STARTPAGE

PEOPLE
MARIE CURIE ACTIONS

Marie Curie Intra-European Fellowships (IEF)
Call: FP7-PEOPLE-2013-IEF

PART B

"MEDIATEMP"

Part B - Table of Contents of Proposal

B1 RESEARCH AND TECHNOLOGICAL QUALITY

Research and Technological Quality

Macroevolutionary biology aims broadly to understand the origin, maintenance and diversification of life on Earth. Initially, this was primarily achieved through the examination of fossil evidence [1, 2] though this approach suffers from inherent limitations due to geographical, temporal and depositional biases in the fossil record [2]. Evolutionary biologists have thus increasingly looked for evidence of macroevolutionary patterns in phylogenetic trees using a wide range of methods (reviewed in [3-5]), especially in trees based on DNA sequences, which enable the origination times of extant taxa to be dated with increasing sophistication.

Macroevolutionary studies of phylogenetic trees in the main aim to identify whether the branching pattern seen in real trees significantly deviates from null expectations, where evolutionary potential (propensity to speciate and avoid extinction) is assumed to be the same for all species and to remain constant through time. These null expectations have typically been translated as meaning that the speciation and extinction rates do not vary across species <u>and</u> time. These expectations are tractable by a variety of analytical methods, reviewed in [4], which revealed that the diversification rate in real phylogenetic trees commonly varies among clades [6]. In some cases, this can be linked to the biological traits the clades possess [7-11] or the habitats and geographical locations they occupy [12-14]. Similarly, the diversification rate has been shown to vary through time [15-18] in various ways: single geological/climatic events coinciding with a burst [19-21] or an absence of diversification [22]; and successive temporal shifts in diversification rate being linked putatively to specific environmental drivers [23-24].

However, the study of molecular phylogenetic trees to detect macroevolutionary patterns also has a number of limitations. First, the trees ideally need to be complete with every extant species represented, which is notoriously difficult to achieve even for medium-sized phylogenies. Second, the origination dates estimated from molecular phylogenies need to be precise and unbiased, both to detect non-random diversification patterns and to potentially link these patterns to geological or climatic events [25]. Third, molecular phylogenies being based on extant species, the inference about extinction is made a lot harder as no direct evidence of extinct species is retained. Early authors claimed that the extinction rate could be estimated from phylogenies, reviewed in [26-28], but in practice the extinction rates calculated were often null or negligibly small [15], when fossil evidence suggests to the contrary that the vast majority of the species that ever lived are now extinct [29]. This has led authors to suggest that extinction rates should not be estimated from molecular phylogenies [30]. However, when rate heterogeneity across lineages is accounted for and periods of diversity declines are permitted, estimates of extinction rate can be obtained from molecular phylogenies [16] that are more plausible and congruent with the fossil record [18].

A fourth, under-appreciated, limitation relates to the null models being used to detect non-random diversification patterns. In essence, the purpose of studying macroevolution is to determine if the fate of species through time (generating new species, resisting extinction) has been influenced by intrinsic factors (such as their respective genotypes and phenotypes) and extrinsic forces (the environment they live in, the other species they coexist with). The alternative is that the fate of all species is determined purely stochastically. Crucially this doesn't necessarily imply that different species must have the same rates of speciation and extinction, or that these will remain constant through time. Species are made of populations of individuals, which may fluctuate in size for purely stochastic reasons (i.e. demographic stochasticity) and thus may reach extinction by accident [31]. We also know that speciation is likely to divide unequally a species [32] and in so doing, cause the two daughter species to have different survival expectancies. The extinction rate may also increase over time, if successive stochastic speciations create increasingly smaller geographic ranges. Finally, it has also been argued that the geographical extent of a species may influence its probability of speciation for stochastic reasons [32-33].

Some advances have been made on these more subtle stochastic ideas, most notably since Stephen Hubbell [34] published his unified neutral theory of biodiversity and biogeography (NTB hereafter). In it, Hubbell assumes that individuals of all species have the same probability of dying and producing offspring, and thus that the dynamics of a community of species is governed mainly by demographic stochasticity. Speciation is also incorporated, but in its original form, the theory assumes that an individual can speciate spontaneously into a new species, and that all individuals irrespective of species have the same probability of doing so. This mode of speciation, nicknamed point mutation, attracted a lot of criticism for being far too unrealistic [35]. In response, a number of alternative modes of speciation were investigated, notably peripatric and random fission [36-37]. However most of the effort was focused on finding analytical expressions for the number of species and the species abundance distribution expected at equilibrium [38-40] and the expected species lifetimes [36-38].

Few studies have directly investigated the impact of Hubbell's ideas on diversification patterns found in phylogenetic trees. Hubbell, himself, commented that the point mutation and random fission modes tend to produce lineage-through-time plots that are different in shape **[34, 36]**, but a formal comparison has not been published. The point mutation mode of speciation tends to generate trees with high values of tree imbalance, not unlike real trees, and significantly more so than the traditional Yule model with constant speciation and no extinction **[41]**. Lastly, spatial and non-spatial simulations of a range of neutral speciation modes **[42]** showed that neutral diversification models mimic the imbalance of real trees but that their tempo of diversification tended to accelerate through time, and more so than any real tree investigated. However, one can expect that using more realistic values of speciation rate and adding a protracted process to speciation **[43]** would bring the tempo of neutral trees within the range of real trees. In short, neutral diversification models (incorporating demographic stochasticity and various speciation modes) can offer an alternative, stochastic, explanation for the diversification patterns found in real trees. The NTB is not unique in this regard. Stochastic models of range size evolution have also been shown to generate realistic patterns of phylogenetic diversification **[32]**.

It is generally accepted that both ecological and stochastic processes are playing a role in diversification **[26]**. The issue, still largely unresolved, is to figure out their relative importance. A direct way to achieve this is by mapping traits onto a reconstructed tree and test whether the evolution of these traits matches shifts in diversification **[3, 44]**. Another, indirect, way is to model diversification purely according to stochastic processes. Any discrepancy between the patterns generated by the models and those observed for a real taxon can then in principle (if all stochastic processes have been captured in the model) be solely attributed to ecological processes, e.g. **[38]**. Much progress remains to be made in this second approach. Most stochastic studies only contemplated a small number of patterns, either phylogenetic **[42]** or macroecological/biogeographical in nature, reviewed in **[45-46]**, but rarely both (*e.g.* **[32]**). Phylogenetic patterns seem quite variable among replicates **[32, 42]** and may not have sufficient power to help distinguish between alternative diversification scenarios. Macroecological patterns may usefully complement phylogenetic patterns in this regard, but their potential remains to be assessed comprehensively. The common paradigm across all stochastic studies has been to infer the expected patterns when the system reaches its equilibrium, without establishing whether natural biotas were indeed at equilibrium. The existence of a limit to diversification is a central question in macroevolution **[47-49]** and no consensus has yet been reached. Some taxa may have reached a limit to their diversification **[50-51]**, while most others appear not to have **[16]**. Stochastic models of diversification could help scientists predict the scale and impact of future extinctions. This project will rigorously test the role that stochastic diversification models may play in understanding biodiversity dynamics both in a phylogenetic and macroecological framework and determine if biota have indeed reached their dynamic equilibrium.

Aim & Objectives.

Our aim is to understand how biodiversity patterns evolve through time and to establish whether current biodiversity patterns in real taxa have reached their equilibrium. First, we will explore exhaustively (Objective 1), via computer simulations, the *macroecological* and *phylogenetic* patterns of biodiversity, produced from a variety of stochastic diversification models. A strong emphasis will be placed (Objective 2) on examining how these patterns change over time from initial to equilibrium conditions and on measuring the time it takes for a given model to reach equilibrium. We will then examine to what extent biodiversity patterns differ between models (Objective 3) to better understand the respective influence of different stochastic processes (population demography, dispersal, speciation). Based on these results, new analytical tools will be developed (Objective 4) and tested on a few systems to identify the stochastic diversification scenario that most closely matches their biodiversity patterns and determine how far they are from reaching equilibrium. Taken together, these developments will determine if the paradigm of predicting regional and global biota at equilibrium is appropriate to understand biodiversity and if stochastic models can explain simultaneously the full range of biodiversity patterns seen in nature.

The work suggested in this proposal (see next subsection) uses concepts and tools from many disciplines in biology (evolutionary biology, population & community ecology, biodiversity, biogeography), mathematics (random processes, probability theory, complex systems), statistics (multivariate and data mining techniques) and computer science (programming, statistical computing).

Appropriateness of Research Methodology and Approach

Objective 1: Exploring the range of biodiversity patterns generated by neutral models of diversification.

The first objective is to explore and quantify the range of patterns generated by stochastic models of diversification, both in terms of phylogenetic (branching times, tree balance and tempo of diversification, spatial distribution of phylogenetic diversity) and macroecological trends (Species Area Relationships, Species Abundance Distribution, beta-diversity). Spatially explicit individual-based simulations (aka cellular automata) will be run on large square grids (1000 x 1000 cells) acting as a continuous torus, following the framework of the unified neutral theory of biodiversity and biogeography [34], with a few modifications to generate a wider range of stochastic diversification scenarios. Hubbell's scheme is chosen over alternative schemes, e.g. **[32, 52]**, due to the wealth of processes considered (birth, death, dispersal, extinction and speciation) and the easy way to simulate the dynamics forward into a cellular automaton (and backward using coalescence, see Objective 4).

The scenarios explored will differ in three essential features: dispersal, speciation rate and speciation mode. In terms of dispersal, an offspring will either be allowed to disperse anywhere on the grid with equal probability (global dispersal), or only disperse in one of the eight cells surrounding the parent cell (adjacent dispersal), or disperse in any of the five concentric layers of cells surrounding the parent cell (close-by dispersal) but with an exponentially decreasing probability for each layer away from the parent cell. In terms of speciation rate, the rate will either be made constant at the individual level (as in Hubbell's theory) or constant at the species level, for an entire run. Finally, in terms of speciation mode, the individuals belonging to the ancestral species will be divvied up between the two daughter species according to the following options. In the point mutation mode of speciation, only one individual becomes a new species and the ancestral species retain all other individuals. Alternatively, the new species ends up with a set number (10 individuals, peripheral isolate mode), a constant fraction (5%, constant fraction mode), a random fraction (random fission mode) or exactly half (equal mode) of the individuals belonging to the ancestral species. When dispersal is non-global, the partition of the individuals will also occur so that the two daughter species do not overlap spatially. Three possible values of speciation rate (at the individual level) will be used: either 10^{-7}, 10^{-6} or 10^{-5} per death-birth-dispersal event. Equivalent values of speciation rate (at the species level) leading to a similar number of species in the system at equilibrium will need to be estimated, so as to make the comparison as fair as possible.

Simulations will be run for a million generations (each made up of a million death-birth-dispersal events), which we know from experience is sufficient for the system to reach its dynamic equilibrium. I already have a computer programme written in C++ that will perform such a simulation for one run in about half an hour with a desktop PC. For each combination of (1) dispersal mode, (2) constancy of speciation rate (at the individual or species level), (3) value of speciation rate and (4) speciation mode, twenty independent simulations will be run to gauge how repeatable the patterns are from one run to another, and how variable the trend is over time within a run. This should provide ample basis to separate regular (and thus informative) patterns from volatile ones. The size of the grid used in the simulations should also help in getting patterns that are consistent.

During each simulation, the computer programme will save on file every extinction event that occurs, noting the time it occurred and the species concerned, and every speciation event, noting the time it occurred, the IDs of the ancestral and daughter species and their respective abundance. Every 10 generations, the abundance and IDs of all species in the metacommunity will also be saved in a different file so that various timelines can be constructed afterwards to estimate when the dynamic equilibrium (in terms of species richness, ecological diversity, evenness, etc.) has been reached. Finally, every 1,000 generations, the position and ID of all individuals in the metacommunity will be saved on file in order to follow how spatial distribution and biodiversity patterns change through time.

