

博士論文（要約）

Genomic analysis for understanding genetic basis of adaptation to novel and variable environments

（ゲノム解析による新規および多様な環境への適応の遺伝的基盤の解明）

生命科学研究科

生態システム生命科学専攻

B3BD3003 玉手智史

General Introduction

Recent environmental changes often caused species's responses such as shifting their geographical distribution and/or adaptation to novel environments in the nature. However, some species cannot respond to rapid environmental changes and lead to extinction, whereas some species such as invaded species can often adapt to novel environments and expand rapidly after the introduction during short-period a time. These indicate that the capacity of a species to adapt to environmental changes differs among species. In general, there are two mechanisms to response to environmental changes, plasticity and physiological tolerance, and adaptive evolution through natural selection. Although both provide the potential of responding to environmental changes, the magnitude of recent climate changes is often larger than those of species' plastic responses, and several researches showed that evolutionary responses through climate-driven selection have often occurred. Thus, adaptive evolution plays a more important role in responses to environmental changes than previously thought.

To what degree a species can adapt to various environments within their habitats is one of important questions for understanding mechanisms of adaptive capacity of the species. Some species have large habitat range with various environments, whereas other species have narrow ranges with uniform environment, and the present habitat ranges might be determined of their historical adaptation. Although habitat ranges are limited by several complex factors, lack of genetic variation within populations or species might be related to failure in expanding their habitat ranges. Why successful invaded species can expand in novel introduced habitat is also a related to problems as to mechanisms of adaptive capacity. Invaded species often expand from small number of founder individuals, and it is unclear whether they can adapt to novel environments with lower genetic variation. Thus, what genetic basis could facilitate adaptation to various environments within their habitat and rapid adaptive responses to novel environments after the introduction are challenging unsolved problems for evolutionary biology and conservation. Recently next generation sequencers have been developed and improved, and a huge genome information is easily available even though non-model species. A whole-genome approach

can utilize overwhelmingly a large number of loci throughout a genome without investigating ecologically important traits that obviously related to phenotypes. Thus, whole-genomic analyses would be effective tools for understanding the genetic basis for adaptive capacity for various species.

In this thesis, I aim to explore the genetic basis conferring specie's adaptive capacity. In chapter 1, I focused on gene duplication event that is one of the source of genetic variation, and estimate the relationship between proportion of duplicated genes (P_D) and their habitat diversity in mammal species. In chapter 2, I focused on the green anole lizard (*Anolis carolinensis*) invaded in Ogasawara Islands and examining genetic basis of rapid adaptation to a novel environment by estimating historical demography and detecting the recent natural selection contributed to the adaptation.

Chapter 1. Contribution of non-ohnologous duplicated genes to high habitat variability in mammals

The mechanism by which genetic systems affect environmental adaptation is a focus of considerable attention in the fields of ecology, evolution, and conservation. However, the genomic characteristics that constrain adaptive evolution have remained unknown. Gene duplication event is one of the source of genetic variation, because duplicated genes often accumulate mutations under relaxed functional constraints, and generate novel genes. A recent study showed that the proportion of duplicated genes in whole *Drosophila* genomes correlated with environmental variability within habitat, but it remains unclear whether the correlation is observed even in vertebrates whose genomes including a large number of duplicated genes generated by whole-genome duplication (WGD). Ohnologs, which are duplicated genes by WGD, are not just ancient duplicated remnants and have different features compared to small-scale duplication (SSD) genes. Ohnolog are likely to be dosage-balanced genes, and gene knockout of an ohnolog causes lethal phenotypes, whereas SSD genes are possibly related to response to biotic stimuli, a property that could be important for adaptive traits. Here, I used fully sequenced mammalian genomes that experienced WGD in early vertebrate lineages and focus on the relationship between both ohnolog and SSD genes and habitat variability. To identify duplicated genes, small-scale duplication (SSD) genes and lineage specific lost genes in mammalian genomes, I conducted an

