, that is external changes to the system such as gene knockdown experiments.

Data measured under interventions ("interventional data") is ubiquitous in biology, raising the demand for estimation methods based on such data. Under observational (that is, non-interventional) data, the identifiability of causal models is limited by the so-called Markov equivalence. We show to which extent interventional data improves the identifiability of causal models. For this aim, we extend the notion of Markov equivalence of DAGs to the interventional case and present a graph theoretic characterization of the corresponding equivalence classes. Furthermore, we consider Gaussian causal models and address the problem of calculating the maximum likelihood estimator from interventional data. We present a generalization of Chickering's Greedy Equivalence Search algorithm to interventional data that makes regularized maximum likelihood estimation computationally feasible. We demonstrate the performance of this algorithm on simulated as well as real-world data.

Selection for altruism through random drift in variable size populations

Bahram Houchmandzadeh
(Université Joseph-Fourier)

Altruistic behavior is defined as helping others at a cost to oneself and a lowered fitness. The lower fitness implies that altruists should be selected against, which is in contradiction with their widespread presence in nature. Evolutionary dynamics is a competition between deterministic selection pressure and stochastic events due to random sampling from one generation to the next. We show here that an altruistic allele extending the carrying capacity of the habitat can win by increasing the random drift of "selfish" alleles. In other terms, the fixation probability of altruistic genes can be higher than those of selfish ones, even though altruists have a smaller fitness. Moreover when populations are geographically structured, the altruists advantage can be highly amplified and the fixation probability of selfish genes can tend toward zero. The above results are obtained both by numerical and analytical calculations. Analytical results are obtained in the limit of large populations. The theory we present does not involve kin or multilevel selection, but is based on the effect of random drift in variable size populations. The model is a generalization of the original Fisher-Wright and Moran models where the carrying capacity depends on the number of altruists.

Family size statistics

David Kessler
(Bar-Ilan University)

We examine the problem of family size statistics (the number of individuals carrying the same surname, or the same DNA sequence) in a given size subsample of an exponentially growing population, subject to birth, death and mutation. We find there are two regimes. If the growth rate is larger than the mutation rate, then the distribution is fat-tailed, whereas otherwise it is a power-law cut-off exponentially for large family size. We apply this to surname statistics for Norway, finding good agreement. We also apply our results to the problem of haplotype statistics for human mtDNA, deriving the (average) growth rate. We also apply our results to the statistics of number of species in genera, obtaining reasonable values for origination times of families.

Genome size evolution: challenging intuition with modelling

Carole Knibbe
(Université Lyon 1)

Although many complete genomic sequences are available, the evolutionary mechanisms that determine genome size, and in particular the amount of non-coding DNA, are still debated. Forces behind growth, like the proliferation of transposable elements or the creation of new genes by duplication, are well identified, whereas forces behind shrinkage are less clear. The hypothesis that a large genome directly limits the organism's reproduction speed is at odds with several data sets and other, non adaptive, mechanisms have been proposed, like genetic drift and/or a mutational bias causing deletions to be more frequent than insertions. Here, we propose a matrix population model for genome size evolution showing that genome size can also be limited by the spontaneous asymmetrical dynamics of duplications and deletions, which tends to make genomes shrink even if both event types occur at the same frequency. In the absence of selection, we prove the existence of a stationary distribution for genome size even if duplications are twice as frequent as deletions. Then we use numerical simulations to show that Darwinian selection

of the largest genomes cannot overcome this spontaneous dynamics, which is thus a new serious candidate among the forces that may limit genome growth. The complex, counter-intuitive behavior of this minimalist model also exposes the limits of the simple "thought experiment" when considering evolution. List of Authors : (1) Stephan Fischer, (2) Samuel Bernard, (3) Guillaume Beslon, (4) Carole Knibbe.

Statistical estimation of a growth-fragmentation model observed on a genealogical tree

Nathalie Krell
(Université de Rennes 1)

We investigate inference in simple models that describe the evolution in size of a population of bacteria across scales. The size of the system evolves according to a transport-fragmentation equation: each individual grows with a given transport rate, and splits into two offsprings, according to a binary fragmentation process with unknown division rate that depends on its size. Macroscopically, the system is well approximated by a PDE and statistical inference transfers into a nonlinear inverse problem. Microscopically, a more accurate description is given by a stochastic piecewise deterministic Markov process, which allows for other methods of inference, introducing however stochastic dependences. We will discuss and present some results on the inference of the parameters of the system across scales. Real data analysis is conducted on *E. Coli* experiments. This is a joint work with M. Doumic, M. Hoffmann and L. Robert.

Why does the effective size change in population history ?

