COLONIZATION AND ADAPTATION PATTERNS OF THE HOUSE MOUSE (*MUS MUSCULUS DOMESTICUS*) ON THE KERGUELEN ARCHIPELAGO

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Émilie Annie Hardouin

Plön, 6th December 2010

Supervisor:	Prof. Dr. Diethard Tautz
	Prof. Dr. Hinrich Schulenburg
Date of the oral examination:	4th February 2011

Dean

A ma grand-mère, Maïté Dominé

List of content:

Acknowledgements

I am very much indebted to my supervisor Prof. Dr. Diethard Tautz for allowing me to work in his laboratory, letting me develop the Kerguelen project and for all the fruitful discussions we had all along the past 3 years.

Many thanks to Dr. Jean-Louis Chapuis, for his precious help in the organization of the Kerguelen mouse shipping, for his support and for discussions. I am very grateful to Dr. Marc Lebouvier and to the Institut Paul-Emile Victor, program n°136 who helped us a lot in the Kerguelen mouse collection.

I would like to thank all the people who participated in the mouse collection: Dr. Marc Stevens, Dr. Bettine Jansen van Vuuren, Dr. Petra Quillfeldt, Darren Christie, Quiterie Duron, Hélène Demeringo, Alexia Garnier, Marine Pascal, Juan F. Masello, Hendrika van Noordwijk, Keith Springer and to all the Falkland Islanders who helped us to collect samples.

I am very grateful to all the people of my laboratory: Inka Montero, Dr. Angelika Börsch-Haubold, Heinke Buhtz, Conny Burghart, Natacha Hasenkamp, Dr. Meike Teschke and Dr. Anna Büngte for all their help during the past three years. I also want to thank all the people of the Max Planck Institute for all the scientific discussions but also for all the fun we had.

A special thank you to all the proof readers of the thesis for their time and patience with my usage of the English language: Prof. Dr. Diethard Tautz, Inka Montero, Dr. Demetra Andreou, Dr. Angelika Börsch-Haubold, Dr. Anja Schunke and Dr. Arne Nolte. I also would like to thank Natacha Hasenkamp for her help with the German translation of the abstract.

Je voudrais aussi remercier tous les membres de ma famille, en particulier mes parents et mes grand-mères qui m’ont soutenu tout au long de ces 10 années d’études supérieures.

Enfin un grand merci à ma bande de copain d’Angers qui sont toujours là dans les mauvais comme dans les bons moments malgré la distance.

Zusammenfassung

Die Hausmaus (*Mus musculus domesticus)* hat sich im Laufe der Geschichte, von Westeuropa ausgehend, kontinuierlich über die gesamte Welt ausgebreitet. Die Besiedlung eines Großteils der ozeanischen Inseln der südliche Hemisphäre hingegen, geschah innerhalb des vergleichsweise kurzen Zeitraumes der letzten 300 Jahre. Letzteres macht diese Inseln zu einem hervorragenden Modellsystem, um evolutionäre Prozesse in frühen Stadien einer Neubesiedlung zu analysieren und Mechanismen rezenter Anpassung zu verstehen.

Für die vorliegende Untersuchung wurden 24 Mikrosatelliten-Marker, sowie D-loop-Sequenzen von insgesamt 534 Mäusen typisiert. Die Tiere stammten hauptsächlich von den Kerguelen Inseln, zusätzlich aber auch von den Falkland-Inseln, Marion Insel, Amsterdam Insel, Antipodes Insel, Macquarie Insel, Auckland Insel und eine Probe kam aus Süd-Georgien. Trotz des starken Schiffverkehrs über die letzten 100 Jahre zeigen die Ergebnisse, dass ausschließlich die Mäuse die zuerst auf den Kerguelen ankamen, die Hauptinsel (Grande Terre) und den Großteil der kleinen, zugehörigen Inseln besiedelten. Diese Tiere haben sowohl den gleichen D-loop Haplotyp, als auch den gleichen Y chromosomalen Haplotyp. Eine zweite Besiedlung durch Mäuse fand auf den Inseln statt, die weniger eng mit Grande Terre assoziiert sind und wahrscheinlich keine bereits etablierten Mauspopulationen hatten, wie Cimetière Insel und Cochons Insel. Die dortigen Mäuse haben einen anderen mitochondrialen D-loop Haplotypen und unterschieden sich auch in den Autosomen von den Tieren der Hauptinsel. Der Y Haplotyp jedoch entspricht dem der Populationen auf Grand Terre und den assoziierten Inseln, was darauf schließen lässt, dass beide Besiedlungswellen aus verwandten Ursprungspopulationen kamen. Diese Daten deuten daraufhin, dass bereits besiedelte Habitate weiterer Zuwanderung verschlossen sind, möglicherweise durch schnelle Anpassung der Erstbesiedler an die lokalen Bedingungen. Bemerkenswerter Weise wurden auf mehreren untersuchten Inseln mitochondriale Haplotypen gefunden, die durch eine einzelne Mutation aus dem Haupthaplotypen hervorgegangen waren. Dies

weist auf eine ungewöhnlich hohe Mutationsrate oder einen „selective sweep“ im mitochondrialen Genom hin.

Um die genetische Grundlage der Anpassungsprozesse auf den betrachteten Inseln eingehender zu untersuchen, wurde ein genomweiter Mikrosatellitenscreen zur Identifizierung selektiert Loci („selective sweeps“) durchgeführt. Zur Vorauswahl von Kandidaten-Loci wurden zunächst 737 genomweit verteilte Mikrosatelliten in gepoolten Proben von den Kerguelen typisiert und mit europäischen Populationen verglichen. Dann wurden insgesamt 38 ausgewählte Kandidaten-Loci für fünf unterschiedliche Inseln (Kerguelen Inseln, Marion Insel, Auckland Insel, Marion Insel, Antipodes Insel) und sechs Populationen individuell typisiert. So sollten genomische Regionen mit ähnlichen Mustern identifiziert und historisch bedingte Effekte auf diesen „hitchhiking mapping“-Ansatz reduziert werden. Auf diese Weise konnten fünf Mikrosatelliten-Kandidatenloci bestimmt werden, die alle mit jeweils einem Gen assoziiert waren. Einer der Kandidaten hat bekanntermaßen eine Funktion bei Parasiteninfektionen. Ein anderes Kandidatengen kodiert für Kcne1, die Untereinheit eines K^+-Kanals, wobei das Gen für Kcnq1, eine weitere Untereinheit desselben K^+-Kanals, ebenfalls unter den 38 zuvor weniger stringent ausgewählten Kandidatenloci war. Dies deutet darauf hin, dass dieser Kanal für die Mäuse eine Bedeutung für die Adaptation an die neue Umwelt hat. Zusätzliche Experimente sind erforderlich, um die fünf Kandidatengene weiter zu charakterisieren. Dennoch zeigt die vorliegende Studie, dass ein genomweiter Mikrosatellitenscreen ein geeigneter Ansatz zur Identifizierung von Genen darstellt, die an rezenten Anpassungsprozessen beteiligt sind.

Abstract

Starting from Western Europe, the house mouse (*Mus musculus domesticus*) has spread across the globe in historic times. However, most of the southern oceanic islands were colonized by mice only within the past 300 years. This makes them an excellent model for studying the evolutionary processes during early stages of new colonization and for understanding mechanisms of early adaptation.

Twenty-four microsatellite loci and the mitochondrial D-loop sequence were typed in a total of 534 mice mainly from the Kerguelen Archipelago but also from Falkland Islands, Marion Island, Amsterdam Island, Antipodes Island, Macquarie Island, Auckland Islands, and one sample from South Georgia. Although there was heavy ship traffic for over a hundred years to the Kerguelen Archipelago, it appears that only the mice that have arrived first have colonized the main island (Grande Terre) and most of the associated small islands indeed mice shared the same D-loop haplotype as well as the same Y chromosomal haplotype. The second mice invasion has occurred on islands that are detached from Grande Terre (Cimetière Island and Cochons Island) and were likely to have had no resident mice prior to their arrival. They displayed a different mitochondrial D-loop haplotype and were genetically distinct in the autosomes. However, the Y chromosome haplotype was related to the one found in Grande Terre, suggesting that they came from a related source population. These data suggest that an area colonized by mice is refractory to further introgression, possibly due to fast adaptations of the resident mice to local conditions. Interestingly, single step mutational derivatives of one of the major mitochondrial haplotypes were found several times in different southern hemisphere islands, suggesting an unusually high mutation rate or the putative presence of a selective sweep in the mitochondrial genome.

In order to investigate further the genetic basis of adaptation on southern hemisphere islands a genome-wide microsatellite screen for selective sweeps was performed. 737 markers dispersed around the entire genome were typed in pooled samples from Kerguelen Archipelago populations and compared to European populations

in order to pre-selected candidate loci for selective sweeps. A total of 38 pre-selected candidates loci were then individually typed from five different islands (Kerguelen Archipelago, Marion Island, Auckland Island, Marion Island, Antipodes Island) representing six mouse populations in order to identify genomic regions displaying similar patterns and to decrease the impact of the demographic history of the island on the hitchhiking mapping approach used. Five microsatellite candidate loci, all of them associated with a gene, were identified. Interestingly, one of the candidates has a known function during parasite infection. Another candidate gene is a sub-unit of a K+ channel and, surprisingly, the second sub-unit of this K+ channel was picked up during the less stringent pre-selection of the 38 loci, pointing to the importance of this channel for mouse adaptation to their new environment. Additional experiments are required in order to confirm these five candidate genes, but nevertheless this study demonstrates that a genome-wide microsatellite locus screen is a valuable approach to identify genes which are implicated in recent adaptations.

Declaration:

The design of the whole project was done together with my supervisor Prof. Dr. Diethard Tautz. The interpretations of all the different results were acquired during numerous discussions. Practical laboratory work as well as the major parts of the data analysis was conducted by me, with some exceptions:

Chapter 1

Published as a research article in BMC Evolutionary Biology 2010, 10:325, authors' contribution are as follow:
I developed the project, did the laboratory work and wrote the first draft of the manuscript. Jean-Louis Chapuis organized the sample collection on the Kerguelen Archipelago and on the National French Library (Paris), Marc Stevens provided animals from Macquarie Island, Auckland Island and Antipodes Island. Bettine Jansen van Vuuren provided animals from Marion Island and Petra Quillfeldt from the Falkland Islands. Rick Scavetta and Meike Teschke developed the Y-chromosomal primers and provided the information on allele patterns in Cameroon and Europe. Diethard Tautz was the primary supervisor, participated in its design and coordination, and wrote the final manuscript. All authors were involved in writing and data interpretation, read and approved the final manuscript.

Chapter 2

Prof. Dr. Diethard Tautz and I designed the project together. The interpretations of all the different results were acquired during numerous discussions. Laboratory work as well as data analysis were done by myself.

General Introduction

I- MOLECULAR EVOLUTION, A BRIEF INTRODUCTION:

Darwin defined evolution as "descent with modification" meaning a change in the lineage of populations between generations. Nonetheless, Darwin's theory was suffering from the absence of a theory of heredity. Indeed, at that time, most of the heredity theories were blending theories meaning that natural selection would have been much less powerful (Province 2001). The rediscovery of Mendel's ideas at the beginning of the 20^{th} century and the consolidation of both theories gave birth to neo-Darwinism which spread in all the fields of biology and became widely accepted (Province 2001).

Four evolutionary forces are described: mutation, drift, selection and migration. Among them, mutation is the only one which can create novelty. There is a wide variety of mutation types, for example: single amino-acid change, transposition, unequal crossing over, insertion, deletion, duplication. The mutation rate can be estimated from the rate at which detectable new genetic variants arise. Unfortunately, in nature, the majority of mutations is deleterious and is eliminated from the genome prior to detection leading to underestimated mutation rates. Mutation rates vary between organisms. For example, mutation rates are generally high for RNA viruses (10^{-4} mutation rate per nucleotide per replication – Ridley 2004) whereas for humans it was calculated to be 10^{-8} (The 1000 Genomes Project Consortium 2010). Mutation rate also varies between genome regions (Wolfe et al. 1989). For example, the mutation rate of microsatellites (DNA sequences containing a number of short tandem repeat (2-6 bp) sequence) is not uniform and can be affected by microsatellite length: mutation rate increases with an increasing number of repeat units (Ellegreen 2004).

One of the principal forces which drive allele frequencies is natural selection. Different types of selection have been described including directional selection. Positive directional selection can increase the frequency of an allele in the population whereas negative directional selection decreases the frequency of an allele. Balancing selection

results in the maintenance of multiple alleles. If one or several extreme phenotypes are favored at the same time, the selection is called diversifying.

Another force driving allele frequencies is genetic drift which is defined as the random fluctuation of allele frequencies in a finite population due to chance variations in the contribution of each individual to the next generation (Jobling et al. 2004). Genetic drift can lead to the loss or the fixation of an allele just by chance, meaning that it is impossible to predict the behavior of the allele frequency under drift from one generation to another (Hartl and Clark 2007). The effect of genetic drift is larger in smaller populations (Hartl and Clark 2007).

The last evolutionary force which influences the genome is migration. This process refers to the movement of some organism (or their gametes) among populations creating gene flow between populations leading to a limitation of genetic divergence as a result a homogenization of gene frequencies between populations (Hartl and Clark 2007).

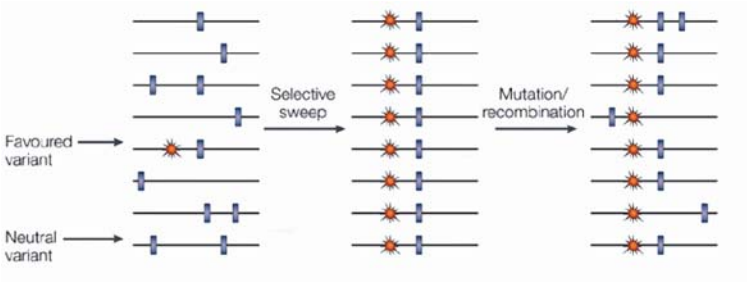


Figure 1: Effect of a selective sweep on the genome (taken from Boffelli et al. 2004)

A challenge for evolutionary geneticists is to differentiate between genetic drift and natural selection. When an advantageous allele appears in a population, natural selection will increase its frequency more or less faster depending of the coefficient of selection. Selection will also increase the frequency of neutral alleles situated near the advantageous allele. This phenomenon is called hitchhiking effect and such a selective event is called selective sweep (Figure 1 and Figure 2 - Maynard Smith and Haigh 1974)

A selective sweep leaves a genetic footprint which can be investigated with different methods such as reduction of heterozygoty, Tajima's D or linkage disequilibrium (Figure 2 – Nielsen 2005). The shape of a selective sweep is mainly determined by the local recombination rate and the coefficient of selection (Maynard Smith and Haigh 1974). Recombination tends to destroy the selective sweep "footprint" whereas depending on the strength of the coefficient of selection, the selective coefficient tends to maintain this region. The selective sweep footprint is gradually lost and the recovery pattern is characterized by an excess of new mutations at low frequency. Therefore, the timeframe in which a selective sweep could be detected is a function of the mutation rate of the investigated region.

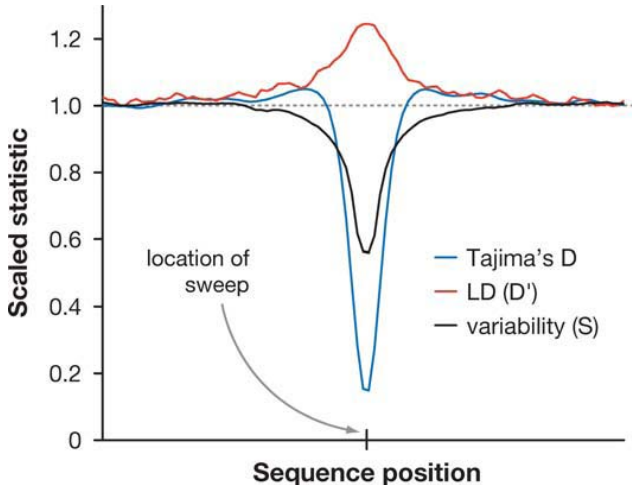


Figure 2: Effect of selective sweep on genetic variation (taken from Nielsen 2005)

Recovery from a demographic change, such as bottleneck or population expansion, results in an excess of rare alleles leading to a comparable genome signature as a selective sweep (Jensen et al. 2005). Hence, the study of selective sweeps should include demographic analysis.

II- WILD MICE AND *MUS LABORATORIUS*:

Laboratory mice have been extensively used in biomedical research, which has resulted in the mouse becoming the most popular model organism for genetic study (for a review see Guénet and Bonhomme 2003). Indeed, mice are easy to keep, breed with a

short generation time of 10-12 weeks and since 2002 the complete reference sequence of the mouse genome is available (Mouse Genome Sequencing Consortium 2002). Moreover, mouse history, phylogeny and evolutionary history are well known (for review see Boursot et al. 1993, and Guénet and Bonhomme 2003) making this animal a suitable organism to understand adaptive mechanisms (Guénet and Bonhomme 2003, Berry and Scriven 2005).

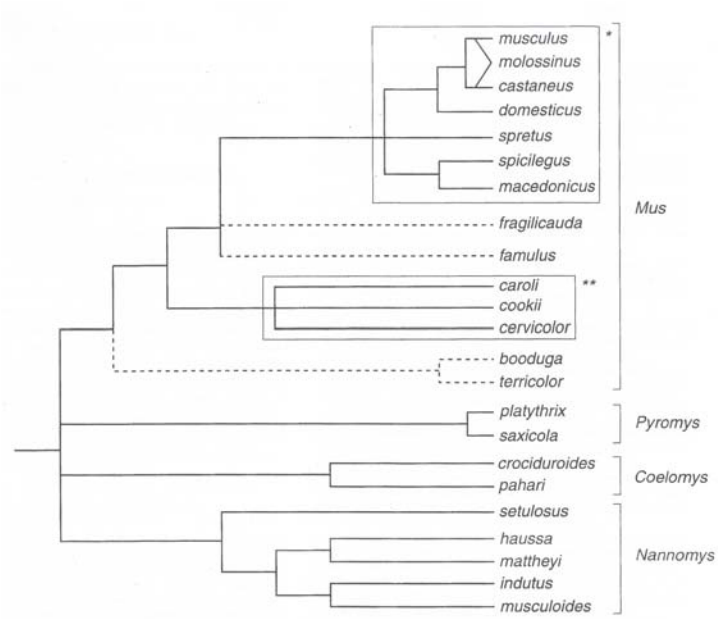


Figure 3: Phylogeny based on molecular data of the subgenera and species in the genus *Mus*. Dashed lines indicate uncertainty in placement of taxa (taken from Tucker 2007)

There are two alternative hypotheses for the origin of wild mice: the first one placed the origin in the northern Indian subcontinent and the second one in west central Asia (Tucker 2007). Molecular data suggest that the initial radiation of *Mus musculus* occurred no more than 0.9 millions years ago and this date of divergence is consistent with the fossil record (Boursot et al. 1993). The house mouse ancestry split into three major sub-species: *Mus musculus musculus* present in northern Asia and Eastern Europe, *Mus musculus castaneaus* in East Asia and *Mus musculus domesticus* in Western Europe (Figure 3 and Figure 4). Hybrid incompatibilities have been described and studied but the three sub-species are not totally genetically isolated and can interbreed. In Europe the hybrid zone occurs between *M. m. domesticus* and *M. m. musculus* and in China between *M. m. musculus* and *M. m. castaneaus*. In Japan, these two sub-species have hybridized

extensively giving rise to *Mus musculus molossinus* (Yonekawa et al. 1988). Only *M. m. castaneus* and *M. m. domesticus* do not have a natural hybrid zone in the wild but they can breed in the laboratory.

Laboratory mouse strains are a mix of different proportions of wild mice, the major contribution being from *M. m. domesticus*, the intermediate contribution of *M. m. musculus* and a small contribution of *M. m. castaneus* (Wade et al. 2002, Wade and Daly 2005, and Sakai et al. 2005). The result is an unnatural genetic constitution and in their review, Guénet and Bonhomme (2003) even proposed to name this new strand: *"Mus laboratorius"*.

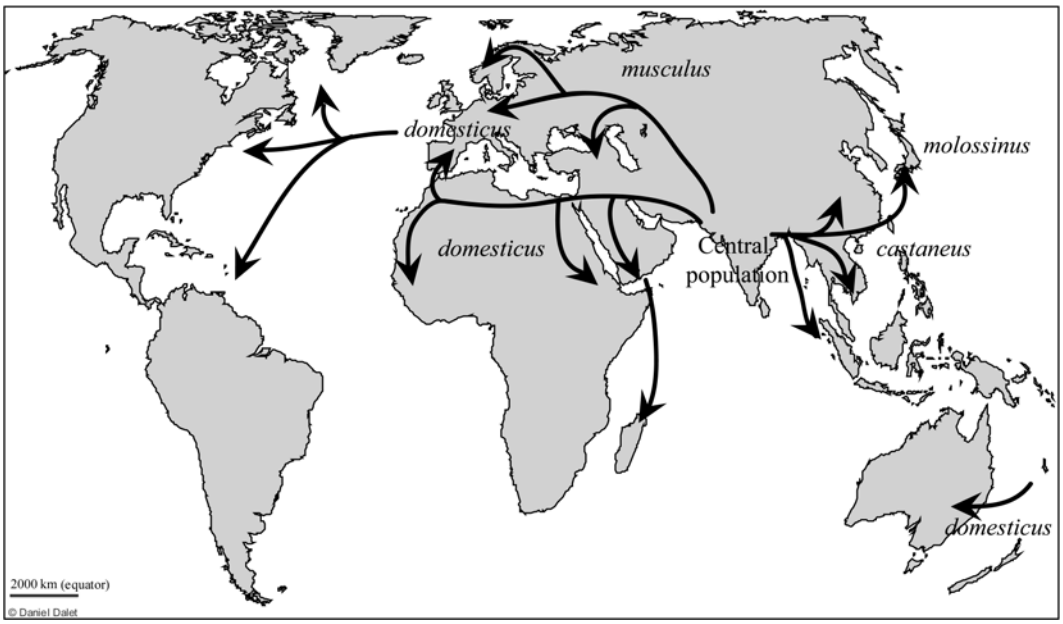


Figure 4: Mice origin and colonization routes in the world. Not shown here, M. *m. castaneus* may have crossed the Pacific Ocean to interbreed with *M. m domesticus* in Southern California (adapted from Morse 2005).