From the file containing the list of all extinctions and speciation that occurred in a given simulation, two types of dated phylogeny will be drawn. The first, named full phylogeny, displays all speciation (branching) events, all extinct branches and all the branches that survived (extant species). The second, named reconstructed phylogeny, is only based on the extant species and their evolutionary history (i.e. the branching events that link them), and therefore does not retain any information on extinction events, nor on speciation events that have an extinct clade as a branch. The full phylogeny is of course the richest in information, but it is the reconstructed phylogeny that ultimately matters the most, since the full phylogeny is not known for any taxon. The following data will be collated from these phylogenies: (1) the origination times (separately for the full and reconstructed phylogenies), (2) the extinction times (full phylogeny only), (3) the waiting times to extinction (length of time between a species origination and its extinction; full phylogeny only), (4) the waiting times to speciation (length of time between two

successive speciations; separately for the full and reconstructed phylogenies), and (5) ages of extant species (since the speciation event that created them; reconstructed phylogeny only). Flexible statistical distributions (such as three and four parameters generalizations of Weibull distribution, but other candidates exist) will then be used to summarise each of these datasets into a few salient summary statistics. In addition, lineage-through-time plots will be analysed to recover their slope and an index of concavity/convexity. Various indices of tree shape, tree topology and tempo of diversification **[4, 53]** will also be calculated. In addition, it will be possible to model speciation as a protracted process **[54]** and hence consider species that have speciated very recently to be still essentially the "same" species without having to rerun the simulation.

The following spatial patterns will also be analysed in detail: species-area relationships (SAR), species abundance distributions (SAD), phylogenetic diversity -area relationships (PDAR) and spatial turnovers in species composition and phylogenetic diversity. Species-area relationships typically follow a triphasic trend **[34]**, where species richness is shown to increase rapidly then more gradually according to a power law, then rapidly again, over the full range of possible areas. A detailed analysis of the triphasic trend shown by SARs (it is not expected under global dispersal though) will be conducted to estimate: (1) the slope of the power law relationship for intermediate areas, (2) the range of areas over which the power law applies, and (3) the magnitude of the deviation from the expectation of the power law at the very small and very large scale respectively. Species abundance distributions will be studied via rank abundance plots calculated for the entire system (*i.e.* the whole grid) and compared to theoretical abundance distribution models, such as the log-series for the point mutation mode [**34**], and the broken-stick for the random fission mode [**40**], to confirm or establish which SAD model best fits the global datasets. We will also analyse how the rank abundance plot for progressively smaller areas deviates from what would be expected from a random draw of individuals from the entire system (calculated exactly using the hypergeometric distribution) to help interpret the patterns found in SARs. The phylogenetic diversity **[55]** (expressed in generations) contained in the entire system and for increasingly smaller areas will also be calculated and this phylogenetic diversity area relationship (PDAR) compared to the SAR above. If the same triphasic trend is also shown, then a simple index describing how quickly PDAR decreases compared to SAR will be calculated. If the trend is different, another approach may be required so as to summarise its most salient features. Finally patterns of beta-diversity will be approached through indices of turnover in species composition (using either presence-absence or relative abundance data) and turnover in phylogenetic diversity **[56, 57]** to assess how quickly spatial turnover increases with geographic distance.

Objective 2: Evolution through time

One can expect that macroecological patterns will continue to change quantitatively until the system reaches its dynamic equilibrium, but that they will stabilise after that. For phylogenetic patterns, the same evolution over time can be expected, except that the signal from early events in the diversification history will be gradually lost, the longer the system spends in its dynamic equilibrium. It is important to confirm these conjectures as they affect how we compare the patterns generated by the various neutral diversification models (third objective).

For each simulation, a timeline will be drawn describing how the number of species in the system fluctuates through time. Two very clear phases can be observed, an ascending phase where species richness increase rapidly over time (in a linear fashion when the speciation rate is constant at the individual level and in a logistic manner when the speciation rate is constant at the species level) followed by a "steady state" phase where the number of species fluctuates stochastically around a particular value. A flexible asymptotic expression (such as the cumulative Weibul function with four parameters) will then be fitted to this timeline to estimate the species richness at equilibrium and how many generations the system takes to approach it (i.e. first time it is within 1% of its dynamic equilibrium). A similar approach will be used to analyse the temporal dynamics of all other biodiversity patterns to assess whether they all seem to converge towards a steady state at approximately the same moment in time. For phylogenetic patterns, we will pay particular attention in assessing whether a steady state actually exists and, if not, what shape their respective dynamics takes (ever increasing, humpbacked?). A complementary analysis will attempt to infer the rates of speciation and extinction from the reconstructed phylogeny (at different points in time) and check whether a dynamic equilibrium (indicated by a null diversification rate towards the present) can be detected. The analytical procedure based on the coalescent that allows for time variable speciation/extinction rates **[16]** and even negative diversification, would be particularly pertinent here.

Objective 3: Comparing patterns generated by the stochastic models

Different datasets will be prepared for three contrasting time periods so that the patterns generated by stochastic diversification models can be compared meaningfully. The three time periods will summarise what is happening when the system (1) is still early in its history of diversification, (2) approaches the steady state but still has some way to go, and (3) is firmly in the steady state. The first two time periods will be set as the number of generations required for the system to have reached 20% and 80% of the number of species found at steady state respectively. For the last time period, the time will be set exactly at three times the number of generations used in the second time period.

For each time period, the phylogenetic and biodiversity patterns considered above in objective 1 will be first analysed on their own to assess how much they individually vary from one neutral diversification model to another, using standard statistical techniques, such as generalised linear models and more robust alternatives such as Multiple Additive Regression Trees (MART) **[58]** and Random Forest **[59]**. Some of these patterns are known to vary with clade size (i.e. the number of extant species), so particular care will be taken at this stage to partition the variation in these patterns into a genuine difference between scenarios (in terms of diversification dynamics) and a spurious difference due to the tree size that these scenarios generate.

Patterns that show a significant amount of variation among the neutral diversification models will then be selected for a combined analysis of all patterns together. At this second stage, multivariate classification (discriminant analyses) and data mining techniques (again MART and Random Forest but as classification techniques this time) will be used to identify which feature of the simulated neutral models (dispersal, speciation rate, speciation mode) has the most impact on phylogenetic and biodiversity patterns. For instance, if one labels models only according to the dispersal mode in the diversification model, the above classification techniques will tell us how confidently (in terms of error rate) it is possible to classify the various dispersal modes on the basis of the phylogenetic and biodiversity patterns observed. They will also tell us which patterns contributed significantly to the performance of the classification. Doing this for different features (and combination of features) will therefore help rank the features in terms of their overall impact on phylogenetic and macroecological patterns. Repeating this sort of analysis using only phylogenetic patterns or only macroecological patterns will also help to determine how complementary these two sets of patterns are to understand diversification. Biodiversity patterns that show little variation among neutral diversification models but consistent differences between time periods would also be very useful to identify, as these may form the basis of new analytical tools aiming at confirming whether a given clade has reached equilibrium.

Objective 4: Devise new analytical tools

Depending on the results obtained in Objective 3, it may be possible to devise new analytical tools that would help (1) pinpoint for a given clade which stochastic diversification scenario most closely matches its biodiversity patterns, and (2) determine whether it has reached equilibrium or how far it is from doing so. At this early stage, it is difficult to anticipate what analytical approach would be most appropriate. It will strongly depend on whether information on spatial biodiversity patterns is required to supplement the information contained in phylogenetic trees. A possible approach may be to use spatially-explicit individual-based backward simulation based on coalescence **[60]** to speed up the generation of thousands of simulations, varying the parameter values and stochastic scenarios each time. The patterns generated by these simulations could then be fed into a Pattern-Oriented Modelling (POM) **[61]** or an Approximate Bayesian Computation (ABC) approach **[62, 63]** to assess the fit between simulated and observed data. We note, however, that a spatially-explicit individual-based coalescence approach will be more complex to implement when the speciation rate is not held constant at the individual level. The new analytical tools (coalescence, POM and ABC) will be tested on a new set of forward simulations (using more realistic parameters of grid size and speciation rate) acting as the "observed" data. The advantage of doing this is that we then know exactly what the target of the coalescence simulations and POM/ABC analyses ought to be in terms of diversification scenario and parameters. This will make it possible to estimate type-I and type-II error rates of these new tools, and also determine the size of phylogenies that start to have reasonable statistical power. Only then would these tools be demonstrated on a few real systems for which good quality data on phylogeny and macroecological patterns exist, such as mammals and birds. Upon completion of the fellowship, these tools would be made available to other researchers as a designated package in R.

Originality and Innovative Nature of the Project

Current research on macroevolution and biotic diversification relies extensively on a null model assuming constant rates of speciation and extinction, across species and through time, to uncover patterns of evolutionary significance, worthy of explanation. More refined stochastic models, such as Hubbell's unified neutral theory of biodiversity and biogeography, have demonstrated that phylogenetic and macroecological patterns very close to those seen in nature can be generated purely stochastically and may not require a specific explanation after all. This project will explore exhaustively for the first time the phylogenetic and macroecological patterns generated by a broad range of stochastic diversification models (not limited to the Neutral theory of Biodiversity and Biogeography). The "total evidence" provided by phylogenies and macroecological patterns will constitute a stronger basis from which to exclude the possibility that a clade has diversified purely stochastically, or failing this to identify the most likely scenario of random diversification that they followed.

Further, this project will depart from the prevailing paradigm of studying biotas at equilibrium and will, for the first time, focus on understanding how biodiversity patterns (not just species richness) evolve over time as biota diversify. This is likely to profoundly change the way biodiversity and diversification are studied in the future, given that an ecological limit to biodiversity may not actually exist.

As well as having an overall focus that is novel, the project also breaks new ground in the details of its methodology (i.e. data mining techniques) that are likely to affect how the dynamic results generated by cellular automata (widely used in ecology and population genetics) are analysed. The data mining techniques envisaged in the project have wide applicability, but, to my knowledge, have not yet been used for total evidence analysis in biodiversity. It therefore represents a welcome heuristic step of pattern triage (to exclude patterns that have little or no predictive power) before using Pattern-Oriented Modelling and Approximate Bayesian Computing to test for a significant fit between simulations and observed data. Finally the anticipated development of a spatially-explicit individual-based backward simulation approach using coalescence promises to speed up considerably simulations of biodiversity dynamics through time at realistic spatial scales. This development would then be useful to other researchers wanting to explore further stochastic biodiversity patterns.

Timeliness and Relevance of the Project

Life on earth is thought to be experiencing a sixth period of mass extinction as a direct consequence of anthropogenic actions. Studying macroevolution and biodiversity patterns are keys to understanding what the consequences of losing so many species will be in terms of loss of genetic material, ecosystem functions, resilience and evolutionary potential. There is still much we do not fully understand and this limits our ability to predict the risks we will face in the future and how best to mitigate them. This project will advance our understanding of the role that stochastic processes play in the diversification of life on earth. If these stochastic processes can be shown to describe closely major patterns of biodiversity, they could offer a first approximation of what may happen in terms of erosion of biotic diversity and loss of evolutionary potential if the level of habitat destruction and environmental change in the world continues unabated.