Olivier Mazet
(INSA de Toulouse)

One aim of population genetics is to reconstruct key aspects of the past demographic history of populations: expansions, contractions (bottlenecks), but also structure with migrations, selection... A structure model may product the

same observed effective population size changing than a bottleneck in panmictic population. We shall show how both could be distinguished, using analytic distribution of the coalescence time.

Localisation in the Bouchaud-Anderson model

Richard Pymar
(University College London)

We study a modification of the parabolic Anderson model where the underlying continuous-time random walk on the lattice \mathbb{Z}^d is replaced with the Bouchaud trap model. We consider the case with Weibull potential field and arbitrary holding time distribution that is bounded away from zero. Under these conditions, we show that the solution of the model is eventually localised at a single site with overwhelming probability.

Spatio-temporal heterogeneity and unstable coexistence

Nadav Shnerb
(Bar-Ilan University)

Niche-based models for ecological populations and communities emphasize the role of deterministic, stabilizing forces, while Hubbell-Kimura neutral models suggest that the dynamics is driven only by demographic stochasticity. Analyzing the longest available records for the dynamics of tropical trees and breeding birds, we show that both stabilizing mechanisms and demographic stochasticity fail to play a dominant role in shaping assemblages over time. Rather, community dynamics in these two very different systems are predominantly driven by environmental stochasticity. Our results highlight the need for a new theory integrating environmental stochasticity with weak stabilizing forces, and suggest that such theory may better describe the dynamics of ecological communities than current neutral theories, deterministic niche-based theories, or recent hybrids.

Stochastic stability for predator-prey model of Holling type II with term refuge

Safia Slimani
(Badji Mokhtar University)

We consider a two-dimensional continuous time dynamical system modelling a predator-prey food chain, and based on a modified version of the Leslie-Gower scheme and on the Holling-type II scheme with term refuge and stochastic perturbation. We prove the existence of equilibrium point and stochastic stability of this point. The proof of stability is based on Lyapunov function.

Genetic hitchhiking and eco-evolution

Charline Smadi
(École Polytechnique)

When a selectively favored mutation occurs and sweeps to fixation, linked genes may increase in their frequencies because they are dragged along with the selected gene. This genetic hitchhiking is responsible for a reduction of diversity around the fixed mutant. We consider a sexually reproducing haploid population and develop a birth and death process with density dependent competition between individuals. Parameter K scales the population size. We study the limit behaviour for large K . We focus on two loci: a neutral locus with selectively equivalent alleles, and a locus fixed for allele A . At time 0, a favorable mutant allele a appears. We aim at understanding the effect of the selective sweep on the neutral diversity. We find two opposite recombination regimes. In the weak one, the recombination probability per reproductive event has order $1/\log K$. In the strong one, it is large with respect to $1/\log K$. In the weak regime, the natural selection on the mutant a drives an increase in frequency of the neutral allele carried by the first mutant. We express it as an explicit function of the ecological parameters, recombination probability and K . In the strong regime, the selective sweep does not modify the neutral diversity.

Integrating knowledge on genomic structure and ancestral diversity to detect outlying genomic regions

Solenn Stoeckel
(INRA Rennes)

The inference of selective events along the genomes may take advantage of all the biological knowledge we learn from a species and its related. In the new generation sequencing area, biologists have the opportunity to know how genes physically hang together and what was the putative ancestral genetic diversity through ancient DNA and the genomes of related species. Our novel approach includes such knowledge within a population genetics model formalized on genotypic frequencies, derived into a system of stochastic differential equations, to provide multidimensional reference distributions of genotypic diversity. A one-class classification method stemming from the statistical learning theory is then implemented to capture regions in the space of genotypic frequencies (input space) where the probability density of these reference distributions lives. By projecting new data sampled and genotyped from real populations in the input space, we are able to detect outlier loci over the investigated genome.

Stochastic modeling of the accumulation of passenger and driver mutations in cancer

Cristian Tomasetti
(Johns Hopkins University)

Important progress has been made in our understanding of cancer thanks to the ever growing amount of data originated by sequencing technologies. One useful approach for better understanding the process of accumulation of somatic mutations in cancer is given by the integration of mathematical modeling with sequencing data of cancer tissues. While it has been hypothesized that some of the somatic mutations found in tumors may occur prior to tumor initiation, there is little experimental or conceptual data on this topic. To gain insights into this fundamental issue, we formulated a mathematical model

for the evolution of somatic mutations in which all relevant phases of a tissue's history are considered. The model provides a way to estimate the in-vivo tissue-specific somatic mutation rates from the sequencing data of tumors. The model also makes novel predictions, validated by our empirical findings, on the expected number of somatic mutations found in tumors of self-renewing tissues. Our results may have substantial implications for the interpretation of the large number of genome-wide cancer studies now being undertaken.