The Western European mouse subspecies *M. m. domesticus* colonized Europe using a Mediterranean colonization route during the Neolithic Age (Figure 5 - Auffray et al. 1990). A study based on paleontological data from Cucchi et al. (2005) demonstrated that mice invaded Europe only 3000 years ago even though human populations settled in Europe since 6000 before Christ (BC). The authors gave several explanations for the late

mouse invasion including an initially limited maritime exchange between the Eastern and Western Mediterranean Basin until the beginning of 1000 BC, leading to little migration. In addition, they suggest a lack of ecological niches for mice before 1000 BC and competition with the wood mouse. Specifically, they suggest that the ecological niches present in Western Europe were initially occupied by the wood mouse and the competition might have favored wood mice until larger and a more commensal environment was developed (Cucchi et al. 2005). Intensification of sea trading in the Bronze Age might have increased passive transport of mice and so played a major role in the colonization (Auffray et al. 1990). Direct proof of human mediated transport was found in the form of the discovery of a mouse mandible in the *Uluburun* shipwreck near the south coast of Turkey (Cucchi 2008). One thousand years ago during the Viking age, mice colonized Scandinavia and Madeira (Gabriel et al. 2010).

In the wake of increased human travel at the beginning of the 16th century, the European mice colonized America, Africa, Australia and New Zealand (Boursot et al. 1993, Guénet and Bonhomme 2003, Searle et al. 2009a). They were also successful in the colonization of small South Hemisphere islands where they were most probably brought by seal hunters and whalers during their journeys (Kidder et al. 1876, Berry et al. 1979, Berry and Peters 1975, Berry et al. 1978, Jansen Van Vuuren and Chown 2007, Searle et al. 2009c). Because of their close relationship with humans, mice have also been studied as a model to infer human colonization history and expansion (Searle et al. 2009, Numone et al. 2010). Jones et al. (in press) found a relationship between mouse genetic diversity and the population size supporting the idea that mice could reflect human population genetics.

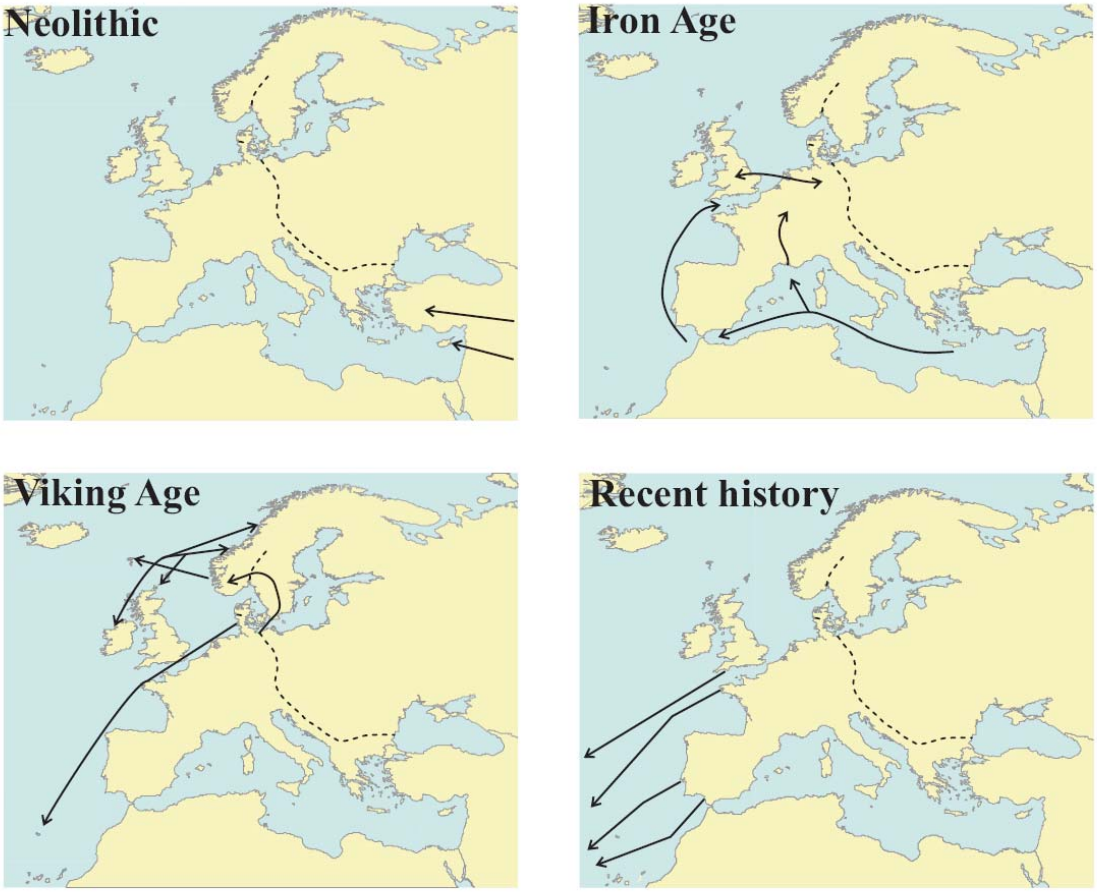


Figure 5: Maps showing possible colonization routes taken by the western house mouse *Mus musculus domesticus*, based on mitochondrial DNA evidence (taken from Gabriel et al. 2010).

In this thesis, I focused on mice living on islands of the Sub-Antarctic area, which are typically islands situated between the 40th and the 60th parallel. In order to understand the colonization patterns and the molecular adaptation of these mice, I compared them with European (France and Germany) mice as reference for an older population and with African (Cameroon) mice as reference for a young colonization. As shown in Figure 6, these 3 populations are genetically separated with no gene flow.

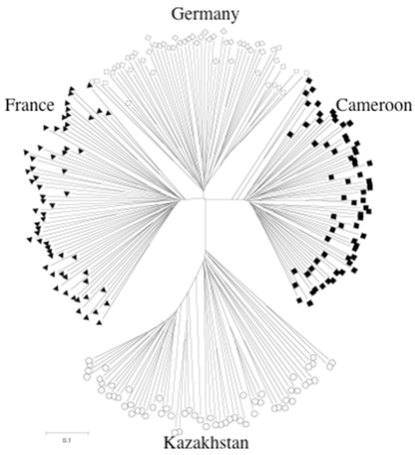


Figure 6: Allele sharing tree based on 204 microsatellites. Mice from Kazakhstan belong to the sub-species *M. m. musculus* and are clearly separate from the *M. m. domesticus* populations (France, Germany and Cameroon) by a longer branch. The populations appear genetically different with no gene flow between them (taken from Ihle et al. 2006).

This repeated and rapid adaptation to Southern Antarctic islands makes mice a tremendous model to study first colonization of this part of the world and also adaptation to this new environment.

III- STUDY SITE: THE SUB-ANTARCTIC AREA:

In 1978, R.J. Berry wrote: "*... the Sub-Antarctic mice must be living close to their physiological limits and hence likely to respond detectably to environmental stresses in ways unnecessary for animals in more temperate environment*." (Berry et al. 1978). This and all the genetic tools already available make the "sub-Antarctic" mice a model of choice to study rapid genetic adaptation. In the next part of the general introduction, I summarize the known information on mice from the sub-Antarctic area and from the Kerguelen Archipelago. As only two genetic studies have been performed on these mice so far, this section will mostly focus on mouse ecology. Understanding their ecology is essential for identifying the challenges the mice faced and how they adapted to it (behavior or genetics).

A°/ Sub Antarctic Area

1- General data

The Sub-Antarctic area consists of several islands localized between the 40th and the 60th parallel (Figure 7 and Table 1). These islands are relatively small and share similar climatic and ecological conditions. The climate is characterized by cold summers and mild winters giving relatively a uniform temperature all along the year. There is high rainfall and strong winds.

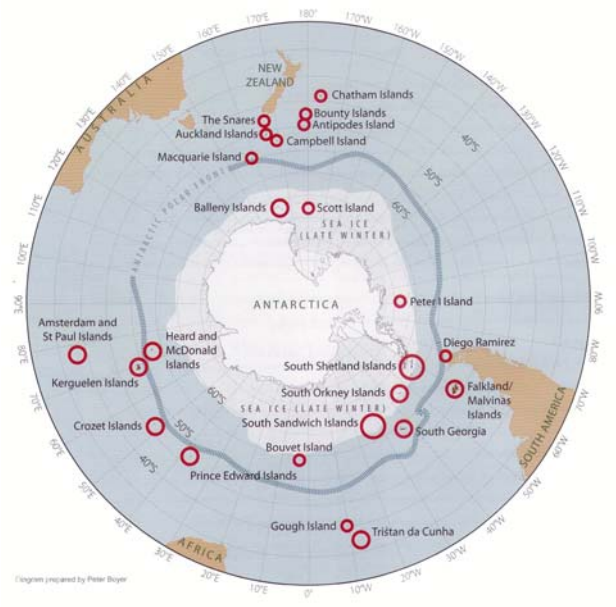


Figure 7: Map of the South Hemisphere, Marion Island is part of the Prince Edwards Islands; the dotted line represents the Antarctic convergence (taken from Lebouvier and Frenot 2007)

Vegetation is rather poor because of the unfavorable climate and the isolation of these islands. Trees are usually absent except on Gough Island, Amsterdam Island and Auckland Island where some small trees and bushes grow. The indigenous terrestrial fauna is also poor and usually limited to invertebrates. However, the maritime fauna, as well as seabird fauna is very rich (Frenot et al. 2005).

Table 1: General information on the islands studied in this thesis (adapted from Frenot et al. 2005, Angel et al. 2009, Learder-Williams 1988)

Island	Latitude	Longitude	Area (km2)	Year of Island discovery	Introduced Animals
Amsterdam Island	37.83 S	77.52 E	55	1522	Norway rats, house mice, cat cattle
Gough Island	40.33 S	9.54 W	64	1505 or 1506	House mice
Antipodes Island	49.68 S	178.77 E	60	1800	House mice
Auckland Island	50.83 S	166.60 E	510	1806	House mice, cats, pigs
West Falklands	51.50 S	60.50 W	4.532	1592	house mice, rat, cats and others
East Falklands	51.50 S	58.50 W	6605		
Marion Island	46.90 S	36.75 E	290	1772	House mice
Kerguelen Archipelago	49.37 S	69.50 E	7200	1772	Black rats, house mice, cats, mouflon, sheep, reindeer, rabbits
South Georgia	54.25 S	37.00 W	4066	1675	Norway rats, house mice, reindeer
Macquarie Island	54.62 S	158.90 E	128	1810	Rats, house mice, rabbits

The islands were free of humans before their "official" discovery (Table 1) and even today most of them are still uninhabited except the Falkland Islands and Tristan da Cunha. However, some islands have a meteorological station or a scientific research station but the human population is rather limited.

2- Mice on Southern Hemisphere islands:

a- Colonization and life:

Mice were very successful in the colonization of the Sub-Antarctic area and today they are known to occur in: Amsterdam Island, St Paul Island, Kerguelen Archipelago, Crozet Archipelago, Marion Island, Antipodes Island, Auckland Island, Macquarie, South

Georgia, Falkland Islands, Tristan da Cunha, and Gough Island (Figure 7 - Kidder 1876, Berry and Peter 1974, Lésel and Derenne 1975, Berry et al. 1978b, Berry et al. 1979, Chapuis et al. 1994, Cuthbert and Hilton 2004, Searle et al. 2009c, Angel et al. 2009). The population abundance on South Georgia is relatively low when compared to the other islands (Frenot et al. 2005). Indeed, the mice occurred in a part of the Nuñez Peninsula which is cut off from the main part of South Georgia by branches of the Esmark Glacier (Berry et al. 1979). This remote situation might have protected them from rats and so from extinction (Berry et al. 1979).

A study from Marion Island showed that the mouse density increased between 1979-80 and 1991-92 (van Aarde et al. 1996) but a recent study showed that the mouse population is becoming stable perhaps due to food limitation and the major die-off during winter which probably limits population size (Ferreira et al. 2006). Interestingly, cold climate is a more important factor to determine mouse survival than food supply, suggesting that seasonal changes in mouse numbers are not dictated by food availability (van Aarde and Jackson 2007). Nevertheless, Frenot et al. (2005) predicted that mouse populations will still continue to increase on the Kerguelen Archipelago, Marion Island, and Macquarie Island. The only mouse predators on these islands are cats and skuas. If the skuas do not limit the mouse population (Mougeot et al. 1998, Schulz and Gales 2004), the cats do (Derenne and Mougin 1976, van Aarde 1980, Quikfeldt et al. 2007, J.-L. Chapuis personnel observation)

b- Mouse diet:

Mice are omnivores and diet studies from Australia and North America showed that they are primarily seed eaters but they also consume a wide variety of plants and animals (see Figure 8 - John and Whitaker 1966, Singleton and Krebs 2007). On Cochons Island from the Crozet Archipelago, mice are apparently feeding mostly on plants (Derenne and Mougin, 1976). However, on Marion Island, Macquarie Island and Guillou Island (Kerguelen Archipelago), invertebrates are the main part of the diet (Copson 1986, Le Roux et al. 2002, Smith et al. 2002). Berry et al. (1979) showed that the South

Georgia mice feed mainly on arthropods and tussock-grass seed. Mice from the Falkland Islands were studied using stable isotopes, which showed that the mouse diet is terrestrial even if some of them have a mixed diet (terrestrial and marine) unfortunately, stomach contents were not sampled (Quillfeldt et al. 2007) .

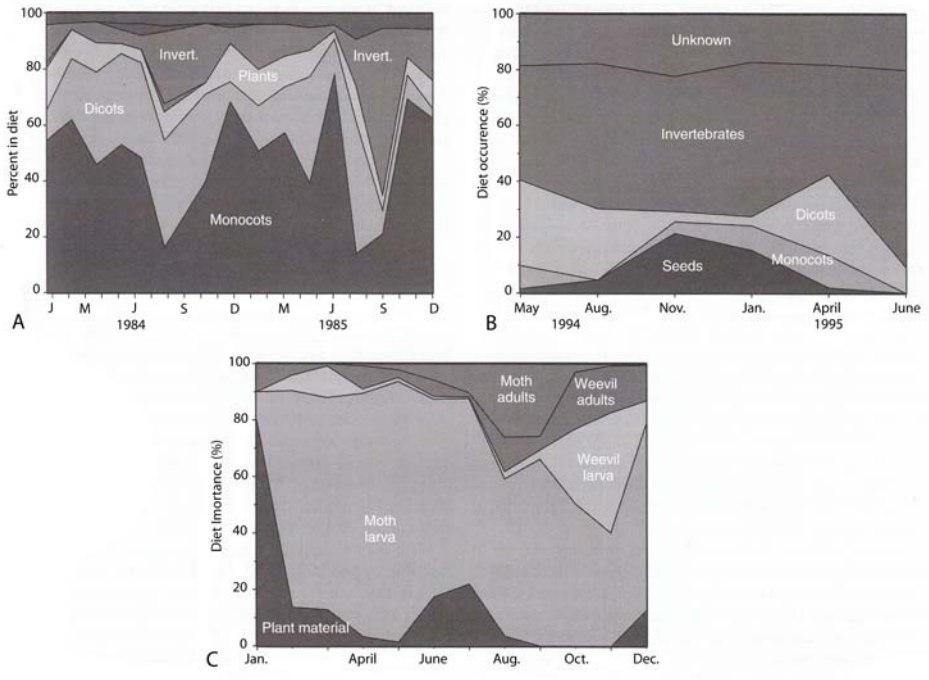


Figure 8: Mouse Diet (A) Mallee Wheatland of Victoria (B) Thevenard Island, Western Australia (C) Marion Island (Sub-Antarctic). (Taken from Singleton and Krebs 2007)

More surprisingly, on Gough Island and on Marion Island, the mice started to attack seabird chicks (Cuthbert and Hilton 2004, Wanless et al. 2007, Jones and Ryan 2010). On Gough Island, it was reported that mice attack and kill healthy chicks up to 300 times of their mass. Videos show up to 10 mice attacking birds, without any appropriate response from the victim. Indeed, if the chick did not die during the first attack, mice will repeatedly feed on them often opening several wounds (Wanless et al. 2007). On Marion Island, although no direct evidence of mice attacking albatrosses was observed, the nature of wounds in dead animals suggests mouse attacks (Jones and Ryan 2010). Wanless et al. (2007) suggested that this behavior might arise when the mice are the only introduced mammals. The Marion Island case reinforced this hypothesis because

the first chick attack by mice was reported a decade after cat eradication (Jones et al. 2010). Bird predation by mice is also suspected on Antipodes Island (Imber et al. 2001). On Cochons Island (Crozet Archipelago), mice have been seen eating cadaver from mammals or birds (Derenne et Mougin 1976). Interestingly, bird attack by mice was only reported on islands where mice are the only introduced mammal. Indeed, there are only 5 Sub-Antarctic islands having this characteristic (Gough Island, Marion Island, Antipodes Island, Australia Island (Kerguelen Archipelago), and St Paul Island) and on 3 of them bird predation by mice is documented or suspected (Angel et al. 2009).

c- Adaptation to Sub-Antarctic environment:

In order to preserve themselves from the cold, mice live in burrows and construct above ground run-away (Lésel and Derenne 1975, Avenant and Smith 2003). Indeed on Marion Island, burrow temperatures seldom drop below 2°C and ground-surface temperature seldom drops to 0°C (Avenant and Smith 2003). Consequently, burrowing behavior might be an important factor for mouse survival on Sub-Antarctic islands. Webb et al. (1997) showed that mice were physiologically adapted to the cold via a reduction in minimum thermal conductance.

B°/ Kerguelen Archipelago:

The Kerguelen Archipelago is situated in the Indian Ocean (48°25'-50°S, 68°25'-70°35'E), around 4000 km from the African and Australian cost. The transport to the archipelago is done by the ship *Marion-Dufresne* four times a year. The rotation takes almost a month and supplies all French Sub-Antarctic territories (Crozet Archipelago, Kerguelen Archipelago, Amsterdam Island and sometime St Paul Island – for a map see Figure 7).

***1- Geology and climate*:**

From volcanic origin (Giret et al. 2003), the archipelago is formed from one main island called Grande Terre (~6675 km^2 with a stretch of 130 km from east to west and of 120 km from north to South) surrounded by around 300 small islands and islets (~540 km^2 – see Figure 11 for a map) making its surface area comparable to the one from Corsica (Figure 9). The archipelago is about 50 million years old.

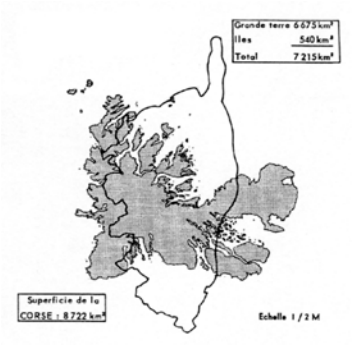


Figure 9: Comparison of the Kerguelen surface when compared to Corsica (taken from Giret et al. 2003)

Situated at the Antarctic convergence (region where the cold Antarctic water sinks below the warmer from the North), the climate is described as oceanic cold and characterized by almost continuous wind which can blow up to 200km/h, no strong winters, 2°C on average for the coldest month, and fresh summers, 8°C on average for the warmest month. Although the rainfall is only 747 mm (350-1479 mm during 1951-2009, Météo-France, Port-aux-Français), it is raining 246 days a year.

***2- History and exploration*:**

In order to understand the mouse colonization and phylogeography, I will, in the next section, briefly describe the general history of the Kerguelen Archipelago. This part is based on: Aubert de la Rüe 1953, Delépine 1964, 1975, 1995 and 2002, and Couesnon and Guyader 1999.

The Kerguelen Archipelago was discovered on the 13th February 1772 by Yves-Joseph de Kerguelen-Trémarec (1734 near Quimper – 1797 Paris). Interestingly, despite his two trips (in 1772 and in 1774), Kerguelen never set foot on the archipelago he

discovered. It was Charles du Boisguéhenneuc who took possession of this new land for France. The 25th December 1776, Captain James Cook (1728 Marton, Middlesbrough – 1779 Hawaii) arrived in the Archipelago and, because of its poor vegetation, baptized it "Desolation Island". During his journey, Cook and his scientists discovered and identified the famous Kerguelen cabbage: "*Pringlea antiscorbutica*". Anderson, the expedition surgeon, observed the presence of many whales and sea elephants and pointed out that this could be useful for France if oil became rare in Europe.

Figure 10: Right: Portrait of the Yves-Joseph de Kerguelen-Trémarec. Left: Portrait of Captain James Cook by Nathaniel Dance-Holland

Indeed the first whaler recorded on the Kerguelen Archipelago arrived in 1792 from Nantucket Island, USA. This journey was followed by several others during the 18th and 19th centuries. From 1786 to 1928, it was estimated that of all boats sent to the Southern Oceans, 46% were American, 19% English, and 3% French. The last American whaling boat observed in the Kerguelen Archipelago was the *Margaret* in 1909.

Many scientific expeditions also stopped at the Kerguelen Archipelago, notably the one from James Clark Ross with the *Erebus* and the *Terror* stayed from the 7th May to the 20th July 1840. Interestingly, Ross reported having seen 700 whaling boats (Delépine 1995); this supports an intensive exploitation of the resources in this part of the world.

At the end of the 19th century, Henry and René-Émile Bossière obtained the permission to exploit the Kerguelen Archipelago for 50 years. This was the start of the first attempt of humans to live on this archipelago. A whaling station in Port-Jeanne d'Arc was built in 1908 and ran until 1926 in which up to 100 people were employed. Shepherds and their families also lived on the archipelago until 1931, this date was the end of the Boissière adventure in Kerguelen.

In 1941, during World War Two, only four ships visited the archipelago, three German corsairs: the *Atlantis*, the *Komet* and the *Pinguin* and one Australian: the *Australia*. All of them stopped for a very short time (few days, only the *Atlantic* stayed for almost a month).

In 1949, the French government reaffirmed its sovereignty in this region. On the 11th December 1949, the *Lapérouse* arrived in the Kerguelen Archipelago with the mission of finding a location for a future scientific base and for a runway. Finally the site of Port-aux-Français was chosen and on the 1st January 1950, a meteorological and a radio station started to work. The runway was never to be constructed. Finally, in 1951, yearly rotations to the archipelago started. The French Austral and Antarctic Territory (TAAF: Terres Australes et Antarctiques Françaises) was created in 1955 and administrated the islands (Crozet Archipelago, Kerguelen Archipelago, St Paul Island, Amsterdam Island and Éparses Islands) and the French Antarctic territory: Terre Adélie. Today, 60 to 120 people, depending on the season, are exclusively living in Port-aux-Français. There is no permanent settlement on the archipelago.

