Graduate School

ED227 MNHN (National Museum of Natural History)

PhD application process

http://enseignementsuperieur.mnhn.fr/fr/enseignement-superieur/doctorat/concours-ed227

http://edu-guides.mnhn.fr/guidedoctoratUK/HTML/#

PhD. Project title

Recostructing detailed admixture histories in human populations having undergone the transatlantic slave trade

Lab

UMR 7206 Éco-anthropologie et Ethnobiologie

Evolutionary Anthropology Team

PhD advisors

Bruno Toupance (bruno.toupance@mnhn.fr)

http://www.ecoanthropologie.cnrs.fr/article542.html?lang=fr

Paul Verdu (paul.verdu@mnhn.fr)

http://www.ecoanthropologie.cnrs.fr/article203.html?lang=fr

PhD. project

Genetic admixture between populations is one of the fundamental mechanisms of the biological evolution of *Homo sapiens*. Population genetics allows, through the study of current genetic diversity, to reconstruct the past history of migrations and admixture in our species (Malick et al. *Science* 2016). In particular, migrations between Europe, Africa and the Americas since the 15th century, in the context of European colonial expansion and the Slave Trade, are the subject of a growing number of studies (Patin et al. *Science* 2017). They made it possible to identify the continental African, American and European, often unknown, geographical origins of enslaved populations. They also highlighted complex processes of asymmetrical sex-specific admixture between female slaves of African descent and men of European descent (Fortes-Lima et al. *AJHG* 2017).

Cape Verde represents an ideal model for this type of historical migrations (Belezza et al. *PLoS Genet* 2013, Verdu et al. *Curr Biol* 2017). This archipelago, off the coast of Senegal, was the first European settlement in sub-Saharan Africa and one of the slave trade hubs in the East Atlantic. It took more than 350 years for European settlers and deported African slaves to people the various islands of Cape Verde. Yet the origins of deported slaves remain poorly understood. The history of founding effects and bottlenecks that have influenced the genetic diversity of each island is also unexplored.

In order to reconstruct these admixture histories, population geneticists have developed inference methods based on allelic frequency spectrum and linkage disequilibrium, based on genomic data (Gravel *Genetics* 2012, Hellenthal et al. *Science* 2014). Although very efficient, these likelihood maximization methods do not take into account realistic models of genetic admixture, recurring over time or asymmetrical according to the sex of the individuals.

In this context, the ANR METHIS project, initiated in 2016, aims to develop methodological tools for the inference of complex admixture processes based on current genetic data. These new tools can then be applied to reconstruct the history of the Cape Verde settlement using newly generated genome-wide data from these islands. Thus, we have developed a novel software ("METHIS") to simulate and analyze genome-wide genetic data under a very large number of models of possible admixture. These models consider admixture processes with varying intensity over time, recurring admixture at each generation, or independent admixture pulses (Verdu et al. *Genetics* 2011). METHIS is coupled with existing methods of Approximate Bayesian Computation (Pudlo et al. *Bioinfo* 2015). This makes it possible to identify the model that produced the simulations closest to the observed data, and to infer the demographic parameters of this model. In parallel, we genotyped 2.5 million SNPs in more than 200 individuals collected between 2010 and 2017 and born in the different islands of the archipelago.

The aim of the PhD thesis is to reconstruct in detail the demographic history of settlement and admixture on each island of Cape Verde. The Ph.D. student will describe the data using traditional population genetics tools, characterizing genetic diversity, levels of differentiation and patterns of admixture within and between each island. He/She will compare these new data with those available publicly (1000 Genomes, African Genome Diversity Project), containing African, European and populations descending from slavery in the Americas. Finally, he/she will use the existing methods of inference as well as the METHIS software to reconstruct the history of each Cape Verde islands from autosomal data.

In parallel, the PhD. student will develop and publish version 2 of the METHIS software. To date, the software makes it possible to study autosomal markers only. He/She will implement in the simulator uniparental (mitochondria, Y chromosome) and biparental (X chromosome) data. In fact, asymmetrical sex-specific admixture processes leave a specific signature in the genetic diversity of these chromosome compartments (Goldberg et al. *Genetics* 2015). The METHIS v2 "sex-linked" software will be used to reconstruct the history of asymmetrical sex-specific admixture in Cape Verde and, more broadly, from public data, in populations resulting from slavery on both sides of the Atlantic.

Publication Strategy

Three articles will be considered for publication during this thesis (years 2 and 3).

A methodological article on the METHIS v2 software and the influence of gender-based mixing processes on genetic diversity. Newspapers: Genetics, Bioinformatics, etc.

Two genome-wide human-data analysis articles aimed at reconstructing the history of complex admixture in Cape Verde during colonization and slavery regarding: 1. the origins and history of the settlement of each island; 2. Sex-specific admixture in Cape Verde and other descendants of slavery.

Targeted Journals: Molecular Biology and Evolution, Proceedings of the Royal Society B, etc.

Presentations at national and international conferences (Evolution, ESEB, SMBE, etc.)

Risk Management Strategy

Cape Verde genetic data is available and the quality control steps have already been successfully completed.

If the inference of the parameters underlining admixture process proves to be difficult in ABC, we will use existing methods implemented in the software TRACTS (Gravel 2012) and GLOBETROTTER (Hellenthal et al., 2014).

The simulation aspects are not risky and will be oriented towards questions on the evolution of the genetic diversity linked to admixture. We will orientate the project to analyse other human populations available in public databases, and can open new projects for other species than *Homo sapiens*. For example, in parallel with this PhD. thesis, we are conducting several collaborations on different populations of hybrid plants in Africa and Europe.

3 Years Schedule

The project is grounded in an ANR funding led by P. Verdu and of which B. Toupance is a collaborator. Computer resources and workplace at the Musée de l'Homme are available. Cape Verde's genomic data are available and have passed the quality controls. METHIS v1 is already available and operational; it will be released in the first half of 2018.

The project is therefore feasible in 3 years; in addition, it benefits from the collaboration of other members of the UMR involved in the ANR.

Year 1: development METHIS v2; description data Cape Verde and history reconstruction with METHIS v1

Year 2: publication METHIS v2; publication History Cape Verde; Reconstruction history of sex-specific admixture with METHIS v2

Year 3: publication sex-specific admixture history of Cape Verde;

Writing and defense of the PhD. thesis

<u>Candidate Profile</u>

Master 2 level or equivalent is expected in population genetics and evolutionary genetics, bio-statistics, and bio-informatics. Marked interest for cultural and biological anthropology will be appreciated

Computer programing skills:

C, Python, R : very good level

English written and spoken:

good level