all-against-all Basic Local Alignment Search Tool (BLAST) search for all protein sequences, and I used the longest sequence for genes with multiple isoforms. For SSD genes, we obtained orthologous of human for each mammalian species from Ensembl database release 63, and if an ortholog corresponded to a human known ohnolog, the gene was defined as a candidate ohnolog. The habitat areas for mammalian species were obtained from the International Union for Conservation of Natural Resources (IUCN) Red List of Threatened Species and literatures. Habitat variability was estimated from the climatic envelope, and habitat diversity was estimated using the Köppen climatic classification. Model selection was applied using linear models to determine which genomic factors affect habitat features. The set of predictors of the explanatory variables that yielded the lowest AIC (Akaike Information Criterion) was selected using a stepwise AIC procedure. To remove any phylogenetic constraints on the relationship between genetic architecture and habitat, phylogenetic trees and phylogenetic contrasts (PICs) were inferred. As a result, I revealed the proportion of SSD genes but not that of WGD genes, is significantly correlated with habitat variability. Moreover, species with low habitat variability have a higher proportion of lost duplicated genes, particularly SSD genes, than those with high habitat variability. These results indicate that species that inhabit variable environments may maintain more SSD genes in their genomes and suggest that SSD genes are important for adapting to novel environments and surviving environmental changes. My findings suggest that proportion of duplicated genes (P_D), in particular the proportion of SSD genes, would be an appropriate index for evaluating species vulnerability to future environmental changes for the conservation of biodiversity.

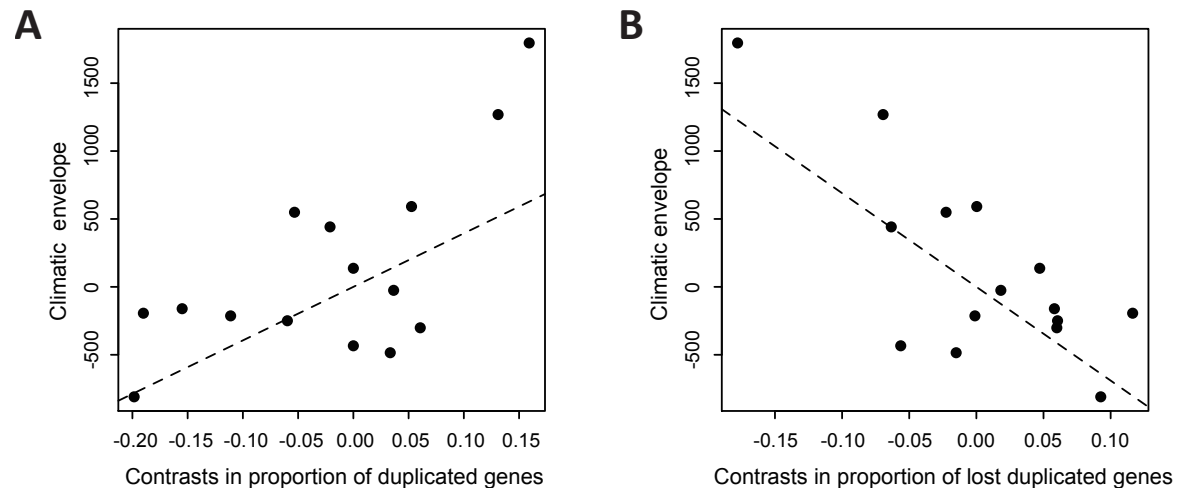


Figure 1-1. Correction between P_D and habitat variability for 16 mammals. The x-axis represents phylogenetic independent contrasts (PICs) in the P_D . The y-axis represents PICs in the climatic envelope estimated from WORLDCLIM.

Chapter 2. Estimating demographic history and recent natural selection for *Anolis carolinensis* introduced into Ogasawara archipelago

Many studies reported that invaded species could rapidly expand and establish to novel environments through adaptive evolution after the invasion. Thus, invaded species could provide a valuable opportunity for natural environment on evolutionary responses. Several important questions as to evolution after invaded species have remained unanswered. First, how the introduced population with small genetic diversity can adapt to new habitat has been called a paradox in invasion biology. Second, many studies on the evolution of invaded populations have examined evolutionary phenotypic changes, but little has been known what and how genetic changes have occurred responding to adaptation to new environments in the invaded populations. Thus it is still unknown what genomic regions have evolved during the expansion of the introduced populations. The green anole (*A. carolinensis*) was introduced into Chichijima, one of Ogasawara Islands in 1960s from southeastern area of the United States, and then expanded rapidly throughout the Chichijima and Hahajima islands. They also caused great impact on native ecosystem by preying native insects and animals. However, it remains unclear