Most European countries have active researchers in the fields of Macroevolution and Macroecology, and the European Research Area is a recognised center of excellence at the forefront of research in these fields, alongside the USA. Internationally, other nations are increasingly active, especially Brazil, Australia, New Zealand, Japan and China. Macroevolution and Macroecology are vibrant and exciting disciplines, attracting many bright young scientists. Fostering research in these fields will therefore ensure that Europe maintains its leading position and keeps attracting foreign students wishing to obtain training in these fields, and thus contribute to boosting the European knowledge economy.

Finally, I am expecting that the personal experience and prestige gained through this project will enable me to move back into a research intensive career that I left reluctantly some 10 years ago when my wife got involved in a serious traumatic head injury, which required me to be a full time carer while holding down a full time job. She is now doing well, completing her own PhD in the next few months, and I welcome this opportunity to engage again in scientific research in Biodiversity and Macroevolution. The project and training that form the fellowship will give me the skills and experience needed to kick-start a successful career as a researcher and propel me on the path to leading an independent research group.

Host Research Expertise in the Field

The Ecole Normale Supérieure (ENS) is a highly prestigious higher education establishment that stands apart from the mainstream framework of the university system in France by its extremely selective recruitment system and high quality training it dispenses. It has trained the French elites for more than two centuries and former students include 8 French winners of the Fields Medal, 12 Nobel Prize winners and half of the recipients of the CNRS Gold Medal. Scientific research at the ENS is organized into seven departments including the Institute of Biology (IBENS) that the candidate will join during the two years of his project. The IBENS gathers 291 people: 74 staff researchers, 78 technical assistants, 69 postdocs and 70 Ph.D students). It has been ranked on the top 100 research institutes in Biology being the first French Institute to appear on the list. The IBENS is a multidisciplinary research unit belonging to the French *Centre National de la Recherche Scientifique* (CNRS, UMR 8197). The scientist in charge of the research project, Dr. Hélène Morlon (CNRS), was just recruited at the IBENS to build an interdisciplinary team in macroecology and macroevolution. She is also (since 2011) an associate member of the Center for Interdisciplinary Research in Biology (CIRB) at the Collège de France (just 10 minutes away from the ENS, walking distance), in the SMILE (Statistical Models for the Inference of Life Evolution) group led by applied mathematician Amaury Lambert. Her research 's research group works very closely with Amaury Lambert's group of theoreticians at the Collège de France (A.L. and H.M. have co-advised several master students and are currently co-advising a Ph.D student). More generally, the ENS is ideally located in downtown Paris allowing easy interactions with researchers at the Museum of Natural History, Paris 6, Paris Sud, and many other institutions. Dr. Hélène Morlon spent three years at the Center for Applied Mathematics at the Ecole Polytechnique, in the team "Modelling Biological Evolution" led by Sylvie Méléard. Interactions with applied mathematicians at the CMAP will be maintained, in particular through the "Mathematical Modelling and Biodiversity" program. Dr. Hélène Morlon also just got awarded an ERC Consolidator Grant that will allow her to recruit several postdocs who will work on problematics relevant to the applicant's proposal.

Quality of the Group/Scientist in Charge

I will work closely with Dr. Hélène Morlon. Dr. Morlon is a CNRS researcher and an associated member of the Center for Interdisciplinary Research in Biology (CIRB) at the prestigious Collège de France. Over the last few years, she has contributed greatly to the field of macroecology and macroevolution. Dr. Morlon has an interdisciplinary background. She was trained in mathematics at the École Normale Supérieure (Cachan), and she now applies her theoretical skills very broadly in the field of ecology and evolution. The impact of her work is recognized, as evidenced by her publications in high profile journals such as Science, PNAS, PLoS Biology and Ecology Letters. She recently received several invitations from prestigious journals (Science, PNAS and Ecology Letters) to contribute regular, perspective and review papers. She also has received many invitations to give keynote talks in Europe, the U.S. and elsewhere and to participate in workshops at the National Center for Ecological Analysis and Synthesis (NCEAS), the National Evolutionary Synthesis Center (NESCENT), the Royal Society in London, and the French Center for Synthesis and Analysis on Biodiversity (CESAB). She is the national representative for the International Biogeography Society in France. She is also on the editorial board of Systematic Biology and Ecology Letters. These invitations attest to the current influence Dr. Morlon is having on the field. Dr. Morlon has recently been awarded several grants, including a Consolidator grant from the European Research Council (ERC), a Chair of Excellence by the French national research agency (ANR), a grant from the French-Berkeley Fund, and a grant from the French embassy in Sao Paulo. Her current research group has a great depth of expertise and experience in statistical computing, mathematical modelling and in phylogenetic and spatial analyses of biodiversity.

With her French national research agency grant, Dr. Morlon has recently recruited two postdocs (Fabien Condamine and Dan Moen). She is also the returning host for Marie-Curie fellow Pierre-Henri Fabre, the main advisor of a Ph.D. student, Jonathan Rolland, and is co-supervising Hannah Salim, a Ph.D. student from UC Berkeley, in the context of a Chateaubriand fellowship. She is regularly supervising master students, as well as hosting internships for master and PhD students. In short, Hélène Morlon has experience supervising and training researchers at all levels and her group is actively publishing in high ranked journals (see below).

Selected recent publications by Hélène Morlon and her group

Moen & Morlon (in press in *TREE*) Why does diversification slow down?
Rolland, **Condamine**, Jiguet & **Morlon** (in press in *PLoS Biol.*) Faster speciation and reduced extinction in the
     tropics explain the mammalian latitudinal diversity gradient.

Morlon *et al.* (under review in *PNAS*) The biogeography of microbial genes for antibiotic production.
Condamine *et al.* (2013) Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16: 72-85
Morlon (2012) Microbial cooperative warfare. *Science* 334: 1184-5
Morlon *et al.* (2012) Explosive radiation of a bacterial species group. *Evolution* 66: 2577–86.
Rolland *et al.* (2012) Using phylogenies in conservation: new perspectives. *Biol. Lett.* 8: 692-4
Morlon *et al.* (2011) Reconciling molecular phylogenies with the fossil record. *PNAS* 108: 16327–32.
*Morlon *et al.* (2011) Spatial patterns of phylogenetic diversity. *Ecol. Lett.* 14:141–49. *Recommended on F1000
*Morlon *et al.(* 2010) Inferring the dynamics of diversification: a coalescent approach. *PLoS Biol*. 8:e1000493
    "Must Read" on F1000
Kembel, Cowan, Helmus, Cornwell, **Morlon**, Ackerly, Blomberg & Webb (2010) Picante: R tools for integrating
    phylogenies and ecology. *Bioinformatics* 26: 1463–4.
Morlon *et al.* (2009) Taking species abundance distributions beyond individuals. *Ecol. Lett.* 12: 488–501.
Morlon *et al.* (2008*)* A general framework for the distance-decay of similarity in ecological communities.
    Ecol. Lett. 11: 904–17.

References Cited

[1] Jablonski *et al.* 1996 *Evolutionary Paleobiology*. University of Chicago Press
[2] Benton & Harper 2009 *Introduction to Paleobiology and the fossil record*. Wiley-Blackwell
[3] O'Meara 2012 *Ann. Rev. Ecol. Evol. & Syst.* 43: 267–85   **[4]** Stadler 2013 *J. Evol. Biol.* 26: 1203–19
[5] Pennell & Harmon 2013 *Annls. NY Acad. Sci.* 1289: 90–105   **[6]** Alfaro *et al.* 2009 *PNAS* 106: 13410–4
[7] Ree 2005 *Evolution* 59: 257–65   **[8]** Biffin *et al.* 2010 *Annls. Botany* 106: 79–93
[9] Goldberg *et al.* 2010 *Science* 330: 493-5   **[10]** Van Bocxlaer *et al.* 2010 *Science* 327: 679–82
[11] Litsios *et al.* 2012 *BMC Evol. Biol.* 12: 212   **[12]** Weir 2006 *Evolution* 60: 842–55
[13] Condamine *et al.* 2012 *Ecol. Lett.* 15: 267–77   **[14]** Baker & Couvreur 2013 *J. Biogeog.* 40: 286–98
[15] Rabosky & Lovette 2008 *Evolution* 62: 1866–75   **[16]** Morlon *et al.* 2010 PLoS Biol. 8: e1000493
[17] Stadler *et al.* 2011 *PNAS* 108: 16145–6   **[18]** Morlon *et al.* 2011 *PNAS* 108: 16327–32
[19] Valente *et al.* 2010 *Proc. R. Soc. B* 277: 1489–96   **[20]** Antonelli & Sanmartin 2011 *Syst. Biol.* 60: 596–615
[21] Wang *et al.* 2012 *New Phytol.* 195: 470–8
[22] Fiz-Palacios & Valcárcel 2013 *Persp. Plant Ecol. Evol. & Syst.* 15: 130–7
[23] Steeman *et al.* 2009 *Syst. Biol*. 58: 573–85   **[24]** Near *et al.* 2012 PNAS 109: 3434–9
[25] Condamine *et al.* 2013 *Ecol. Lett.* 16: 72–85   **[26]** Nee 2006 *Ann. Rev. Ecol. Evol. & Syst.* 37: 1–17
[27] Ricklefs 2007 *TREE* 22: 601–10   **[28]** Purvis 2008 *Ann. Rev. Ecol. Evol. & Syst.* 39: 301–19
[29] Novacek 2001 *The Biodiversity Crisis: Losing What Counts*. The New Press
[30] Rabosky 2010 *Evolution* 64: 1816–24   **[31]** Lande 1993 *Am. Nat.* 142: 911-27
[32] Pigot *et al.* 2010 Syst. Biol. 59: 660–73
[33] Gaston 2003 *The Structure and Dynamics of Geographic Ranges*. Oxford University Press
[34] Hubbell 2001 *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press
[35] Ricklefs 2003 *Oikos* 100: 185–92   **[36]** Hubbell 2003 *Oikos* 100: 193-9
[37] Hubbell & Lake 2003 in *Macroecology: Concepts and Consequences*, pp. 45–63, Cambridge University Press
[38] Allen & Savage 2007 *Ecol. Lett*. 10: 637–46   **[39]** Haegeman & Etienne 2010 *Phys. Rev. E* 81: 031911
[40] Etienne & Haegeman 2011 *Theor. Ecol*. 4: 87–109
[41] Mooers *et al.* 2007 in *Reconstructing Evolution: New Mathematical and Computational Advances*, pp. 147–
    68, Clarendon Press
[42] Davies *et al*. 2011 *Evolution* 65: 1841–50   **[43]** Etienne & Rosindell 2012 *Syst. Biol.* 61: 204–13
[44] Goldberg *et al*. 2011 *Syst. Biol*. 60: 451–65   **[45]** Rosindell *et al.* 2011 *TREE* 26: 340–8
[46] Rosindell *et al.* 2012 *TREE* 27: 203–8   **[47]** Benton & Emerson 2007 *Palaeontology* 50: 23–40
[48] Wiens 2011 *Q. Rev. Biol.* 86: 75–96   **[49]** Cornell 2013 *Biol. Rev*. 88: 140–65
[50] Rabosky 2009 *Ecol. Lett*. 12: 735–43   **[51]** Hardy & Cook 2012 *Am. Nat*. 180: 438–49
[52] de Aguiar et al. 2009 *Nature* 460: 384–7   **[53]** Mooers & Heard 1997 *Q. Rev. Biol.* 72: 31–54
[54] Rosindell et al. 2010 *Ecol. Lett.* 13: 716–27   **[55]** Cadotte *et al.* 2010 *Ecol. Lett.* 13: 96–105
[56] Graham & Fine 2008 *Ecol. Lett.* 11: 1265–77   **[57]** Morlon *et al.* 2011b *Ecol. Lett.* 14: 141–9
[58] Friedman 2001 *Annls. Stat.* 29: 1189-232   **[59]** Breiman 2001 *Machine Learning* 45: 5-32
[60] Rosindell *et al.* 2007 *Ecol. Lett.* 10: 586–95
[61] Grimm & Railsback 2012 Philos. Trans. R. Soc. Lond. B 367: 298–310
[62] Jabot & Chave 2009 *Ecol. Lett.* 12: 239–48 **[63]** Csilléry et al. 2010 *TREE* 25: 410–8

B2 TRAINING

Clarity and quality of the research training objectives for the researcher.

