SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Lyon, November 8-9-10, 2017, Centre Otelia

https://project.inria.fr/aiem2017/

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Program

Lyon, 8 to 10th of November 2017 https://project.inria.fr/aiem2017/

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Program

Wednesday, November 8th

Chair: Guillaume Achaz

•13h30 – 13h50 ————————Welcome Coffee————————
•13h50 – 14h00 Opening
•14h00 – 14h15 Ingrid Lafontaine “Analyse fonctionnelle et évolutive des gènes à « duplex ARN » chez les levures”
•14h20 – 14h35 Fabien Pierriel “Evolution of hydroxylases in ubiquinone biosynthesis: variations in number and regio-selectivity”
•14h40 – 15h15 Thomas Heams “Une ingénierie du vivant est-elle possible ?”
•15h20 – 15h35 Pablo Jensen “Emergence of new scientific disciplines”
•15h40 – 16h20 ————————Coffee Break————————

Chair: Sylvain Charlat

•16h20 – 16h35 Priscila Biller “Evolutionary applicability of computational evolvability”
•16h40 – 16h55 Fanny Pouyet “Recombination and selection”
•17h00 – 17h15 Guillaume Laval “Genetic evidence for greater human adaptation out of Africa over the last 100,000 years”
•17h20 – 17h35 Thibault Latrille “A population-genetic model of the Red Queen dynamic of recombination”
•17h40 – 17h55 François Blanquart “Viral genetic variation accounts for a third of variability in HIV-1 set-point viral load in Europe”
•18h00 – 19h00 ————————Poster session/Wine————————
•19h00 – 20h00 ————————Dinner————————

Thursday, November 9th

Chair: Angeles de Cara

•09h00 – 09h15 Michael Blum “Genome scans for local adaptation using principal component analysis”
•09h20 – 09h35 Ivan Scotti “Analysis and modelling of micro-geographic divergence in tree populations”
•09h40 – 09h55 Alexis Simon “Selection on hybrid genotypes and Fisher’s Geometric model”

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

•10h00 – 10h15 Maud Duranton Are incidental islands less likely to introgress? Insights from haplotype-resolved genomes in European sea bass.
•10h20 – 10h35 Ivan Paz Vinas “how geomorphologic resistance affects the spatial distribution of genetic diversity”
•10h35 – 11h15 ————————Coffee Break————————

Chair: Guillaume Beslon

•11h20 – 11h35 Verónica Miró Pina “Chromosome painting”
•11h40 – 11h55 Stéphanie Bedhomme “Plasmid and clonal interference in experimental evolution”
•12h00 – 12h15 Charles Rocabert “Phenotypic Noise and the Cost of Complexity”
•12h20 – 12h35 Maud Gautier “Characterisation of biased gene conversion in mouse recombination hotspots”
12h35 – 14h00 ————————Lunch————————

Chair: Catherine Breton

•14h00 – 14h15 Anabelle Haudry “Tempo and mode in genome size evolution and transposable elements content in fly genomes”
•14h20 – 14h35 Rousselle Marjolaine “Impact of GC-biased gene conversion on the adaptive substitution rate in fowls”
•14h40 – 14h55 Vincent Liard “Robustness and Evolvability: quantitative definitions in the context of modeling and simulation”
•15h00 – 15h15 Bastien Boussau “Combining relaxed clocks with gene transfers to date species trees”
•15h20 – 15h35 Celine Scornavacca “gene tree species tree reconciliation with transfer and incomplete lineage sorting”
•15h40 – 16h20 ————————Coffee Break————————

Chair: Renaud Vitalis

•16h20 – 16h35 Vincent Castric “Evolutionary novelty in a receptor-ligand interaction”
•16h40 – 16h55 Thomas Brom “Maintenance of gametophytic self-incompatibility system in spatially structured population”
•17h00 – 17h15 Gabriel Marais “Evolutionary stasis of the pseudo-autosomal region in Strepsirhine primates“
•17h20 – 17h35 Gilles Didier “Detecting molecular basis of phenotypic convergence”
•17h40 – 17h55 Nathanaelle Saclier “Life history traits impact the nuclear rate of substitution but not the mitochondrial rate in Isopods”
•18h00 – 19h00 ————————Poster session/Wine————————
•19h00 – 20h00 ————————Dinner————————

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Friday, November 10th

Chair: Anamaria Necsulea

•09h00 – 09h15 Gilles Fischer “The evolution of the temporal program of genome replication”
•09h20 – 09h35 Ivan Junier “What evolution has to say about the physics of DNA?”

•09h40 – 09h55 Victoire Baillet “Impact of DNA methylation on the rate and spectrum of mutation in Arabidopsis thaliana”
•10h00 – 10h15 Aurélien Chateigner “disentangling the complex web of interactions underlying biomass production”
•10h20 – 10h35 Julien Fouret “Towards understanding of anti-viral response in bats: comparative positive selection analysis combined with functional network”
•10h35 – 11h15 ————————Coffee Break————————

Chair: Eric Tannier

•11h20 – 11h35 Lucie Etienne “Evolutionary-conserved antagonism of the housekeeping Smc5/6 complex by mammalian Hepatitis B virus”
•11h40 – 11h55 Stéphanie Jacquet “First genetic evidences that NTCP evolution has been driven by hepadnaviruses in bats and primates”
•12h00 – 12h15 Angeles de Cara “Targets of selection and the evolution of a butterfly mimicry supergene”
•12h20 – 12h35 Yves Clément “Genome-wide enhancer – target gene regulatory maps in two vertebrate genomes”
•12h40 – 12h55 Maeva Mollion “Impact of domestication on the distribution of fitness effects in several plant species”
•13h00 – 14h00 ————————Lunchbox————————

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Talk Session

Title : **Analyse fonctionnelle et évolutive des gènes à « duplex ARN » chez les levures.**

Authors : Jules Gilet, Romain Conte, Lionel Benard et **Ingrid Lafontaine**

Address :
UMR 7141, Institut de Biologie Physico-Chimique, Paris
UMR 8226, Institut de Biologie Physico-Chimique, Paris

Abstract :
Il existe chez la levure S. cerevisiae, des gènes adjacents transcrits de façon convergente dont les ARN messagers peuvent former des structures en double hélice au niveau de leurs extrémités 3'. Ces structures sont potentiellement responsables d'une régulation post-transcriptionnelle de leur expression.
Nous avons en effet observé que la majorité des ARNmim (pour ARNm interagissant avec d'autres ARNm) expriment des gènes impliqués dans la réponse au stress. Les données de transcriptomique (ChiP-Seq) montrent que, lors de la réponse au stress, les ARNmim ne s'associent ni aux facteurs responsables de la répression traductionnelle ni à ceux responsables de l'inactivation des ARNm par leur recrutement dans des agrégats comme les p-bodies. Ces facteurs lient préférentiellement la région 3' simple-brin des ARNm, ce qui explique comment ces ARNmim leur échappent, par la formation de duplex ARN à leurs extrémités 3'. Cette propriété permet donc le maintien de leur traduction en condition de stress.
Nous avons également montré que la conservation de l'orientation relative de ces paires de gènes dans 34 génomes de levures Saccharomycotina est plus importante que celle des gènes ne formant pas de paires ARNmim, ce qui conforte notre hypothèse fonctionnelle. Les travaux de validation expérimentales sont en cours.

Title : **Evolution of hydroxylases in ubiquinone biosynthesis: variations in number and regio-selectivity**

Authors : **Fabien PIERREL**

Address :

Abstract :
Ubiquinone (UQ) is a key molecule for cellular bioenergetics, conserved from proteobacteria to humans. UQ acts as a liposoluble electrons and protons carrier in respiratory chains that build an electrochemical gradient which is used by ATP synthase to synthesize ATP which represents most of the cellular energy. UQ biosynthesis has been studied only in a few model organisms, thus the diversity of UQ biosynthesis pathways is largely unknown [1]. Biosynthesis of UQ requires three hydroxylation reactions on contiguous positions of an aromatic ring. In Escherichia coli, each of three UQ flavin monooxygenases (FMOs), called UbiF, UbiH and UbiI, hydroxylates a single position of the aromatic ring. This “three hydroxylation reactions/three proteins” pattern has been accepted as a paradigm in UQ biology.

Using a phylogenetic analysis, we showed that UbiF, UbiH and UbiI are detected only in a small fraction of proteobacteria and we identified two new types of UQ FMOs: UbiM, which is distributed in α-, β- and γ-classes and UbiL, which is restricted to α-proteobacteria [2]. Remarkably, the ubiL and ubiM genes were found in genomes with less than three UQ hydroxylase-encoding genes and we demonstrated using biochemical approaches that UbiL from Rhodospirillum rubrum and UbiM from Neisseria meningiditis hydroxylate respectively two and three positions of the aromatic ring during UQ biosynthesis. Thus, bacteria containing only one or two UQ hydroxylases have developed generalist enzymes that are able to catalyze several steps of UQ biosynthesis. Overall, bacteria have evolved a large repertoire of hydroxylase combinations for UQ biosynthesis, including pathways with either three specialist enzymes (as in E. coli) or pathways with one or two generalist enzymes of broader regio-selectivity (as in N. meningiditis and R. rubrum) [2]. I will discuss how the emergence of broad regio-selectivity UQ hydroxylases may be related to genome reduction events in several proteobacterial species and how the composition of the UQ biosynthetic pathway has evolved since its appearance in an ancestral proteobacterium, more than two billion years ago.