A stochastic model of tumor growth based on branching processes

Pierre Vallois
(Université de Lorraine)

We consider a treatment of radiotherapy applied to a tumor. The aim is to quantify the effect of this treatment. We define a model of tumor growth based on branching processes which takes into account heterogeneous damages. We are able to determine numerically the mean number of cancer cells at any time.

Posters

Balance between absorbing and positive fixed points in resource consumption models

Hilla Behar
(Bar-Ilan University)

The effect of resource usage on economic growth has been studied in multiple models. However, the generic effect of improving resource usage efficacy through improved technical skills has not been studied in detail. We here analyze a model incorporating resource usage by capital and the parallel production of technical skill in order to study the effect of improving the efficacy of resources usage with advanced technologies. We show that a practically inevitable result of such a model is that improving the resource usage efficacy leads to a lower steady-state level of resources. A surprising conclusion from ordinary differential equations realization of the model is an extreme sensitivity to parameters, where a small parameter change can lead to an irreversible state through a hysteresis mechanism between a scenario of a collapse of the economy and a scenario of sustainable economy. This sensitivity is lost when spatial stochastic simulations are performed. In the stochastic regime the two scenarios coexist, with different fractions of the lattice residing in each state. Changing parameters smoothly changes the fraction of lattice sites in each state. The transition between the collapsed economy and the sustainable one is not symmetrical. Escape from the collapsed situation can only occur through diffusion from neighboring sustained lattice sites. On the other hand, the collapse can occur even in the absence of diffusion. This difference leads to diffusion dependent capital growth, where an optimal capital is obtained for middiffusion values. Such a transition may actually be generic phenomena in ecological and economic systems.

Evolution of a community of preys and predators in a rare mutation framework

Manon Costa
(École Polytechnique)

We consider an evolutionary model of a community of preys and predators evolving in the same environment. We are interested in the evolution of the interaction between preys and predators. Our main interest is to understand the impact of natural selection on the shape of the community and the survival of prey types as well as predators types. We model the dynamics as a two types birth and death process. Each prey and predator is characterised by a phenotypic trait that influences the predation and is impacted by natural selection. We want to predict the direction of selection on these traits and understand how the interaction network is modified through time. We consider the evolution in a rare mutation time scale; meaning that the demographic evolution is much faster than the mutational one. We prove that under good scaling, the process converges to a pure jump process representing the fast invasions of mutants and the ecological equilibria in between. We focus on the question of the existence and the stability of these equilibria.

This is joint work with Céline Hauzy, Nicolas Loeuille et Sylvie Méléard.

A damage spreading transition in a stochastic host-pathogen system

Yael Fried
(Bar-Ilan University)

One of the leading proposals for solving the biodiversity problem is the Janzen-Connell (JC) hypothesis, suggesting that the abundance of a species is limited by a host-specific exploiter. We will analyze a spatially explicit host-pathogen model, looking for the coexistence conditions under stochastic dynamics. Above the standard extinction transition associated with the failure of the pathogen to invade, we present another damage spreading transition, marking the point where macroscopic clusters of host individuals disappear.

Beyond its practical significance, this transition is apparently a generic landmark along the axis of decreasing stochasticity, if the deterministic dynamics support cycles or quasicycles.

Simulation model of a tropical foliar epidemic disease at plant scale: case of black sigatoka on banana

Clara Landry
(CIRAD)

Black Sigatoka (BS), caused by the fungal pathogen *Mycosphaerella fijiensis*, is considered as the most destructive foliar disease of bananas. Due to important damages on yield, the BS integrated management in particular by using resistant varieties appears essential. To better understand the pathogen dynamics and to identify the most effective resistance components, a mechanistic simulation model of BS was designed. The model developed in discrete time at plant scale, describes, without spatialization and under optimal epidemiological conditions, the lesions development during several crop cycles. Two sub-models are defined: the first one describes simply the banana growth in a deterministic way; the second one describes the complete and detailed epidemic cycle by integrating stochasticity. Infectious cycle data were collected in both controlled and natural infestation conditions on susceptible and resistant cultivars. We performed a sensitivity analysis to quantify the impact of model parameters (element of the epidemic cycle) and their interaction on the model output. Estimation of the model parameters was realized in a Bayesian framework using MCMC (Markov Chain Monte Carlo) methods such as the Metropolis-within-Gibbs algorithm. The posterior distributions of the parameters and residual variance were obtained.