This project, as well as undertaking ambitious and exciting scientific research, has been planned with the explicit goal of bringing the applicant to a position to develop independent research programmes on biodiversity and macroevolution, across the full spectrum of its subdisciplines (field ecology, molecular phylogenies, macroecology and biotic diversification). I am already a highly competent ecologist with advanced statistical and computing skills and considerable experience in field ecology, macroecology and biogeography. I have no doubt about my competence to complete the current project. However, there are a number of areas where formal and informal training would considerably strengthen my skills and enable me to undertake research at the cutting edge. Formal training will be directly provided by the host organisation, in the form of undergraduate and graduate (master) courses in mathematics, in particular on random processes and probability theory. I have already a good understanding of these disciplines, which I have gleaned haphazardly. A more formal and extensive coverage of these mathematical disciplines and their application would empower me to tackle future analytical challenges more effectively. These fields play an important part in understanding and predicting (1) the shape of phylogenetic trees (branching processes), (2) and the dynamics of stochastic systems. They are used extensively to simulate complex systems efficiently without having to trace the movement and behaviour of every individual. They would be essential tools to simplify at a later stage (see Objective 4 in B1) the spatially-explicit individual-based simulations that are used at the beginning of the project. Being able to simplify simulations is important to predict the biodiversity patterns expected at realistic spatial scale, because running cellular automata with billions of individuals is not a viable option. It will be particularly interesting to learn the mathematics underpinning the coalescent, as this approach (running backward in time from the present) promises to be very efficient (only extant species need to be traced backward and the number of individuals to cope with diminish steadily backward in time too). The scientist in charge, Dr. Morlon, is an expert in this field and has contributed major developments in the use of coalescence for studying macroevolution. Interacting with her and her team will be very important for me to build these new skills into my portfolio

To date, I have only used desktop computers to cover my computational needs. Access to the computer cluster of the host organisation will therefore give me an invaluable opportunity to learn how to use such resources efficiently. I would also be keen to gain some experience with parallel computing, as this becomes ever more essential to cope with genomics data and large simulations. The host organisation has a very active computer science department that will be at hand if more formal training proves necessary.

Informal training in applied mathematics, statistical computing and macroevolution will be gained through working daily alongside other researchers at the ENS (more specifically in the *Mathematical Eco-Evolution* group) and by attending regular research seminars at the host institution and nearby scientific institutions (Collège de France, Université de Paris-Sud, Muséum National d'Histoire Naturelle). By developing these skills I will be better equipped to undertake modern macroevolution and biodiversity research and will have taken full advantage of the internationally recognised expertise in modelling stochastic biological processes in the Parisian area. A number of complementary and transferable skills that would be useful to develop to achieve an independent research career are also detailed below.

Relevance and Quality of Additional Research Training as well as of Transferable Skills Offered

I have extensive experience of working in a research-led university as an academic already, but my contract over the last eight years has been exclusively devoted to teaching, with no provision at all for research. I have been devising research projects and providing project supervision to students for many years (on a variety of biodiversity topics) as part of my teaching duties but have not had the flexibility in my contract or the support from my department to engage in research more actively.

In addition to training in generic research skills that are provided through informal workshops at the host institution (on topics such as: Project planning, Time management, Intellectual property, Managing staff, Accounting, etc.), the two main areas of training I would most benefit from are grant and publication writing.

I have been successful in a number of grant applications over the years, however, my current contract as a teaching-only academic makes me ineligible to apply for mainstream research grants. Demonstrated success in obtaining funding through competitive grants would greatly strengthen any researcher's CV. I would therefore pursue the opportunity of preparing grants, under the supervision and mentorship of Dr Morlon, the scientist in

charge, and experience the whole grant writing process, from identifying funding opportunities and defining the aims to writing proposals, seeking feedback, putting together realistic budgets and responding to the referees' assessments.

I specifically want to gain further experience and mentorship in writing up scientific publications. As part of the fellowship there is scope for writing up research papers at regular intervals. The expected results will provide opportunity for at least four research papers and one or two review papers (based on the literature reviews conducted at the beginning of the fellowship). Dr Morlon has an impressive track record of producing influential papers in top journals. I am sure that working with her will prove to be very productive and will shape the way I approach scientific publishing. Her group has an ethos of collaborating on each other's papers, which is another aspect to becoming a productive scientist I wish to gain experience of.

Measures taken by the host for providing quantitative and qualitative mentoring/tutoring.

Dr Morlon, the scientist in charge, has currently two postdocs, Fabien Condamine and Daniel Moen, who she is training to become independent scientists, as mentioned in section B1. Fabien Condamine just got granted his IOF Marie-Curie proposal. Dr. Morlon is the returning host for Marie-Curie fellow Pierre-Henri Fabre, currently in Harvard in Jonathan Losos' lab. She is also supervising or co-supervising a number of PhD & Masters students. New postdocs are encouraged to develop their research with a great deal of independence through regular meetings with their mentor. Dr Morlon has gained a lot of experience while working in close succession in several US labs, which she is now imparting to those working with her. Dr Morlon will mentor me individually on a weekly basis, will maintain momentum and ensure I make the most of the opportunity given to me in this fellowship and rapidly improve my publication output. Her lab holds weekly research meetings to discuss progress and new ideas, which again will be a forum for me to learn and get feedback quickly from peers. The field of stochastic processes is vast and while I have identified three undergraduate and master courses that I will follow during this fellowship, this by no means limit me to only take these courses. With Dr Morlon's guidance, I will determine what gaps remains in my knowledge and fill them in the most time effective manner either by formal training or informal one-on-one tutoring.

Dr. Morlon is also working in close partnership with Prof. Amaury Lambert at the Collège de France, who is an expert in the mathematical analysis of stochastic processes in ecology and evolution. Visits arranged to see Prof. Lambert, will be invaluable to help me develop the new analytical tools required for objective 4.

A program of seminars and workshops is run by both the ENS and the IBENS to broaden the horizon of its staff and students. The ENS has an excellent track record of training researchers, at all levels of their career, who often go on to take research positions in academia. More generally, the ENS has a large human resource department that provides various personal development and training workshops which I would also be able to pick and choose from depending on time constraints and what Dr. Morlon and I decided was most appropriate once I was in post.

B3 RESEARCHER

Research Experience

I have experience of working on all aspects of biodiversity patterns and processes, from the local (alpha and beta diversities, maintenance processes) to the regional and even global scale (biogeography, diversification processes, biodiversity patterns). My interests cover therefore numerous disciplines: population, community and ecosystem ecology, evolution, taxonomy, systematics, macroecology and biogeography.

CURRENT POSITION (2008-PRESENT)

Senior Teaching Fellow (permanent) University of York, Biology Department

QUALIFICATIONS

PhD, Tropical Insect Biodiversity

2000, Université Libre de Bruxelles, Belgium, Highest distinction "1st Class"

BSc, Zoological Sciences (Hons)

1991, Université Libre de Bruxelles, Belgium, Highest distinction "1st Class"

HONOURS AND AWARDS

1994-98	Belgian National Science Foundation, competitive PhD scholarship (only 10 awarded nationwide across all scientific disciplines) *Action de renforcement du potentiel scientifique des institutions fédérales.* Held at the *Institut royal des Sciences Naturelles de Belgique* € 125,000
1993	Belgian Fund for Scientific Research in Industry and Agriculture grant awarded to an outstanding graduate for pilot study focused on tropical biodiversity € 12,000
1991	Graduated as top student, BSc Zoological Sciences (Université Libre de Bruxelles)
1991	Paul Brien Prize, best honours thesis in Zoology (Université Libre de Bruxelles)

PREVIOUS EMPLOYMENT

2006 - 08	*Teaching Fellow*, Biology Department, University of York, UK
2004 - 05	*Lecturer in Animal Ecology*, (Temporary) University of Aberdeen, UK
2002 - 03	*Arthropod Scientist and Laboratory Manager*, *Smithsonian Institution*'s Monitoring and Assessment of Biodiversity (SI/MAB) Program, Gamba, Gabon
2001	*Rapid Assessment Training Program Manager*, *Conservation International*, Washington DC & Eastern Kanuku Mountains, Guyana
Jan 2001	*Field Course instructor*, four weeks, for *Wildlife Conservation Society* National Training Program, Ivimka field station, Papua New Guinea.
Dec 1999	*Field Course instructor*, four weeks, for *Wildlife Conservation Society* National Training Program, Lababia, Papua New Guinea.

CAREER BREAK

In June 2002 my wife was impaled in the head in a boating accident in Gabon where I was working at the time. As a result, she suffered from serious brain injuries with dire complications (e.g. gangrene of the brain), which took her many years to recover from. She needed many operations to initially save her life and later reconstruct her face, and many hospital appointments for a long program of rehabilitation. We eventually came back to the UK (my wife is British) after a few traumatic months juggling the initiation of my wife's program of rehabilitation with the completion of my contract for the Smithsonian Institution. Pursuing short-term research contracts as a tropical biologist at this stage of my career was no longer a feasible option for us. I took a temporary position for one year as a lecturer at the University of Aberdeen to cover the sabbatical leave of a wildlife ecologist, which then led to a further 9 month contract to coordinate temporarily the MSc program. I was asked many times by her medical team to stop working to become her full time carer. Instead I took a job as a teaching fellow as it was the only permanent job on offer that could provide the continuity of medical care that she needed for her recovery. This accident was traumatic for her, but also curtailed my career as I became the sole income earner. All

my time outside of work was spent getting the care she needed at home, and when necessary take her to new experts for further medical treatment (Belgium to have her skull re-contoured, Australia for spinal rehabilitation). After 7 years, my wife not only learned to walk and read again, manage her fatigue and migraines, but also got a fellowship to return to science and will be completing her own PhD shortly. She is now well enough for me to resume a career as a research scientist. My career break from research has been difficult and terribly frustrating. I have devised and supervised short undergraduate and master projects, with little opportunity or freedom to convert these into publications. Nevertheless, it gave me a chance to experiment with different fields of research and helped me define more clearly the research area I find most stimulating and rewarding, now that long field trips in the tropics are out of the question. This period of forced reflection has highlighted to me that macroevolution, diversification and biodiversity patterns at large scale are the disciplines I want to move into. Indeed if I am successful in this grant I will happily give up a well-paid permanent position with regular annual salary increases for a 2-year fellowship to pursue the opportunity to conduct world class research again.

RECENT FUNDING

I have recently applied for three grants and been successful in all applications to pursue innovative ideas taking advantage of advances in new technologies (acoustic bat monitoring, next generation sequencing on museum specimens). These 3-year projects totalling £48,441 are coming soon to fruition, will change techniques used in their respective fields and have scope for further research. In particular, the *Oribius* project below would lend itself very well to a detailed study of the role that mountains play in the generation and preservation of diversity in the tropics. I would be keen to pursue this research further after the Marie Curie Fellowship.