1. Aussel L, Pierrel F, Loiseau L, Lombard M, Fontecave M, Barras F (2014) Biosynthesis and physiology of coenzyme Q in bacteria. Biochim. Biophys. Acta 1837: 1004-1011

2. Pelosi L, Ducluzeau A-L, Loiseau L, Barras F, Schneider D, Junier I, Pierrel F (2016) Evolution of ubiquinone biosynthesis: multiple proteobacterial enzymes with various regio-selectivities to catalyze three contiguous aromatic hydroxylation reactions. mSystems 1: e00091-00016

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Une ingénierie du vivant est-elle possible ?**

Authors : **Thomas Heams**

Address :

Abstract :

Title : **Emergence of new scientific disciplines**

Authors : **Pablo Jensen**

Address :

Abstract :
The description of large temporal graphs requires effective methods giving an appropriate mesoscopic partition. Many approaches exist today to detect “communities”, ie groups of nodes that are densely connected (Fortunato, 2010), in static graphs. However, many networks are intrinsically dynamical, and need a dynamic mesoscale description, as interpreting them as static networks would cause loss of important information (Holme and Saramaki, 2012; Holme 2015). For example, dynamic processes such as the emergence of new scientific disciplines, their fusion, split or death need a mesoscopic description of the evolving network of scientific articles. Here, we present a new approach that distinguishes real trends and noise in the mesoscopic description of social data using the continuity of social evolutions. To be able to follow the dynamics, we compute partitions for each time slice, but to avoid transients generated by noise, we modify the community description at time t using the structures found at times t-1 and t+1. We show the relevance of our method on the analysis of a scientific network showing the birth of a new subfield, wavelet analysis.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Evolutionary applicability of computational evolvability**

Authors : **Priscila Biller**

Address :

Abstract :
Mathematical models for evolving molecular mechanisms often don't include a theoretical assessment of their "evolvability", understood here as their capacity to adapt to new environments in a reasonable amount of time. A framework derived from machine learning, seeing evolution as a learning algorithm, has been proposed to fill this gap (Valiant, 2009). It provides a solid ground in the same vein as in computer science, in which computational complexity theory answers if a problem can be computable with finite resources available (e.g. time, memory, storage).
So far the computational evolvability framework has been studied in the machine learning field under a theoretical perspective, with several results proving which classes of functions organisms can learn by evolution or not. But given its potential relevance in molecular evolution, here we provide a link between these two fields by examining together theoretical arguments and experiments with Aevol, an in silico evolution framework of digital organisms. We compare the performance of evolvable and non-evolvable mechanisms---binary conjunctions and parity functions, respectively---under diverse evolutionary scenarios generated with an adapted version of Aevol. Depending on the individual location in the fitness landscape, we show that the so-called "evolvable" mechanisms can struggle or even be unable to evolve to the target behavior, contrary to the "predictable and inexorable" evolution predicted by theoretical results. Concerning non-evolvable mechanisms, we show that it is possible to make them "evolvable" by slightly changing how the fitness computation is done. More specifically, to evaluate the fitness, inputs are sampled in order to compare the individual behavior with the ideal outputs. We show that parity functions can be evolvable in the same scenarios as conjunctions just by replacing uniform by binomial sampling in the fitness evaluation. In conclusion, we define better the boundaries and implications of the evolvability theory when brought closer to biological objects.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Recombination and selection**

Authors : **Fanny Pouyet**

Address :

Abstract :
Recombination mediates the impact of selection at neighbouring neutral loci, but its exact effect on patterns of genomic diversity is still largely unknown in the human genome. To better examine the relationship between selection and recombination in 1000G and SGDP full genome data sets, we have focused on the number of derived alleles per individual (INDA), a simple statistic that is not affected by demography. Interestingly, after controlling for mutation rate, INDA increases almost linearly with the log of local recombination rate in all individuals and populations examined so far. This observation is best explained by increased effects of background selection (BGS) in regions of lower recombination. Remarkably, the effect of BGS extends over non-coding and non-transcribed regions, implying it affects the whole genome. While we expect a minimal effect of BGS in high recombination regions, biased gene conversion towards G and C alleles (gBGC) becomes important and greatly affects INDA. BGS and gBGC also affect the site frequency spectrum in complex ways, leading to potentially biased inferences of demography and positive selection. We use forward simulations to check that BGS can lead to observed diversity in populations having different demographic histories, and we identify a new set of SNPs that is minimally affected by BGS and gBGC, and which is thus best suited to infer the demographic history of human populations.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Genetic evidence for increased human adaptation out of Africa during the last 100,000 years**

Authors : **Guillaume Laval**1,2,3,*, Etienne Patin1,2,3 and Lluis Quintana-Murci1,2,3

Address :
1 Human Evolutionary Genetics, Institut Pasteur, 75015 Paris, France
2 Centre National de la Recherche Scientifique, URA3012, 75015 Paris, France
3 Center of Bioinformatics, Biostatistics and Integrative Biology, Institut Pasteur, 75015 Paris, France

Abstract :
The detection of the legacy of past selection in the human genome has proved crucial for identification of genes underlying the broad morphological and physiological diversity observed across human populations, and for increasing our understanding of the genetic architecture of adaptive phenotypes. Over the last decade, genomic scans for the signatures of positive selection have provided a flurry of promising candidate loci. However, there is a poor overlap among candidates, false discovery rates remain high, and functional evidence for their putative adaptive nature have been obtained for only a handful of genes (e. g., the rs4988235 in the LCT region). In this context, the extent to which classic positive, Darwinian selection has participated in recent human adaptation remains highly-debated1-3, and a fundamental parameter of our adaptive history, the number of true positively-selected loci in different human populations, remains uncharacterized. Furthermore, although it has been suggested that the number of selected events might be higher in non-African populations4, it remains unclear whether this reflects higher adaptation in populations migrating to the new environments of Eurasia, or that the known bottleneck that these populations underwent accentuated the molecular signals interpreted as resulting from positive selection4.
Here, instead of using the classical bottom-up strategy to investigate human adaptation (i.e., from candidate loci to adaptive history), we adopted a top-down approach, where we first estimated, using Approximate Bayesian Computation5 (ABC), a number of key parameters of human adaptive history (top) to subsequently delineated candidate loci based on the estimated parameters (down). We applied this ABC method based on the simulations of whole genomes sequencing data according to validated demographic models6,7, to the 1,000 Genomes (1000G) data8, on which we jointly estimated the number of positively-selected loci (X) per population, and the means of the distributions of the intensity (S) and age (T) of selection.
We subsequently assessed the accuracy of the ABC method through whole genome simulations (mimicking the features of the 1000G data i. e., ~3Gbp of DNA sequences in ~100

individuals per population) used as empirical data where values of adaptive parameters are known. For both X, S and T, we found unbiased and highly accurate joint estimations. Because background selection is known to mimic, to some extent, positive selection signatures, we corrected the method for this confounding factor and found virtually unchanged accuracies, regardless the demographic scenario considered (similar accuracy in African, European and Asian populations).
We then jointly estimated X, S and T in each 1000G population separately and we found substantial shift toward higher values of the approximated posterior distributions of X outside Africa, mainly in Asia (e. g., XAFR=27[CI:13-42], XEUR=57[CI:35-81] and XASI=70[CI:44-98]). Our results showed twice more adaptive events in Europeans and Asians, relative to Africans (simulation-based statistical tests, P=1x10-5), and this when correcting for the bottleneck occurred during the Out-of-Africa exodus. Interestingly, this increase in number of positively selected loci outside Africa was not accompanied by drastic changes of the intensity of selection (S and T found to be similar across populations). Finally, we performed a classical genome wide scan and, by refining candidate regions on the basis of our X estimates, we highlighted a top-candidate region detected in all European populations only, which encompasses functional variants modulating the sensitivity to UV-induced melanoma. This selection signal, ~1-1.2Mb long, suggests that the recent Pleistocene climatic warming (~20-10Kya) may have modulated human sensitivity to UV-induced melanoma in Europe.
Our study jointly estimates, for the first time, the number of loci targeted by recent, ongoing positive selection in the human genome and the underlying average intensity and age of selection. A future direction will be to update this approach in order to investigate more subtle modes of selection (e. g., soft selection on standing variation, adaptive introgression from ancient hominids). More generally, the proposed ABC method can easily be updated to infer adaptation rates of other model and non-model organisms, pacing the way to broad-scale investigation of the impact of reproductive modes, domestication, generation time, or species size on adaptive evolution.

References

1. Hernandez, R.D. et al. Classic selective sweeps were rare in recent human evolution. Science 331, 920-4 (2011).
2. Pritchard, J.K. & Di Rienzo, A. Adaptation - not by sweeps alone. Nature Reviews Genetics 11, 665-667 (2010).
3. Pritchard, J.K., Pickrell, J.K. & Coop, G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr Biol 20, R208-15 (2010).
4. Coop, G. et al. The Role of Geography in Human Adaptation. Plos Genetics 5(2009).
5. Beaumont, M.A., Zhang, W.Y. & Balding, D.J. Approximate Bayesian computation in population genetics. Genetics 162, 2025-2035 (2002).
6. Grossman, S.R. et al. Identifying recent adaptations in large-scale genomic data. Cell 152, 703-13 (2013).

7. Grossman, S.R. et al. A composite of multiple signals distinguishes causal variants in regions of positive selection. Science 327, 883-6 (2010).
8. Auton, A. et al. A global reference for human genetic variation. Nature 526, 68-74 (2015).

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **A population-genetic model of the Red Queen dynamic of recombination**

Authors : **Thibault Latrille**

Address :

Abstract :
In humans and many other species, recombination events cluster in narrow hot spots distributed across the genome, whose location is determined by the Zn-finger protein PRDM9. Hot spots are not shared between closely related species, suggesting that hot spots are short-lived. To explain this fast evolutionary dynamics of recombination landscapes, an intra-genomic Red Queen model, based on the interplay between two antagonistic forces, has been proposed. On the one hand, biased gene conversion, mediated by double-strand breaks, results in a rapid extinction of hot spots in the population. On the other hand, the resulting genome-wide depletion of recombination induces positive selection favoring new Prdm9 alleles recognizing new sequence motifs across the genome and restoring normal levels of recombination. This Red Queen scenario is currently the reference model for explaining the fast turnover of recombination landscapes. Thus far, however, it has not been formalized as a quantitative population-genetic model, fully accounting for the intricate interplay between biased gene conversion, mutation, selection, demography and genetic diversity at the PRDM9 locus.