3- House mice on the Kerguelen Archipelago:

Mice are one of the seven terrestrial mammal species still living on the archipelago. In contrast to rabbits (*Oryctolagus cuniculus*), cats (*Felis silvestris*), sheeps (*Ovis aries*), mouflons (*Ovis aries musimon*) and reindeers (*Rangifer tarandus*), rats (*Rattus rattus*) and the house mouse were not voluntarily introduced (Chapuis et al. 1994). The exact date of mouse introduction is unknown, however given the historical record from this archipelago, it is estimated that the mice came at the beginning of the 19th century (Kidder 1876). Although evidence of mouse introduction on islands before its "official" introduction was described in Madeira (Förster et al. 2009), (when it was proved that the Viking introduced mice before the Portuguese invasion), this is rather improbable considering the localization of the Kerguelen Archipelago.

The first documented observation of mice living on the archipelago was reported by Dr. J. H. Kidder, naturalist in the Transit-of-Venus expedition in 1876. He observed that: "*the only mammal found on the island is the common mouse (Mus musculus), which abounds everywhere, and was doubtless imported by one the early sealers. It builds its nest in holes in sand-banks, lining it with dried grass-stems or bits of oakum, and appears to feed mostly on grass-seed*". The other paper mentioning the mice was published in 1975 by Lésel and Derenne. It is described that the mice occur on the main island (Grande Terre) and perhaps on other islands. The authors wrote: "*The mouse was introduced in this way* (by whalers and sealers) *and it now occurs over the whole of Grande Terre. The species is perfectly adapted to its surroundings and lives in burrows, following the vegetation zone in its distribution*". These two papers are very important because they allow the dating of the mice introduction as well as the first observations of their rapid adaptation.

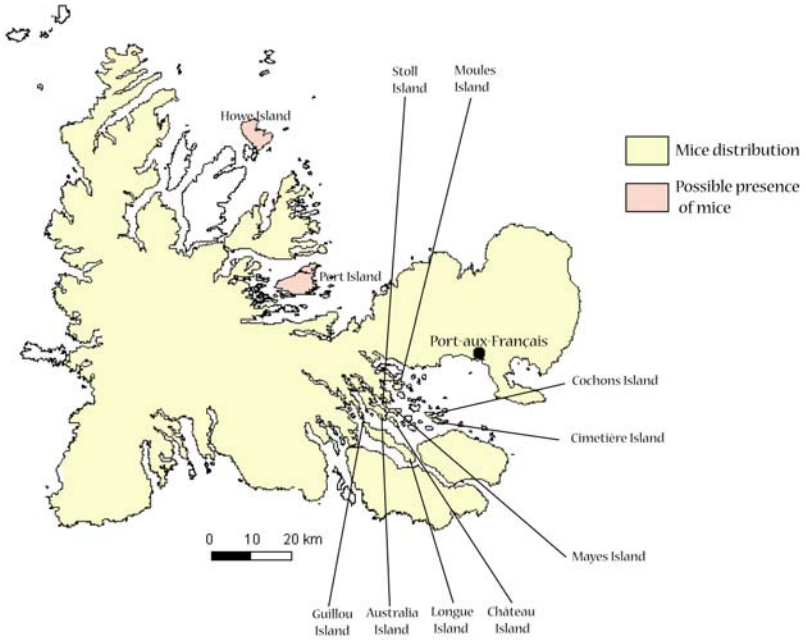


Figure 11: Mouse distribution on the Kerguelen Archipelago (Chapuis et al. 2002 and J.-L. Chapuis personal observation)

Today, the mice are distributed all over Grande Terre and in some islands from the Morbihan Gulf (Figure 11 - Chapuis et al. 2002 and J.-L. Chapuis personal observation). Pisanu et al. (2002) demonstrated that mouse populations on the Kerguelen

Archipelago are going through a succession of yearly population expansion, during austral summer, and contraction, during austral winter. Today the mice are still expanding and colonizing new territories on the archipelago; one example of this phenomenon is Cimetière Island. Indeed mice on this island were observed from 2002 onwards (J.-L. Chapuis, personal observation). It is not unlikely that other islands could also have been colonized. The way of dispersal is unknown, however it is most probably a natural event even if human mediated transport cannot be excluded. Although the understanding of mouse colonization of Grande Terre is relatively easy, it is difficult to understand how the mice can migrate from one island to another. It has been reported that mice are able to swim (Randall et al. 1999) and direct observation of rats swimming from one island to another has been reported (Russell et al. 2005). However, considering the water temperature on these latitudes (5°C) it seems that active swimming cannot explain mouse migration. Moreover, since there are no trees on the archipelago, there is no possibility of rafting between islands. New mouse introductions from La Réunion could almost (but never totally) be excluded.

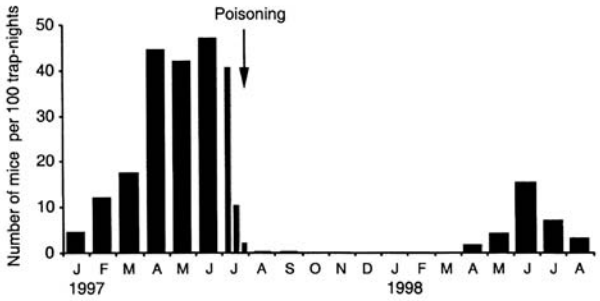


Figure 12: Index of mouse abundance in Cochons Island in 1997-1998 during the operation of rabbit poisoning (taken from Chapuis et al. 2001)

Because invasive species are damaging the Kerguelen environment, there are attempts to eradicate them. In 1992, 1994, and 1997, attempts to eradicate rabbits were performed respectively on Verte Island, Guillou Island and Cochons Island. Chapuis et al. (2001) used the first generation of anticoagulant: chlorophacinone. During this treatment, the Cochons Island mouse population was monitored (Figure 12) and mouse abundance dropped during several months before it regenerated again the year after. This suggests that the rabbit poisoning could have influenced the genetic variability of the

mouse population from Cochons Island but also from Guillou Island where the same procedure was applied (no mice are living on Verte Island).

IV- AIMS OF THE STUDY:

Chapter 1: I investigated and described mouse genetic variability from sub-Antarctic regions using mitochondrial and nuclear markers. These data allowed me to search for mechanisms and patterns of colonization.

Chapter 2: A genome-wide screen to investigated signatures of positive selection on Kerguelen Archipelago mice was performed using microsatellites. The candidate regions found in this first step were investigated also in other islands from the Sub-Antarctic area in order to identify genomic regions of parallel adaptation. Finally, a set of five candidate genes which could putatively have a role in mouse adaptation to the Sub-Antarctic environment were identified.

Chapter I:

House mouse colonization patterns on the sub-Antarctic Kerguelen Archipelago suggest singular primary invasions and resilience against re-invasion

I- INTRODUCTION:

Island colonization dynamics are of general interest in evolutionary biology, both with respect to understanding adaptive radiations, as well as for tracing migration patterns. In this context it is of particular interest to ask whether a single colonization can already result in a new established population that is refractory to further invasions, or whether multiple independent invaders are required before a new stable population can be established. This question can be particularly well studied in cases of recent island colonization, since this provides insights into the early phases of establishment and adaptation in a population context. The spread of the house mouse (*Mus musculus* L.) across many oceanic islands in contemporary times constitutes an excellent model system in this respect (Guénet and Bonhomme 2003, Berry and Scriven 2005).

Mus musculus originated on the Indian subcontinent within the past million years and there are currently at least three recognized subspecies: *M. m. domesticus, M. m. musculus, M. m. castaneus* (Boursot et al. 1993). *M. m. domesticus* invaded Western Europe about 3,000 years ago (Cucchi et al. 2005) and then colonized the rest of the world (i.e. Africa, America and Australia / New Zealand) mostly in the wake of increased human travel across the globe that started in the 16th century (Boursot et al. 1993, Guénet and Bonhomme 2003, Searle et al. 2009a). They were also very successful in colonizing isolated islands, such as those of the Southern Ocean (Berry and Peters 1975, Berry et al. 1978, Berry et al. 1979, Jansen van Vuuren and Chown 2007, Searle et al. 2009c), where

they were brought by whaling ships making stops during their journeys or went for seal hunting.

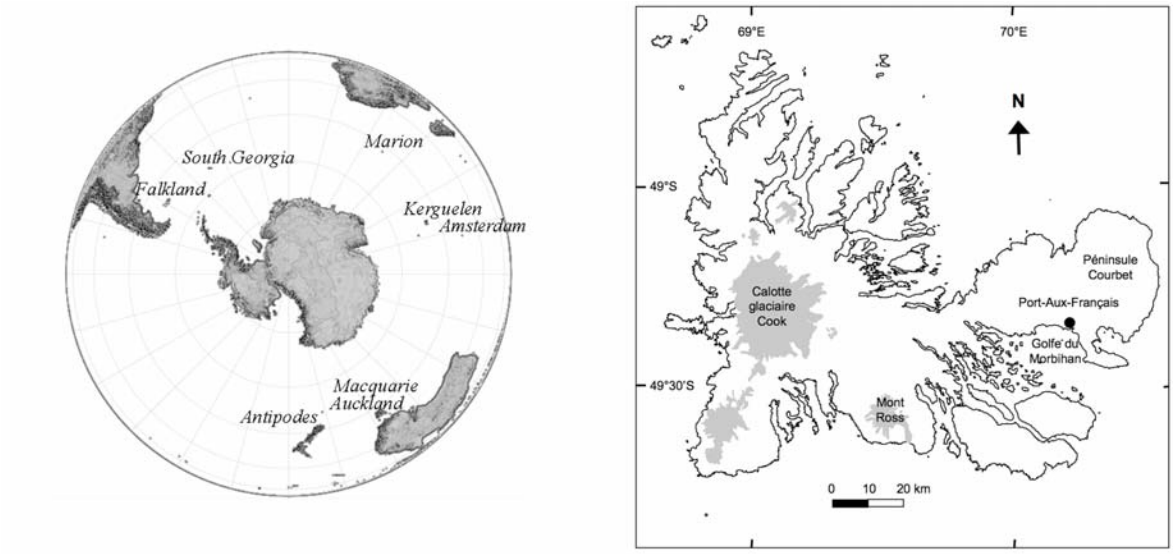


Figure 1: Locations of the Southern Ocean islands assessed in this study (left) and map of the Kerguelen Archipelago (right). The sampling sites in the Kerguelen Archipelago are all around the Morbihan Gulf and the research station at Port-aux-Français except Port-Couvreux, Pointe du Morne and the Cap Ratmanoff (see further details in Figure 5).

The Kerguelen archipelago was discovered on the 12th of February 1772 by Yves-Joseph de Kerguelen-Trémarec. It is situated about 4,000 km away from the African and the Australia coast (Figure 1), has a large main island of 6,500 km² called Grande Terre, and approximately 60 small islands (1–200 km²) surrounding it (Figure 1b). The climate is Oceanic cold, characterized by cold summers (8°C on average), no rigorous winters (2°C on average), by strong wind and mean annual rainfall of 747 mm (350-1479 mm during 1951-2009, Météo-France, Port-aux-Français). There was never an extended human settlement on the archipelago, but since 1951, there is a research and weather station with a continuous turnover of about 60 to 120 people per year.

The house mouse was most likely introduced to the Kerguelen at the beginning of the 19th century (Kidder 1876, Lésel and Derenne 1974, Chapuis et al. 1994), but certainly not before 1772, since it is too far away from the continents to have been a destination for ship traffic in previous times. During the high times of whale and seal hunting, there was

heavy boat traffic in this area, with a large potential to bring additional mice. Today, the mice have colonized all of Grande Terre as well as many of the small islands of the Morbihan Gulf (Pisanu et al. 2001, Le Roux et al. 2002).

With this defined history, as well as extensive data on the genetic diversity of the relevant source populations (Western Europe), we have an excellent test case to study population genetic consequences of island invasion, the subsequent spread to further islands and patterns of re-invasion.

II- METHODS

A°/ Mouse samples

Population samples from Cologne-Bonn (Germany), Massif Central (France) and Cameroon were described previously (Ihle et al. 2006). For these we had applied a sampling scheme that took care to sample the genetic variation within an area of about 50 km diameter (i.e. trapping sites were at least 300 meters away from each other). Hence we consider these samples to reflect the local population diversity. Additional samples from Schleswig-Holstein (Northern Germany) were trapped in 2006 using the same scheme. In contrast, the mice from Paris (n=20) were caught within the confinements of the garden of the National French Library (BNF) in 2009, i.e. not following the extended sampling scheme above. The mice in the BNF are living in a space of around 1 ha at the center of the national library building. They are separated from other populations outside the BNF by poisoning. Hence, these are considered to represent a single sample from a local population, not necessarily reflecting the diversity in the extended area.

The Kerguelen Archipelago samples were caught mainly in the Morbihan Gulf area including several islands and the adjacent Grande Terre (see Table 1 for details). Again it was not possible to apply the extended sampling scheme in this case. Instead, the sampling followed the scheme described in Chapuis et al. (2001). All the mice in the Kerguelen Archipelago were captured in non-inhabited area except around the research

station in Port-aux-Français. The mice were trapped using a line system, with three parallel lines 40m away from each other and a length of approx. 100m each with 34 traps along the line (1 trap every 3m). Mice from Port-aux-Français (n=41), Guillou Island (n=79), Cochons Island (n=69), Isthme Bas (n=38), Mayes Island (n=18), La cabane dite "Jacky" (n=29), Cimetière Island (n=28), Australia Island (n=24), Port-Jeanne d'Arc (n=16), Cap Ratmanoff (n=8), Sourcils noirs (n=5), Port-Couvreux (n=4) and Pointe du Morne (n=1) were trapped in 2008 and 2009. Mice from Moules Island (n=12) and Stoll Island (n=4) were captured in 2005. Other Mayes Island (n=57) and Australia Island (n=4) mice were trapped in 1996. Amsterdam Island (n=3) samples were collected in December 2007. Marion Island mice (n=18) were caught at two localities across the island, namely at the Meteorological Station and at Mixed Pickel Cove in 1990 (n=6) and 2004 (n=12) (Jansen van Vuuren and Chown 2007). Macquarie Island (n=12), Antipodes Island (n=18) and Auckland Island (n=13) mice where caught in 2005-2006 (Searle et al. 2009c). Additional samples from Macquarie Island (n=28) from 2005 were used. Falkland Islands samples from New Island (n=12) were caught in 2006 (Quillfeldt et al. 2008) and 2010 (n=18). 425 Samples from the other Falkland Islands namely Saunders Island (n=4), Steeple Jason Island (n=5), East Falkland (n=2) and West Falkland (n=3) and a mouse from South Georgia were caught in 2008/2009.

B°/ D-loop sequencing

DNA was extracted using salt extraction. The D-loop was amplified using the primers 5'- CATTACTCTGGTCTTGTAAACC and 5'- GCCAGGACCAAACCTTTGTGT. The reactions were carried out in 10µL final volume with the following cycling parameters: 95°C for 15 minutes followed by 35 cycles of 95°C for 30s, 60°C for 1.30 min, 72°C for 1min and 15 min at 70°C for elongation time. Exo-Sap purification (USB Corp.) was performed with the following incubation: 37°C for 20min and 80°C for 20min. The cycle sequencing reaction parameters were 96°C for 1 min followed by 29 cycles of 96°C for 10s, 55°C for 15s and 60°C for 4min. The sequences generated were visualized using CodonCodeAligner Ver. 2.0.1 (CodonCode 440 Corp.) BioEdit ver.7.0.9.0 (Hall 1999) and MEGA ver. 4 (Tamura et al. 2007). The

haplotype data file was calculated using DnaSP 4.50.3 (Rozas et al. 2003). The network was calculated using the Median Joining method and drawn with Network ver. 4.5.1.0 (Fluxus Technology Ltd), taking care that missing data did not affect the network (Joly et al. 2007). The sequences were submitted to Genbank and are available under accession numbers HQ185258 to HQ185282.

C°/ Microsatellite typing

From a previously described set of 1,000 microsatellites (Thomas et al. 2007, Teschke et al. 2008), we chose 18 (Chr01_25, Chr02_01, Chr03_21, Chr03_24, Chr04_31, Chr05_15, Chr05_45, Chr07_38, Chr08_11, Chr09_20, Chr11_64, Chr12_05, Chr13_22, Chr14_16, Chr16_21, Chr17_09, Chr18_08, Chr19_08) which were known to be polymorphic in the German and French populations. Six Y-chromosomal loci which we found to be polymorphic in the German and French populations were also typed for all island samples where more than 8 males were available. Primer sequences used to type the Y-chromosome were: Y6 aaccaccactatcttcattc and acagagtatacgtacgtgtg, Y12 cccaatctaggcatttaatt and attcaccattctccagtgtg, Y21 accatcagatgatcaccaagtgc and tccagcattcaatggtacaggct, Y22 tcatggtagacaccatggcaac and tcagttttctaggtggaggggtg, Y23 acctcactcaggatgatgccctc and agcctgtgcgcacgtgtgtg, Y24 tctgggggtttcgggtggagcct and gcatcacagctgaggctctgtgg. Forward primers were labeled with FAM or HEX dye on the 5' end. The reactions were carried out in 5µL final volumes using 10ng DNA template using a multiplex 460 PCR kit (Qiagen). The PCR conditions were: 95°C for 15min followed by 28 cycles at 95°C for 30s, 60°C for 1.30min, 72°C for 1.30 min with a final extension at 72°C for 10 min. PCR products were diluted 1:20 in water. 1µL of this dilution was added to 10µL of HiDi formamide and 0.1µL of 500 ROX size standard (Applied Biosystems). The denaturation step was performed with the following incubation times: 90°C for 2min and 20°C for 5min. The alleles were analyzed using GeneMapper ver. 4.0 software (Applied Bioscience). The distances for the allele sharing tree were calculated using MSA3.15 (Dieringer 2003). The tree was generated using R and drawn using MEGA4 (Tamura et al. 2007). Structure was analyzed using Instruct (Gao et al. 2007)

because this method does not assume Hardy Weinberg equilibrium within loci. The run parameters were as follow: 2 chain number, a burn-in period of 100,000 simulations followed by a run length of 2,000,000 MCMC simulations and ten iterations for each *K* (number of clusters). To draw the structure diagram the softwares CLUMPP (version 1.1.2 - Jakobsson and Rosenberg 2007) and Distruct (Rosenberg, 2004) were used. The PCA was generated using the software Genetix 4.03 (Belkhir et al. 2004).

III- RESULTS

A°/ Mitochondrial Data

We sequenced 834 bp of the mitochondrial control region (D-loop) from all samples and found that all haplotypes grouped within the known *M. m. domesticus* haplotypes (Figure 2a), i.e. belong to this subspecies. This was already known for some of the islands (Jansen van Vuuren and Chown 2007, Searle et al. 2009c) and we show here that it is also the case for the Kerguelen Archipelago, Amsterdam Island, Falkland Islands and South Georgia. Hence, the source populations of the mice colonizing the small Southern Ocean islands came most likely from Western Europe, or via Atlantic Islands and North America which were colonized by Western European mice.

For the Kerguelen Archipelago mice, we identified two major haplotypes. One is a very common one that was previously found in Western Europe, Cameroon and USA and occurs also on other Southern Ocean islands including the Falkland Islands and Auckland Islands. We find it on Grande Terre as well as adjacent islands (colored yellow in Figure 2a). The second major haplotype is very different from the first one and was previously found in Portugal and in the USA, but is also related to a haplotype known from Cameroon. Within the Kerguelen Archipelago, it is restricted to two small neighboring islands in the Morbihan Gulf, namely Cochons Island and Cimetière Island (colored dark pink in Figure 2a).

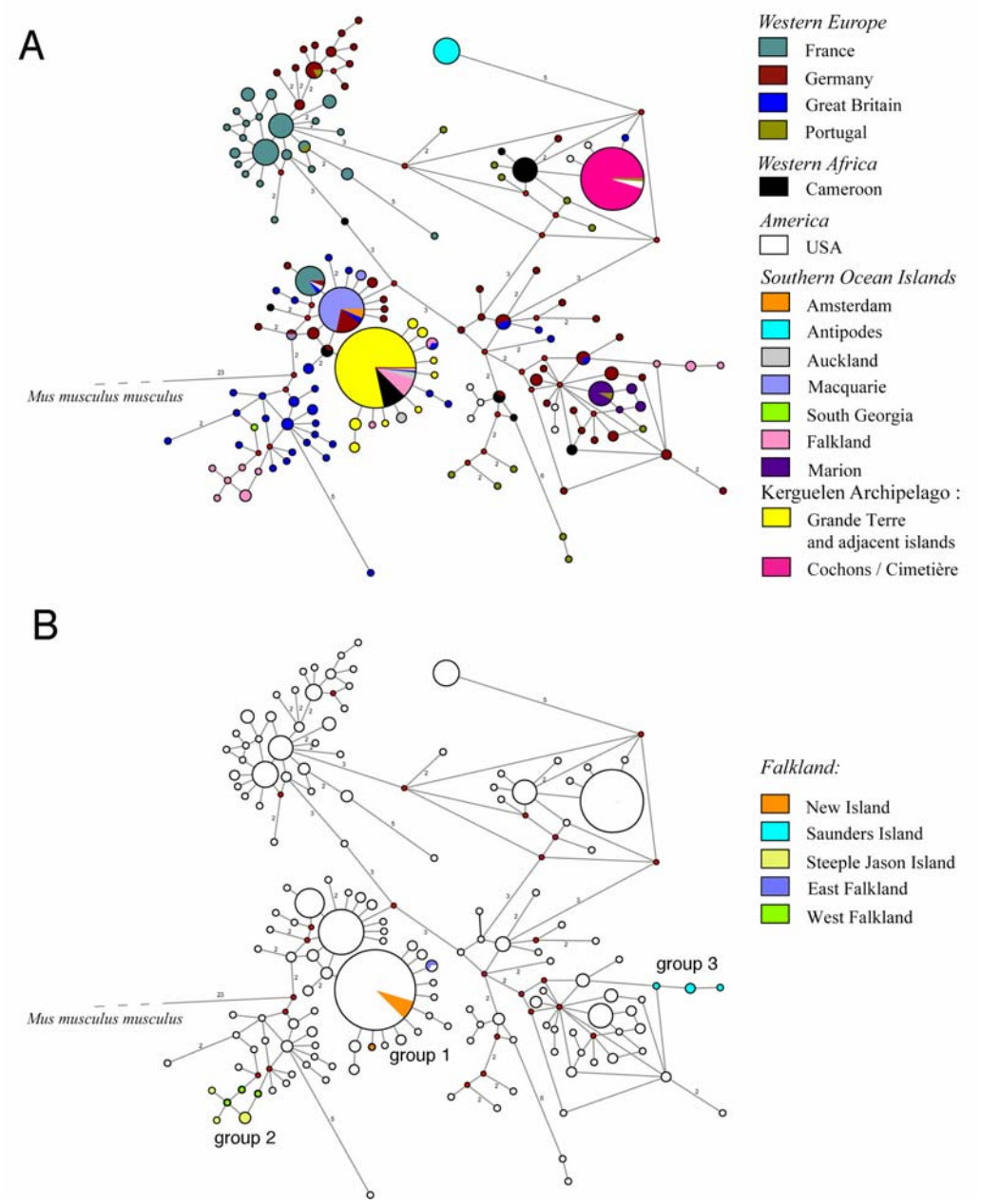


Figure 2: D-loop haplotype networks calculated using Median Joining for *M. m. domesticus* samples with *M. m. musculus* as outgroup. The size of the circles represents the frequency of the respective haplotype in our sample. Each node is one mutational step away from the next node, numbers indicate the cases where more than one step is required to join the nodes. Small red circles indicate branch splits. (A) General network including all published sequences that are related to the Kerguelen haplotypes. (B) Same network as in (A), but only with the Falkland samples highlighted.