2011	People's Trust for Endangered Species, UK (Irwin, Missa) Improving foraging habitat for bats in urban spaces	£ 18,315
2010	BBSRC/NERC *Syntax Initiative*, UK. (Hofreiter, Missa) Systematics of *Oribius* weevils in New Guinea.	£ 22,126
2010	Internal pump-priming funding, University of York, UK (Missa, Dytham) Real-time EcoObs bat detectors to support PhD student work.	£ 8,000

TEACHING/ACADEMIC EXPERIENCE

I have nine years' teaching experience at undergraduate and graduate level, first at the University of Aberdeen (Jan 2004 – Sept 2005, 2/3 Teaching 1/3 Research) then at the University of York (2006-present, 100% teaching). I have delivered a total of approx. 100 hours of lecture material, 160 hours of practical material and 25 days of residential field course. The broad topics covered are: Ecology, Biological Conservation, Wildlife Management, Invertebrate Biology, Mathematics for Biologists and Advanced Statistics with R. I have the highest teaching load in Biology and yet regularly achieve ratings over 4.5 (out of 5) from the students and contribute to the department maintaining its consistent ranking in the top 3 biology departments in national league tables for the quality of its teaching. I have supervised 27 Honours and 17 MSc students, on topics as diverse as ground beetle biodiversity in Scotland to simulating the diversification process according to neutral dynamics. The combination of strong ecological framework with sound statistics, programming and scientific rigour has led to good outcome for my students often going on to do PhDs at top universities in the UK. I am often approached by international PhD candidates to supervise projects, but my current position does not allow me to do this.

REFEREED PUBLICATIONS

Basset, Y., Cizek, L., Cuenoud, P., Didham, R.K., Guilhaumon, F., **Missa, O.**, Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J., Tishechkin, A.K., Winchester, N.N., Roubik, D.W., … 24 others & Leponce, M. (2012). Arthropod diversity in a tropical forest. *Science* 338(6113): 1481-4.

Irwin, N.R, Bayerlova, M., **Missa, O.** & Martinkova, N. (2012) Complex patterns of host switching in New World arenaviruses. *Molecular Ecology* 21(16): 4137-50.

Missa, O., Basset, Y., Alonso, A., Miller, S.E., Curletti, G., De Meyer, M., Eardley, C., Mansell, M.W., Wagner, T. (2009) Monitoring arthropods in a tropical landscape: relative effects of sampling methods and habitat types on trap catches. *Journal of Insect Conservation* 13(1): 103-18.

Basset, Y., **Missa, O.**, Alonso, A., Miller, S.E., Curletti, G., De Meyer, M., Eardley, C., Mansell, M.W., Novotny, V., Wagner, T. (2008) Faunal turnover of arthropod assemblages along a wide gradient of disturbance in Gabon. *African Entomology* 16(1): 47-59

Basset, Y., **Missa, O.**, Alonso, A., Miller, S.E., Curletti, G., De Meyer, M., Eardley, C., Lewis, O.T., Mansell, M.W., Novotny, V., Wagner, T. (2008) Changes in arthropod assemblages along a wide gradient of disturbance in Gabon. *Conservation Biology* 22(6): 1552-63.

Basset, Y., **Missa, O.**, Alonso, A., Miller, S.E., Curletti, G., De Meyer, M., Eardley, C., Lewis, O.T., Mansell, M.W., Novotny, V., Wagner, T. (2008) Choice of metrics for studying arthropod responses to habitat disturbance: one example from Gabon. *Insect Conservation and Diversity* 1(1): 55-66.

Basset, Y., Corbara, B., Barrios, H.,Cuenoud, P., Leponce, M., another 23 authors, **Missa, O.**, another 20 authors (2007) IBISCA-Panama, a large-scale study of arthropod beta-diversity and vertical stratification in a lowland rainforest: rationale, study sites and field protocols. *Bulletin de l'Institut royal des Sciences Naturelles de Belgique. Entomologie* 77(1): 39-69.

Basset, Y., Mavoungou, J.F., Mikissa, J.-B., **Missa, O.,** Miller, S.E., Kitching, R.L., Alonso, A. (2004) Discriminatory power of different arthropod data sets for the biological monitoring of anthropogenic disturbance in tropical forests. *Biodiversity and Conservation* 13(4): 709-32.

Basset, Y., Novotny, V., Miller, S.E., Weiblen, G.D., **Missa, O.**, Stewart, A.J.A. (2004) Conservation and biological monitoring of tropical forests: the role of parataxonomists. *Journal of Applied Ecology* 41(1): 163-74.

Novotny, V. & **Missa, O.** (2000) Local versus regional species richness in tropical insects: one lowland site compared with the island of New Guinea. *Ecological Entomology* 25: 445-51.

Tursch, B., **Missa, O.** & Bouillon, J. (1992) Studies on Olividae XIV. The taxonomic structure of *Oliva oliva* (auct.). *Apex* 7(1): 3-22.

NON-REFEREED PUBLICATIONS

Missa, O. (2005) The unified neutral theory of biodiversity and biogeography: alive and kicking. *Bulletin of the British Ecological Society* 36(2): 12-17.

Alonso, L.E., Schulenberg, T.S., Radilofe, S. & **Missa, O.** (Eds.) (2002) *A Biological Assessment of the Réserve Naturelle Intégrale d'Ankarafantsika, Madagascar*. Conservation International. RAP Bulletin 23.

Montambault, J.R. & **Missa, O.** (Eds.) (2002) *A Biodiversity Assessment of the Eastern Kanuku Mountains, Lower Kwitaro River, Guyana*. Conservation International. RAP Bulletin 26.

Missa, O. & Montambault, J.R. (2002) An Ecological, Socio-economic, and Conservation overview of the Kanuku Mountain Region, Southern Guyana. Pp.25-32 in: Montambault, J.R. & Missa, O. (Eds.) (2002) *A Biodiversity Assessment of the Eastern Kanuku Mountains, Lower Kwitaro River, Guyana*. Conservation International. RAP Bulletin 26.

Finch, D., Hinds, W., Sanderson, J. & **Missa, O.** (2002) Avifauna of the Eastern Edge of the Eastern Kanuku Mountains, Lower Kwitaro River, Guyana. Pp.43-46 in: Montambault, J.R. & Missa, O. (Eds.) (2002) *A Biodiversity Assessment of the Eastern Kanuku Mountains, Lower Kwitaro River, Guyana*. RAP Bulletin 26.

PAPERS IN PREPARATION (to be submitted in the next three months)

Missa, O. An inordinate fondness for weevils? A paragon of diversity at local, regional and global scales. *Proc. R. Soc. B* or *Biology Letters*

Missa, O. Effective specialization of weevils in the canopy of a tropical rainforest. *Journal of Tropical Ecology*

Missa, O. Complementarity and efficiency of sampling methods to survey insect diversity in the canopy of tropical forests *Biodiversity and Conservation*

Irwin, N.R. , Hugall, A.F., Moritz, C., Baldauf, S. & **Missa, O.** Evolutionary history of tube-nosed fruit bats from Australia *Molecular Phylogenetics and Evolution*.

OTHER SKILLS

In The Field: Sampling techniques for terrestrial arthropods, including canopy fogging, sticky-trap, UV light-trap, Malaise trap, single rope canopy access, radio-tracking, harp-trapping, bat mist netting and ringing, bat echolocation detection, remote camping and logistics for large research teams.

Analytical: Expert user of R, programmer using C & Python. Advanced Statistics (phylogenetically informed, spatial and multivariate statistics) and Bioinformatics (Genome Assembly).

Other: Insect alpha taxonomy, molecular techniques, scientific illustration, macrophotography of museum insect specimens. First aid trained.

Media: *Discovery Channel* 30 minute documentary about my PhD research on tropical weevils.

PROFESSIONAL ACTIVITIES

Presenter at several international conferences.

Referee for several international journals: *Aquatic Biology*, *Diversity*, *Journal of Insect Science*, *Bulletin of Entomological Research*, *Journal of Biogeography*, *Ecological Entomology*, *Methods in Ecology & Evolution*, *Journal of Applied Ecology*, *Insect Conservation and Diversity*, *Tropical Zoology*, *Journal of Insect Conservation*, *Biotropica.*

Internal Examiner, University of York. MSc Research thesis:

Bunke, M. (2010) The end-to-end approach to marine ecosystem modelling.

External Examiner of two PhD theses:

Mercado Cardenas, A. (2012) Ecology of beetles in a Panamanian tropical forest with taxonomic notes on Curculionidae and Histeridae. McGill University, Canada.

Delsinne, T (2007) Structure des assemblages de fourmis le long d'un gradient d'aridité situé dans le Chaco sec Paraguayen. Université Libre de Bruxelles, Belgium.

Professional Affiliations

British Ecological Society (2007 – present)

Research results including publications, teaching etc.

TROPICAL BIODIVERSITY

PhD on weevil diversity in the canopy of a New Guinean rainforest. I designed, funded through successful awards and grant competitions an intensive ecological study of the weevil fauna living in the canopy of a 1km^2 area of lowland rainforest in New Guinea. After 4 years of field work in the same area aided by three local assistants, I was able to demonstrate the sheer scale of arthropod diversity in tropical forests. I had collected in this single km^2 a total of 16,500 specimens belonging to 1,168 species of weevils (family Curculionidae) and showed through detailed mathematical analysis using new extrapolation techniques that the number of species likely to have been present in this patch of forest was much higher, in the region of 3,500 species. To put this into perspective, only 1,200 species of weevils had been previously recorded on New Guinea, the second largest island in the world. My work also showed that weevils at the adult stage are not very closely associated to the trees they reside on, which makes their apparent rarity (40 % of the species were represented by single individuals) all the more puzzling. I was also able to estimate the weevil diversity to be approximately 40,000 species for the island and 800,000 for the entire world (of which, only about 52,000 are currently described). Weevils can therefore rightly claim to be the most diverse group of multicellular organisms on Earth. My work also looked at how diversity was distributed spatially, and how complementary different sampling methods were. I submitted three papers, one went out to review in *Science,* before eventually being rejected, the other two were accepted subject to revision just before my wife's major accident (see details above in "career break") and I never got round to send the corrections in time. This dataset has not been superseded and still remains the most diverse collection from a single taxon anywhere in the tropics. The publications are now being reworked and will be submitted in the next few months.

Scientific expedition to the Eastern Kanuku Mountains, Guyana. I led for a month a combined field course and scientific expedition with a team of international experts to a remote region of Guyana, the Eastern Kanuku Mountains. The expedition sponsored by *Conservation International* (CI) led to the first rapid biodiversity assessment (on plants, bats, other mammals and birds) despite coinciding with the 9/11 terror attack on the World Trade Center and several experts not being able to join us. On the basis of the scientific report I edited and wrote in part on the expedition (published by CI in its RAP Bulletin), CI's country office for Guyana put pressure on the national government to support conservation measures in the region and engaged directly with local communities to help them preserve their land.

Monitoring the arthropod fauna in four habitats along a disturbance gradient in Gabon. I was contracted by the Smithsonian Institution (the National Museum of Natural History in Washington DC) to conduct over one year a detailed assessment of the arthropod fauna living in four habitats under contrasting amount of human disturbance in Gamba, Gabon sponsored by Shell International and Shell Gabon. I was responsible for leading a team of eight parataxonomists to sample arthropods across twelve sites, sort 20 taxa into morphospecies and identify the rest to family or ordinal level. I was also in charge of curating the collections until they could be sent to international experts for formal identification. Indeed, several taxonomic papers (incl. descriptions of new species) were generated from this project. Six ecological and conservation papers were written based on the results of this project; I am a first author or co-author on all of them. The main message of the project is that for tropical arthropods there is little evidence that habitats along disturbance gradients represent neat series of impoverished habitats derived from older forests. Instead, it seems that entire suites of species are being replaced as habitats are modified. The most rewarding conservation strategy may therefore be to manage tropical landscapes as mosaics of habitats (encompassing pristine and disturbed ones) rather than focus solely on pristine sites. This project formed part of a larger inventory effort of the Gamba region and indeed was so large, the US ambassadors for Cameroon, Gabon, and the Congo came to visit the laboratory while I was running it.