Here, we propose a population-genetic model of the Red Queen dynamic of recombination. This model was implemented as a Wright-Fisher simulator, allowing exploration of the behaviour of the model (in terms of the implied mean equilibrium recombination rate, diversity at the PRDM9 locus, or turnover rate) as a function of the parameters (effective population size, mutation and erosion rates). In a second step, analytical results, based on self-consistent mean-field approximations, were derived. These analytical results reproduce the scaling relations observed in the simulations, offering key insights about the detailed population-genetic mechanisms of the Red Queen model. Empirical fit of the model to current data from the mouse and humans suggests both a high mutation rate at PRDM9 and strong biased gene conversion on its targets.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Viral genetic variation accounts for a third of variability in HIV-1 set-point viral load in Europe**

Authors : **François Blanquart**

Address :

Abstract :
HIV-1 set-point viral load—the approximately stable value of viraemia in the first years of chronic infection—is a strong predictor of clinical outcome and is highly variable across infected individuals. To better understand HIV-1 pathogenesis and the evolution of the viral population, we must quantify the heritability of set-point viral load, which is the fraction of variation in this phenotype attributable to viral genetic variation. However, current estimates of heritability vary widely, from 6% to 59%. Here we used a dataset of 2,028 seroconverters infected between 1985 and 2013 from 5 European countries (Belgium, Switzerland, France, the Netherlands and the United Kingdom) and estimated the heritability of set-point viral load at 31% (CI 15%–43%). Specifically, heritability was measured using models of character evolution describing how viral load evolves on the phylogeny of whole-genome viral sequences. In contrast to previous studies, (i) we measured viral loads using standardized assays on a sample collected in a strict time window of 6 to 24 months after infection, from which the viral genome was also sequenced; (ii) we compared 2 models of character evolution, the classical “Brownian motion” model and another model (“Ornstein–Uhlenbeck”) that includes stabilising selection on viral load; (iii) we controlled for covariates, including age and sex, which may inflate estimates of heritability; and (iv) we developed a goodness of fit test based on the correlation of viral loads in cherries of the phylogenetic tree, showing that both models of character evolution fit the data well. An overall heritability of 31% (CI 15%–43%) is consistent with other studies based on regression of viral load in donor–recipient pairs. Thus, about a third of variation in HIV-1 virulence is attributable to viral genetic variation.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Natural selection would have selected principal component analysis!**

Authors : **Michael Blum**

Address :

Abstract :
Abstract: In population genomics, I show that variations around principal component analysis (PCA) provide efficient approaches for mapping genes involved in natural selection. In several biological scenarios ranging from local adaptation to adaptive introgression, PCA generates less false discoveries and is faster than most statistical approaches currently implemented. I conclude that with large amount of data, statistical methods should be more data-driven than model-driven.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Analysis and modelling of micro-geographic divergence in tree populations**

Authors : **Ivan Scotti**

Address :

Abstract :
Microgeographic adaptation can occur when the effects of directional selection are not wiped out in spite of strong gene flow. In such circumstances, traits and genetic loci that are the target of selection can show adaptive divergence, against the backdrop of very little differentiation at other traits or loci. Through this mechanism, large genetic pools spanning across multiple habitats may maintain adaptive genetic diversity and population structure. How common could such events be in the genome of large populations of long-lived organisms, and how strong is selection underlying them? We addressed these questions by analysing patterns of genomic divergence in eight alpine, Mediterranean and tropical tree species, by screening matching pairs of closely-related populations sampled along multiple types of steep ecological gradients. Our results show that a very small subset of the genome may respond to such selective pressure, but that selection can be locally very strong.

Title : **Selection on hybrid genotypes and Fisher's Geometric model**

Authors : **Alexis Simon** (1,2)*, Nicolas Bierne (1,3), and John J. Welch (2)

Address :
1 : OREME station méditerranéenne, 2 rue des chantiers 34200 Sète, France.
2 : Department of Genetics, University of Cambridge, Downing St. Cambridge, CB23EH, UK.
3 : Université de Montpellier, Sète, France; Institut des Sciences de l'Evolution, CNRS-UM-IRD, Montpellier, France.
* Author for correspondence: alexis.simon@umontpellier.fr

Abstract :
Fitness landscape models can aid the study of speciation, hybidization and admixture. Here, we study a class of fitness landscapes based on Fisher's geometric model, which generates a rich variety of epistatic effects, with few parameters. We develop a random walk approximation, and derive simple predictions for the relative fitness of any hybrid individual. We show that the model predicts observed differences between different classes of hybrid, and makes novel predictions about selection on heterozygosity within crosses. We test these predictions with F2 and backcross data from Mytilus mussels, and published data from hybrid zones and controlled crosses of Zea, Populus, Senecio, Mus, Teleogryllus and Drosophila. The model's major predictions are supported. Finally, we develop a general model of genetic incompatibilities, and compare its predictions to Fisher's model. We show that the two modeling approaches yield many of the same predictions, but only when the dominance relations of individual incompatibilities are assigned in a particular way.

Title : **Are incidental islands less likely to introgress? Insights from haplotype-resolved genomes in European sea bass.**

Authors : **Maud Duranton,** François Bonhomme et Pierre-Alexandre Gagnaire**.**

Address :
Université de Montpellier, , France; Institut des Sciences de l'Evolution, CNRS-UM-IRD, Montpellier, France., Bat 24 1 er étage, Place Eugène Bataillon 34095 Montpellier cedex 05 France
Author for correspondence: maud.duranton@umontpellier.fr

Periods of geographic isolation may promote divergence of particular genomic regions that experience increased rates of lineage sorting through linked selection. Whether these so-called incidental islands of divergence play a role in the speciation process remains largely unknown. Therefore, we need to evaluate the extent to which incidental islands resist to gene flow upon secondary contact. We focus on the European seabass (Dicentrarchus labrax), which is genetically subdivided into two partially reproductively isolated lineages, one living in the Atlantic and the other in the Mediterranean Sea. We used haplotype-resolved whole-genome sequences to directly identify Atlantic tracts introgressed into Mediterranean genomes and show that they occur at highly variable frequency across the seabass genome. We also reconstruct ancestral genomes prior to gene flow, and show that incidental islands that emerged during allopatric episodes have been disproportionately resistant to erosion by secondary gene flow. We conclude that genomic islands first emerged in allopatry through the effect of linked selection acting on a heterogeneous recombination landscape. Upon secondary contact, preexisting islands were strongly remolded by differential introgression, revealing variable fitness effects among regions involved in reproductive isolation.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **How geomorphologic resistance affects the spatial distribution of genetic diversity**

Authors : **Ivan Paz Vinas**

Address **:**

Abstract :
Describing, understanding and predicting the spatial distribution of genetic diversity is a central issue in biological sciences. In dendritic landscapes such as riverscapes, it is generally predicted that neutral genetic diversity should increase downstream, a result that has recently been validated across taxonomic groups. This spatial pattern of diversity is notably driven by processes such as downstream-biased dispersal, increases in habitat availability downstream, and/or upstream-directed colonization. Although these results hold true for organisms living in surface waters, it remains unknown whether these results also apply for groundwater organisms (e.g. organisms living in the hyporheic zones underlying rivers).

Here, we first used pattern-oriented genetic data simulations to explore how geomorphologic resistance (a feature expected to deeply influence genetic diversity in groundwater organisms) affects the spatial distribution of genetic diversity. Second, we used machine-learning approaches (i.e. random forest classification models) to assess whether simulations ruled by geomorphologic resistance may be discriminated from simulations ruled by other competing processes (i.e. increase in habitat availability downstream and upstream-directed colonization). Third, we applied approximate Bayesian computation procedures to infer the most probable processes having generated the spatial patterns of genetic diversity observed for three populations of Proasellus walteri, an isopod species living in the hyporheic zone. Finally, we compared results obtained through the analysis of data simulated under different models to demonstrate the importance of considering the dendricity of riverscapes when conducting model-based genetic inferences for aquatic organisms, even when observed data has been only obtained for a single linear river.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Chromosome painting**

Authors : **Verónica Miró Pina**

Address **:**

Abstract :
Recently it has become feasible to detect blocks in the chromosome that share a common ancestry. The lengths of these IBD blocks (identical-by-descent) can be used to infer the demographic history of populations, when the mutation rates are to low to provide information about recent events (Ringbauer et al, 2017).

Our goal is to derive a neutral model for the distribution of IBD block lengths in a chromosome in order to contribute to the development of neutrality tests based on haplotype composition.

We consider a Moran model, where at time 0, all individuals of a haploid population have their unique chromosome painted in a distinct colour. At birth events, due to recombination (modelled as a single crossing-over), the chromosome of the newborn is a mosaic of its two parental chromosomes. Our goal is to describe the mosaic of colours that is fixed in the population.