In addition to these major haplotypes, we identified eight new haplotypes in the Kerguelen Archipelago, which are only a single or two mutational steps apart from the first major Kerguelen haplotype. It is therefore likely that these have arisen on the Kerguelen Archipelago. A comparable pattern of one major haplotype with several single step derivatives is also known for the mice that have colonized Madeira (Förster et al. 2009).

To further assess whether this is a general pattern on small islands, we have more closely inspected the haplotype distribution on the Falkland Islands. We find three closely related haplotype groups (Figure 3b). The first group is identical to the Grand Terre haplotypes on the Kerguelen and occurs in East Falkland, as well as on New Island, which is in the far West of the archipelago. The second is related to haplotypes known from Great Britain and occurs on Steeple Jason Island, which lies far out in the North-West, as well as on West Falkland. The third is related to haplotypes known from Germany and Great Britain and is found on Saunders Island, which is very close to the Northern part of West Falkland. None of these islands has disparate sets of haplotypes, thus confirming the notion of single primary colonizations. On the other hand, the colonization pattern and history across the archipelago is apparently complex, since geographic proximity within the archipelago does not correlate with haplotype similarities. More intensive sampling will be required to unravel this further.

B°/ Microsatellite data

A total of 18 unlinked autosomal microsatellites were typed for all samples and heterozygosities as well as average number of alleles were calculated for each sampling location (Table 1). All island samples show reduced heterozygosities (0.43 on average) and lower average allele numbers (2.9 on average) when compared to the standard samples from the European mainland populations Cologne-Bonn, Germany (0.84 /11.7) and Massif Central, France (0.86 / 12.1). Such a reduced heterozygosity and lower allele numbers on the islands is in principle in line with the assumption of a colonization

Table 1: Genetic parameters for 18 microsatellites loci and D-loop sequences

			autosomal microsatellite loci				mitochondrial D-loop	
area	**location**		**N**	**H_{exp}**	**H_{obs}**	**average number of alleles per locus**	**N**	**number of haplotypes**
Germany		Schleswig-Holstein	-	-	-	-	9	6
		Cologne - Bonn	43	0.84	0.61	11.7	44	26
France		Paris	20	0.46	0.47	3.0	17	1
		Massif Central	46	0.86	0.75	12.1	62	22
Cameroon		Kumba	46	0.61	0.48	6.7	58	9
Kerguelen Archipelago	Grande Terre	Port-aux-Français	41	0.48	0.44	4.1	28	3
		Jacky	29	0.48	0.49	3.3	21	3
		Isthme Bas	38	0.48	0.46	3.9	36	1
		Cap Ratmanoff	8	0.46	0.49	2.7	6	1
		Pointe du Morne	1	0.28	0.56	1.6	1	1
		Port-Couvreux	4	0.47	0.49	2.6	4	2
		Port-Jeanne d'Arc	16	0.42	0.42	3.3	15	3
		Sourcils Noirs	5	0.38	0.43	2.3	5	1
	Morbihan Gulf	Moules Island	12	0.33	0.37	2.4	10	1
		Stoll Island	4	0.38	0.48	2.1	4	1
		Australia Island	28	0.43	0.38	3.5	27	1
		Mayes Island	71	0.41	0.36	4.0	71	2
		Guillou Island	79	0.36	0.34	2.4	78	1
		Cochons Island	69	0.36	0.35	2.4	65	1
		Cimetière Island	28	0.38	0.37	2.5	27	1
other Southern Ocean islands		Marion Island	18	0.56	0.51	4.3	18	4
		Amsterdam Island	3	0.49	0.54	2.6	3	1
		South Georgia	1	0.16	0.31	1.3	1	1
		Antipodes Island	18	0.44	0.51	3.1	17	1
		Macquarie Island	40	0.42	0.38	3.3	38	3
		Auckland Island	13	0.42	0.39	3.2	13	2
	Falkland Islands	New Island	12	0.44	0.41	3.2	30	2
		Steeple Jason	5	0.33	0.33	2.3	5	3
		Saunders Island	4	0.55	0.44	3.0	4	3
		East Falkland	2	0.49	0.64	2.3	2	1
		West Falkland	3	0.48	0.49	2.7	3	3

bottleneck, but the situation is more complex. It is known that local inbreeding and communal nesting occurs in natural populations of the house mouse (Berry and Bronson 1992), which can lead to local reduction of genetic diversity (Ihle et al. 2006). The sampling scheme for the German and French standard populations took account of this effect and took samples from an extended area to obtain a measure of the average genetic diversity in the extended area (Ihle et al. 2006). But the sampling on the islands could not be done in this way, since many are actually smaller than the areas considered for the standard sampling protocol. To compare the island results with an equivalent sampling scheme on the mainland, we typed samples that were all caught in the garden of the National Library of France in Paris. Average heterozygosity (0.47) and average allele number (3.0) is indeed also lowered for these and thus more comparable to the island samples.

Previous studies found evidence that there could be differences between male and female mediated gene flow patterns (Jones et al. 1995, Förster et al. 2009). We have therefore also typed six Y-chromosomal microsatellites for those island samples where eight or more males were available (Table 2). For five of the Y-chromosomal loci we find only one major allele at most locations on the Kerguelen Archipelago, suggesting that only one Y chromosomal haplotype has been involved in the colonization (Table 2). Only Y24 on Cimetière Island is fixed for a different allele, but this is a secondary effect, since these mice are derived from Cochons Island (see below) that harbor this allele at low frequency. Additional alleles are also found at other loci, but most of these are only a single mutational step away from the major allele and have thus likely been generated after the colonization. This explains also the diversity of alleles at locus Y22, since this locus appears to be generally hypervariable, suggesting a particularly high mutation rate. Interestingly, even this hypervariable locus has only a single major allele on Cochons Island and Cimetière Island, indicative of very recent colonization or strong bottleneck effects.

Table 2: Distribution of Y-chromosomal microsatellite alleles at six loci. Top row numbers refer to allele length, the other numbers refer to the frequency of the respective allele (in percent). Alleles for the Cameroon samples are only included in the top part for comparison. The alleles for loci Y23 and Y 22 are not congruent with the major haplotype that is found on the Kerguelen Islands.

	Y6			Y12							Y21							Y24						
	120	122	124	118	124	129	132	135	137	140	295	316	318	320	321	322	324	373	392	393	395	397	399	403
Marion (8)		75	25						100		75				25			71	29					
Antipodes (12)	100					100								25		67	8					100		
Macquarie (21)		100						10	90					95	5								95	5
Cameroon (23)	95		5		93		7				5	95						5		95				
Port-aux-Français (28)	100				100							100								86	14			
Jacky (13)	100				100							100								92	8			
Isthme Bas (29)	100				100							100								100				
Port-Jeanne d´Arc (11)	100				100							90	10							100				
Moules (8)	100				100							50	50							100				
Australia (15)	100			7	93							93	7							100				
Mayes (52)	100			2	96	2						100								92	4	4		
Guillou (38)	100				100							100								100				
Cochons (36)	100				100							100								80	20			
Cimetière (16)	100				100							100									100			

	Y23						Y22																	
	315	317	319	321	323	329	239	253	255	257	259	261	263	265	267	269	271	273	275	277	279	281	291	294
Marion (8)						100		25				13	50	13										
Antipodes (12)			100				67						25				8							
Macquarie (21)		100								5		62	24	9										
Port-aux-Français (28)				93	7									4		4	18	40	21	11		4		
Jacky (13)				100												38	15	15	23	8				
Isthme Bas (29)				93	7											10	3	14	14	31	24		3	3
Port-Jeanne d´Arc (11)				100														100						
Moules (8)	50			50															14	57	29			
Australia (15)				100							7			7		40	13	20	13					
Mayes (52)				100						10	10			2	2	35	4	25	8	2				
Guillou (38)				100													16	84						
Cochons (36)			3	97										97	3									
Cimetière (16)				100									6	94										

The Y-chromosomal allele patterns from the other islands that were typed (Marion Island, Antipodes Island and Macquarie Island) are very distinct from the ones that we found in the Kerguelen, with almost no overlap in the major alleles. Thus, they represent distinct Y chromosomal haplotypes. All of the loci considered here were previously typed for the German and French populations and also showed a high diversity of alleles there (not shown). Intriguingly, however, for four of the loci the major alleles found in the Cameroon population correspond to the major alleles in the Kerguelen (included in the upper part of Table 2). This suggests that there is some relationship of the Kerguelen mice to the Cameroon mice, albeit not necessarily a direct one, since the Cameroon population represents a new colonization by itself.

C°/ Population relationships

To assess the population structure and relationships on the basis of the 18 autosomal microsatellites, we produced an allele sharing tree and run a PCA analysis (Figure 3). In the allele sharing tree, we find a coherent assignment of most populations and samples to distinct clades (Figure 3a). The sole exceptions are population samples from the Cologne- Bonn and Massif Central areas that are represented in multiple clades, likely reflecting their high diversity. With additional markers their genetic clustering was readily recovered in a previous study (Ihle et al. 2006).

Two major clades are apparent within the Kerguelen Archipelago. The Cochons/Cimetière island samples are very different from all the other islands, although they appear to be somewhat associated to the Cameroon/German clade. Among the other islands, the Guillou Island samples form a single distinct clade and the island pair Australia/Mayes a separate mixed clade (Figure 3a). The Grande Terre samples as well as Moules Island and Stoll Island are mixed among each other, without clear distinction.

The PCA analysis is largely congruent with the allele sharing tree, but shows a stronger distinction of the two Kerguelen groups and no particular association of the

Cochons/Cimetière island samples with the Cameroon/German clade (Figure 3b). On the other hand, it provides less resolution within each of the groups.

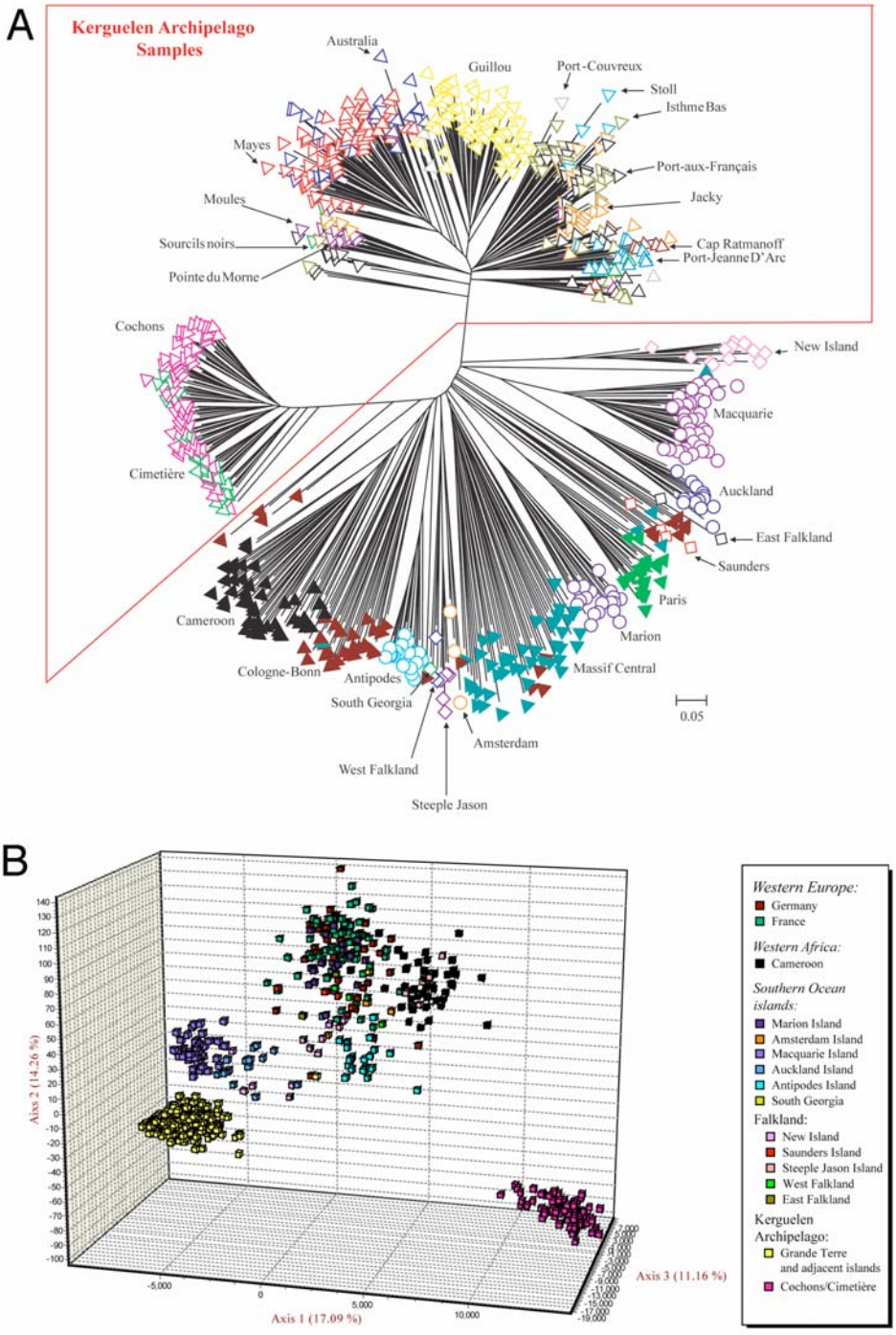


Figure 3 : Population structure based on autosomal microsatellite loci. (A) Neighborjoining tree based on the calculation of the proportion of shared alleles calculated for all individuals. Samples from the same location share the symbol/color pattern. (B) PCA plot with three axes displayed. Every square represents an individual, color patterns match the ones in (A).

To study the population structure within the Kerguelen Archipelago further, we conducted an individual-based cluster analysis with the program Instruct (Gao et al. 2007). To assess the possible number of clusters *K*, we performed runs with increasing numbers of *K* and recorded the likelihoods. A plateau was reached for *K* = 10 to 15, depending on the run, but the assignment of individuals to clusters was very unstable for these values, indicating that the lower *K* values reflect the true structure better. In Figure 4 we plot therefore only the results for values of *K* ranging from 4 to 8, alongside the number of runs that gave consistent assignments to clusters. We find that a value of *K* = 6 appears to be stable and we therefore use this for evaluation.

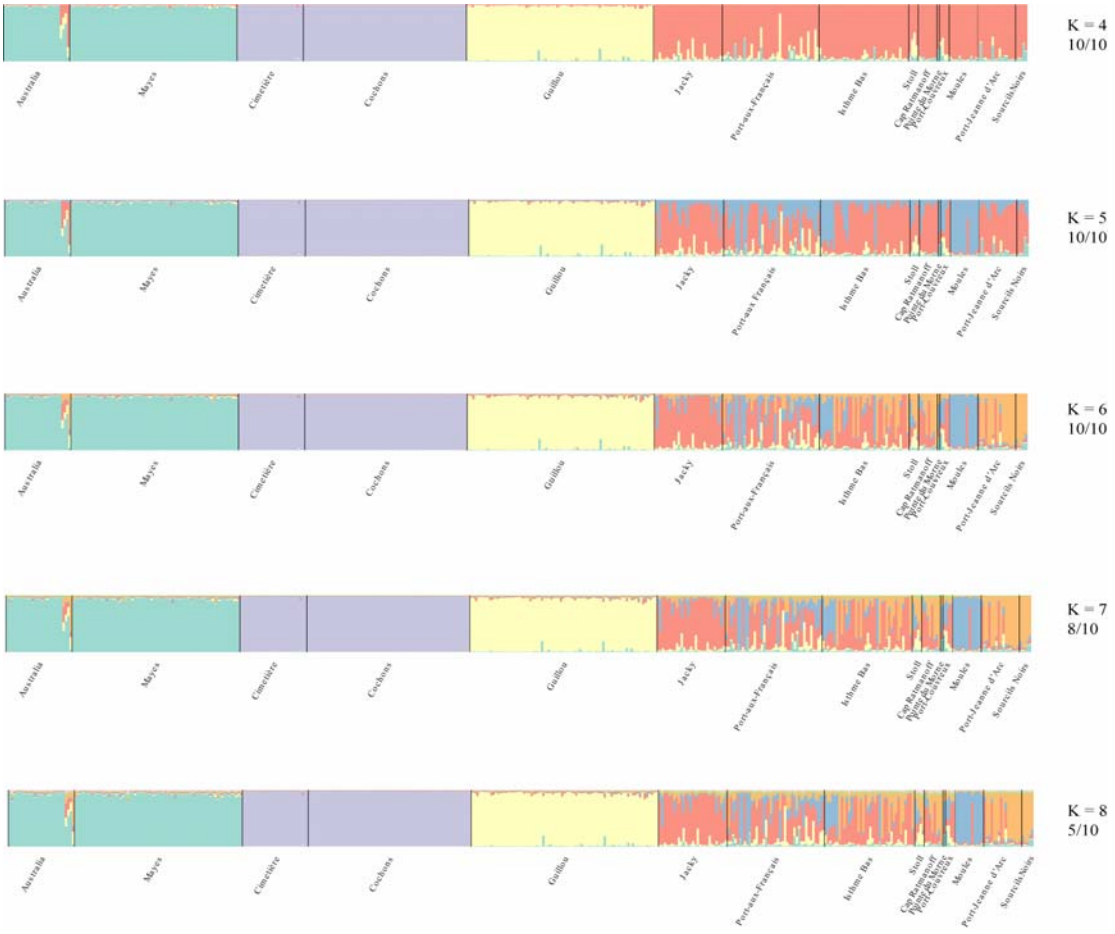


Figure 4: Structure analysis within the Kerguelen Archipelago. Only the results for the hypothesis of between 4 - 8 population groups (K = 4 to K = 8) are shown, represented by different colors. Each vertical bar represents a single individual, as well as its likelihood to belong to a given population group. The numbers below the K values represent the number of times that the same pattern was obtained in 10 independent runs of the program.

The structure analysis results thus confirm the pattern seen on the allele sharing tree. The island pairs Australia/Mayes and Cochons/Cimetière each form a single cluster and Guillou Island forms a clear separate cluster. The remaining locations are much more intermixed, only Moules Island is relatively homogeneous, although Moules-like genotypes appear to occur also in other locations. Interestingly, Stoll Island is mixed into Grande Terre populations but given the geographical location and its proximity to the main island (about 20 m), the mice might have originated from there (Figure 4). Figure 5 provides a summary diagram showing the genetic structure of all sampling sites on the Kerguelen Archipelago. Note the clear distinction of the two mitochondrial haplotype groups and the clear structure results for Guillou Island and Moules Island, as well as the island pairs Mayes/Australia 220 and Cochons/Cimetière.

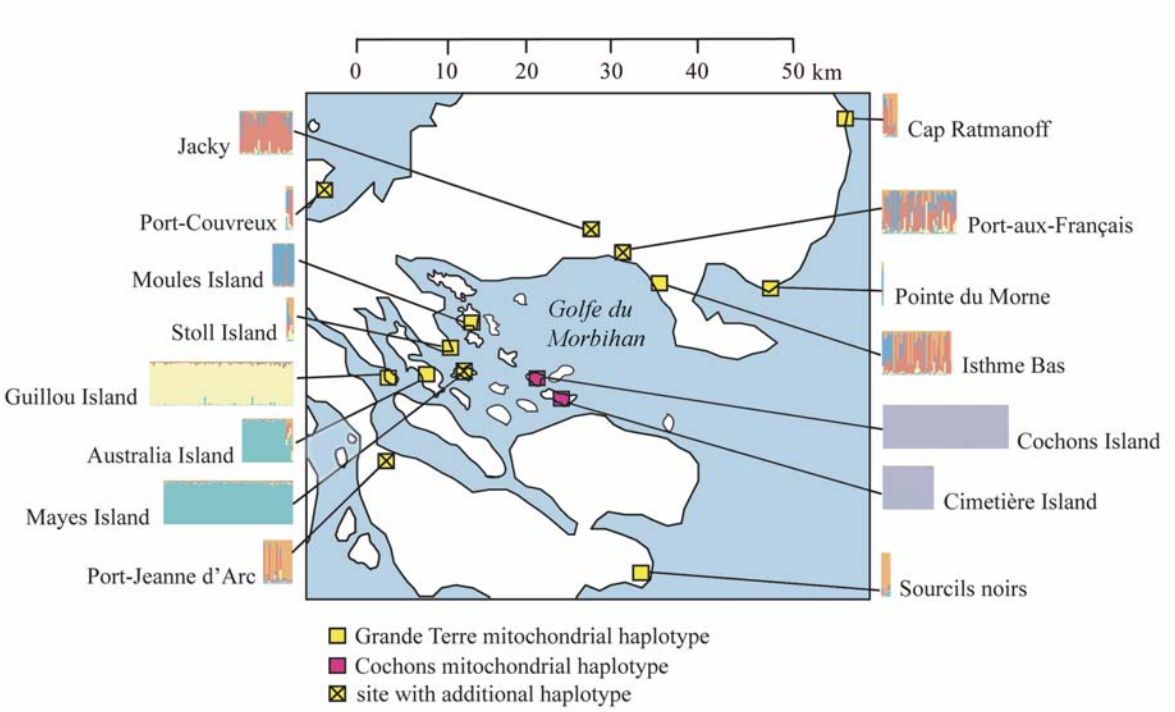


Figure 5: Summary of population structure analysis and mitochondrial haplotype distributions across the Kerguelen Archipelago.