Arthropod diversity in a tropical forest in Panama. Recently published in *Science*, I was involved in the analysis of the largest and most detailed arthropod dataset ever assembled for a tropical forest. I was also invited to take part in the field work of this project, but had to decline to look after my wife, who was convalescing at the time. I was able to offer guidance and expertise on how to extrapolate sensible estimates of arthropod diversity from the samples (similar to my PhD work). In this article we showed that a single 60 km^2 area of rainforest is estimated to harbour conservatively as many arthropod species as an entire temperate country, such as the UK. In this paper, I also created a new method derived from first principles to test the statistical difference between several species accumulation curves. In future, researchers will therefore be able to compare simultaneously several communities (instead of just two) and determine if they are different in terms of diversity.

COMPUTATIONAL APPROACHES

Developing and running a very successful course on Advanced Statistics using R for Masters and PhD students. Five years ago, I developed a new course for Master students on how to use R for Advanced Statistics. The course covers all the essential statistical techniques that a research student ought to know about: classical linear models, generalised linear models, mixed effects models, generalised additive models and multivariate techniques of ordination and clustering. This course rapidly became popular among the students. Virtually all new ecology PhD students are now also taking this course in their first year. A former master student, who moved to the University of Edinburgh, also asked permission to use my teaching material (lectures and practicals) to share it with his fellow PhD students. I am also locally known as the expert able to troubleshoot R programming issues.

Simulating large landscapes to test the reliability of species richness estimators. I ran a project with a group of four computational biology masters students on assessing the performance of extrapolation techniques used to estimate how many species live in a habitat from a set of samples. We generated different types of landscapes (spatially homogeneous or heterogeneous) and showed that the asymptote of species accumulation curves cannot be used indiscriminately to obtain a reliable estimate of species richness (this is only valid in homogeneous environments). Instead, it is more rigorous to limit one's extrapolation to the spatial extent of the original samples. These results will be published as supplementary material in a forthcoming paper on weevil diversity at local, regional and global scale, to support the methodological approach.

Simulating the diversification processes according to Hubbell's neutral theory of biodiversity and Biogeography. I co-supervised with Prof. Calvin Dytham two different master projects on this topic. The first looked at the effect of local vs global dispersal on the phylogenies generated by stochastic diversification (assuming point mutation mode of speciation). The second expanded the scope of the first project by considering alternative modes of speciation. The preliminary results from these projects convinced me of the potential of the approach and form the basis of the current fellowship proposal.

REGIONAL AND GLOBAL BIODIVERSITY PATTERNS

Comparative biogeography of three weevil genera in New Guinea. Master project based on recorded localities in the literature for three weevil genera from New Guinea. After obtaining the coordinates and elevation of all these localities, the student was able to show that the three genera differed substantially in their altitudinal profile of diversity. One genus appeared more diverse in the lowlands but after correcting for the different areas of land available at different elevations, it was shown that mid-elevations harboured more species than their area

explains. The other two genera showed even more clearly a peak of diversity at mid elevation, suggesting that mountains may promote speciation or buffer species from extinction in periods of climate change. Only a detailed molecular phylogeny study would help distinguish between these two alternative explanations.

Next generation sequencing techniques to resolve the complex biogeography of *Oribius* weevils. In collaboration with Prof. Michael Hofreiter, we have obtained funding for a preliminary study on the potential of using next generation sequencing techniques and DNA target capture to reconstruct the evolutionary history of *Oribius* weevils, a genus of serious agricultural pests in New Guinea with an estimated 200+ species. We are due to received genomic data on a first set of species, which after analysis (de novo genome assembly) should allow us to devise an efficient protocol to capture the whole mitochondrial genome and a few nuclear markers from museum specimens. This work is currently on going.

Host associations in New World arenaviruses. I was invited to contribute to this research on the evolution of host association between arenaviruses and their hosts, mostly rodents. I was able to offer my R programming skills to retrieve host geographic range data, and suggested the use of a phylogenetically informed Mantel test, recently developed to test for significant relationships between virus genetic distance, host genetic distance and geographical proximity of hosts. The results obtained were consistent with the idea that arenaviruses switch hosts readily and, worryingly for the pathogenic clade, randomly across the rodent tree. This paper was published in *Molecular Ecology* last year.

Global biodiversity of mammals and amphibians. A number of undergraduate and master projects have been undertaking looking at (1) the distribution of diversity among the terrestrial ecoregions of the world, (2) the traits associated with evolutionary success (family richness) and (3) the global distribution of endangered species.

BIODIVERSITY OF BATS

Managing habitats for bats in an urban environment Funded by the PTES (see above) and an equipment grant at the University of York together with the York City Council and the Yorkshire Wildlife Trust, we were able to utilise new bat echolocation technology and deploy eleven bat detectors all night throughout the bat season (April to Oct) to test a simple idea. Does leaving the ground vegetation to grow through the season in public green spaces result in better foraging grounds for urban bats? The city council agreed to keep some green areas short through regular mowing or grazing, while leaving nearby areas to grow. The vast dataset (hundreds of thousands of audio files that can be assigned to species using automated software) is revealing patterns of activity that were not expected or seen before. Apart from the numerous publications that this study will generate, the results will give clear management advice to local councils trying to implement European law for protected bat species.

Ecology of tropical bats in Papua New Guinea. I have gained experience in conducting bat surveys using harp traps and mist nets, how to handle live individuals and take non-lethal biopsies from them and how to radio track bats at night in a rain forest.

Urban bat ecology. I helped devise a new standardised protocol for surveying bats in cities that lends itself to rigorous statistical analysis. Results showed that a widely cited statistical association of bat activity around white street lights may actually be an indirect correlation with other environmental features, rather than indicate bats true feeding preference.

Independent thinking and leadership qualities.

I have demonstrated initiative throughout my career from as early as designing and successfully gaining funding for my own PhD to finding ways to keep publishing science despite the limitation imposed by my current position and personal circumstances as a full time carer.

Shortly after my PhD I was hired by *Conservation International* to cover the maternity leave of the Rapid Assessment Program Manager. I was put in charge of devising a field course for local conservationists from the Guianas and then lead an international scientific expedition to the Eastern Kanuku Mountains in Guyana. My leadership abilities were tested to their limit when the 9/11 terror attack to the world trade center occurred, putting a freeze to all international flights from and to the US, effectively blocking some of our experts from joining us. The head office of CI was keen to see us all abandon the expedition and return safely. After arguing that we would be stuck in Georgetown, the capital, anyway, I received the support of CI's country office and led the expedition to a successful conclusion (see above). A year later, I was hired by the Smithsonian Institution for my organisational skills in the field, my ability to communicate fluently in French and English, and for managing and training a team of eight parataxonomists with no prior experience of working with arthropods in Gabon.

Again, despite a major disaster in my personal life, I completed my contract and ensured that the project reached its intended outcome. More recently, in my teaching fellow position at the University of York I identified a gap in the teaching of master students and seized the initiative of putting together an Advanced Statistics course using R as a platform. My leadership and organisational skills are also in evidence every year when I run a two week field course on Biodiversity techniques for Master students on the island of Majorca. I have to manage everything from flights to first aid kits, carry water to the site daily, arrange bike and vehicle hire, deal with inter-personal issues, manage budgets, while ensuring the students are safe and learn to implement short projects in the field.

I have always been the initiator in my research projects, seeking to collaborate with other researcher with close or complementary interests to mine. This fellowship will use and develop the leadership, organisational and creative skills I already have and focus them to start a career in research once more.

Match between the *fellow*'s profile and project.

The current project involves running extensive computer simulations, analysing biodiversity patterns in classical and innovative ways and developing analytical tools informed by the results we obtain. This project builds on my strengths in computational approaches, extensive knowledge of biodiversity patterns at all scales and mathematical understanding. I have eleven years' experience of working with R, as a platform to conduct sophisticated statistical analyses and as a programming and simulation environment. I have not found a numerical problem yet that I could not solve in R. I know biodiversity patterns intimately and the theories that have been proposed to explain them. I have also strong mathematical skills that I have developed on my own over the years. Calculus, matrix algebra, probability and combinatorics do not faze me. I actually enjoy using maths to crack problems. The project is also a direct extension of two projects I have run with master students.

Potential for reaching or reinforcing a position of professional maturity.

I have a depth of experience in many fields and subdisciplines. My mathematical skills have always been apparent but have never been actively nurtured or pushed to their limits. I want the opportunity to develop them further, and to be mentored by a world class scientist in her field to help me reach my full potential. I graduated the top of my class, won a competitive scholarship, embarked on an ambitious PhD, and had 3 papers in review, before my wife's traumatic brain injury (see above); my research time but not enthusiasm has been curtailed.

I am creative as demonstrated by my ability to successfully apply for grants (despite not being able to apply for mainstream funding), the range of topics I have devised for students to work on, and my contributions to the analysis and write up of papers during my wife's illness. Recently, in just a few weeks I was able to contribute significantly to papers that were published in *Science* and *Molecular Ecology.* This fellowship will therefore enable me to reach professional maturity by filling the research gap I have unwillingly suffered in my career to date. This project will provide me with many opportunities to publish papers, acquire new mathematical skills and develop new statistical tools, which will have impact in the field by enabling better predictions of biodiversity patterns.

Potential to acquire new knowledge.

In my current position in the Ecology section of the Biology department I am mathematically the person that people often come to for help with statistics for their PhD students, their projects, or general advice. I have devised and taught a course in R to MSc and PhD students. All my programming skills (in R, C++, Mathematica, Python) have been self-taught, all my advanced mathematics has also been self-learned. During my time as a teaching fellow I have had to acquire new knowledge in all kinds of areas outside my main interests to teach, write, and deliver new courses. This fellowship has been designed to build on my past knowledge and increase my skills in areas that will not only provide me with multiple avenues of exciting research but also will give me new highly desirable transferable skills in creating computer modelling tools. In my research career I have never had the opportunity to be in an academic community where mathematics is at the core of research. I have previously always worked in relative isolation, perhaps due to me being self-assured and independent. Working in Dr. Morlon's group and that of Dr. Lambert's group at the Collège de France with their focus on applied mathematics will immerse me in like-minded people that can only deepen my understanding and broaden my skill set especially in model building. I will also interact with members of the *Mathematical Eco-Evolution* group at IBENS led by Regis Ferrière, who include mathematicians, physicists and programmers. Although I am presently working in an academic environment, I strongly feel the need to be in a group with strong quantitative skills to quicken the pace of my return to science.

B4 IMPLEMENTATION

Quality of infrastructures/facilities and international collaborations of host.

The recent recruitment of Dr. Hélène Morlon at the IBENS comes with great office space, ideally located in downtown Paris. My project being mathematical and computational, I do not have specific lab requirements. The IBENS has an efficient computer cluster and an efficient bioinformatics platform. Funds will also be available to me through Dr. Morlon's ERC Consolidator grant and ANR grant (Chaire d'Excellence). The group holds weekly lab meetings. In addition, I will benefit from the proximity of the National Museum of Natural History and the University Paris 6 in Paris, the University of Paris-Sud, and the Collège de France where many conferences are hosted.