This is joint work with Amaury Lambert and Emmanuel Schertzer

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Plasmid and clonal interference in experimental evolution**

Authors : **Stéphanie Bedhomme**

Address **:**

Abstract :
Plasmids are nucleic acid molecules that can drive their own replication in a living cell. Plasmid replication and plasmid gene expression consume cellular resources and carrying plasmids usually incur fitness costs. But many plasmids carry genes that can be beneficial under certain conditions, allowing the cell to endure in the presence of antibiotics, toxins, competitors or parasites. Horizontal transfer of plasmid-encoded genes can thus instantaneously confer differential adaptation to local or transient selection conditions. This conflict between cellular fitness and plasmid spread sets the scene for multilevel selection processes. We have engineered a system to study the short-term evolutionary impact of different synonymous versions of a plasmid-encoded antibiotic resistance gene. Applying experimental evolution under different selection conditions and deep sequencing allowed us to show rapid local adaptation to the presence of antibiotic and to the specific version of the resistance gene transferred. We describe the presence of clonal interference at two different levels: at the within-cell level, because a single cell can carry several plasmids, and at the between-cell level, because a bacterial population may contain several clones carrying different plasmids and displaying different fitness in the presence/absence of antibiotic. Our data suggest that, for a given gene and selection pressure, the localization on a plasmid strongly affects the evolutionary dynamics.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Phenotypic Noise and the Cost of Complexity**

Authors : **Charles Rocabert**

Address **:**

Abstract :
The phenotype of an organism results from a complex non-linear cascade of developmental, physiological and regulatory processes, formalized by the concept of genotype-to-phenotype map. An increasing number of experimental studies demonstrate the existence of phenotypic noise, possibly modulated by mutations in the genotype-to-phenotype map. Theoretical models demonstrated that under stabilizing selection, when the population is at a fitness optimum, phenotypic noise is deleterious and minimized by evolution. Phenotypic noise could however be positively selected when the population is under stressful conditions, or under directional selection, far from the fitness optimum. It was thus suggested that during an adaptation event, phenotypic noise would increase under directional selection, and then be reduced when the selection becomes stabilizing. Theoretical predictions made on the evolution of phenotypic noise under directional selection are mostly based on single character models and partial experimental observations on the expression of a single gene. Yet, in 1930, R.A. Fisher predicted with his geometric model the existence of a cost of phenotypic complexity, where beneficial mutations become increasingly harder to fix when the number of phenotypic characters increases. Is there also a cost of complexity on the phenotypic noise? To address this question, we extended Fisher's geometric model by adding an evolvable phenotypic noise. Based on an analysis of phenomic data, we allowed for evolvable correlations between noise levels on the various characters in this model. We show that a cost of complexity indeed makes phenotypic noise useless under directional selection, except if noise correlations between phenotypic characters are evolvable. We then show that such phenotypic noise evolves towards a flattened, one-dimensional configuration, with elevated noise in the direction of the fitness optimum, and minimized noise in all other directions. In this case, phenotypic noise facilitates evolution towards the fitness optimum, and significantly compensates for the cost of complexity, that usually drastically reduces the probability to fix beneficial mutations for complex phenotypes. Our results suggest that a non-isotropic phenotypic noise could be exploited by evolution, and call for further experiments to assess the functional nature of phenotypic noise.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Characterisation of biased gene conversion in mouse recombination hotspots.**

Authors : **Maud Gautier**

Address **:**

Abstract :
Meiotic recombination ensures the proper segregation of homologous chromosomes and is thus essential to the fertility of many eukaryotes. This process also facilitates adaptation by leading to the formation of new combinations of alleles and is therefore generally viewed as an evolutionary advantage.

Paradoxically, recombination hotspots, where recombination occurs, also host two forms of potentially deleterious meiotic drive: GC-biased gene conversion (gBGC) and double-strand break induced biased gene conversion (dBGC). gBGC results from a bias in the repair of AT:GC mismatches that favours the fixation of AT -> GC mutations, regardless of their impact on the fitness. This mechanism therefore plays a major role in the evolution of base composition in the vicinity of recombination hotspots and is likely to increase the genetic load. As for dBGC, evidence shows that it leads to the conversion of recombigenic alleles by non-recombigenic ones, hence leading to the self-destruction of hotspots. This process considerably influences the evolution of the landscape of recombination and is expected to damage fertility in the long term.

Even though these two phenomena are tightly intertwined and possibly tricky to dissociate, the role they both play in the evolution of genome architecture makes them crucial to understand. In order to characterise and directly quantify these events in the house mouse, we performed a sperm-typing study in this species. Our observations show that dBGC reinforces gBGC in cases of hybridisation between long-term structured populations. Our data further allow a direct quantification of segregation biases (b) in mice. In particular, we observe that b is much weaker in mice (0.05) than in humans (0.36). However, given their larger effective population size (Ne), the population-scaled intensity of gBGC is stronger in mice than in humans. These observations are consistent with the hypothesis that there might be selective contraints on b to limit the negative impact of gBGC in species with large Ne

Title : **Tempo and mode in genome size evolution and transposable elements content in fly genomes**

Authors : **Anabelle Haudry**

Address **:**

Abstract :
While the evolutionary mechanisms driving Eukaryote genome size evolution are still debated, content in transposable elements (TEs) appears to be crucial. Beside phylogenetic inertia, demography and life history traits are expected to affect the ability to regulate TEs activity through their impact on the efficacy of selection, resulting in both intra- and interspecific variation. Because flies exhibit a three-fold variation in genome size, they appear as a great model to address this question. Based on whole-genome sequencing data, the accumulation of TEs in the genomes was modelled on the phylogeny of closely related Drosophila species. The content in transposable elements appears highly correlated to genome size evolution among these species. A strong phylogenetic signal was found on the evolution of both genome size and TE content, and a genome contraction was detected in D. melanogaster subgroup. Besides, we identified several TE insertions evolving under constraints suggesting possible ancestral domestication/exaptation events. Deciphering the mechanisms at play in TE accumulation in a population, a species —at different evolutionary scales, represents a major challenge to tackle the longstanding question of genome size evolution.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Impact of GC-biased gene conversion on the adaptive substitution rate in fowls**

Authors : **Marjolaine Rousselle**

Address **:**

Abstract :
Biased gene conversion towards GC (gBGC) is a recombination-linked mechanism that can increase the probability of fixation of GC alleles around recombination hot-spots, independently of their fitness effects. As such, it is thought to mimic the effects of positive selection and to potentially cause an over-estimation of the adaptive substitution rate, as shown in primates. In birds, recent studies showed that gBGC affects the dN/dS ratio very differently from primates, and as such it may have different effects on the estimations of adaptive substitution rates. To explore this question, we used transcriptomic data of six species of fowls (Galloanserae), each represented by ten individuals, to compute dN/dS ratios, site frequency spectra and proportions of adaptive substitution rates (obtained via a Mc Donald & Kreitman-like test) for each type of mutations (AT→GC, GC→AT and G↔C/A↔T) and three categories of genes depending on their GC content. We show that gBGC strongly increases the synonymous substitution rate towards GC, but not the non-synonymous substitution rate, possibly due to an interaction with selective effects, resulting in a decrease of the dN/dS ratio when the recombination rate increases. In spite of this effect on the dN/dS ratio, we denote a slight increase of the proportion of adaptive substitutions due to gBGC. Additionally, we reveal an unexpected effect of GC content on GC conservative substitution rates, which could suggest a potential mutagenic effect of recombination.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Robustness and Evolvability: quantitative definitions in the context of modeling and simulation**

Authors: **Vincent Liard**, Jonathan Rouzaud-Cornabas, Guillaume Beslon

Address **:**

Abstract :
Robustness and evolvability concepts are knowing an increasing interest among evolutionary biology. However they are currently used merely as qualitative concepts which often lack precise quantitative definition and measures. In-silico experimental evolution comes with the ability to implement ``impossible experiments'' and to quantify ``impossible measures''. For example, it is possible to produce a virtually infinite number of offspring of a given organism, hence to locally sample its fitness landscape. This leads to a quantitative estimation of robustness and evolvability which can be further substantiated by defining it in the frame of well known quantitative genetics models, such as Fisher's Geometric Model (FGM).

We define the robustness of a given individual as the probability that an offspring has the same fitness as its parent. This property can then be split into mutational robustness (proportion of offspring that undergo a mutation but keep their ancestral fitness) and reproductive robustness (proportion of offspring that undergo no mutation during reproduction). Similarly, evolvability can be defined as the expectation of fitness improvement during reproduction while ``anti-robustness'' can be defined as the expectation of fitness loss. All these measures could be evaluated either in the context of in-silico experimental evolution or in the context of quantitative genetics models such as FGM. We claim that precisely defining such measures in the context of modelling and simulation can be a first step toward their experimental assessment in the context of real populations.

We implemented these measures in the Aevol model: by sampling a very large numbers of offspring of a given organism (up to 10 millions), we are able to quantify its robustness, evolvability and anti-robustness. Furthermore, we are able to test whether or not these properties impact subsequent evolution of the population (e.g. fitness but also later-on variations of robustness, evolvability and anti-robustness themselves). Based on these definitions and measures we propose that robustness can be selected on the long term due to positive feedback (robustness promoting robustness) but that, on the opposite, evolvability cannot be selected on the long term due to negative feedback: evolvability indeed promotes mutations that later-on reduce evolvability.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Combining relaxed clocks with gene transfers to date species trees**

Authors : **Bastien Boussau**

Address: :

Abstract :
The molecular clock hypothesis and its relaxed offspring have provided powerful ways to date clades in the tree of life. Combined with dated fossils of ancient organisms, they give insight into the diversification of life on Earth. Yet, in clades where dated fossils are scarce, their predictions are highly uncertain and imprecise. This is particularly the case during the first few billion years in the history of life when all life was unicellular and left little fossilized remains. In this talk I present a new method to improve dating estimates provided by molecular clocks when dated fossils are rare. This method is based on the analysis of Lateral Gene Transfers (LGTs).
LGTs provide a way to order in time two nodes of a species tree with respect to each other, because transfers necessarily occur between contemporaneous species. Recently new methods have been developed to identify LGTs at the genomic scale (Szollosi et al. 2013), but these LGTs have yet to be used to date phylogenies. Here I present a method that combines LGT-based relative time constraints with molecular clock models to date phylogenies. I present results obtained on simulations and discuss the information contained in relative constraints versus absolute calibrations. This new method is implemented in RevBayes.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **Gene tree species tree reconciliation with transfer and incomplete lineage sorting**

Authors : **Celine Scornavacca**

Address :

Abstract :
Gene trees and species trees can be discordant due to several processes. Standard models of reconciliations consider macro-evolutionary events at the gene level: duplications, losses and transfers of genes. However, another common source of gene tree-species tree discordance is incomplete lineage sorting (ILS), whereby gene divergences corresponding to speciations occur “out of order”. However, ILS is seldom considered in reconciliation models. In this paper, we devise a unified formal IDTL reconciliation model which includes all the abovementioned pro-cesses. We show how to properly cost ILS under this model, and then give a fixed-parameter tractable (FPT) algorithm which calculates the most parsimonious IDTL reconciliation, with guaranteed time-consistency of transfer events. Provided that the number of branches in contiguous regions of the species tree in which ILS is allowed is bounded by a constant, this algorithm is linear in the number of genes and quadratic in the number of species. This provides a formal foundation to the inference of ILS in a reconciliation framework.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Evolutionary novelty in a receptor-ligand interaction**