IV- DISCUSSION

A°/ Primary colonization

Our data are compatible with the notion of an initial colonization of the Kerguelen Archipelago by a small group of mice, at the minimum the genetic equivalent of two females and one male. This can be inferred from the presence of two major mitochondrial haplotypes, as well as a single major Y-chromosomal haplotype. However, the distinct placement of the Cochons/Cimetière island samples in the allele sharing tree and the PCA analysis, as well as the presence of a single mitochondrial halotype only (i.e. no single step derivatives), suggests that these mice are in fact derived from a second more recent colonization event. Intriguingly, however, there is a similarity of the Y-chromosomal haplotypes of these mice with the rest of the Kerguelen mice. This seems unlikely to have occurred by chance, since the loci we have typed are generally polymorphic in the Western European mice. The three other island samples in our study (Macquarie Island, Antipodes Island and Marion Island) have indeed different major Y-chromosomal haplotypes (Table 2). Given that the Y-chromosomal haplotype from Cameroon is closely related to the Kerguelen haplotype, one could propose that both, the first and the second colonization came from Cameroon, which was itself colonized from Western Europe. The fact that both major mitochondrial haplotypes in the Kerguelen Archipelago are also identical or closely related to haplotypes found in Cameroon supports this notion. However, we are not confident that such a direct connection exists. Although historical ship journeys are known to have stopped both at Cameroon and the Kerguelen Archipelago, for example the German scientific expedition “Deutsche Tiefsee” in 1898 from La Valdivia (Chum 1903), these usually have had several additional stops on other islands and it is very difficult to trace how many boats went on to the Kerguelen Archipelago and the routes they took. Thus, it seems also possible that other Atlantic islands or Atlantic harbors of the USA, where most whaling boats that went to the Kerguelen Archipelago came from (the first whaling expedition known came from Nantucket Island (USA) in 1792 - Delépine 1995), share the allele patterns with the Cameroon population and could thus have been the source population for a secondary

invasion on Cochons/Cimetière Islands. Again we note that the two major mitochondrial haplotypes found in the Kerguelen occur also in the USA. More intensive sampling of the USA locations needs to be done before this question can be answered in a satisfactory way. Still, it remains noteworthy, and also unexpected, that two separate primary invasions on the Kerguelen have come from related source populations. Although our study is focused on the Kerguelen Archipelago, we also identified interesting patterns for mice on the Falklands Islands. Both in the mitochondrial haplotype analysis, as well as in the allele sharing tree, different locations in the archipelago can be molecularly differentiated, and therefore may be regarded as different populations. Three mitochondrial haplotype groups were detected, whereby two of them are shared between disparate islands. Interestingly, West Falkland and Steeple Jason Island, which are about 40km away from each other, share not only the mitochondrial haplotypes but are also grouped together in the allele sharing tree (Figure 3a). On the other hand, New Island and East Island, which share also the mitochondrial haplotypes, are very different in the allele sharing tree. The Falkland Islands have been regularly visited by boats from different nations (e.g. England, Spain, France etc.) and even today a population of around 3,000 Falkland Islanders lives there. Thus, the geographical location (near South America) and the presence of an extended human population should have increased the number of potential colonization events in the Falklands Islands. Still, it appears from our limited data that the different colonizations that have occurred on different islands of the archipelago may also have been resilient to re-invasions.

B°/ Subsequent spread

As expected, the population and allele patterns found within the Kerguelen Archipelago allow some general conclusions on the fate of populations after initial colonization. First of all, we note that these mice have retained a certain amount of genetic diversity. The heterozygosity values, as well as the average number of alleles, are comparable to the sample that we caught within a single deme in Europe (Paris). Since the mice that came with the first ship would likely represent the deme from the harbor where the ship started, we can assume that the mice entering these ships had a similarly

reduced diversity (when compared to the diversity across demes in the French and German populations). Hence, there may have been only little additional loss of genetic diversity during the ship passage and after colonization. In population genetic terms this means that the mouse population would have quickly expanded after arrival on the Kerguelen Archipelago, which would have prevented further loss of genetic diversity due to drift in small populations. Mice generally go through successions of population expansions and contractions between seasons every year (for Guillou Island, see Pisanu et al. 2002) suggesting that their life history patterns are well compatible with such a scenario. There were further colonization cycles within the Kerguelen Archipelago, namely the ones that lead to the colonization of the islands in the Morhiban Gulf. The islands Guillou, Mayes and Australia are close to Grande Terre (<500 m) and initial colonization might have occurred by animals that drifted there, or were transported by humans. Active swimming, as it was directly observed for rats (Russell et al. 2005) can also not be excluded, but seems less likely for small rodents due to the cold water temperatures (about 5°C in summer). Nevertheless, many small islands close to Grand Terre harbor mouse colonies. The initial colonizers on the small islands would have quickly expanded and retained much of their genetic variation, although the allelic patterns are sufficiently distinct to make them genetically separable from the Grande Terre population. The Grande Terre samples, on the other hand, are not genetically distinct from each other, suggesting that they are connected by continuous gene flow. This shows at the same time that very little re-invasion of the smaller islands appears to occur, since their genetic distinctness appears to be maintained (i.e. not subjected to the high levels of gene flow that occur on Grande Terre). In contrast to the islands close to Grand Terre, Cochons Island and Cimetière Island are located further away in the Morbihan Gulf and these are the ones where we indeed see a different pattern, namely a secondary invasion by mice not coming from Grand Terre (see above). They harbor only a single mitochondrial haplotype with no additional mutational variants (Figure 2) and also only a single major allele at the hypervariable Y-chromosomal locus Y22 (Table 2). This implies that the colonization has occurred later than that of the rest of the Archipelago. Indeed for Cimetière Island mice have only been recorded from 2002 onwards and it is possible that they were inadvertently transferred from the neighboring

Cochons Island by humans. Another possibility could be a natural migration, since the distance between the islands is only tenths of meters a low tide. There is evidence to suggest that the island was frequently visited for whaling and fishing activities around 100 years ago when the first mice could have arrived, although cauldrons used for extracting fat from penguin can be found only on Cochons Island today. On Cochons Island the mouse population was also inadvertently affected by a rabbit eradication program using poison from 1992 to 1997 (Chapuis et al. 2001). This could have resulted in a bottleneck and could thus explain the lowered genetic diversity. The same eradication program was also conducted on Guillou Island and could be the reason for low genetic diversity on this island as well as the different cluster in the structure and the allele sharing tree compared to other Kerguelen samples. Apart from the Cochons/Cimetière Islands case, we have no evidence for secondary successful colonization across the entire Kerguelen Archipelago, although new mice must have frequently arrived every year during whaling times. In other island mice colonization studies, it was found that although mitochondrial patterns similarly suggest only a single invasion, there could still be continued male mediated gene flow (Jones et al. 1995, Förster et al. 2009). However, given that we have only one major Y-chromosomal haplotype throughout the archipelago, this seems unlikely for the Kerguelen Archipelago. Hence, we can conclude that it must be difficult for newly arriving mice to invade the already occupied territory in the Kerguelen. Thus, our findings of single primary invasions and resilience to re-invasions corroborate the studies by Searle et al. (2009a, b, c), which have suggested that the phylogeographic patterns seen for mouse populations reflect ancient human movements, with only little disturbance by later movements. The successful experimental introduction of house mice into the Scottish Isle of May (Berry et al. 1991) does not contradict this conclusion, since in this case the mice came from another Scottish island with similar ecology, i.e. are expected to have had the same environmental adaptations at the time where they arrived.

C°/ D-loop mutation rate

We identified several new mitochondrial haplotypes, mainly in Kerguelen, but also on the Falklands, Marion Island and Macquarie Island, most of which are only one step away from the major resident haplotype (Figure 2). These can be expected to have arisen only after colonization of the respective islands. We can therefore estimate a mutation rate based on the colonization time of approximately 200 years ago. A single mutation among 834 bp is equivalent to 0.12% sequence divergence which, when divided by 200 years, gives a mutation rate of 6×10^{-6} per year. This is a factor of 150 higher than the estimate of 4×10^{-8} per year for the intraspecific mutation rates of the same D-loop region suggested by Geraldes et al. (2008) and Rajabi-Maham et al. (2008), which is already higher than the interspecific rate. The dependence of such estimates on the coalescence times considered is a well known pattern in various taxa (Ho et al. 2005), although the reasons for this a still disputed (Woodhams 2006, Galtier et al. 2009). The sequencing of the mitochondrial genomes of laboratory derived strains that were established about 100 years ago indeed suggests a 10-15 times higher mitochondrial mutation rate among such recently derived lineages, although no new mutations were found in the D-loop region (Goios et al. 2007). But even taking this rate into account, our estimate is still a factor of 10 times higher, suggesting that another process must play a role. This could be selective sweeps caused by advantageous mutations elsewhere in the mitochondrial genome and providing a new adaptation in the respective matriline. For humans it has been suggested that such mutations do indeed occur and have specifically been fixed in individuals of populations living at higher latitudes indicative of providing an adaptation to the colder climates (Ruiz- Pesini 2004).

D°/ Adaptation and genetic isolation

The ecological situation of the mice on the Kerguelen Archipelago is very different from Western European conditions, both with respect to the cold climate, as well as food conditions and the virtual absence of human settlements. Still, mouse densities can become very high, at least in regions where they have no predators (Angel

et al. 2009). Also, it has been shown for mice on sub-Antarctic islands that they have changed their preferred diet from plant seeds to macroinvertebrates for most of the year (Le Roux et al. 2001, Smith et al. 2002). All of this indicates that mice are likely to be locally adapted to these conditions. This could explain why it is so difficult for newly arriving mice to invade the existing populations. They would not only have problems to become integrated into the existing social structure, but would also have to compete with better adapted competitors. Alternatively, this may be a simple statistical effect, given that newly arriving mice would usually be few in numbers and resident mice form a large population. Thus, even if newly arriving mice mate successfully with the resident mice, the new alleles and haplotypes that they carry might not rise to sufficient frequencies to make an impact on the overall pattern. On the other hand, given that the single colonization pattern appears to be consistently found on all small islands, it seems more likely that local adaptation plays a role as well. Interestingly, the colonization of the much larger and ecologically diverse New Zealand Island is characterized by multiple invasions, including different sub-species (Searle et al. 2009c). Thus, it seems possible that the mouse populations on small islands can become more quickly ecologically and genetically isolated than mouse populations on larger islands and thereby have a higher propensity to eventually form a new subspecies or species, possibly enhanced through the fast formation of new chromosomal races (Britton-Davidian 2000).

Chapter 2

Microsatellite genome wide screen to find selective sweeps for adaptation to the Sub-Antarctic environment

I- INTRODUCTION

One challenge of evolutionary biology is to understand the genetic mechanisms of phenotypic adaptation. The most common approaches are candidate gene studies, in which a list of putative genes potentially involved in adaptations are tested (for a review see Jorgensen et al. 2009) and genome wide scans, which do not use an *a priori* hypothesis and so are able to identify new genes that have recently been subject to positive selection (for a review see Oleksyk et al. 2010).

One of the most common approaches to identify loci under selection is "hitchhiking mapping" (Harr et al. 2002, Schlötterer 2003). Under positive selection, an allele will increase its frequency within population until it reaches fixation. During this process, the particular allele will drag neighboring genomic regions to fixation as well. This phenomenon is called hitchhiking and the resulting reduction of variability at loci linked to the positively selected one is known as selective sweep (Maynard Smith and Haigh 1974). Subsequently, recombination will break the haplotype blocks and by this, after a certain amount of time, the signatures of the selective sweep will be lost. Parameters like selection intensity, recombination rates, mutation rates, and population size are important for the strength of the hitchhiking effect (Maynard Smith and Haigh 1974, Kaplan et al. 1989, Wiehe and Stephan 1993). The hitchhiking mapping approach can only identify regions which have experienced very recent and very strong positive selection on new mutations. The advantages and weaknesses of the hitching approach are well understood (Thornton et al. 2007). The main problem of such an analysis is that certain events in the history of demography such as a strong bottleneck for example can mimic the polymorphism produced by selection (Jensen et al. 2005, Hermisson 2009)

leading to the necessity of taking demographic history into account in the approach of hitchhiking mapping.

Single nucleotide polymorphisms (SNP) and microsatellites are two types of markers used in genome wide association studies. Markers must be chosen carefully. SNP screens have increased considerably during the past decade and numerous studies have been based on this marker (for example see Nielsen et al. 2005, Sabeti et al. 2007, Oleksyk et al. 2010 for a review). However, many of the SNPs were initially identified using a SNP discovery process which leads to loci with high frequency alleles being preferentially chosen. Typically, the SNPs are ascertained by direct sequencing in a relatively small population and then typed in much broader samples resulting in ascertainment bias which has to be taken into consideration during further analysis. Microsatellites have also been used as markers for selective sweep screens (Schlötterer 2001, Harr et al. 2002, Payseur et al. 2002, Kauer et al. 2003, Thomas et al. 2007, Teschke et al. 2008). Microsatellites are typically composed of short tandem repeats (2-6 bp - Tautz and Renz 1984) and are very polymorphic (Tautz 1989). Mutation rates among microsatellites are not uniform; indeed, their mutation rates increase with an increasing number of repeat units (Ellegreen 2004). Thanks to their high degrees of polymorphism, microsatellites are less subject to ascertainment bias and became one of the favorite markers for population genetic studies, identity testing, and genome mapping (Tautz 1989, Schlötterer 2001, Harr et al. 2002, Payseur et al. 2002, Ellegreen 2004, Thomas et al. 2007, Teschke et al. 2008, Jones et al. 2010).

House mice are best known as commensal pest species all over the world (Global Invasive Species Database). Indeed they have the capability to survival in an enormous variety of environmental conditions, from frozen carcasses in cold storage (Laurie 1946) to wild living in Africa (Ihle et al. 2006). Here, I focused on a particular case, the adaptation to Sub-Antarctic and Oceanic Cold islands in the southern hemisphere. Five islands were used in the study, three Sub-Antarctic (Kerguelen Archipelago, Macquarie Island and Marion Island) and two Oceanic Cold islands (Antipodes Island and Auckland Island). As already discussed in the general introduction, mice were probably introduced

at the beginning of the 19th century from Europe (Chapuis et al. 1994, Hänel and Chown 1999, Kidder 1876, Lésel and Derenne 1974, Searle et al. 2009c). The climate on these islands is described as oceanic cold. This type of climate is known to have no rigorous winter (2°C on average for the coldest month), cold summer (8°C on average for the warmest month), strong wind, and rainfall. There is no extended human settlement on any of these islands with the exception of research stations. These mice have been described to live close to their physiological limits and hence are likely to respond detectably to environmental stresses (Berry et al. 1978). This tremendous change in environmental conditions is predicted to induce behavioral changes but also to increase selective pressures on the genome, leading to allele fixation at adaptive trait loci.

A genome-wide screen using microsatellites (Thomas et al. 2007) to search for signatures of selective sweeps was performed on Kerguelen Archipelago mice and led to the identification of 38 pre-selected loci. Because the Kerguelen Archipelago mouse genomes have low levels of heterozygosity (Chapter I), a fixed region in the genome could be due to their demographic history (i.e. the strong bottleneck that mice experienced during their immigration to these islands) rather than to a selective sweep. The pre-selected loci were tested on 5 different islands representing 6 mouse populations in total. All the different island mouse populations used in this study were defined as different genetic populations with no gene flow between them (Chapter I). Looking at these different islands would therefore potentially allow the identification of genomic regions involved in parallel adaptations. Using this approach, 5 genes lying in regions which displayed extreme patterns of selective sweeps and which could putatively play a role in mouse adaptation to the Sub-Antarctic Area were identified.

A°/ Mouse samples

Mouse sampling sites are depicted in Figures 1 and 2. Samples from Cologne – Bonn (Germany, n=46) and Massif Central (France, n=46) were previously described by Ihle et al. (2006). The Kerguelen Archipelago samples were caught around the Morbihan Gulf (Figure 2A). Mice from Port-Jeanne-d'Arc (n=14), Port-aux-Français (n=41), La cabane dite "Jacky" (n=29), Guillou Island (n=30), Isthme Bas (n=6), Cimetière Island (n=28) and Cochons Island (n=30) were caught in 2008 and 2009. Mice from Australia Island (n= 4) and Mayes Island (n=57) were trapped in 1996. The sampling scheme for the Kerguelen Archipelago mouse collection is described in Chapuis et al. (2001). The 3 mice from Amsterdam Island were trapped in 2007. Marion Island mice (n=18) were caught in the Meteorological station and at the Mixed Picked Cove in 1990 and 2004 (Jansen Van Vuuren & Chown 2007). Macquarie Island (n=12), Antipodes Island (n=18), and Auckland Island (n=13) were trapped in 2005-2006 (Searle et al. 2009c). Other mouse samples from Macquarie Island caught in 2005 (n=28) were also used.

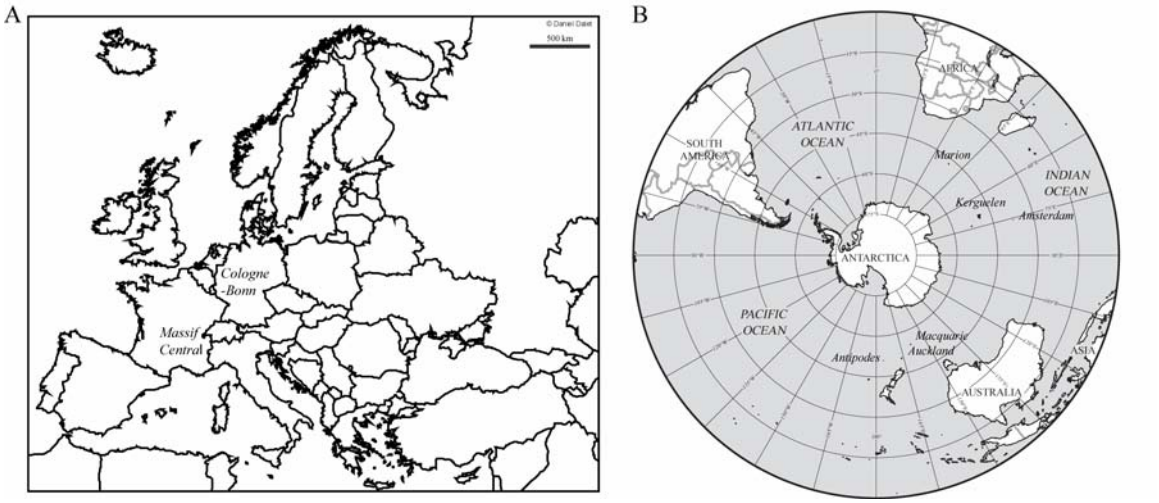


Figure 1: A- Map of the European sites sampled. B- Localization of the southern hemisphere islands named in this chapter.

B°/ Genome scan

The genome scan was performed using 960 microsatellite loci (Thomas et al. 2007). For the Kerguelen Archipelago, 30 individuals from Mayes Island, Guillou Island, and Cochons Island were pooled as well as 6 individuals from Isthme Bas, 4 from Australia Island, and 3 from Port-Jeanne d'Arc (Figure 2A and 2B). For Amsterdam Island (Figure 1B), DNA was pooled for 3 individuals. The DNA concentration for each individual was first adjusted to 100 ng/µL and subsequently pooled at a concentration of 10 ng/µL. Single individuals from Kerguelen and from European populations were typed individually in order to obtain a reference for the typical allele pattern of the locus. All PCR reactions were carried out in a 10 µL final volume using 30 ng of pooled DNA applying the following cycling protocol: 95°C for 15 minutes followed by 28 cycles of 95°C for 30s, 60°C for 1.30 min, 72°C for 1.30 min and 10 min at 70°C for elongation time. PCR products were diluted 1:20 in water. 1µL of this dilution was added to 10µL of HiDi formamide and 0.1µL of 500 ROX size standard (Applied Biosystems). The denaturation step was performed with the following incubation times: 90°C for 2min and 20°C for 5min. GeneMapper v4.0 software (Applied Bioscience) was used to visualize the data. All generated pool patterns were analyzed by eye for differences in allelic peak patterns based on pairwise comparisons between the Kerguelen Archipelago and the European (Massif Central, France and Cologne-Bonn, Germany) populations. The complexity of peaks is assumed to represent the degree of polymorphism of the population (example of peak patterns can be found in Figure 4), i.e. a complex pattern is interpreted as a signature for high polymorphism. Thirty eight loci showing a high polymorphism in European populations and a reduced polymorphism in the Kerguelen Archipelago populations were pre-selected and investigated further. Data from the European populations were previously generated by Teschke et al. (2008).

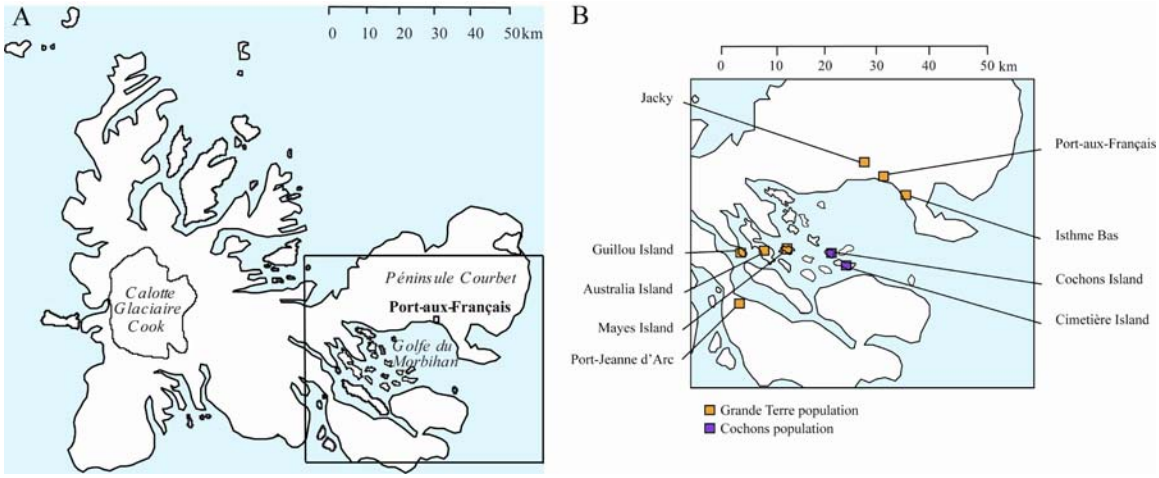


Figure 2: A- Map of the Kerguelen Archipelago. B- Map of the Morbihan Gulf with localization of the sampling sites assessed in this study. Populations on the orange sites have previously been shown as belonging to the "Grande Terre" population and in blue as belonging to the "Cochons" population (Chapter 1).