Influence of a spatial structure on phenotypic evolution

Hélène Leman
(École Polytechnique)

Recently, a number of stochastic studies have been developed to understand precisely the phenotypic evolution of a population. In this presentation,

we will focus on the influence of a spatial structure on this evolution. Indeed, the environment is considered to play a key role in Darwinian evolution. In particular, some ecological researches conclude that a heterogeneous environment can induce some diversification within species, as observed in the example of the Darwinian finches. First of all, we will describe the stochastic model, which is an individual based model with spatially dependent coefficients. Some simulations of the model, which correspond to ecological observations, lead us to want to analyse it. A rescaling process and its limits in large population help us to understand precisely the model. So, in a second part, we will be interested in the properties of that limit which is a solution to a nonlinear partial differential equation. We will focus on the case of a dimorphic population to find the parameters which characterize the phenotype best adapted to the environment.

Soil pathogens and plant diversity

Magnus Lindh
(Umeå University)

A difficult problem in evolutionary ecology is the high plant diversity found in for example tropical forests, where hundreds of species can coexist in only a few hectares, despite only a few apparent resources. It is believed that soil pathogens promote diversity in plant communities. We present a spatial individual based stochastic model where the success of a seedling depends on the distance and the density of the parent tree. Each grid cell in the individual based model corresponds to a standard crown diameter of an established tree. The seed production of the parent tree in turn depends on time since maturation, initially seed production is increasing as the parent tree is establishing, but in the long run seed production is going down due to soil pathogens. The soil pathogens are typically very persistent and can be thought of as always present in the soil. As the density of seedlings increase the specie specific soil pathogens also increase, and this eventually increases seedling mortality. We are interested in finding out: 1) How many different trees with their respective soil pathogens could coexist in a given area, and 2) If soil pathogens could promote diversity as suggested in the literature.

Coding multi-type continuous time Galton-Watson forests

Thi Ngoc Anh Nguyen
(Université d'Angers)

Coding branching forests from integer valued random paths allows us to preserve the information on the genealogy of the population modeled which is partly lost in the associated branching processes. T. Harris coded a Galton-Watson forest by a random walk through a bijection while its multi-type version is coded by L. Chaumont and R. Liu recently by several multidimensional random walks. We will here generalize the result of L. Chaumont and R. Liu by giving multidimensional compound Poisson processes which code a multi-type continuous time Galton-Watson branching forest. A result on the law of the number of leaves will then be presented as a direct application.

Coupling a biophysical and genetic model for the study of connectivity in marine invertebrates forming a metapopulation

Mariana Padron
(UPMC)

Landscape connectivity results from the interactions between demographic processes and environmental filtering. At sea, most species, and particularly marine benthic invertebrates exhibit a life cycle with a dispersive larval stage, which promotes connectivity between distant populations. The degree of connectivity among populations forming a metapopulation has direct consequences for species evolution and their capacity to adapt to changes in the environment. Understanding the complex processes driving seascape genetics of benthic invertebrates requires the use of tools that integrate information regarding the demographic and genetic linkages resulting from the larval exchange among populations. Although different modeling approaches have been developed recently to incorporate genetic mixing due to connectivity, the incorporation of demographic effects, such as local transient extinction or recruitment limitation, linked to stochastic connectivity has not yet been accomplished. Here, we present a novel model that will assess the effect of demographic stochasticity on allele frequencies in marine populations of benthic invertebrates. This

model couples a demographic metapopulation model using realistic connectivity matrices and a genetic model including genetic drift. This approach provides a new tool for predicting patterns of population genetics structure, enabling the testing of our understanding of demographic stochasticity induced by connectivity against existing empirical genetic data.

Non-linear evolutionary dynamics in partially asexual populations

Katja Reichel
(INRA Rennes)

Partially asexual organisms are able to reproduce both sexually and asexually (clonally). They occur throughout all the earth's biomes and often play ecologically important roles (colonizers, ecosystem engineers, parasites, invasives). We are interested in the effect of partial asexuality on evolution.

Using a discrete-time Markov chain model based on the composition of a single population of fixed size out of different genotypes, we are able to study the effect of mutation, drift and different rates of asexuality c (constant or cyclic) through time, visualized in de Finetti diagrams. Already for the "classic" case of a single locus with two different alleles, we found that the short-term dynamics within partially asexual populations are not linear dependent on c . Moreover, the population genetics of cyclic asexuals, such as daphnia or aphids, cannot be treated as either equivalent to complete sexuality, complete asexuality or equal to an average rate. Thus, the microevolution of partially asexual organisms constitutes a distinct case: the paths leading to neutral as well as adaptive changes in their genome have a different shape.