Authors: Maxime Chantreau, Céline Poux, Xavier Vekemans, Sylvain Billiard, **Vincent Castric**

Address **:**

Abstract :
Understanding how new phenotypes arise its a fundamental -yet difficult- question in biology because functional diversification of even simple biological systems requires coordinate evolution of many different interacting partners. In most cases, we have limited theoretical understanding and functional knowledge of the conditions and rate at which phenotypic novelties can arise and be maintained in natural populations. The highly diversified genetic system controlling self-incompatibility in outcrossing Arabidopsis species is controlled by a single genomic region that contains two genes expressed in pollen and pistils, that form a molecular lock-and-key system. The many distinct recognition specificities that segregate at this locus are each encoded by a specific combination of the lock and the key, but the conditions under which new specificities can arise has remained a central challenge in the field. We used a phylogenetic approach to reconstruct the most recent common ancestor of two currently segregating self-incompatibility specificities in Arabidopsis halleri. We then « resurrected » the ancestral receptor by expressing it into transgenic Arabidopsis thaliana plants in which we successfully transferred the self-incompatibility system. We then phenotypically compared recognition specificities and showed that one of the two contemporary descendent alleles has retained an identical specificity as the ancestor, while the other one has diverged functionally. Our results provide the first direct evidence that the rate of allelic turnover at the self-incompatibility locus is low. We are currently trying to identify the amino-acid residues that have allowed this asymmetrical shift in recognition specificity.

Title: **Maintenance of gametophytic self-incompatibility system in spatially structured population**

Authors: **Thomas Brom**

Address **:**

Abstract :
Many flowering plants exhibit a mate choice based on pollen rejection. This self-incompatibility system enables hermaphroditic plants to recognize their own pollen and prevent its germination on their own pistil. In gametophytic self-incompatibility system, self-recognition depends on a multiallelic locus, the S locus. A mutation in the S locus can disable the expression of recognition specificities in pollen, pistil or both and allowed for self-reproduction. Self-compatible mutants benefit of a better propagation because they can self-pollinate and do not reject or be rejected by any potential mate. However, individuals born from self-fecundation can suffer inbreeding depression, decreasing their fitness. Previous theoretical works show that in panmictic populations the maintenance of self-incompatibility system is possible only in a reduce parameters range of self-pollination rate and inbreeding depression. Some authors made the hypothesis that spatial structuration of populations can be a key to understand the puzzling evolution and maintenance of self-incompatibility system. We present an individual based model of hermaphroditic plants with gametophytic self-incompatibility system in a spatially structured metapopulation. We assessed the resistance of this system to the invasion by self-compatible mutants under various conditions of spatial structuration of the population. We showed that spatial structuration can have important effects on the inbreeding depression value required for self-incompatible system to resist to the invasion by a self-compatible mutant. These effect occur mainly through effective size of the population and effective allelic diversity.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : Evolutionary stasis of the pseudo-autosomal region in Strepsirhine primates

Authors : Rylan Shearn1, Emilie Lecompte2, Corinne Regis1, Guillaume Douay3, Brigitte Crouau-Roy2, **Gabriel Marais**1

Address **:**
1 Laboratoire Biométrie et Biologie Evolutive, CNRS / Univ. Lyon 1, France
2 Laboratoire Evolution et Diversité Biologique, CNRS / Univ. Toulouse, France
3 Zoo de Lyon, France

Abstract :
Based on studies of a few model organisms, mammalian sex chromosomes appear to have originated from a pair of autosomes that developed a sex-determining gene just prior to the divergence of placentals and marsupials. Since then, at several moments throughout evolutionary history, vast regions of the Y chromosome have suddenly stopped recombining with the X, forming the so-called evolutionary strata. Two such strata (4 and 5) have originated in the history of Catarrhini (Apes and Old World Monkeys) sex chromosomes, and now only a very small pseudoautosomal region continues to recombine between X and Y in those primates. Several hypotheses have been proposed to explain this trend observed in mammals and in other animal and plant lineages. One of them posits that sexually antagonistic genes may favor the suppression of recombination, for example to genetically link male-beneficial/female-detrimental genes to the Y chromosome. However, empirical evidence clearly telling apart these hypotheses are missing. Here we focus on Strepsirrhine primates (lemurs, galagos, pottos and lorises), which contrast with other studied primate so far in that they tend to show less or no sexual dimorphism. We hypothesized that if male-female differentiation were associated with the degree of X-Y recombination suppression, Strepsirrhines should show evidence of less recombination suppression compared to Catarrhini. However, very little is known about the sex chromosomes of Strepsirrhines. Using Illumina short-read sequencing technology, we have sequenced a male and a female genome in about 10 species covering the main Strepsirrhine lineages. We have then mapped the reads onto publicly available reference genomes of a few Strepsirrhine primates (using the human X to scaffold the Strepsirrhine X chromosome) and computed the male:female ratio to identify on the X chromosome the X-specific region and the pseudoautosomal region. We have found that no recombination suppressing event has occurred in the Strepsirrhines since the diversification of all primates about 65 millions years ago. All studied Strepsirrhines harbor a large pseudoautosomal region including the genes that are in PAR1, strata 4 and 5 in humans. We discuss the implications of this finding for understanding the birth of strata evolutionary in sex chromosomes.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Detecting molecular basis of phenotypic convergence**

Authors: **Gilles Didier**

Address **:**

Abstract :
Convergence is the process by which several species independently evolve similar traits.
This evolutionary process is not only strongly related to fundamental questions such as the predictability of evolution and the role of adaptation, its study also may provide new insights about genes involved in the convergent character. We focus on this latter question and aim to detect molecular basis of a given phenotypic convergence.
After pointing out a number of concerns about detection methods based on ancestral reconstruction, we propose a novel approach combining an original measure of the extent to which a site supports a phenotypic convergence, with a statistical framework for selecting genes from the measure of their sites. First, our measure of ``convergence level'' outperforms two previous ones in distinguishing simulated convergent sites from non-convergent ones.
Second, by applying our detection approach to the well-studied case of convergent echolocation between dolphins and bats, we identified a set of genes which is very significantly annotated with audition-related GO-terms. This result constitutes an indirect evidence that genes involved in a phenotypic convergence can be identified with a genome-wide approach, a point which was highly debated, notably in the echolocation case (the latest articles published on this topic were quite pessimistic).
Our approach opens the way to systematic studies of numerous examples of convergent evolution in order to link (convergent) phenotype to genotype.

Titre : **Life history traits impact the nuclear rate of substitution but not the mitochondrial rate in Isopods.**

Authors: **Nathanaelle Saclier**

Address **:**

Abstract :
DNA sequences evolve at different rates among species. Because there are widespread correlations between rate of molecular evolution and generation time, longevity and metabolic rate, life history trait changes are speculated to be a major source of molecular rate variation. However the relative contribution of each trait to this variation is poorly understood. Here, we try to disentangle the different hypotheses by using isopods that have independently made multiple transitions from surface to subterranean environments. Importantly, during these subterranean transitions species evolve higher longevity, a longer generation time and a lower metabolic rate. We assembled the transcriptomes and the mitochondrial genomes of 13 pairs of closely related isopods, with each pair composed of one surface and one subterranean species. Based on a total of 382 nuclear and 12 mitochondrial orthologous genes, we found that subterranean species have a lower rate of nuclear synonymous substitution while the mitochondrial rate remained unchanged. This unexpected result suggests that the rate of molecular evolution of these two genomes are subject to different factors. We propose that this rate decoupling between genomes comes from different replication processes. In Isopods, the nuclear rate appeared to be influenced by the generation time alone, unlike the mitochondrial rate which is independent of life history traits.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Titre : **The evolution of the temporal program of genome replication**

Authors: **Gilles Fischer**

Address **:**

Abstract :
To ensure the completion of genome doubling before cell division, eukaryotic chromosomes initiate DNA replication from multiple sites, termed replication origins. All origins activate at different times during S-phase, some firing early and others late, and have variable efficiencies, some origins activating in a majority of cells while others fire more rarely. The combination of origin location, firing time and efficiency creates, at population-level, a specific temporal program of genome replication. The now-available knowledge of temporal programs of genome replication for a dozen of yeast species raises many important questions regarding the mechanisms of replication timing remodeling during evolution. Indeed, comparative analyses revealed a nearly complete evolutionary conservation of replication programs between closely related species, in marked contrast with the comprehensive replication reprogramming observed between more distantly related species. To address this issue, we generated genome-wide replication timing profiles for ten related yeast species that cover a continuous evolutionary range from closely related to more diverged species. Comparative analysis of replication timing revealed that the replication program linearly evolves with increasing evolutionary divergence between these species. We found that the evolution of the timing program mainly results from a high evolutionary turnover rate of the cohort of active replication origins and revealed the rules that govern the birth and death, or conservation, of active replication origin over evolutionary time.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **What evolution has to say about the physics of DNA?**

Authors: **Ivan Junier**

Address **:**

Abstract :
Can evolutionary studies of transcriptional regulation unravel novel, un-exploited aspects of the physics of DNA? With this question in mind, I will discuss the unsolved problem of the engineering of transcriptional regulatory networks, which was raised more than 50 years ago by François Jacob and Jacques Monod.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Impact of DNA methylation on the rate and spectrum of mutation in Arabidopsis thaliana**