C°/ Candidate validation

All the sampling sites cited below are depicted in Figure 2B. The pre-selected candidates identified through the genome wide screen were typed individually in the Kerguelen populations as well as in the European populations. In order to look for parallel adaptation in the Southern Hemisphere, the pre-selected candidate loci were also typed from two other sub-Antarctic islands (Marion Island and Macquarie Island) and in two other Oceanic Cold islands (Auckland Island and Antipodes Island). Amplifications were carried out in 5µL final volumes using 10ng DNA template using a multiplex PCR kit (Quiagen). The PCR conditions were: 95°C for 15min followed by 28 cycles at 95°C for 30s, 60°C for 1.30min, 72°C for 1.30 min with a final extension at 72°C for 10 min. PCR products were diluted 1:20 in water. 1µL of this dilution was added to 10µL of HiDi formamide and 0.1µL of 500 ROX size standard (Applied Biosystems). The denaturation step was performed with the following incubation times: 90°C for 2min and 20°C for 5min. Samples ran on a 3730 DNA Anayser sequencer from Applied Biosystem. The alleles were analyzed using GeneMapper v4.0 software (Applied Bioscience). The gene ontology analysis was done using WebGestalt2 (Zhang et al. 2005, Duncan et al. 2010).

D°/ Statistics

Estimation of genetic diversity was calculated fro each population using MSA3.15 (Dieringer 2003) and the ln*RH* values (Kauer et al. 2003) were calculated using the following formula:

$$LnRH = \ln \frac{\left(\frac{1}{1-H(loc1, pop1)}\right)^2 - 1}{\left(\frac{1}{1-H(loc1, pop2)}\right)^2 - 1}$$

H = heterozygosity

loc = locus

pop = population

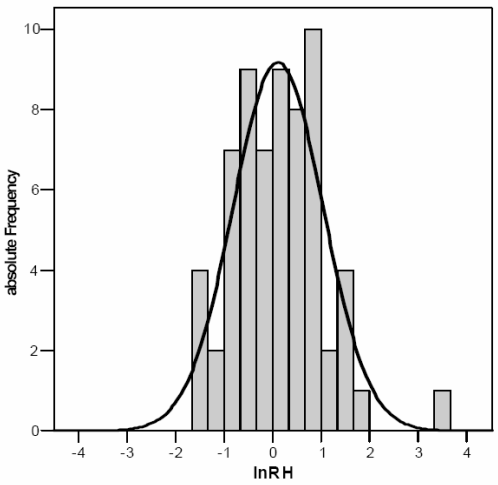


Figure 3: reference data set of 64 neutral loci genotyped between two *M. m. domesticus* populations namely France and Germany (taken from Thomas 2006)

The ln*RH* statistic allows the comparison of heterozygosity values between 2 different populations. These values have been calculated between each island and Europe (France and Germany). Unfortunately, a reference distribution of ln*RH* values made with neutral loci between Europe and each of the islands was not available so the significance of the microsatellite loci typed could not be calculated with full confidence. However, I compared the values obtained with the ln*RH* statistics with a reference data set made up of 64 neutral markers typed in the French and the German populations (collected by Ihle et al. 2006). This reference data set does not significantly deviate from a normal distribution (p=0.724 in a Kolmogorov-Smirnov test). Estimation of the mean (0.0875) and standard deviation (0.8584) of these 64 neutral loci (Thomas 2006) were used in order to identify outliers. The final candidate loci for selective sweeps were in the 0.1% ($3\sigma = 2.58$) of the tail of the reference distribution for the 6 island populations studied.

E°/ Sequencing of Il13rα2

Given the allele frequency pattern of Il13rα2, a strategy of exon sequencing was followed in order to find putatively adaptive non-synonymous mutations. Pools of 10 individuals for each of the eight populations (Massif Central, Cologne-Bonn, Grande Terre, Cochons Island, Auckland Island, Antipodes Island, Macquarie Island, and Marion Island) were made by adjusting each individual DNA sample to 50ng/µL in order to obtain a pool with a final concentration of 5ng/µL. The reactions were carried out in a 10 µL final volume with the cycling parameters as following: 95°C for 15 min followed by 35 cycles of 95°C for 30s, X°C (see Table 1 for the annealing temperature) for 1.30min, 72°C for 1min and 15 min at 70°C for elongation time. Annealing temperature was adjusted for all primer pairs. Exo-Sap purification was performed with the incubation time of 37°C for 20 min and 80°C for 20min. The cycle sequencing parameters were: 96°C for 1 min followed by 29 cycles of 96°C for 10s, 55°C for 15s and 60°C for 4min. Sequences were obtained using a 3730 DNA Analyser sequencer from Applied Biosystem The sequences generated were visualized using CodonCodeAligner Ver. 2.0.1 (CoconCode Corp.) and BioEdit ver.7.0.9.0 (Hall 1999).

Table 1: List of primer sequences and annealing temperatures used for the sequencing of Il13rα2

Forward primer name	Forward primer 5'->3'	Reverse primer name	Reverse primer 5'->3'	Annealing temperature	PCR product (bp)
Il13ra2_1_F	attactccccagaaaagccct	Il13ra2_1_R	ctgaggaatgttgtggcactagag	66°C	615
Il13ra2_2_F	actcacacaggaatgtgttcacag	Il13ra2_2_R	acagtctagtaggacacaggt	66°C	318
Il13ra2_3_F	ctggacatgaaacaagagtgtctg	Il13ra2_3_R	gatttagcccattagcagtgactc	66°C	419
Il13ra2_4_F	ataatcccacccaacaagccaa	Il13ra2_4_R	tgagggactggacgacagcct	62°C	392
Il13ra2_5_F	ttccctggtatgagcaaagctc	Il13ra2_5_R	cctctggctatttcaggaacacc	62°C	389
Il13ra2_6_F	gactccacatcttagcctagagag	Il13ra2_6_R	catggctcaaggggcacagtt	66°C	746
Il13ra2_7_F	tggatagtgaagtcagtggtcac	Il13ra2_7_R	ggaatcaggtgatgggcatttgg	66°C	323
Il13ra2_8_F	aactccttactgaggacactacc	Il13ra2_8_R	agctcatgttcctatcacagagtc	66°C	635
Il13ra2_9_F	gtgctctgtactaatcctgacag	Il13ra2_9_R	ggtttgcctatactcctcacagtg	66°C	361
Il13ra2_10_F	gcttctgggtttgcactaccatcg	Il13ra2_10_R	tgcgtgtagcaaatagtaggtgca	66°C	717
Il13ra2_11_F	gtccactagatttggccttctgga	Il13ra2_11_R	agtccaccttgattggcaagca	66°C	470
Il13ra2_12_F	tgctacccaatagccccagtt	Il13ra2_12_R	cagtgtaaagtggtggacctt	66°C	328
Il13ra2_13_F	actgccactccccaaaatgtgggt	Il13ra2_13_R	ttgtggcatctgtgcattgac	66°C	439
Il13ra2_14_F	gacttggctgtatctcggtagg	Il13ra2_14_R	gtaggctctaagggaacactggtg	66°C	660

III- RESULTS

A°/Genome scan

Table 2: Summary of the microsatellite loci screened and their localization in the mouse genome

Chromosome	No of loci screened	No of loci successfully amplified for the populations Massif Central, Cologne-Bonn, Mayes Island, Guillou Island and Cochons Island	No of pre-selected loci in the Kerguelen Archipelago mice	No of candidate loci
1	31	19		
2	66	51	3	
3	100	76	4	1
4	42	25	1	
5	49	36	4	
6	35	29	2	
7	83	69	6	
8	40	27	2	
9	31	23	1	1
10	74	63	2	
11	83	66	3	
12	14	13		
13	46	35		
14	30	27		
15	37	21	1	
16	29	24	2	1
17	35	30	1	
18	15	13	1	1
19	24	15	1	
X	96	75	4	1
Total	960	737	38	5

A genome-wide screen was obtained using a set of 960 microsatellite loci (Thomas et al. 2007). Kerguelen populations from six sites around the Morbihan gulf (see Figure 1), namely Guillou Island, Mayes Island, Cochons Island, Australia Island, Isthme bas, and Port-Jeanne d'Arc, as well as a population from Amsterdam Island were typed and compared to European populations (Teschke et al. 2008, Thomas et al. 2007). A total of 737 microsatellite loci were successfully amplified for the samples of the European populations, Guillou Island, Mayes Island, and Cochons Island (Table 2). The population pooling was of 30 individuals except in the populations where this number of mice was not available (Amsterdam Island, Port-Jeanne d'Arc, Isthme Bas, Australia Island). The comparison between Kerguelen Archipelago and the European populations was done by

eye using a pairwise comparison method (Thomas et al. 2007). Loci presenting a reduction of variability in the Kerguelen Archipelago when compared to Massif Central and Cologne-Bonn were pre-selected. Examples of these patterns are shown in Figure 4. A total of 38 loci around the mouse genome were pre-selected and investigated further (Table 2).

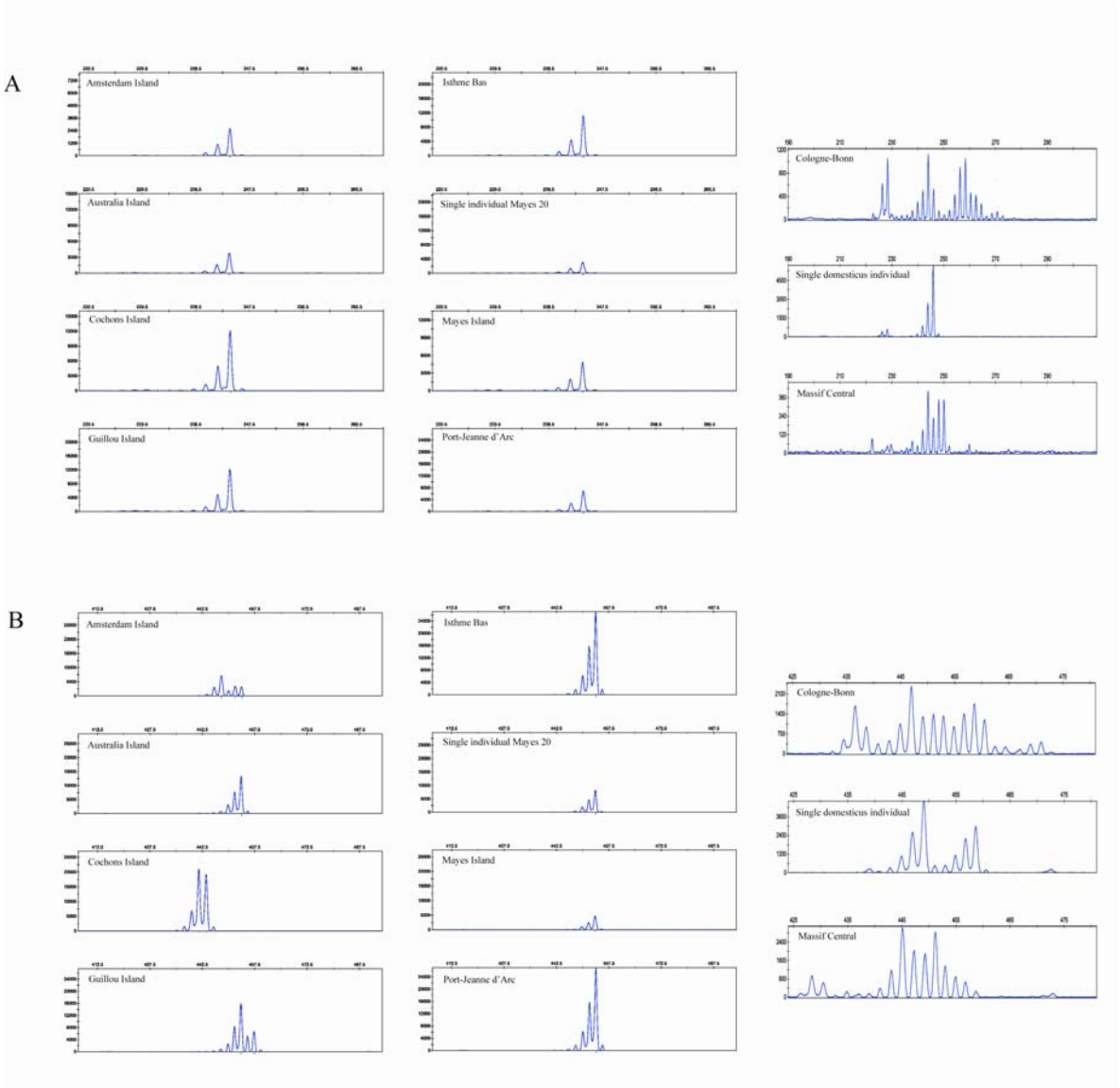


Figure 4: Example of the output from the microsatellite genome wide screen. **A**: locus X_80: From the diagram, this locus appears to be fixed for all the Kerguelen populations investigated whereas the polymorphism is higher in the European populations. **B**: locus Chr 15_13: The diagram is showing a lower polymorphism in the Kerguelen than in the European populations. Note that the Cochons Island allele is different from the rest of the Archipelago giving some hint for parallel adaptation.

B°/ Pre-selected loci

The 38 pre-selected loci were individually typed in different Kerguelen populations, namely: Port-aux-Français, Isthme Bas, la cabane dite "Jacky", Mayes Island, Australia Island, Guillou Island, Port-Jeanne d'Arc, Cochons Island, and Cimetière Island. Chapter 1 demonstrated that 2 different populations in a population genetic sense were living on the Kerguelen Archipelago. Accordingly, the Kerguelen sites are separated into these 2 genetic populations for the rest of the chapter. "Grande Terre" represents the mice captured in Isthme Bas, Australia Island, Port-Jeanne-d'Arc, Port-aux-Français, Mayes Island, La cabane dite "Jacky", and Guillou Island and "Cochons" the ones captured in Cochons Island and Cimetière Island. It also has been described that the archipelago has a rather low genetic variability (Chapter 1) and in order to distinguish between the effect of demography and selection, mice from other Southern Hemisphere islands (Marion Island, Macquarie Island, Auckland Islands, and Antipodes Island) were included in the study. All the island sites listed above were described as different populations in the population genetic sense (Chapter 1), allowing to look for parallel adaptation between all locations. If the same pattern is found on genetically different populations, this is more likely to be explained through selection rather than drift. The island mice were grouped into 6 populations, namely: Grande Terre, Cochons, Macquarie Island, Antipodes Island, Auckland Island, and Marion Island. The summary of the genetic data can be found in Table 3. The genetic diversity of the European mice (13.2 alleles per locus on average) is higher than that found on the islands (3.0 on average). This results is expected because the island mice have lower polymorphism (Chapter 1) and also because the pre-selected loci were chosen for their low degree of variability in the Kerguelen Archipelago when compared to Germany and France.

Table 3: Heterozygosity and average number of alleles per locus per island population calculated with the 38 pre-selected loci.

	N	Hexp.	Standard deviations Hexp.	Hobs.	Standard deviations Hobs.	Average number of allele per locus
Europe	92	0.82	0.08	0.52	0.16	13.24
Cochons Island	58	0.24	0.23	0.23	0.23	2.11
Grande Terre	205	0.28	0.21	0.20	0.16	3.97
Macquarie Island	40	0.42	0.21	0.38	0.22	3.03
Auckland Island	13	0.38	0.26	0.37	0.28	2.58
Antipodes Island	18	0.39	0.24	0.38	0.29	2.95
Marion Island	18	0.52	0.20	0.42	0.21	3.50

All microsatellites screened are located not more than 5,000 bp away from the start codon of neighboring genes and in 90% of the loci the distance is less than 2,000 bp, allowing a direct association to a gene (Thomas et al. 2007). In order to investigate the relationship between the candidates and whether one pathway in the pre-selected candidates is over-represented, a gene ontology (GO) analysis was performed (Figure 5). Of the 38 pre-selected loci, one was found not to be associated with a gene and 3 genes did not have an annotation. The reference used for the analysis was made with the gene associated with the microsatellite genome wide screen. The Gene Set of Analysis Toolkit was set to show the 10 most significant categories. From the 737 genes, 637 were annotated and used as reference. No categories were found enriched in molecular function and cellular component (Figure 5). However three categories in biological process (localization p-value=0.0293; establishment of localization p-value=0.022; transport p-value=0.0293) were enriched in the pre-candidate genes when compared to the reference data set of genes (Figure 5). No chromosome position enrichment was found.

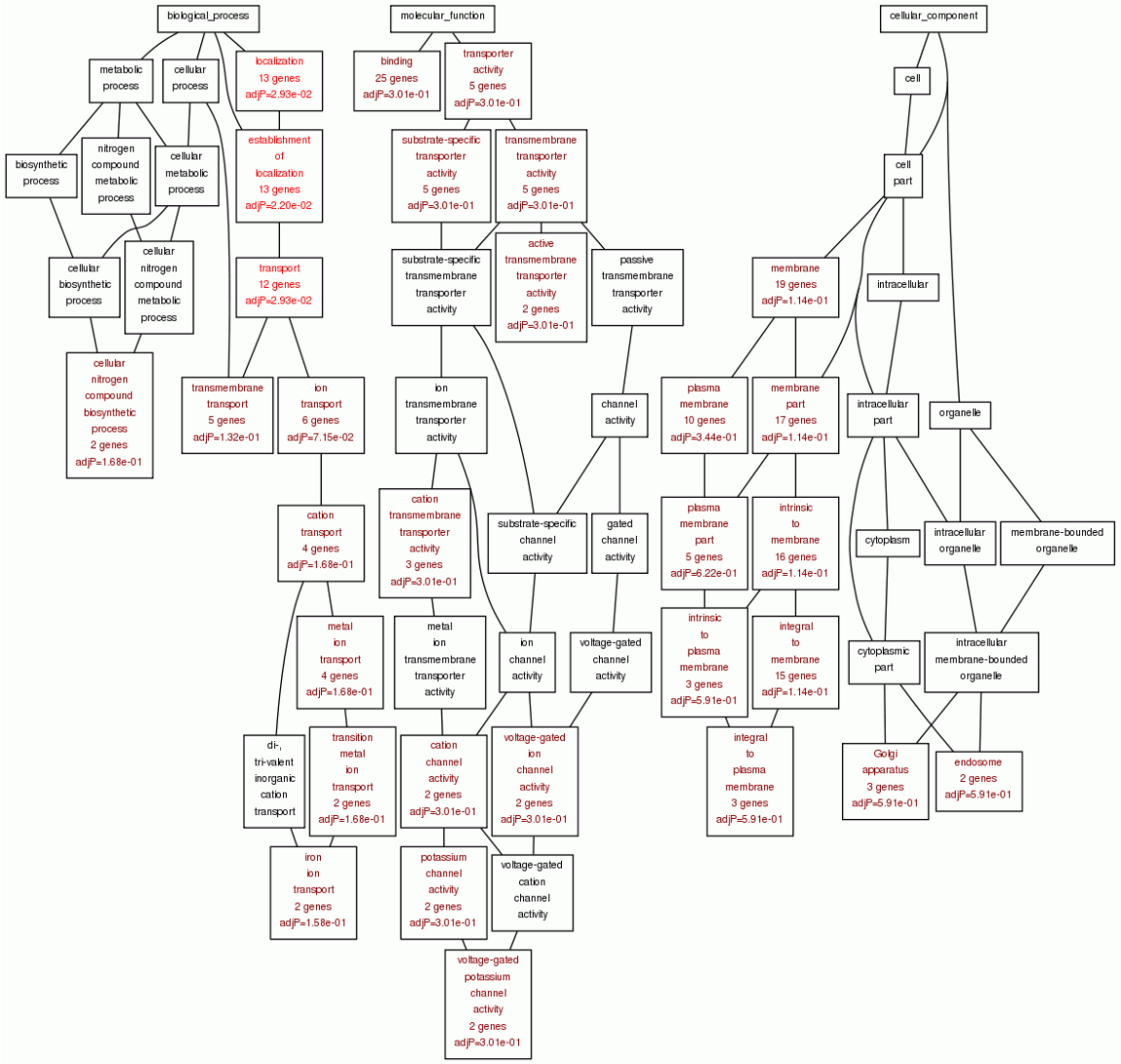


Figure 5: Enriched DAG (Directed Acylic Graph) under biological process, molecular function and cellular component. Categories in red are enriched, in brown are the categories from the top 10 with a p-values lower than 0.05 and in black are the parent categories of the top 10 categories.

Table 4: ln*RH* values calculated for all comparisons between islands and Europe. The highlighted numbers are values which fall within the 0.1% of the tails of the reference distribution. In case of all 6 populations showing an extreme ln*RH* value (in the 0.1% cut-off), the locus was selected as a candidate for a selective sweep.