Reconstructing graphs associated with pedigree of cultivars

Chaker Sbai
(Université d'Angers)

The objective of this work is to develop mathematical models and computer tools applied to the reconstruction of genealogy of vegetatively propagated plant cultivars, in order to optimise their crossbreeding.

The genealogy of the diploid individuals considered in this study consists, most of the time, in a network including loops which is called a pedigree. The available material consists in genotyping data for these cultivars, which are either internal nodes or terminals in the pedigree, for three to six generations. Then, the question which arises is: what is the most likely network linking all these individuals ?

We propose to tackle this question developing mathematical models (probabilistic) linked to graph theory.

A graphical interface was created and linked to the mathematical models in order to resolve relationship problems between individuals.

This is a joint work with Richard Pymar, Loïc Chaumont and Valéry Malécot.

List of participants

Étienne **Adam** (École Polytechnique)
Linda **Allen** (Texas Tech University)
Nicolas **Bacaër** (IRD Bondy)
Vincent **Bansaye** (École Polytechnique)
Hilla **Behar** (Bar-Ilan University)
Yossi **Ben Zion** (Bar-Ilan University)
Julien **Berestycki** (UPMC)
Samuel **Bernard** (Université Lyon 1)
Romain **Biard** (Université de Franche-Comté)
Sylvain **Billiard** (Université Lille 1)
Małgorzata **Bogdan** (Wroclaw University)
Simon **Boitard** (INRA Jouy-en-Josas)
Nicolas **Champagnat** (INRIA Nancy)
Loïc **Chaumont** (Université d'Angers)
Stéphane **Chrétien** (Université de Franche-Comté)
Camille **Coron** (Université Paris-Sud)
Manon **Costa** (École Polytechnique)
Luca **Dall'Asta** (Politecnico di Torino)
Jean-François **Delmas** (École des Ponts ParisTech)
Yves **Dumont** (CIRAD)
Abdelkarim **Ed-Darraz** (Université Cadi Ayyad)
Mikael **Falconnet** (Université d'Évry)
Déborah **Ferré** (Université d'Angers)
Yael **Fried** (Bar-Ilan University)
Rodolphe **Garbit** (Université d'Angers)
Dan **Goreac** (Université Paris-Est Marne-la-Vallée)
Piotr **Graczyk** (Université d'Angers)
Simona **Grusea** (INSA de Toulouse)
Iqbal **Hamada** (University of Saida)
Frédéric **Hamelin** (Agrocampus Ouest)
Beata **Hat-Plewinska** (Institute of Fundamental Technological Research, Poland)
Alain **Hauser** (University of Bern)
Patrick **Hoscheit** (INRA Jouy-en-Josas)
Bahram **Houchmandzadeh** (Université Joseph-Fourier)
David **Kessler** (Bar-Ilan University)
Carole **Knibbe** (Université Lyon 1)

Nathalie **Krell** (Université de Rennes 1)
Amaury **Lambert** (UPMC)
Clara **Landry** (CIRAD)
Bruno **Le Cam** (INRA Angers)
Christophe **Lemaire** (Université d'Angers)
Hélène **Leman** (École Polytechnique)
Magnus **Lindh** (Umeå University)
Sébastien **Loustau** (Université d'Angers)
Valéry **Malécot** (Agrocampus Ouest)
Salha **Mamane** (Université d'Angers)
Olivier **Mazet** (INSA de Toulouse)
Sylvie **Méléard** (École Polytechnique)
Hans **Metz** (Leiden University)
Éric **Miqueu** (Université de Bretagne-Sud)
Thi Ngoc Anh **Nguyen** (Université d'Angers)
Mariana **Padron** (UPMC)
Étienne **Pardoux** (Université d'Aix-Marseille)
Todd **Parsons** (UPMC)
Sylvain **Poggi** (INRA Rennes)
Besnik **Pumo** (Agrocampus Ouest)
Richard **Pymar** (University College London)
Landy **Rabehasaina** (Université de Franche-Comté)
Katja **Reichel** (INRA Rennes)
Mathieu **Richard** (École Polytechnique)
Natalia **Sapoukhina** (INRA Angers)
Chaker **Sbai** (Université d'Angers)
Nadav **Shnerb** (Bar-Ilan University)
Safia **Slimani** (Badji Mokhtar University)
Charline **Smadi** (École Polytechnique)
Terry **Speed** (University of California)
Solenn **Stoeckel** (INRA Rennes)
Elizabeth **Thompson** (University of Washington)
Christian **Tomasetti** (Johns Hopkins University)
Pierre **Vallois** (Université de Lorraine)
Lioudmila **Vostrikova** (Université d'Angers)