Authors: **Victoire Baillet**

Address **:**

Abstract :
DNA methylation is an epigenetic modification that is pivotal in ensuring proper genome function and integrity, notably through the silencing of transposable elements (TEs). However, as spontaneous deamination of 5-methylcytosine (5mC), which can lead to C>T transitions, is more frequent than that of unmethylated C, DNA methylation is also inherently mutagenic. This higher mutability of 5mC has indeed been proposed to explain the depletion in CpG dinucleotides in mammalian genomes, which are typically methylated at these sites except in so-called CpG islands. Despite this well-characterized effect of DNA methylation, we still lack a comprehensive view of its impact on the whole mutation spectrum in any given organism. Here, we present an in-depth assessment of the mutations that have accumulated in a population of epigenetic recombinant inbred lines (epiRILs) with almost identical genomes but heritable losses of DNA methylation at multiple TE loci in the flowering plant Arabidopsis thaliana. As anticipated, the epiRILs have accumulated numerous transposition events, that yet belong to only a small fraction of the TE families defined as mobile at the species level. Similarly, although the rate of C>T transitions is lower in the epiRILs that in classical Mutation Accumulation (MA) lines, it is at least three times lower than that expected based on the reduced number of 5mC in the epiRILs compared to the MA lines. Along with other hypotheses that are being analyzed, this reduction can be explained by the contribution of transcription-coupled DNA repair. This and other findings will be discussed.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Disentangling the complex web of interactions underlying biomass production**

Authors: **Aurélien Chateigner**

Address **:**

Abstract :
Short-rotation poplar coppice is a promising source of lignocellulosic biomass for biofuels production. Yet, one needs to study its variability to understand the underlying genetics. Furthermore, the current studies remain limited due generally to narrow genotypic screening and the involvement of complex genetic determinisms difficult to tackle.
Our study aims at disentangling the complex web of interactions underlying biomass production and quality by addressing the different angles of the problem with a system biology approach, specifically involving polymorphism, expression and phenotype evaluated within natural black poplar (Populus nigra) populations.
480 trees (240 cloned genotypes, repeated twice) from an existing field experiment based in INRA Orléans, France, selected to be representative of the genetic structure of P. nigra in Western Europe and with known phenotype for biomass yield and quality, were sequenced for their mRNA. After calling of the polymorphisms from the RNA-seq data and an extensive analysis of genes expression, a linear mixed-model approach enabled the partition of genetics and environment influence on gene expression and thus to identify 20330 potentially interesting genes. This study of high power and the reconstruction of genes networks gave us insights on 2 groups, representing 6993 genes, highly correlated with wood growth and 1 group, representing 104 genes, involved in local adaptation.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Towards understanding of anti-viral response in bats: comparative positive selection analysis combined with functional network**

Authors: **Julien Fouret** (1, 2), Noémie Aurine (2), François Enchéry (2), Marie Guinier (1), Magali Roche (1), Marc Bailly-Bechet (3), Branka Horvat (2), Catherine Legras-Lachuer (1, 4)

Address **:**
1 : Viroscan3D, Trévoux, France
2 :International Center for Infectiology Research (CIRI), INSERM U1111, CNRS UMR5308, University Lyon 1, ENS Lyon, Lyon, France
3 :Université Côte d'Azur, INRA, CNRS, ISA, France (4) : Ecologie Microbienne, CNRS UMR5557, University Lyon 1, Lyon, France

Abstract :
Bats are reported to be the reservoir of numerous viruses responsible for emerging diseases including Nipah virus (NiV), SARS-CoV or Ebola. NiV infection induces a severe zoonotic disease associated with vasculitis, acute respiratory syndrome and lethal encephalitis in a large spectrum of mammalian species, including: human horse and pig. Conversely NiV infection in its natural host the megabat Pteropus giganteus is asymptomatic. Besides, bats have been reported to have very unique biological features among mammals: the capacity of flight, longevity, increased metabolism, hibernation, echolocation as well as ability to safely host many viruses highly pathogenic for many other mammalian species. In this work, we want to identify pathways that have specifically evolved in bats and that could possibly explain the absence of virus-induced pathogenicity in bats. A comparative approach based on the detection of positively selected genes in bats (asymptomatic infection) versus a background of mammals (symptomatic infection) showed that, out of 19,200 shared and conserved genes in mammalian species, 800 are under positive selection only in megabats. Most of these selected genes are more likely related to other megabats' specific features rather than their ability to control the viral infection. This is confirmed by the fact that, after gene set enrichment analysis on identified 800 genes, immune pathways were not found significantly enriched, in contrast to other pathways related to metabolism. To overcome these limitations, we developed a methodology, based on data mining, to highlight genes under positive selection that are functionally highly connected to genes already described as the major player in NiV infection. This work led to the identification of 27 candidate genes not yet described to play a role during NiV infection and/or persistence in bats. Further bioinformatics analysis combined with functional immune-virological study should lead to a more comprehensive view of anti-viral response in bats.

Title: **Evolutionary-conserved antagonism of the housekeeping Smc5/6 complex by mammalian Hepatitis B virus.**

Authors: Fabien Filleton*, Fabien Abdul*, Laetitia Gerossier, Alexia Paturel, Janet Hall, Michel Strubin, **Lucie Etienne**§

Address **:**
1 :CIRI, Centre International de Recherche en Infectiologie, Lyon, France
2 :INSERM, U1111
3 :Ecole Normale Supérieure ENS de Lyon
4 :CNRS, UMR5308
5 :University of Lyon, Lyon I, UMS3444/US8 BioSciences Gerland
6 :Labex ANR ECOFECT of Université de Lyon
7 :Department of Microbiology and Molecular Medicine, University Medical Centre (C.M.U.) / University of Geneva, Rue Michel-Servet 1, 1211 Geneva 4, Switzerland
8 :CRCL- UMR Inserm 1052 - CNRS 5286, 151 cours Albert Thomas, 69300 Lyon, France

Abstract :
The Smc5/6 complex is, together with condensin and cohesin, one of the three related structural maintenance of chromosomes (SMC) complexes identified in eukaryotes. These complexes play distinct and essential roles in chromosome organization, segregation and topology. The Smc5/6 complex also acts as a restriction factor against human hepatitis B virus (HBV). It binds to the HBV episomal DNA genome to block viral transcription. HBV counteracts this defence mechanism by expressing the regulatory HBx protein, which interacts with and targets the Smc5/6 complex for degradation. Here, we examined whether the antiviral function of the Smc5/6 complex impacted on its evolutionary history and whether HBx-mediated degradation of the complex is conserved among mammalian hepatitis B viruses.
We performed phylogenetic and positive selection analyses of the six Smc5/6 subunits, as well as the Smc subunits of cohesin and condensin. We found that these genes have been highly conserved in primates with signatures of genetic conflict present only in Smc6. We also functionally tested the HBx proteins encoded by diverse mammalian hepatitis B viruses naturally infecting bats, rodents or primates. Despite substantial sequence divergence, all proteins efficiently degraded Smc5/6 in various cell lines and rescued the replication defect of an HBx-deficient HBV in primary human hepatocytes.
These findings point to an evolutionary conserved requirement for Smc5/6 inactivation by HBx during HBV infection in mammals. They further suggest that, despite having evolved with their host for millions of years, hepatitis B viruses have not been dominant drivers of Smc5/6 protein adaptation in mammals. It will therefore be interesting to determine whether Smc5/6

has restriction activities against other DNA viruses that may have driven some of its adaptation.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : **First genetic evidences that NTCP evolution has been driven by hepadnaviruses in bats and primates**

Authors : **Stéphanie. Jacquet**1,2,3, L. Etienne2,3,*, D. Pontier1,3*

Address
1 :Laboratoire de Biométrie et Biologie Evolutive (LBBE), UMR 5558, UCBL1, CNRS
2 : Centre International de Recherche en Infectiologie (CIRI), CNRS UMR5308, INSERM U1111, UCBL1, ENS-Lyon
3 :LabEx ECOFECT

Abstract :
Human hepatitis B virus (HBV) is a major global health problem, affecting more than 350 million people worldwide. In non-human primates and bats, which harbor hepadnaviruses closely related to HBV and is admitted as potential source of HBV, the pathogenicity of HBV-like viruses is unclear. Pathogenic viruses can affect the evolution of their host genomes. In particular, host molecular interactions with pathogenic viruses can lead to genetic “arm-races”, which can be witnessed by signatures of diversifying selection at the host-pathogen interfaces. In cellular proteins serving as entry receptors for pathogenic viruses, positive selection marks can be found at the exact sites of virus-host interactions. The sodium taurocholate cotransporting polypeptide (NTCP) is the main cellular receptor for HBV entry in hepatocyte cells. Here, we determine whether the evolution of bat and primate NTCP has been driven by hepadnaviruses, and used this as a proxy to understand if HBV-like viruses have been pathogenic in bats and primates.
For that purpose, we used a comparative approach based on the detection of positive selection and other genetic innovations (i.e. recombination and gene expansion) in bat and primate NTCP. Briefly, primate and bat NTCP orthologous sequences were retrieved from public databases. Sequences of bat (9 sp.) and primate NCTP (22 sp.) were aligned using PRANK and phylogenetic trees were obtained using PhyML. Gene-wide, lineage and site-specific positive selection analyses were carried out for bats and primates separately, using HYPHY and PAML packages. Our evolutionary analyses revealed that both bat and primate NTCP proteins have mostly evolved under purifying selection. However, we found that models allowing certain codons to evolve under positive selection are highly supported in both primate and bat NTCP. In primates, the amino acids with the strongest signal of positive selection are located in the HBV-binding region of NTCP. In bats, two positively selected sites were found near to the region of interaction. Overall, these results suggest that HBV could be the main driver of NTCP adaptation, at least in primates, and that hepadnaviruses have probably impacted primates’ fitness during their evolution. The evolutionary history of NTCP in bats was herein inferred

from only nine bat species and thus does not allow us to conclude whether or not hepadnaviruses have been pathogenic during bat evolution. To improve the power of our positive selection analyses and to characterize the evolutionary history of NTCP in bats and primates, we are currently increasing bat and primate species sampling through de novo sequencing of NTCP. Our evolutionary assumptions will be further investigated at the population level, which allow to decipher potential recent/ongoing genetic conflicts.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title : T**argets of selection and the evolution of a butterfly mimicry supergene**

Authors : **Angeles de Cara** 1, Annabel Whibley 2, Florence Piron-Prunier 3, Mathieu Joron 4