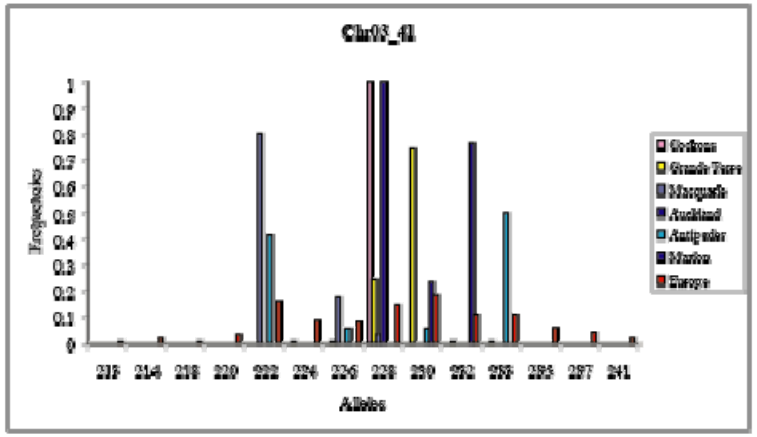
	ln*RH* (Cochons/ Europe)	ln*RH* (GrandeTerre/ Europe)	ln*RH* (Macquarie/Europe)	ln*RH* (Auckland/ Europe)	ln*RH* (Antipodes/ Europe)	ln*RH* (Marion/ Europe)
Chr02_48	-4.56	-7.08	-4.14	-2.74	-1.82	-3.65
Chr02_56	-2.71	-2.83	-5.94	-3.06	-3.99	-1.35
Chr02_58	-7.60	-3.12	-1.63	-2.88	-1.93	-2.49
D3Mit42	-0.69	-0.26	-2.13	-0.17	-0.77	-3.51
D3Mit155	-7.55	-3.67	-2.05	-2.63	-3.22	-4.16
Chr03_41	-7.66	-3.82	-4.09	-6.00	-2.64	-3.92
Chr03_49	-4.01	-1.64	-1.93	-4.31	-1.91	-1.18
Chr04_21	-5.69	-7.66	-5.93	-1.75	-4.26	-1.20
Chr05_14	-2.36	-3.11	-0.30	-4.97	-1.96	-2.40
Chr05_28	-7.01	-2.76	-1.74	-1.49	-2.35	-2.01
Chr05_31	-6.52	-3.52	-2.88	-1.60	-1.46	-0.41
Chr05_32	-0.49	-3.85	-1.48	-3.93		
D6Mit139	-4.97	-6.72	-3.23	-5.30	-4.09	-2.28
D6Mit333	-0.06	-1.53	-1.32	-3.71	-3.87	-0.01
D7Mit347	-6.43	-3.28	-1.27	-1.38	-2.42	-1.79
Chr07_18	-3.53	-1.95	-0.86	-2.71	-1.74	-1.44
Chr07_27	-1.99	-4.55	-1.26	-0.86	-1.74	-2.38
Chr07_02	-3.82	-4.45	-3.46	-2.01	-3.47	-3.87
Chr07_38	-2.91	-2.06	-3.65	-2.38	-4.15	-3.01
Chr07_58	-5.39	-1.74	-0.97	-3.80	-4.16	0.01
Chr08_02	-3.56	-2.37	-3.43	-2.95	-3.73	-1.82
Chr08_09	-6.72	-5.76	-2.57	-0.81	-4.04	-3.13
Chr09_09	-3.91	-4.69	-4.14	-6.26	-3.73	-3.17
Chr10_59	-3.33	-3.91	-3.73	-2.96	-2.09	-1.79
Chr10_73	-2.50	-2.32	-5.43	-4.76	-3.37	-1.21
Chr11_35	-5.45	-6.71	-1.06	-1.70	-2.15	-0.30
Chr11_57	-5.27	-2.50	-0.38	-0.27	-1.74	-1.62
Chr11_66	-7.87	-5.58	-3.49	-3.68	-4.67	-2.11
Chr15_13	-4.61	-6.10	-4.10	-4.37	-1.51	-3.38
Chr16_03	-7.04	-2.69	-2.63	-5.45	-3.08	-2.84
Chr16_16	-6.40	-3.71	-3.03	-1.99	-2.90	-3.37
Chr17_32	-3.18	-7.72	-5.66	-0.12	-4.03	-2.19
Chr18_08	-4.83	-5.71	-3.86	-2.99	-6.94	-2.91
Chr19_17	-2.88	-4.88	-3.44	-3.00	-2.22	-1.92
X_06	-6.64	-5.26	-5.07	-2.79	-1.15	-0.92
X_80	-7.15	-6.33	-6.78	-5.60	-4.40	-5.96
X_92	-7.85	-5.32	-2.97	-5.42	-6.66	-1.90
ChrX_127	-2.04	-5.69	-3.29	-3.10	-5.20	-0.94

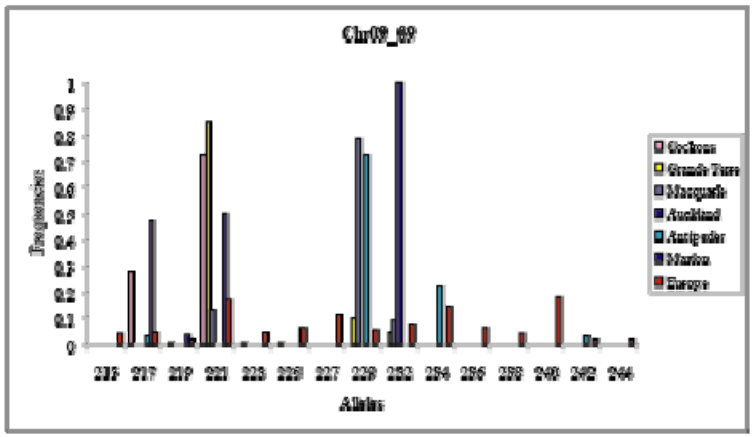
The ln*RH* values for the pre-selected loci were calculated between the island populations and Europe (Table 4). The ln*RH* values of Grande Terre and Cochons are smaller than the other 4 populations. Because the genome scan was performed on Kerguelen Archipelago mice, more loci are expected to have extreme values for these 2 populations. Note that only negative ln*RH* values were obtained, because the genome scan was performed in order to find a decrease in variability in the Kerguelen Archipelago mice, i.e. no loci presenting a higher degree of variability in the Kerguelen Archipelago populations when compared to Europe were chosen. It is also interesting to note that for all the loci typed, 82.7% of the ln*RH* values, between the island and Europe populations, are in the 5% tail of the reference distribution validating the pre-selection procedure. Five loci were in the 0.1% cut-off for all the 6 island populations and were chosen as candidates loci and assessed in more detail (see Table 4 for a data summary of these loci).

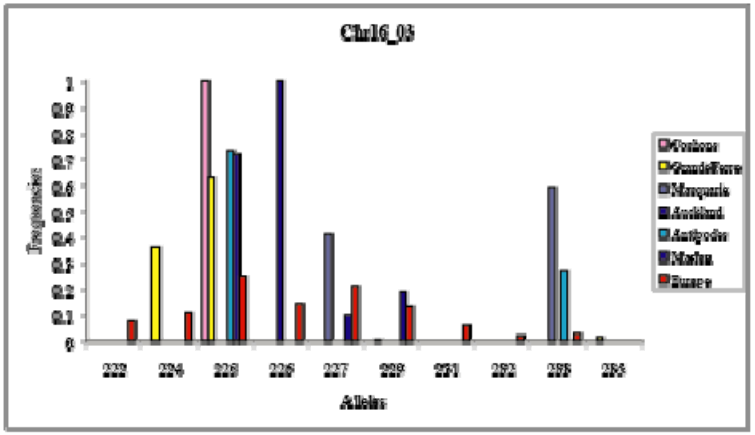
The allele frequency spectrum of the five candidate loci is presented in Figure 6. At each of these loci, the European populations show a high degree of polymorphism while the island populations have a reduced one or are even fixed (Table 5 and Figure 6). It is interesting to note that the island populations could share the same allele at high frequencies. For example, all the island populations studied here are fixed or almost fixed for the same allele 244 at the ChrX_80 locus (Table 5 and Figure 6). Given the diverse population genetic origin of all these six island populations (Chapter 1), this makes this locus a particularly interesting candidate for parallel adaptation to the southern hemisphere area. Some of the loci presented have a more diverse pattern i.e. each of the island populations have different alleles in high frequencies, for example Chr03_41. The fact that a locus repeatedly presents signs for hitchhiking (positive selection), even if the allele in high frequency is different between islands, makes it also a strong candidate for parallel adaptation. It has to be pointed out, that the microsatellites themselves are not expected to be under selection but serve only as markers for selective sweeps in the respective genomic region. Genome localizations of these loci on the mouse chromosomes are represented in Figure 7.

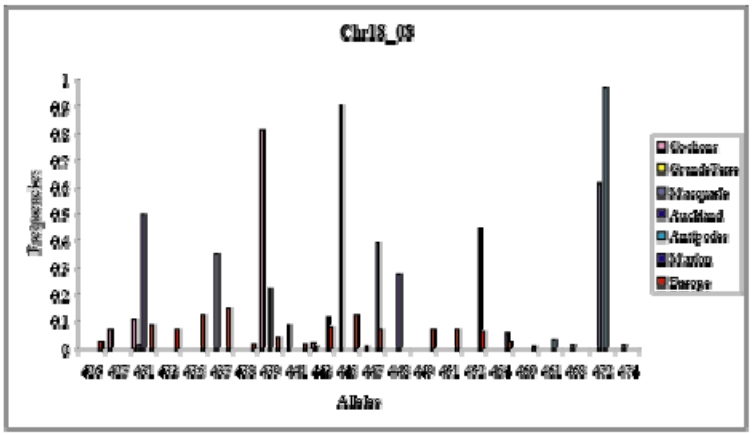
Table 5: Locus information, heterozygosity and ln*RH* values of the candidates loci found in the genome wide screen and their associated genes.

		Candidate loci				
	Marker name	Chr03_41	Chr09_09	Chr16_03	Chr18_08	X_80
Loci informations	chromosome	3	9	16	18	X
	Physical position (bp)	18320630	88222471	92346246	20668805	6251145
	Gene associated	Dnajc19	Nt5e	Kcne1	Dsg3	Il13ra2
	number of exons	6	9	2	19	14
Heterozygosity	Europe	0.89	0.90	0.84	0.92	0.86
	Cochons Island	0.00	0.40	0.00	0.33	0.00
	Grande Terre	0.39	0.26	0.48	0.19	0.04
	Macquarie Island	0.34	0.36	0.49	0.51	0.03
	Marion Island	0.37	0.54	0.45	0.67	0.00
	Auckland Island	0.00	0.00	0.00	0.66	0.08
	Antipodes Island	0.61	0.44	0.41	0.07	0.20
ln*RH* values	Cochons Island - Europe	-7.66	-3.91	-7.04	-4.83	-7.15
	Grande Terre - Europe	-3.82	-4.69	-2.69	-5.71	-6.33
	Macquarie Island - Europe	-4.09	-4.14	-2.63	-3.86	-6.78
	Marion Island - Europe	-3.92	-3.17	-2.84	-2.91	-5.96
	Auckland Island - Europe	-6.00	-6.26	-5.45	-2.99	-5.60
	Antipodes Island - Europe	-2.64	-3.73	-3.08	-6.94	-4.40









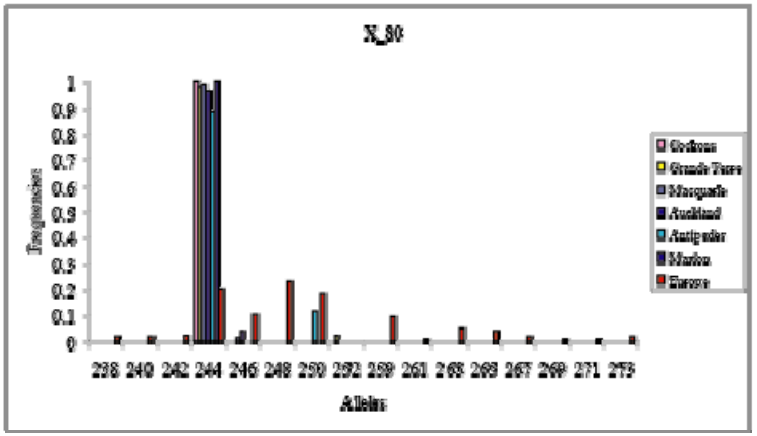


Figure 6: Allele frequency spectrum for the 5 candidate loci. The European populations display always higher degrees of polymorphism while the Southern Hemisphere mice show a reduction of variability at these loci.

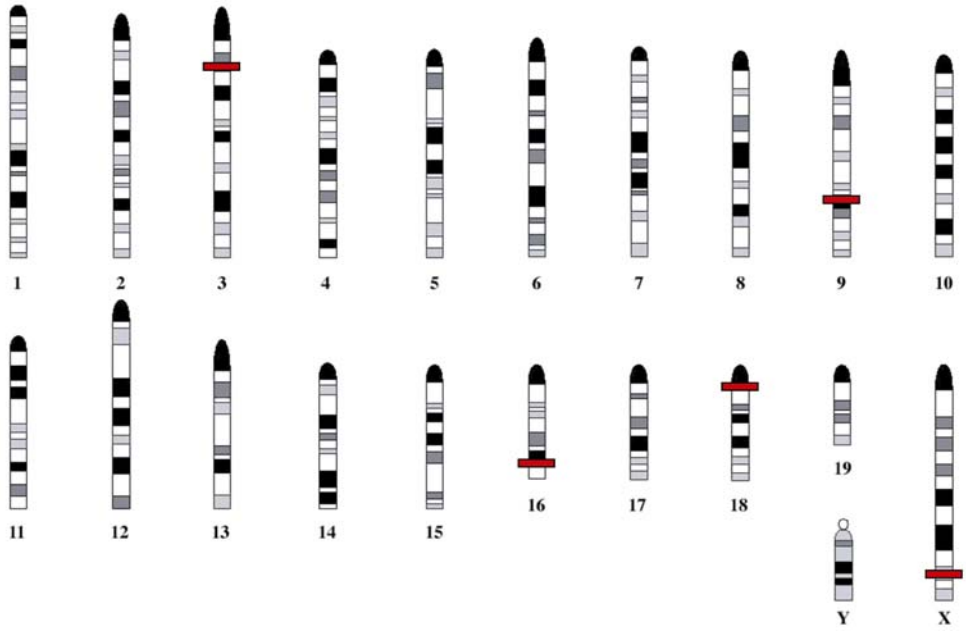


Figure 7: Distribution of the candidate loci throughout the mouse genome. All of the candidate loci are displaying extreme ln*RH* values (in the 0.1% of the tail distribution of the reference dataset)

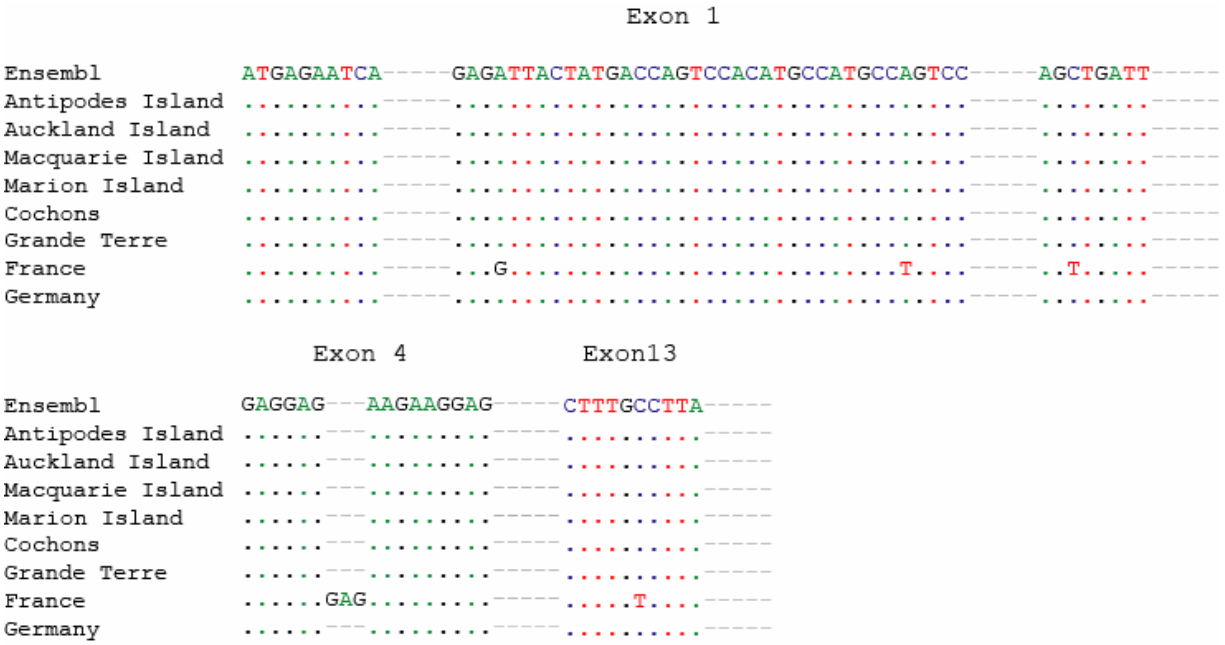


Figure 8: Sequence alignment of the Il13rα2 gene associated with the locus microsatellite X_80. Only mutations are shown.

C°/ X_80 and Il13rα2

Since the locus X_80 showed extreme patterns in term of allele frequencies and ln*RH* values (Table 5 and Figure 6), sequencing was performed for the 14 exons of the gene associated to this locus (Figure 8). Interestingly, only the French mice were polymorphic, but all of the mutations are localized in the non-transcript part of the exons. The microsatellite X_80 allele frequencies were investigated in more detail for the European populations (Figure 9) and surprisingly, heterozygosity for German mice (0.88) was higher than for the French mice (0.71).

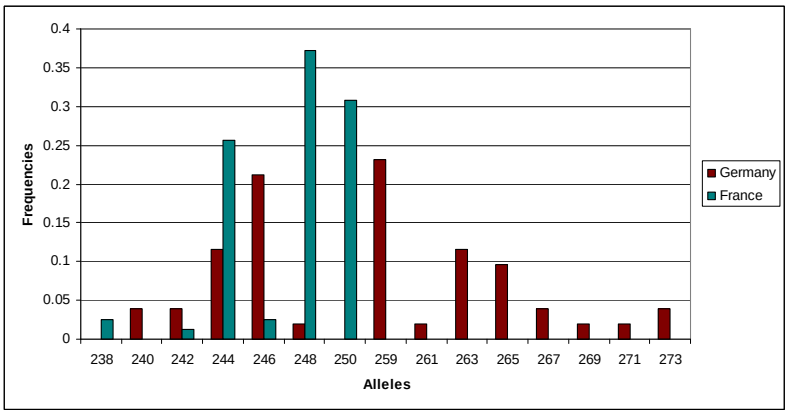


Figure 9: Allele frequency spectrum of the French and German mice at locus X_80.

IV- DISCUSSION

The Kerguelen Archipelago mice are an attractive model for studying recent adaptation. Indeed, on all these islands, the mice have been challenged by an extreme environment: feral life style, change in food availability, and an all year long colder climate. Because mice are living near their physiological limits, they are likely to have responded detectably to the environmental changes they were subjected to (Berry et al. 1978). By screening for selective sweeps, I pre-selected loci which diverge between populations more than expected by chance. However, because of the low polymorphism of the Kerguelen Archipelago mice, it is difficult to differentiate demographic history (i.e. bottlenecks) from signatures of selection. Parallel evolution has been described as common across replicate populations adapting to new environments and is predicted to have a probability under natural selection nearly twice as large as under neutrality (Orr

2005). Because life condition are similar on islands from the Sub-Antarctic area, 4 others islands (Marion Island, Auckland Island, Macquarie Island, and Antipodes Island) were also investigated allowing the discovery of candidate loci where parallel selection could have occurred.

Using the approach of pooling DNA to conduct a microsatellite genome scan (Thomas et al. 2007), I was able to genotype the Kerguelen Archipelago mouse populations for 737 microsatellites. This screening method reduced the amount of sequencing effort involved in identifying signatures of selective sweeps at the genomic scale by pre-selecting loci which presented a reduction of variability of the Kerguelen Archipelago mice when compared to European mice. As pooled DNA was used, I was not able to calculate allele frequencies. Only the degree of polymorphism of the investigated loci was assessable but due to the microsatellite slippage the polymorphism was not quantifiable. One single individual from the Kerguelen Archipelago and from the European populations each were genotyped for every microsatellite allowing us to assess the microsatellite slippage on a genotype plot, but even with this control it is not possible to determine the exact number of alleles at each locus. This and the non uniform mutation rate of microsatellites lead to the impossibility of pre-selecting loci using a statistical method, but inspection by eye allows a good pre-selection.

All the pre-selected loci were individually typed for the Grande Terre, Cochons Island, Cologne-Bonn, and Massif Central populations as well as for the other southern hemisphere islands (Macquarie Island, Auckland Island, Antipodes Island and Marion Island) in order to look for parallel adaptation. It was not possible to calculate an unequivocal significance of the ln*RH* values for the pre-selected loci because a reference distribution made with neutral loci between every island and the continent was not available. Unfortunately, generating the data was not advisable because of sequencing cost and limited samples from some islands. Instead, the distribution of the ln*RH* values of 64 neutral loci calculated between the France and Germany populations (collected by Ihle et al. 2006) was used as a reference. Since, these 2 populations, as well as the Sub-Antarctic island populations, belong to the sub-species *M. m. domesticus*, the reference

distribution gives an idea about the ln*RH* value distribution allowing the identification of loci in the distribution's tail. Thus, we could identify five candidate loci which presented extreme patterns in the six Sub-Antarctic island populations. Because all these island mice differ genetically and each of them has its own source population (Chapter 1), the presence of a single allele at high frequencies for the selected loci is unlikely to be due to genetic drift.

The candidate locus X_80 provided the best pattern of a selective sweep with almost all the island populations being fixed for the same allele (Table 5 and Figure 5). Therefore, the exons of the Il13rα2 gene associated with the candidate locus X_80 were sequenced. Unfortunately this strategy failed to discover putatively adaptive non-synonymous mutations in island mice. Nevertheless, the sequences of all populations except France are conserved and, in its sequence France only has mutations in the non-coding part of the exon.

The interleukin 13 receptor alpha 2 (Il13rα2) is know to limit IL-13 (Interleukin 13) effector function in vivo (for a review see Wynn 2003). IL-13 was described to have many functions such as regulation of gastrointestine parasite expulsion, airway hyperresponsiveness, allergic inflammation, tissue eosinophilia, mastocytosis, IgE Ab production, goblet cell hyperplasia, tumor cell growth, intracellular parasitism, tissue remodeling, or fibrosis (Wynn 2003). Studies in animal models demonstrated that T lymphocytes and cytokines have a key function in determining the outcome of parasitic infection and that IL-13 is playing a crucial role in these processes (for a review see Sher and Coffman 1992). This gene could be putatively adaptive in the fight against parasites that the mice encountered on the islands. The most commonly recorded parasites of wild house mice are 3 helminthes: *Syphacia obvelata* (21 records in 14 countries) *Taenia teaniformis* (17 records in 11 countries) and *Aspiculuris tetraptera* (15 records in 12 countries – Tattersall et al. 1994). *S. obvelata* and *A. tetraptera* are routinely found in laboratory mice (Pritchett 2007). It is interesting to note that on the Kerguelen Archipelago, only one gastrointestinal helminth was found: the nematode *S. obvelata* in the intestine (Pisanu et al. 2001). This nematode is expanding in the mouse population.

Indeed, in former studies Cochons Island mice were free of parasites (Pisanu et al. 2001) but in 2009, an *S. obvelata* was found on Cochons Island and Cimetière Island (Pisanu, Kalbe, and Hardouin, personal observation). In Macquarie Island, *S. obvelata* was found as well as the cestode *Rodentolepis fraternal* (Moro et al. 2003). The low diversity in the helminth community of these island animals, when compared to the continent, is probably due to the lower number of host founders and to their isolation or distance from the neighboring and infected populations (Dobson, 1988).

Chr09_09 is associated with the ecto-5'-nucleotidase gene (Nt5e) also called CD73. The expression of CD73 is conserved across various species. While CD73 is expressed on almost every cell type, higher amounts of expression are found in colon, kidney, and brain (Thompson et al. 2004). CD73 contributed to the protective effect of hypoxia in the inflamed intestinal mucosa and this effect is mediated by the regulation of the adenosine signaling through its tissue specific receptor (for a review see Sotnikov and Louis, 2010). CD73-/- mice cannot produce interferon α A even in response to inflammation (Sotnikov and Louis, 2010).