how genetic variation at the founder population was small, and what genomic regions have evolved during the establishment of the introduced populations of *A. carolinensis* in Ogasawara Island. The purpose of this chapter is to estimate genetic variation and demographic history of *A. carolinensis* populations introduced into Ogasawara Islands, and to detect what genes have been subject to selective sweep or positive selection after the introduction under the estimated genetic variations of the founder populations. To estimate the demographic history and detect recent natural selection after the invasion, I conducted coalescent simulations and sliding window analysis using the whole-genome sequences of both the introduced Ogasawara and native Florida populations, and also compared the morphological traits among the populations. The results indicated that the estimated initial and the present effective size of the introduced population were 28 and 229, respectively, and no migration between the introduced and native population have occurred. Morphological analysis showed that lower hindlimb length was longer in the introduced populations than the native populations and the head width was larger in Hahajima islands than other populations. Twenty candidate genes that had been subject to selective sweep or positive selection were detected and among the detected candidate genes, 6 genes have functions of muscle development and contraction, three have metabolic process, growth factors and temperature perception. The results indicated that although the effective number of introduced population was small, several genes had been subject to selection and might be related to adaptation to the introduced environments. Some of genes might contribute to evolution of hindlimb length which might be related to adaptation to change in microhabitat uses. Other gene has function of metabolic processes, which might be related to change in food habit in the introduced populations.

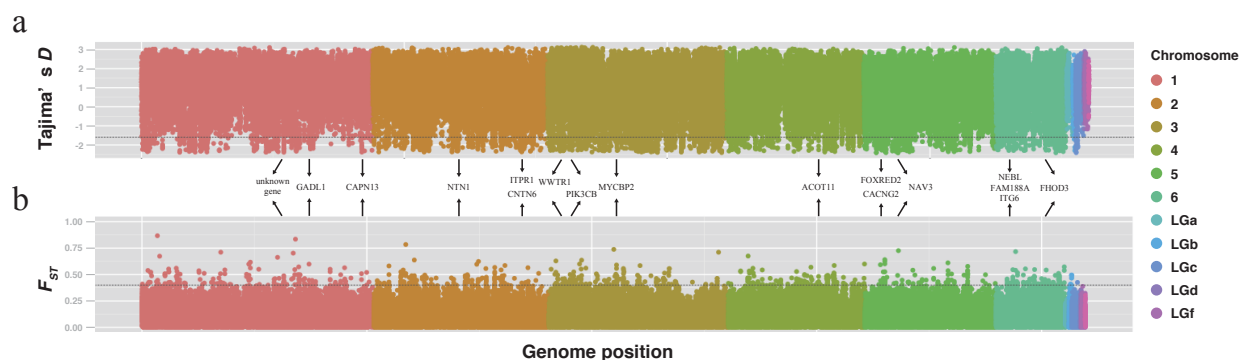


Figure 2-3. Genome wide sliding window analysis (window size = 10kb, sliding size = 5kb) of Tajima's D and F_{ST} in both introduced population. Each color indicates the difference of chromosome 1-6 and LGa-f. $P < 0.05$. P values were determined based on the null distribution produced by the coalescent simulation. (a) Tajima's D values in each window of Chichijima population. The horizontal dashed line indicates the threshold at lower $P < 0.05$. (b) F_{ST} values in each window of Chichijima population. The horizontal dashed line indicates the threshold at higher

General discussion

In chapter 1, I revealed that the species inhabit variable environments maintained higher P_D in their genome especially SSD genes in mammals. My results suggest that the proportion of gene duplications, not ohnolog but particular SSD genes, in a genome are important for adapting to various environments within their habitat and it becomes an effective signature of species adaptive capacity. In chapter 2, I revealed the historical demography of Ogasawara islands, and detected morphological changes and some genes are related to recent rapid expansion. My results indicate that *A. carolinensis* might expand responding to the introduced environments with adaptive evolution despite of the low initial population size due to bottleneck effect. Recent rapid environmental changes cause serious effects for various species, but my results suggest that the species which maintains higher P_D in their genome would be robust for environmental changes, and species also might lead to a rapid adaptive evolution in spite of the small initial population size, if the founder genetic variation is not too small. Therefore, I propose that we could evaluate the adaptive capacity of a species by estimating the genetic basis of the species such as P_D and the historical effective population sizes. Since the whole genome information expected to be available for many species, genomic structure of the species revealed by whole-genome analysis would provide variable information for conservation of biodiversity.