Address :
1 :Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Montpellier 5, France
2 : John Innes Centre, Norwich, United Kingdom
3 : Museum National d'Histoire Naturelle, Paris, France
4 : Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Montpellier, France

Abstract :
Butterflies' mimicry is an anti-predation strategy, relying on mimicking the lifestyle of chemically-defended prey in terms of colour-patterns, wing shape, flight, behaviour, habitat choice, life-history, and certain niche characteristics. Heliconius numata is polymorphic for up to 7 differentiated forms which combine multiple characters enhancing resemblance to distinct Melinaea species. Polymorphism is maintained by a supergene architecture, involving polymorphic, adjacent inversions. Three chromosome types coexist within population. Mimicry alleles are under directional selection for specific resemblance, but morphs mate disassortatively, suggesting deleterious variation associated with inversions. Understanding adaptation therefore requires deciphering how multiple regimes of selection act on those allelic classes. Based on 65 resequenced genomes, we looked for sites associated with phenotype, indicating the multiple putative targets of selection. Segments with reduced diversity may point to selective sweeps for mimicry or selection against deleterious mutations, or both. We combined analytical models that establish the conditions for polymorphism to emerge (under disassortative mating and balancing selection) with statistical inferences and resequence data. By combining neutrality tests within and between non-recombining allelic classes, and recently-developed methods testing for positive and background selection, we establish whether supergene evolution involves the coupling of positive mutations enhancing mimicry, or deleterious effects minimized by homozygote avoidance.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Title: **Genome-wide enhancer – target gene regulatory maps in two vertebrate genomes**

Authors: **Yves Clément** & Hugues Roest Crollius

Address:
Labo Dynamique et Organisation des Génomes (DYOGEN), Institut de Biologie de l'ENS (UMR 8197), Ecole normale supérieure, 46 rue d'ULM, 75005 Paris
Corresponding author: yclement@biologie.ens.fr, hrc@ens.fr

Abstract:
Enhancers are regulatory regions that drive the fine-scale regulation of gene expression in various tissues and life stages. Because they can regulate genes at a long distance (>100 kbps), identifying their target genes is challenging.
We previously developed a method to identify evolutionary conserved enhancers of target genes using comparative genomics in vertebrates. We applied here this method to human and zebrafish in two separate analyses and identified ~50,000 enhancers targeting ~17,000 genes in zebrafish and ~1,300,000 enhancers targeting ~18,000 genes in human. Among these, the core set of human-zebrafish conserved regulatory interactions outline genes regulated by enhancers dating back to the ancestral vertebrate, and which are associated with development and brain functions. Moreover, we find that enhancer – target gene interaction distances scales remarkably well with genome size, suggesting the absence of physical constraints associated with physical distances. Finally, we highlight for all human genes a positive association between regulation complexity and expression breath, from which we derive a model where gene expression in a particular context is driven by one or more specific enhancers.
We next studied enhancer evolution following gene duplication and found that conservation of enhancers between duplicated genes gradually with time, while the combined number of enhancers increases with time. Remarkably, both patterns are consistent with the evolution of gene expression profiles. Finally, both conservation of ancestral enhancers and acquisition of new enhancers are biased towards one gene, showing that genes are not equal within duplicated pairs.
Together, these results provide key insights into the mechanisms and evolution of gene regulation in vertebrates.

Title : **Impact of domestication on the distribution of fitness effects in several plant species**

Authors: **Maeva Mollion**

Address:

Abstract:
The amount of adaptation and selection and how these shape genome diversity is a widely studied question in biology, and more data is available especially in the field of population genomics.
Domestication offers a unique situation to study adaptation and its genomic consequences. It entails strong selection for a few specific traits but plant domestication is also often believed to incur a fairly intense genetic bottleneck that in turn can bring a substantial load of slightly deleterious mutations (sometimes termed the cost of domestication).
We analyse a set of 10 species where individuals from the wild relatives and domesticated gene pool were re-sequenced using RNA-seq (ARCAD project, collaboration with Sylvain Glemin).
We use a new method, named polyDFE (developed by Paula Tataru, Thomas Bataillon, and Sylvain Glemin), to infer the distribution of fitness effects from polymorphism. This method allowed us to infer the proportion of several class of mutations (ranging from highly deleterious to highly adaptive) within each species.
allowing us to know more about the impact of domestication on the DFE of a plant species.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

Poster Session

Lyon, 8 to 10th of November 2017 https://project.inria.fr/aiem2017/

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 1

Title :

Authors : **Rik Verdonck**

Address :

Abstract :

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 2

Blanc Guillaume

Title: **A Glimpse of Nucleo-Cytoplasmic Large DNA Virus Biodiversity through the Eukaryotic Genomics Window**

Authors : **Blanc Guillaume**

Address :

Abstract :
The nucleocytoplasmic large DNA viruses (NCLDV) are a group of extremely complex double-stranded DNA viruses, which are major parasites of a variety of eukaryotes. Recent studies showed that certain eukaryotes contain fragments of NCLDV DNA integrated in their genome, when surprisingly many of these organisms were not previously shown to be infected by NCLDVs. We performed an update survey of NCLDV genes hidden in eukaryotic sequences to measure the incidence of this phenomenon in common public sequence databases. A total of 66 eukaryotic genomic or transcriptomic datasets—many of which are from algae and aquatic protists—contained at least one of the five most consistently conserved NCLDV core genes. Phylogenetic study of the eukaryotic NCLDV-like sequences identified putative new members of already recognized viral families, as well as members of as yet unknown viral clades. Genomic evidence suggested that most of these sequences resulted from viral DNA integrations rather than contaminating viruses. Furthermore, the nature of the inserted viral genes helped predicting original functional capacities of the donor viruses. These insights confirm that genomic insertions of NCLDV DNA are common in eukaryotes and can be exploited to delineate the contours of NCLDV biodiversity.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 3

Fichant Gwennaele, Yves Quentin

Title : **TnpY1, a superfamily of single strand DNA scissors**

Authors : **Yves Quentin**, Patricia Siguier, Mike Chandler and **Gwennaele Fichant**

Address :
Laboratoire de Microbiologie et Génétique Moléculaires, UMR5100, Centre de Biologie Intégrative (CBI)
Centre National de la Recherche Scientifique (CNRS), Université de Toulouse, UPS, Toulouse F-31062, France.

Abstract :
Some mobile genetic elements have the ability to target one of the template strands during DNA replication. Well-documented examples in bacteria are insertion sequences IS608 and ISDra2 both members of the IS200/IS605 family. They use obligatory single-stranded DNA intermediates for excision and insertion and encode a transposase, TnpA, which recognizes subterminal secondary structures at the IS ends. Similar structures, REP (Repeated Extragenic Palindromes), are present in many bacterial genomes individually or in clusters. Proteins, related to the IS200/IS605 transposase, called RAYT or TnpAREP have been identified and could be responsible for REP proliferation. These proteins belong to the HUH endonuclease superfamily, share a conserved HuH/Tyrosine domain responsible for catalysis and are involved in processes of single strand DNA editing requiring cleavage and ligation of ssDNA. Members of this protein family are found in all three domains of life.
A genome-wide analysis of sequences similar to TnpAIS200/IS605 and TnpAREP in archaeal and bacterial species revealed a large number of family members. Based on sequence similarity, these can be arranged into several distinct classes and subclasses. Subclass 1.1 includes sequence similar to transposase of IS200/IS605 while proteins of the other subclasses do not present typical features of IS sequences. Each subclass is characterized by specific additional sequence domains which might be involved in protein/DNA or protein/protein interactions. More than 24% of the species analyzed encode at least one of these genes. They are even “overrepresented” in some phyla. There appears to be a positive correlation with the genome size as observed for accessory genes. Their patchy taxonomic distribution is consistent with a dissemination by multiple horizontal gene transfers followed by gene loss. The genes, of a closely related subset of subclasses, are flanked by typical REP sequences in a structure call a REPtron. The differential taxonomic distribution and the diversity of domain arrangements of the proteins belonging to different subclasses suggest

that they adapt to different cell physiology and raises the possibility that their domestication in cell function related to single strand DNA edition.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 4

Cécile Fruchard

Title :

Authors : **Cécile Fruchard**

Address :

Abstract :
Although rarer than in animals, separate sexes (dioecy) have evolved in ~15,600 angiosperm species (~6% of all species, ~20% of crops). Only two master sex-determining genes have been identified in dioecious plants so far, namely in persimmons and asparagus. In the vast majority of dioecious species, those genes are unknown, even in the well-studied papaya and Silene latifolia. Sequencing Y chromosomes is a necessary step towards identifying sex-determining genes. However, this sequencing of Y chromosomes remains one of the greatest challenges of current genomics. In sharp contrast with the >300 fully sequenced eukaryotic genomes, only a handful of Y chromosomes (mostly animal ones) have been sequenced to date. The non-recombining Y chromosome tends to accumulate repeats (transposable elements and amplicons) which makes virtually impossible the assembly using only short-read sequencing technologies. We aim at sequencing the Y chromosome of S. latifolia, a well-studied dioecious plant, in which three large sex-determining regions have been described but which represents a real challenge as the Y, in this species, is 550 Mb-long and probably comprises a very large fraction of accumulated repeats. We obtained ~95% pure S. latifolia Y DNA using flow cytometry. Sequencing of this chromosome relies on combining both short-read Illumina paired-end sequencing and PacBio plus MinION, two 3rd generation sequencing technologies providing long (>1Kb) to extra-long (>50Kb) reads to improve assembly. Assembly and annotation are ongoing. Scaffolding will be performed using different independent data sources (RNA-seq, Y-linked BACs, physical map of the Y). Annotation will be done with the help of a reference transcriptome that we have assembled previously and in which a number of Y-linked genes have already been identified.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 5

Frédéric Brunet

Title : **Unusual high rate of gene retention after four whole genome duplications among vertebrates: case studies of seven gene families**

Authors : **Frédéric Brunet**

Address :

Abstract :

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 6

Elise Kerdoncuff

Title : **Detection of strong decline in populations by a genomic approach.**

Authors : **Elise Kerdoncuff**

Address :