Dr. Morlon spent more than five years in the United States, in four different institutions (University of California Berkeley, University of Pennsylvania, the University of Oregon, and University of California at Merced). During this time she developed a large network of international collaborations with leading researchers in ecology and evolution, in particular through working groups at the National Center for Ecological Analysis and Synthesis (NCEAS), the Santa Fe Institute, and the Royal Society in London. Back in France since March 2011, she has both maintained her existing collaborations in the U.S. and expanded her network to Brazilian and European researchers. In the US, she has been granted a France-Berkeley fund (collaborator Charles Marshall), she has been invited to participate to a working group at the National Evolutionary Synthesis Center (NESCENT), and she is hosting a Chateaubriand Ph.D. candidate from Berkeley (Hannah Salim) for five months in 2013. In Brazil, Dr. Morlon has been granted a project to work with Tiago Quental (University of Sao Paulo). Over the last year in France, Dr. Morlon has organized two conferences attracting researchers from across North America and Europe, and she participates to a working group of the French Center for Synthesis and Analysis on Biodiversity (CESAB). I will benefit from this dynamic collaborative environment, be introduced to researchers that come through the lab, and have the opportunity to attend research conferences and work parties that I have missed out on in the last years.

Practical arrangements for the implementation and management of the research project.

At the beginning of the fellowship, I will purchase a top of the range computing workstation for my exclusive use, to work more efficiently and flexibly on the simulation and analytical components of the proposal. The Bioinformatics platform at the IBENS will provide advice on what computer specifications to choose for optimal cost/performance ratio.

Dr. Morlon will provide mentoring in mathematical analysis and backward simulation using coalescence, the two main training objectives for this project. At initial stages of the fellowship and during times of focused work on the project, Dr. Morlon and I will meet regularly, approximately once a week. At later stages (such as manuscript preparation), we will have an individual meeting every two weeks to ensure that adequate progress is being made. In addition, the laboratory as a whole has weekly meetings, and this presents a time in which general ideas for research can be discussed. Because the proposed research does not involve collecting empirical data, the project presents no difficulties with respect to obtaining permits, buying or borrowing equipment, or travel. My primary research needs will be computational and I will have access to the computer cluster at the ENS when necessary. Possible collaborations for the last objective (developing new analytical tools) will be fostered by collaborations with other postdoctoral researchers at, or collaborating with, the Center for Applied Mathematics at the Ecole Polytechnique (Dr. Morlon's former institution) and the College de France, through video conference and email correspondence and attendance at annual meetings of professional societies. In particular, I plan to attend the annual meeting of the European Society for Evolutionary Biology (ESEB) and take advantage of being in France to attend the Evolutionary Biology Meeting, held annually in Marseilles, France.

Feasibility and credibility of the project, including work plan.

The four objectives in this fellowship form a neat progression of computer simulations, pattern analyses and tool development towards the ultimate goals of (1) understanding how biodiversity patterns evolve through time and (2) establishing whether current biodiversity patterns in real taxa have reached their equilibrium.

The first objective is to run large spatially-explicit individual-based simulations according to stochastic models of diversification. I have expertise in running these kinds of simulation, through two master projects I co-supervised with Prof. Calvin Dytham at the University of York. We both contributed to writing the C++ code of these simulations. All that is required to start generating the complete set of simulations for Objective 1 is fairly minor

tweaks in the code to implement the close-by dispersal mode and a constant rate of speciation at the species level. All other elements of the stochastic models are already implemented. One simulation run over a million generations takes about 30 minutes to complete. In the fellowship we propose to explore the biodiversity patterns generated by these simulations across 90 different scenarios (3 dispersal modes x 2 levels of speciation x 3 speciation rates x 5 "division" modes during speciation). One complete set of these 90 scenarios would therefore take about 2 days to produce and the C++ code will be adapted so that the programme cycle through these scenarios automatically without human intervention. Generating the 20 replicates of each stochastic scenario thus should only take 40 days. Using a computer, especially purchased for my fellowship, that is faster than what I have had access to so far, is likely to reduce this estimate further. Conservatively, I have estimated that the completion of all these simulations would be 2 months instead in the Gantt chart below, with an additional 1 month leeway (Table 1), to account for possible hiccups or developments. I have also expertise in analysing the files generated by these simulations using R, the statistical computing platform of choice for diversification analyses. I have 11 years experience of using R as an expert user/programmer and 5 years of teaching Masters students how to use R for advanced statistical analysis at the University of York. I have written several R scripts already that will (1) infer the complete and reconstructed phylogenies from the extinction/speciation files up to any point in time (i.e. generation) along the simulation, (2) draw summary timelines (in terms of species richness, ecological diversity index, evenness index, etc.) and (3) draw the rank abundance plot for the entire system and any continuous area within it. So, the backbone necessary to analyse the simulations is already there. While running these simulations I therefore intend to spend in parallel the first four months reviewing the literature on macroecological and phylogenetic patterns. These are fields I know already well, so this is just to make sure that I have not missed any macroecological or phylogenetic patterns in the literature and that all the indices or measures that have been proposed can take part in our analyses. Alongside this literature review, I will continue writing R scripts with the view to automate the analysis of each pattern of interest and how they evolve through time. A number of R packages devoted to phylogenetic data or diversification analysis (*ape, apTreeshape, caper, cladoRcpp, DDD, diversitree, geiger, iteRates, laser, paleotree, phytools, TESS, treePar, treeSim, treeSimGM*) will be exploited to that purpose. Objective 1 is expected to be completed by the end of the 5th month, possibly 6th month in the fellowship.
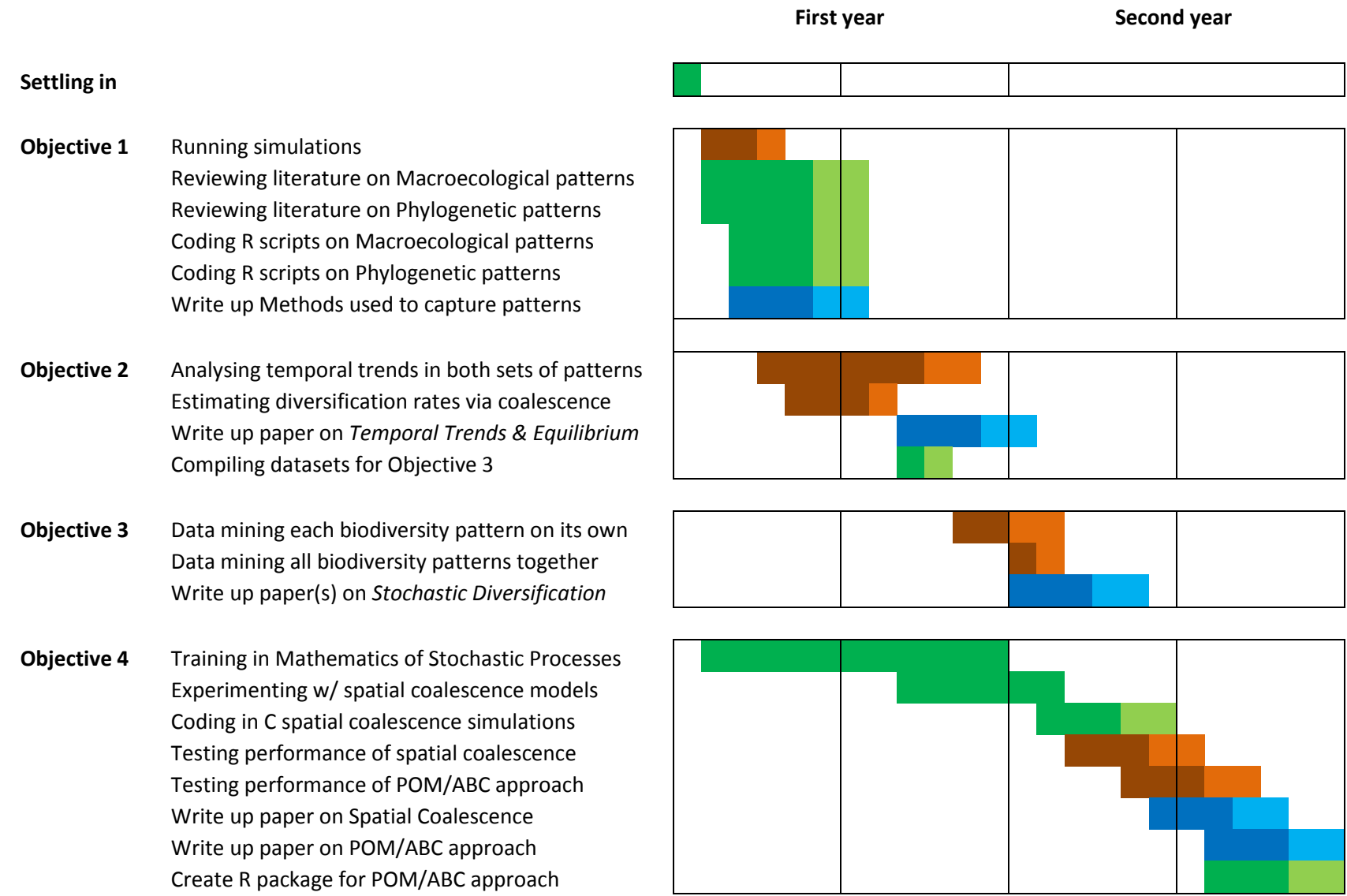
The temporal trend found in the biodiversity patterns (second objective) will be analysed using the scripts developed above, in three blocks (Phylogenetic patterns, Spatial patterns and Species abundance distribution models) across all 1800 simulations (90 scenarios x 20 replicates). The complete analysis of temporal trends are estimated to take about 4 months of continuous computing time and six months have therefore been allocated for its completion, with an extra 2 months of leeway. Estimating the speciation and extinction rates via a coalescence approach for all simulations will use the R scripts coded by Dr Morlon, my host, who invented the method (see B1). Completing these estimations should take no longer than two months and will be carried out in parallel to the temporal trend analyses using the ENS computer cluster. Writing a first paper on the temporal trends could then start at around month nine and be ready for submission at the end of the first year.

The three separate data files required to analyse through data mining (third objective) how different the patterns are between diversification scenarios in three time periods (early in the diversification, close to equilibrium and at equilibrium) will take very little time to compile, just a matter of days. Mining the data found in these files will proceed smoothly and quickly, thanks to a number of dedicated R packages (*rpart*, *gbm*, *randomForest*). Assessing how much each biodiversity pattern vary among the 90 diversification scenarios will take about 2 months, and assessing how each aspect of the stochastic diversification influence biodiversity patterns overall about 1 month. Writing a second paper on stochastic diversification and its impact on biodiversity patterns could then be started at the beginning of the second year and ready for submission around month 15.

In order to prepare for the fourth objective, devising new analytical tools, I will take a number of undergraduate and graduate courses on offer at the host institution, on the mathematics of stochastic processes, during the first year of the fellowship. Around month 9, I will start experimenting with backward simulation using coalescence, the technique most likely to be of relevance, under the guidance of my mentor, Dr. Hélène Morlon, who is an expert in this field. Based on the results of objective 3, we will have a good idea of the patterns most likely to be useful for diagnosing stochastic scenarios at the beginning of the second year. Coding the final coalescent simulation approach in C could then start around month 14 and be completed within 3 month. Developing the rest of the new analytical approach (Pattern-Oriented Modelling/Approximate Bayesian Computation framework) and testing its performance would hopefully be complete around month 18 (end of third quarter). This leaves the last quarter of the fellowship for writing papers on the new analytical approach and for putting together an R package implementing it.

Table 1: Gantt chart of activities relating to each objective. Intensive computational activities are highlighted in brown/orange, writing up activities in blue and all other activities (Training, Literature review, Computer coding) in green. The months highlighted in dark shade give an idea of the likely completion (conservatively) of each activity and the months highlighted in light shade are the available leeway that would still keep the whole project on track.