Abstract :
Since the explosion of human activities, a lot of species are threatened with extinction. However, 95% of described species do not have a conservation status : we don't know their extinction risk. That lack of information is due to the methods used to assessed the conservation status. These methods cannot be generalized to all species. We thus developed a model that describe the DNA of a sample of individuals in a population evolving under different demographic scenarios. Using this model, we have studied features of DNA sequences under a constant size or a strong population decline. We have developed a statistical test based on proportion of long linkage blocks, which is very sensitive to population decline. This test will be able to assign extinction risks to many of the unassessed species

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 7

MARTIN Guillaume

Title : **Insight into the hybrid mosaic structure of cultivated banana genomes**

Authors : **G. Martin**1,2, C. Cardi1,2, G. Sarah2, C. Jourda1,2,3, S. Ricci2,4, C. Jenny1,2, X. Perrier1,2, A. D'Hont1,2, J.C. Glaszmann1,2, N. Yahiaoui1,2

Address :
1 CIRAD, UMR AGAP, F-34398 Montpellier, France;
2 AGAP, Univ Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France.
3 UMR PVBMT, CIRAD, Saint-Pierre, La Réunion, France;
4 CIRAD, UMR AGAP, F-97130 Capesterre Belle-Eau, Guadeloupe

Abstract :
Cultivated bananas are derived from hybridization events among species and subspecies of the genus Musa that differentiated in various regions and archipelagos of tropical South-East Asia. These hybridizations generated diploid and triploid inter(sub)specific hybrids with impaired fertility. Some of them, having seedless parthenocarpic fruits, were selected by humans and further dispersed through vegetative propagation. The genomes of these domesticated bananas are expected to have gone through few/very limited rounds of recombination and to be organized in a mosaic of large blocks of sequences from different ancestral origins.
To characterize mosaic genome structures of diploid bananas, we generated a set of SNPs from RNAseq data of 24 selected seedy (wild) and parthenocarpic (cultivated) diploid banana accessions. We applied multivariate (COA) and SNP clustering approaches to assign alleles to ancestral banana groups. Group assigned alleles were then used to locally infer the ancestral origin of genomic regions along the 11 banana chromosomes.
The results showed instance of relatively simple hybrid mosaic structures for some accessions derived from two ancestral groups that corresponded to two Musa acuminata subspecies. Other cultivated accessions showed more complex mosaics and involved more than two ancestors. Interestingly, the results also suggested that at least one unknown ancestral group, in addition to the four main groups already identified; contributed to the studied cultivated bananas. Globally, this study showed that some of the domesticated banana cultivars originated from a more complex history of hybridization events than previously thought. Deciphering the mosaic structure of banana genomes will improve our understanding of banana domestication and help breeding programs in their strategies to reconstruct improved hybrids.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 8

Aurélien COTTIN

Title : **Evaluation of methodologies for the characterization of plant mosaic genomes**

Authors : **A. Cottin**1,2, B. Penaud1,2, G. Martin1,2, J. Santos1,2, F. Curk3, A. D'Hont1,2, J.C. Glaszmann1,2, N. Yahiaoui1,2, M. Gautier4

Address :
1 CIRAD, UMR AGAP, F-34398 Montpellier, France
2 AGAP, Univ. Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France
3 INRA, UMR AGAP, F-20230 San Giuliano, France
4 INRA, UMR CBGP, F-34988 Montferrier-sur-Lez Cedex, France

Abstract :
Hybridization events between species and subspecies are considered as major evolutionary steps, possibly contributing to the advent of new phenotypes. These events are widespread in several crop species and are expected to produce genomes with a mosaic structure of sequence blocks from different ancestry. Characterizing the inter(sub)specific mosaic structure of crop plant genomes that result from recent hybridization events can help understanding how they were formed, their domestication history, and possibly the ancestral origin of phenotypic traits. With the development of NGS genotyping technologies, several population
genomics approaches have been proposed to infer the ancestry of genome segments, by comparing polymorphism patterns across individuals along chromosomes. However, these Local Ancestry Inference (LAI) methods have mainly been developed for applications in animal models, and human most particularly. They are based on assumptions which do not always fit plant models due to more complex genome structures (e.g. different ploidy levels, variable heterozygosity levels within species) or different reproductive systems (e.g., vegetative propagation, selfing). In this context, there is a need to evaluate available methods on plant models.
To that end, we developed a small and flexible R program to simulate data under a wide variety of scenarios representative of plant model characteristics. We use this tool to evaluate two main types of LAI methods: exploratory approaches (based on multivariate analysis) and full probabilistic approaches (based on Hidden Markov Model). First results will be presented and discussed.

POSTER 9

Yasmine Mansour

Title : **How transposable elements shape mosquito genomes**

Authors : **Yasmine Mansour**[1,2,3*], Mickaël Hamouma[1,2,3], Annie Chateau[1,2] and Anna-Sophie Fiston-Lavier[3]

Address :
1 : Laboratoire d´Informatique de Robotique et de Microélectronique de Montpellier (LIRMM) - Université de Montpellier, CNRS, Montpellier, France.
2 : Institut de Biologie Computationnelle (IBC), Montpellier, France.
3 :Institut des Sciences de l'Evolution de Montpellier (ISEM) Université de Montpellier, CNRS, Montpellier, France.
*Corresponding author: yasmine.mansour@umontpellier.fr

Abstract :
Mosquitoes are human infectious disease vectors that have been extensively studied, not only because of the high genetic diversity their genomes manifest, but also for their remarkably strong capacity of fast adaptation, such as climate changes or insecticide resistance. While several studies of genes known to be involved in adaptation help to shed light on the putative role of transposable elements (TEs) in such evolution process, the impact of TEs on mosquito genome structure and evolution are still poorly tackled (Assogba et al., 2016).
Here, we carry out the study on TE abundance and distribution in mosquito genomes. In March 2017, a research group ended up with the first chromosome-length scaffolds in Culex pipiens quinquefaciatus and Aedes aegypti. We decided to start focusing on the new version of the Cx. pipiens genome assembly (CpipJ3) (Dudchenko et al., 2017). We started developing a new tool to estimate the recombination rates along chromosomes based on Marey maps (Fiston-Lavier et al., 2010) (Rezvoy et al., 2007). Our tool includes a statistical-based approach for the detection of the heterochromatin boundaries that automatically re-adjusts estimates in regions with a depletion of fitness between the polynomial and the data. After assessing the veracity of the tool with experimental data from Anopheles gambiae (Sharakhova et al., 2010), we estimated the recombination rate along the Cx. pipiens new assembly.
On the other hand, we annotated individual TE insertions in Cx. pipiens. We built a Culex specific TE library, a set of canonical sequences representative of TE families in this genome. We then annotated them combining results from homology-based (TEfam database:

https://tefam.biochem.vt.edu/) and signature-based approaches. We reported a high diversity with TE families from the three main types of TEs (DNA, LTR, non-LTR).
The annotation of individual TE insertions in CpipJ3 reveals a higher TE content compared with previous studies (33% instead of 29% for CpipJ2). Our results also showed a nonhomogenous distribution of TEs along the Cx. pipiens chromosomes with an enrichment of TEs in the heterochromatin. In-depth analysis of the TE organization is currently in process. Our results should help explaining the Cx. pipiens genome structure but also assessing the quality of the new release of the assembly.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 10

Guillaume Louvel

Title : Estimate gene duplications rates and their link to vertebrate diversification

Authors : **Guillaume Louvel,** Hugues Roest Crollius

Address :
Labo Dynamique et Organisation des Génomes (DYOGEN), Institut de Biologie de l'ENS (UMR 8197), Ecole normale supérieure, 46 rue d'ULM, 75005 Paris
Corresponding author: glouvel@biologie.ens.fr, hrc@ens.fr

Abstract :
Several mechanisms have been hypothesised to explain species divergence, from genomic incompatibilities to divergent selection pressures (1). Given the current availability of full genomes for many non-model organisms sampling various branches of the vertebrate phylogeny, we can now combine genomic data with patterns of speciation from more complete phylogenies (2). As Ohno (3) initially postulated, duplicated genes are good candidates for generating functional novelty and adaptation. We first dated duplications using dS (synonymous substitution rate) calculations, in order to obtain a fine estimation of the rate of gene duplication through time and lineages. This method however is sensitive to multiple bias (fast evolving branches, quality of gene alignments, etc). We are currently working on improving these dating method, and aim towards more sophisticated models of gene evolution that could estimate duplication ages (either adapting existing models or developping one). Our broader aim is to compare duplication rates with the diversification rates of taxons, and to assess the role of gene duplication in evolution.

[1] J. A. Coyne and H. A. Orr, "The evolutionary genetics of speciation." _Philosophical transactions of the Royal Society of London. Series B, Biological sciences_, vol. 353, no. 1366, pp. 287–305, Feb. 1998.

[2] E. Lewitus and H. Morlon, "Natural Constraints to Species Diversification," _PLOS Biology_, vol. 14, no. 8, p. e1002532, Aug. 2016.

[3] S. Ohno, _Evolution by Gene Duplication_. Springer-Verlag, 1970.

SMBE regional meeting in Lyon, Interdisciplinary Approaches for Molecular Evolution

POSTER 11

Adrian Davin

Title : **Gene transfers, like fossils, can date the tree of life**

Authors : **Adrian Davin**

Address :

Abstract :
Biodiversity has always been predominantly microbial and the scarcity of fossils from bacteria, archaea and microbial eukaryotes has prevented a comprehensive dating of the tree of life. Here we show that patterns of lateral gene transfer deduced from the analysis of modern genomes encode a novel and abundant source of information about the temporal coexistence of lineages throughout the history of life. We use new phylogenetic methods to reconstruct the history of thousands of gene families and demonstrate that dates implied by gene transfers are consistent with estimates from relaxed molecular clocks in Bacteria, Archaea and Eukaryotes. An inspection of discrepancies between transfers and clocks and a comparison with mammal fossils show that gene transfer in microbes is potentially as informative for dating the tree of life as the geological record in macroorganisms.