	First year	Second year
Settling in		
Objective 1	Running simulations	
	Reviewing literature on Macroecological patterns	
	Reviewing literature on Phylogenetic patterns	
	Coding R scripts on Macroecological patterns	
	Coding R scripts on Phylogenetic patterns	
	Write up Methods used to capture patterns	
Objective 2	Analysing temporal trends in both sets of patterns	
	Estimating diversification rates via coalescence	
	Write up paper on *Temporal Trends & Equilibrium*	
	Compiling datasets for Objective 3	
Objective 3	Data mining each biodiversity pattern on its own	
	Data mining all biodiversity patterns together	
	Write up paper(s) on *Stochastic Diversification*	
Objective 4	Training in Mathematics of Stochastic Processes	
	Experimenting w/ spatial coalescence models	
	Coding in C spatial coalescence simulations	
	Testing performance of spatial coalescence	
	Testing performance of POM/ABC approach	
	Write up paper on Spatial Coalescence	
	Write up paper on POM/ABC approach	
	Create R package for POM/ABC approach	

The work schedule is ambitious but achievable due to the preliminary work already completed for objective one and my extensive experience in programming, statistical computing and biodiversity patterns. There is therefore no doubt that the first three objectives will be met. The only objective that is more speculative, will be tackled later in the fellowship after I have received formal training in the mathematics of stochastic processes and guidance on how to use a coalescence approach for backward simulation (more efficient computationally than the traditional forward approaches).

The work plan has been devised with great care to avoid potential bottlenecks. The computational analyses for each objective do not overlap temporally. Likewise, the preparation of several papers on the results generated will take place in different periods, regularly through the fellowship.

Practical and administrative arrangements, and support for the hosting of the fellow.

Support for foreign researchers is very strong at the ENS, which provides many services for the large international community of students and researchers at the university. The ENS is a member of Campus France, an organization that helps visiting researchers in all capacities. The free services of Campus France assist visiting scientists with tasks that are typically complicated and difficult to do prior to arrival, including obtaining a working-scientist visa and finding housing. Campus France also provides many free services to scientists during their stay in France, including help with legal issues, health insurance, and learning the French language. My wife will gladly take up the opportunity to improve her French language skills and our daughter who will be 3 years old at the commencement of this fellowship will need a preschool place and integration into a wider community. While Campus France will help researchers find any type of accommodation, Paris also has the Cité Internationale Universitaire de Paris, a large campus of residencies for visiting international students and researchers. This community of international scientists can greatly aid visiting scientists in their initial adjustment to France.

French is my mother tongue, but I have not used it much over the last 15 years after emigrating from my native Belgium. I welcome the opportunity to use French again, daily in the workplace and as a medium to communicate Science (see section B5 Impact). The French administration (for pension, school fees, income tax, etc.) will be totally new to me, and I appreciate the access to a service like Campus France to help me and my family integrate to living in a new country.

B5 IMPACT

Impact of competencies acquired during the fellowship on the future career prospects of the researcher

Improving my already substantial skills in computer programming (R, C and Python language) and statistical computing will allow me to tackle future computational challenges with greater efficiency. Learning to use a computer cluster for the first time at the host institution, and parallel computing (particularly useful for bioinformatics and genome assembly) suits the same purpose. Gaining new skills in mathematical analysis of stochastic processes, which play an important part in modern biology (population ecology, population genetics, evolutionary biology), will likewise broaden and deepen the future studies I can pursue. These new skills added to those I already have (in field ecology, tropical biology, taxonomy, biodiversity, macroecology) and the ones I am currently developing (in molecular phylogenetics, Next Generation Sequencing, Genome Assembly) in a separate project with Prof. Michael Hofreiter (see B3) effectively mean that I will be able to design and supervise biodiversity and macroevolution projects, across the full spectrum of disciplines, at any spatial or temporal scale.

Having worked independently with minimal supervision through most of my career, the experience of working as part of a dynamic and productive lab will be invaluable for me to learn how to instil and foster a strong collaborative spirit in a team of my own. The transferable skills that I will develop and strengthen during this fellowship that are not mathematical or computational will be in terms of grant writing, project design and implementation, conference presentation and scientific publishing, all of which will help me lead a successful research group of my own in the future.

Contribution to career development, or re-establishment where relevant.

My contract for the last eight years has been 100% teaching. The research I have published in my spare time, I have done via international collaboration or through datasets I developed in contracts before my current position. Recently I have sought funding from non-mainstream funding (the only ones available to me due to my 100% teaching contract) to be able to develop more interesting research for my honours and MSc student projects. This fellowship will therefore enable me to firstly develop myself mathematically and computationally in both a formal setting through course work and through close mentorship and supervision from one of the world's leading experts, Dr. Morlon, in this field. Mentorship in an emerging field at the intersection of computational statistics, mathematical analysis and biodiversity theory in a vibrant academic institution where global talents are concentrated will be a springboard for me to return to research. I am still connected and collaborate with many international scientists study tropical insect biodiversity (e.g. the *Science* article in Dec 2012), but I need to establish new connections in the field of research I have decided to focus on in the future. Moving to and working in the laboratory at the center of current mathematical innovations in the field can only help re-establish me strongly in research. The project has been thought out carefully. A pilot study ensured that preliminary programming is in place to start working on this project without delay. Unlike many other modelling projects, this fellowship will be fail safe and outcome including multiple publications, which is vital to my medium career development are guaranteed. The only way I currently see of returning to research and re-establishing a career in science is via a fellowship. I have a family to support but the formal training and above all clear and regular mentorship that I would gain, will speed up the process for me to re-establish a full scientific research career.

Benefit of the mobility to the European Research Area.

Over my career, I have worked for museums (in Belgium and the US), international NGOs (Conservation International and Wildlife Conservation Society), and biology departments at academic institutions (Universities of Aberdeen and York) and have partnered with local governments (York City Council) too. This has given me a wide set of experiences of working with people from different background and with different aspiration. I have experience of working in Belgium (my home country), the UK, Australia, Gabon, the US (briefly in Washington DC and Hawai'i) and Guyana. I have never previously worked in France. The ENS will be a genuine different working experience for me. I have also never previously worked in a quantitative group and I am eager to see how different the scientific culture is, how people approach problems and work towards finding solutions. My experience of field ecology and museum collections will contrast with those of my colleagues and I hope to share my skills and knowledge to make the working environment beneficial to all.

On a personal level my Belgian French speaking family do not speak English and my British in-laws do not speak French, it is vital therefore that our daughter is bilingual. I am particularly keen therefore that she uses the opportunity of mobility across Europe to help her fluency in both languages.

Development of lasting cooperation and collaborations with other countries.

The likelihood of developing long-term collaboration between French research and myself (a French-speaking Belgian) is high at the end of this fellowship. Not only will I develop lasting collaboration with the host, Dr. Morlon, leader in the field of coalescence and macroevolution due to the nature of this fellowship, but also with the international members of her lab. I would like to invite to France during my fellowship Prof. Calvin Dytham (UK) with who I initiated the preliminary project to develop new ventures focusing on the role of dispersal. I would also invite James Rosindell (UK), who I know personally to collaborate on applying coalescence to a spatial context.

I would also facilitate collaboration between the large international consortium of 30 researchers from 20 countries that was involved in the *Science* paper and the French mathematical group. This consortium is producing ever increasing detailed datasets, combining phylogenetic information with biodiversity patterns across the world. My connections could help forge strong links between field ecologists and mathematicians from the host institution at the end of this fellowship that would be beneficial to all concerned.

Contribution to European excellence and European competitiveness regarding the expected research results.

This project fulfils a main objective of the Seventh Framework Programme for Research and Innovation in Environment (Biodiversity) by focussing on understanding the dynamics of biodiversity. This project is ground breaking and has no equivalent in the world. The results would confirm that Morlon's lab in particular and Europe more generally are leading in the field of Macroevolution and Macroecology. The development of rigorous tests of phylogenetic and biodiversity patterns in a single stochastic framework would be highly influential and cited as it has never been attempted before.

Testing finally whether biotas have reached their equilibrium will determine if the prevailing paradigm of predicting biotas from a position of dynamic equilibrium is appropriate to understand biodiversity. Again, research on this question has broad implications as it affects how biodiversity would be predicted to respond to perturbation, in the face of increasing anthropogenic disturbance.

The new analytical tools developed in the fellowship will be released freely for other researchers to use in a user friendly package in R. The production of such packages has been shown to foster new, unexpected, collaborations with like-minded researchers across Europe and the world. Long-term synergies can therefore be anticipated through the production and maintenance of such an R package.

Impact of the proposed outreach activities.

I have been involved in many outreach activities as a teaching fellow hosting practical experience for school children and the general public, and giving talks to the public. The discovery channel also made a 30-minute program on my PhD research. For this project the most appropriate outreach at the moment would be to design and write a 1 hour seminar for a lay audience that I could deliver in various places, e.g. amateur natural history/scientific societies, the national museum of natural history, or science open days. The seminar would introduce the main hypotheses put forward by scientists to account for how many species exist on the planet? Is it limited or could it increase indefinitely. What is the evidence from the fossil record telling us? What insights can stochastic models of diversification offer on the issue? I will be able to explain where my research fits in testing these theories using computer simulation, demonstrate how this is done, what the simulations look like and what the results show. I will also be able to explain why it is useful to know this, to help us predict the consequences of global environmental change. I enjoy engaging with the public and having an open forum at the end of the seminar to interact with people and answer their questions more fully.

I would also be eager to engage schools in France and develop a computer-based activity that teachers could take advantage of either to illustrate the use of computer programming (using R) to tackle interesting scientific questions, or to help discuss the concepts of equilibrium in a science class. I will need to put some effort in finding the right forum to achieve this, as I am not familiar with the French school system.

B6 ETHICS ISSUES **(No page limit)**

The current fellowship proposal has no ethics issue. The work proposed is almost exclusively numerical and based on computer simulations. Towards the end, a few datasets on real phylogenies and biodiversity patterns on plants and animals will be analysed. When it comes to select the datasets to analyse, we will either choose datasets that have been released into the public domain and are accessible freely online, or we will contact the researchers who put the datasets together to discuss the possibility of using their datasets.

ETHICS ISSUES TABLE

Research on Human Embryo/Foetus	YES	Page
Does the proposed research involve human Embryos?		
Does the proposed research involve Foetal Tissues/Cells?		
Does the proposed research involve human Embryonic Stem Cells (hESCs)?		
Does the proposed research on human Embryonic Stem Cells involve cells in culture?		
Does the proposed research on Human Embryonic Stem Cells involve the derivation of cells from Embryos?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	YES	

Research on Humans	YES	Page
Does the proposed research involve children?		
Does the proposed research involve patients?		
Does the proposed research involve people not able to give consent?		
Does the proposed research involve adult healthy volunteers?		
Does the proposed research involve Human genetic material?		
Does the proposed research involve Human biological samples?		
Does the proposed research involve Human data collection?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	YES	

Privacy	YES	Page
Does the proposed research involve processing of genetic information or personal data (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?		
Does the proposed research involve tracking the location or observation of people?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	YES	

Research on Animals	YES	Page
Does the proposed research involve research on animals?		
Are those animals transgenic small laboratory animals?		
Are those animals transgenic farm animals?		
Are those animals non-human primates?		
Are those animals cloned farm animals?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	YES	

Research Involving Developing Countries	YES	Page
Is any material used in the research (e.g. personal data, animal and/or human tissue samples, genetic material, live animals, etc.): a) Collected and processed in any of the ICPC countries?		
b) Exported to any other country (including ICPC and EU Member States)?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	YES	

ENDPAGE

PEOPLE
MARIE CURIE ACTIONS

**Marie Curie Intra-European Fellowships (IEF)
Call: FP7-PEOPLE-2013-IEF**

PART B

"MEDIATEMP"