Chr 18_08 is associated with the desmoglein 3 gene (Dsg3). Dsg3 is part of the desmosome, an intracellular adhesive junction of epithelia cells (Buxton and Magee 1992). Desmosomes are adhesive junction containing clusters of specialized cadherines: 4 isoforms of Dsg and 3 desmocollins, all of them being part of the cadherin superfamily of cell-cell adhesions molecules (Amagai 2010). Dsg1 and Dsg3 are the major isoforms of the skin and mucous membrane. They are targeted by IgG in the autoimmune disease called pempigus (Amagai 2010).

The microsatellite Chr16_03 is predicted to be associated with the potassium voltage-gated channel subfamily E member 1 (Kcne1) gene. Kcne1 together with Kcnq1 forms a K+ channel (see Wangemann 2002 for a review). Mutations in Kcne1 or lack of Kcne1 or Kcnq1 in engineered mice have been associated with deafness (Vetter et al. 1996, Casimiro et al. 2001, Lee et al. 2000, Letts et al. 2000, for a review see Wangemann 2002). The Kcne1/Kcnq1 channel was also described as a K+ transporter in

cardiac myocytes and it plays a key role in the repolarization phase of the cardiac action potential (Barhanin et al. 1996, Sanguinetti et al. 1996, Varnum et al. 1993). Mutation in Kcne1/Kcnq1 could cause diverse forms of long QT-syndrome which is a prolongation of the action potential (Roden 2001, Wangemann 2002). Interestingly, Kcnq1 is associated with the locus Chr07_58 which was one of the pre-selected candidates and which is in the 0.1% tail distribution for Cochons, Auckland Island, and Antipodes Island. Grande Terre is in the 5% of the tail distribution for this locus. The fact that these 2 genes that are physically localized on 2 different chromosomes but together form a potassium channel were picked up by the microsatellite screen is rather a good indication of their putative function in a specific adaptation of the south hemisphere mice.

The last candidate was Chr03_41 associated with the DnaJ (Hsp40) homolog, subfamily C, member 19 (Dnajc19) gene, also called Tim14. The mitochondria contain around 1000 proteins in which only one eighth is encoded by the mitochondrial genome, the rest have to be imported from the cytosol to the mitochondrial matrix (Sickmann et al 2003, Reinders et al. 2006). Protein import to the mitochondria is particularly complicated because this organelle has 2 membranes. Preproteins pass the outer mitonchondrial membrane using the Tom complex and the inner membrane using the Tim complex (van der Laan et al. 2010 for a review). Tim14 associated with Tim16 is required to be at the inner membrane of the mitochondria, probably in close contact to the Tim23 complex to mediate the transport of preproteins (Mokranjac et al 2007). The presence of Tim14 is essential for the viability of yeast cells (van der Laan et al. 2010).

The microsatellite genome scan allowed the pre-selection of 38 candidate loci which could putatively be adaptive for life in South Hemisphere islands. Five microsatellite loci were found in the 0.1% of the tail distribution of the reference data for the 6 mouse populations studied, making their associated gene a good candidate for selective sweep. From the 5 five final candidate, the X_80 locus associated with the Il13rα2 gene was showing an extreme pattern of selective sweep. If no mutation in the coding regions of the gene were found, a putatively adaptive mutation in a regulatory region localized in the intron of the gene or elsewhere in the genome can not be excluded.

Because of its extreme pattern and its described function in the fight against parasite infections, this gene is one of the more promising candidates. Interestingly, the 2 sub-units of a K+ channel were found to have extreme pattern of selective sweep, Kcne1 even being associated with one of the 5 candidates loci. It is difficult to understand why this potassium channel could be selected for, but the fact that the 2 sub-units were found in this genome wide-screen makes it a strong candidate. Of course, the 5 candidate genes have to be studied in more detail in order to confirm if positive selection is acting on them. The main problem is the relative low polymorphism of the island mouse genomes (Chapter 1) leading to the constitution of big haplotype blocks and so the microsatellite locus could be associated to more than one genes. The follow-up experiment to this project would be to type other microsatellites near the candidate loci or to look into the region using SNP data, allowing the identification of the haplotype block in which our candidates are located. If the candidate status of these genes could be confirmed, functional experiments could follow to assess the effectiveness of the alleles that were under selection. If ideally, the detection of selected gene by statistical methods should be verified by functional approaches, such experiments are usually challenging and unfortunately, the candidate genes found have several functions and it will be difficult to understand why they have been selected for and which environmental feature(s) was involved.

Bibliography:

Amagai M: **Autoimmune and infectious skin diseases that target desmogleins.** *Proc. Jpn. Acad*. 2010, **86**:524-536.

Angel A, Wanless RM, Cooper J: **Review of impacts of the introduced house mouse on islands in the Southern Ocean: are mice equivalent to rats?** *Biological Invasions* 2009, **11**:1743-1754.

Aubert de la Rüe E: **Les terres australes**, *Presses universitaires de France*, 1953.

Auffray J, Vanlerberghe F, Britton-Davidian J: **The house mouse progression in Eurasia: a palaeontological and archaeozoological approach.** *Biological Journal of the Linnean Society* 1990, **41**(1-3), 13-25.

Avenant NL and Smith VR: **The microenvironment of the house mice on Marion Island (Sub-Antarctic)**. *Polar Biology* 2003, **26**:129-141

Avenant NL and Smith VR: **Seasonal changes in age class structure and reproductive status of house mice on Marion Island (Sub-Antarctic)**. *Polar Biology* 2004, **27**:99-111.

Barhanin J, Lesage F, Guillemare E, Fink M, Lazdunski M, Romey G : **KVLQT1 and IsK (minK) proteins associate to form the IKs cardiac potassium current**. *Nature* 1996, **384**:78-80.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F : 1996-2004 **GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations.** Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France)

Berry RJ, Bronson FH: **Life history and bioeconomy of the house mouse**. *Biological reviews of the Cambridge Philosophical Society* 1992, **67**:519-550.

Berry RJ, Scriven PN: **The genus *Mus* as a model for evolutionary studies**. *Biological Journal of the Linnean Society* 2005, **84**:332-347.

Berry RJ, Peters J, Van Aarde RJ: **Sub-Antarctic House mice: colonization, survival and selection**. *Journal of Zoology, London*, 1978b, **184**:127-141.

Berry RJ, Bonner WN, Peters J: **Natural selection in House mice (*Mus musculus*) from South Georgia (South Atlantic Ocean)**. *Journal of Zoology, London*, 1979, **189**:385-398.

Berry RJ, Peters J: **Macquarie Island House mice: A genetical isolate on a Sub-Antartic island.** *Journal of Zoology, London* 1975, **176**:375-389.

Berry RJ, Triggs GS, King P, Nash HR, Noble LR: **Hybridization and gene flow in house mice introduced into an existing population on an island**. *Journal of Zoology, London,* 1991, **225**:615-632.

Boffelli D, Nobrega MA, Rubin EM: **Comparative genomics at the vertebrate extremes**. *Nature Reviews Genetics* (2004), **5**:456-465.

Boursot P, Auffray J-C, Britton-Davidian J, Bonhomme F: **The evolution of the house mice.** *Annual Review of Ecology and Systematics* 1993, **24**:119-152.

Britton-Davidian J, Catalan J, da Graça Ramalhinho M, Ganem G, Auffray JC, Capela R, Biscoito M, Searle JB, da Luz Mathias M: **Rapid chromosomal evolution in island mice.** *Nature* 2000, **403**:158.

Buxton RS, Magee AI: **Structure and interactions of desmosomal and other cadherins**. *Seminars in Cell Biology* 1992, **3**:157-167.

Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A: **GENECODIS: A web-based tool for finding significant concurrent annotations in gene lists**. *Genome Biology* 2007, **8**(1):R3

Casimiro MC, Knollmann BC, Ebert SN, Vary JC Jr, Greene AE, Franz MR, Grinberg A, Huang SP, Pfeifer K**: Targeted disruption of the KCNQ1 gene produces a mouse model of Jervell** and L. *Proc. Natl. Acad. Sci. USA 98* 2001, 2526-2531.

Chun Carl: **Aus den Tiefen des Weltmeeres**. 2. umgearbeitete und stark vermehrte Auflage. Gustav Fischer, Jena 1903 (Online: edoc-Server der Humboldt-Universität zu Berlin)

Chapuis J-L, Boussès P, Barnaud G: **Alien mammals, impact and management in the French subantarctic islands.** *Biological Conservation* 1994, **67**:97-104

Chapuis J-L, Frénot Y, Lebouvier M : **Une gamme d'îles de référence, un atout majeur pour l'évaluation de programmes de restauration dans l'archipel de Kerguelen.** *Revue Écologie (Terre Vie),* supplément 9, 2002.

Chapuis J-L, Le Roux V, Asseline J, Lefèvre L, Kerleau F : **Eradication of rabbits by poisoning on three islands of the subantarctic Kerguelen Archipelago**. *Wildlife Research* 2001, **28**:323-331.

Copson GR: **The diet of the introduced rodents *Mus Musculus* L. and *Rattus rattus* L. on Subantarctic Macquarie Island**. *Australia Wildlife Research* 1986, **13**:441-445.

Couesnon P and Guyader A: **Histoire postale des terres australes et antarctiques françaises**. 1999.

Cucchi T: **Uluburun shipwreck stowaway house mouse: molar shape analysis and indirect clues about the vessel's last journey**. *Journal of Archaelogical Science* 2008, **35**: 2953-2959.

Cucchi T, Vigne J-D, Auffray J-C: **First occurrence of the house mouse (*Mus musculus domesticus* Schwarz & Schwarz, 1943) in the Western Mediterranean: a zooarchaeological revision of subfossil occurrences.** *Biological Journal of the Linnean Society*, 2005, **84**:429-446.

Cuthbert R and Hilton G: **Introduced house mice Mus musculus: a significant predator of endangered and endemic birds on Gough Island, South Atlantic Ocean?** *Biological Conservation*, 2004, 117:483-489.

Delépine G: **Les Allemands aux Kerguelen durant les deuxième guerre mondiale**. *T.A.A.F.*, N°26, Janvier-Mars 1964.

Delépine G: **Toponymie des terres australes françaises**. 1975.

Delépine G: **Les îles australes françaises.** Éditions Ouest France, Rennes 1995.

Delépine G: **Histoires extraordinaires et inconnues dans les mers australes**. Éditions Ouest France, Rennes 2002.

Derenne P and Mougin J-L : **Données écologiques sur les mammifères introduits de l'île aux Cochons, archipel Crozet (46°06'S, 50°14'E)**. *Mammalia* 1976, **40**:21-53

Dieringer DS, Schlötterer C: **Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets.** *Molecular Ecology Notes* 2003, **3**:167-169.

Dobson AP: **The potential of parasites to control introduced mammals**. *Conservation Biology* 1988, **2**(1):31-39.

Duncan, D.T., Prodduturi, N., Zhang, B. (2010). **WebGestalt2: an updated and expanded version of the Web-based Gene Set Analysis Toolkit**. *BMC Bioinformatics*, 11(Suppl 4):P10.

Ellegren H: **Microsatellites: simple sequences with complex evolution**. *Nature reviews genetics* 2004, **5**:435-445.

Ferreira SM, van Aarde RJ, Wassenaar TD: **Demographic responses of house mice to density and temperature on sub-Antarctic Marion Island**. *Polar biology* 2006, **30**:83-94.

Frénot Y, Chown SL, Whinam J, Selkirk PM, Convey P, Skotnicki M, Bergstrom DM: **Biological invasion in the Antarctic: extent, impacts and implications**. *Biological review* 2005, 80: 45-72.

Förster DW, Gündüz I, Nunes AC, Gabriel S, Ramalhinho MG, Mathias ML, Britton-Davidian J, Searle JB: **Molecular insights into the colonization and chromosomal diversification of Madeiran house mice**. *Molecular Ecology* 2009, **18**:4477-4494.

Gabriel SI, Jóhannesdóttir F, Jones EP, Searles JB: **Colonization, mouse-style.** *BMC Biology* 2010, **8**:131

Galtier N, Nabholz B, Glémin S, Hurst GD: **Mitochondrial DNA as a marker of molecular diversity: a reappraisal.** *Molecular Ecology* 2009, **18**:4541-4550.

Gao H, Williamson S, Bustamante CD: **A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data**. *Genetics* 2007, **176**:1635-1651.

Geraldes A, Basset P, Gibson B, Smith KL, Harr B, Yu HT, Bulatova N, Ziv Y, Nachman MW: **Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes**. *Molecular Ecology* 2008, **17**:5349-5463.

Giret A, Weis D, Grégoire M, Mattielli N, Moine B, Michon G, Scoates J, Tourpin S, Delpech G, Gerbe M-C, Doucet S, Ethien R and Cottin J-Y: **L'archipel de Kerguelen: les plus vielles îles dans le plus jeune ocean**, based on the article published in Géologues, 2003, **137**:15-23.

Goios A, Pereira L, Bogue M, Macaulay V, Amorim A: **mtDNA phylogeny and evolution of laboratory mouse strains.** *Genome Research* 2007, **17**:293-298.

Global Invasive Species Database (http://www.issg.org/database)

Guénet J-L, Bonhomme F: **Wild mice: an ever-increasing contribution to a popular mammalian model.** *Trends in Genetic* 2003, **19**:24-31.

Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symposium Series* 1999, **41**:95-98.

Hänel C & Chown, SL: **An introduction guide to the Marion and Prince Edward Island Special Nature Reserves. Fifty years after annexation**. Department of Environmental Affairs and Tourism, Pretoria, 1999.

Hartl DL, Clark AG: **Principles of population genetics**. *Sinauer Associates, Inc. Publishers* Fourth Edition, 2007.

Harr B, Kauer M, Schlötterer C: **Hitchhiking mapping: a population-based fine-mapping strategy for adaptive mutations in *Drosophila melanogaster***. *PNAS* 2002, **99**(20)12949-12954.

Hermisson J. 2009. **Who believes in whole-genome scans for selection?** *Heredity* 2009, **103**(4): 283-284.

Ho SYW, Phillips MJ, Cooper A, Drummond AJ: **Time dependency of molecular rate estimates and systematic overestimation of recent divergence times**. *Molecular Biology and Evolution* 2005, **22**:1561-1568.

Ihle S, Ravaoarimanana I,Thomas M, Tautz D: **An analysis of signatures of selective sweeps in natural populations of the house mouse.** *Molecular Biology and Evolution* 2006, **23**:790-797.

Imber MJ, Bell BD, Bell EA: **Antipodes Island birds in Autumn 2001**. *Notornis* 2005, **52**:125–132.

Jakobsson M, Rosenberg NA: **CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure**. *Bioinformatics* 2004, **23**: 1801-1806.

Jansen van Vuuren B, Chown SL: **Genetic evidence confirms the origin of the house mouse on sub-Antarctic Marion Island**. *Polar Biology* 2007, **30**:327-332.

Jensen JD, Kim K, Bauer DuMont V, Aquadro CF, Bustamante CD: **Distinguishing Between Selective Sweeps and Demography Using DNA Polymorphism Data.** *Genetics* 2005, 170(3), 1401-1410.

Jobling MA, Hurles ME, Tyler-Smith C: **Human evolutionary genetics: origins, peoples and disease**. *Garland Publishing* 2004.

John O and Whitaker JR**:** ***Mus musculus, Peromyscus maniculatus bairdi and Peromyscus leucopus*** **in Vigo County Indiana**. *Journal of Mammalogy* 1966, **47**(3):473-486.

Jones CS, Noble LR, Jones JS, Tegelström H, Triggs GS, Berry RJ: **Differential male genetic success determines gene flow in an experimentally manipulated mouse population**. *Proceedings of the Royal Society* 1995, **260**:251-256.

Jones MGW, Ryan PG: **Evidence of mouse attacks on albatross chicks on Sub-Antarctic Marion Island**. *Antarctic Science* 2010, **22**(1):39-42.

Jones EP, Jensen J-K, Magnussen E, Gregersen N, Hansen HS, Searle JB: **A molecular characterization of the charismatic Faroe house mouse**. *Biol J Linn Soc*, in press.

Joly S, Stevens MI, van Vuuran BJ: **Haplotype networks can be misleading in the presence of missing data**. *Systematic Biology* 2007, **56**:857-862.

Jorgensen TJ, Ruczinski I, Kessing B, Smith MW, Shugart YY, Alberg AJ: **Hypothesis-driven candidate gene association studies: practical design and analytical considerations**. *American Journal of Epidemiology* 2009, **170**(8):986-993.

Kaplan NL, Hudson RR, Langley CH: **The ``Hitchhiking Effect'' Revisited.** *Genetics* 1989, **123**(4), 887-899.

Kauer MO, Dieringer D, Schlötterer C: **A microsatellite variability screen for positive selection associated with the "out of Africa" habitat expansion of *Drosophila melanogaster***. *Genetics* 2003, **165**:1137-1148.

Kidder JH: **The natural history of Kerguelen Island**. *The American Naturalist* 1876, **10**:491.

Laurie EMO: **The reproduction of the House Mouse (Mus musculus) living in different environments**. Proceedings of the Royal Society 1946, **133**:248-281.

Leader-Williams N: **Reindeer on South Georgia**. *Cambridge University Press* (1988), 319 p.

Lebouvier M, Frenot Y: **Conservation and management in the French sub-Antarctic islands and surrounding seas**. *Papers and Proceedings of the Royal Society of Tasmania 2007*, **141**: 23-28.

Lee MP, Ravenel JD, Hu RJ, Lustig LR, Tomaselli G, Berger RD, Brandenburg SA, Litzi TJ, Bunton TE, Limb C, Francis H, Gorelikow M, Gu H, Washington K, Argani, P, Goldenring JR, Coffey RJ, Feinberg AP: **Targeted disruption of the Kvlqt1 gene causes deafness and gastric hyperplasia in mice.** *J. Clin. Invest.* 2000, **106**:1447-1455.

Lésel R and Derenne P: **Introducing animals to Iles Kerguelen**. *Polar Record* 1975, **17**(110): 485-494.

Le Roux V, Chapuis J-L, Frenot Y, Vernon P: **Diet of the house mouse (*Mus musculus*) on Guillou Island, Kerguelen archipelaga, Subantartic.** *Polar Biology* 2002, **25**:49-57.

Letts VA, Valenzuela A, Dunbar C, Zheng QY, Johnson KR, Frankel WN: **A new spontaneous mouse mutation in the Kcne1 gene.** *Mamm. Genome* 2000, 11:831-835.

Matthewson DC, van Aarde RJ, Skinner JD**: Population biology of house mice (Mus *musculus* L.) on sub-Antarctic Marion Island**. *South African Journal of Zoology* 1994, **29**:99-106.

Maynard Smith J and Haigh J: **The hitchhiking effect of a favourable gene**. *Genet. Res.* 1974, **23**:23-25.

Mokranjac D, Berg A, Adam A, Neupert W, Hell K: **Association of the Tim14-Tim16 subcomplex with the TIM23 translocase is crucial for function of the mitochondrial protein import motor**. *The Journal of Biological Chemistry* 2007, **282**(25):18037-18045.

Moro D, Lawson MA, Hobbs RP, Thompson RCA: **Pathogens of house mice on arid Boullanger Island and Subantarctic Macquarie Island, Autralia**. *Journal of Wildlife Disease* 2003, **39**(4):762-771.

Morse HC: **Building a better mouse: one hundred years of genetics and biology**. *The mouse in Biomedical research 2nd Edition* 2007, **1**(1):1-11.

Mougeot F, Genevois F, Bretagnolle V: **Predation on burrowing petrels by the brown skua (*Catharacta skua lönnbergi*) at Mayes Island, Kerguelen**. *The Zoological Society of London* 1998, 244:429-438.

Nielsen R: **Molecular signatures of natural selection.** *Annual Review of Genetics* 2005, **39**:197-218.

Nogales-Cadenas R, Carmona-Saez P, Vazquez M, Vicente C, Yang X, Tirado F, Carazo JM, Pascual-Montano A: **GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information**. *Nucleic Acids Research* 2009; doi: 10.1093/nar/gkp416

Nunome M, Ishimori C, Aplin KP, Tsuchiya K, Yonekawa H, Moriwaki K, Suzuki H: **Detection of recombinant haplotypes in wild mice (*Mus musculus*) provides new insights into the origin of Japanese mice.** *Molecular Ecology* 2010, **19**:2474-2489.

Oleksyk TK, Smith MW, O'Brien SJ: **Genome-wide scans for footprints of natural selection**. *Phil. Trans. R. Soc. B* 2010: **365**:185-205.

Orr HA**: The probability if parallel evolution**. *Evolution* 2005, **59**(1)216-220.

Payseur BA, Cutter AD, Nachman MW: **Searching for Evidence of Positive Selection in the Human Genome Using Patterns of Microsatellite Variability**. *Molecular Biology And Evolution* 2002, **19**:1143-1153.

Pisanu B, Chapuis J-L, Durette-Desset M-C : **Helminths from introduced small mammals on Kerguelen, Crozet, and Amsterdam Islands (Southern Indian Ocean)**. *Journal of Parasitology* 2001, **87**:1205-1208.

Pisanu B, Chapuis J-L, Durette-Desset M-C, Morand S: **Epizootiology of *Syphacia obvelata* from a domestic mouse population on the subantarctic Kerguelen archipelago**. *Journal of Parasitology* 2002, **88**:645-649.

Prager EM, Sage RD, Gyllensten U, Thomas WK, Hubner R, Jones CB, Noble L, Searle JB, Wilson AC: **Mitochondrial DNA sequence diversity and the colonization of Scandinavia by house mice from East Holstein**. *Biological Journal of the Linnean Society* 1993, **50**:85-122.

Pritchett KR: **Helminth parasites of laboratory mice**. *The mouse in Biomedical research 2nd Edition* 2007, **2**(22):551-564.

Province WB: **The origins of theoretical population genetics**. *The University of Chicago Press,* 2001.

Quillfeldt P, Schenk I, McGill RAR, Strange IJ, Masello JF, Gladbach A, Roesch V, Furness RW: **Introduced mammals coexist with seabirds at New Island, Falkland Islands: abundance, habitat preferences and stable isotope analysis of diet.** *Polar Biology* 2008, **31**:333-349.

Rajabi-Maham H, Orth A, Bonhomme F: **Phylogeography and postglacial expansion of *Mus musculus domesticus* inferred from mitochondrial DNA coalescent, from Iran to Europe**. *Molecular Ecology* 2008, **17**:627-641.

Randall C. (Ed): **Vertebrate pest-management – a guide for commercial application** (Extension Bulletin E-2050). East lansing, Michigan State University 1999.

Reinders J, Zahedi RP, Pfanner N, Meisinger C, Sickmann A: **Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics**. *J. Proteome Res*. 2006, **5**:1543–1554.

Ridley M: **Evolution** third edition. *Blackwell Publishing* 2009.

Roden DM: **Pharmacogenetics and drug-induced arrhythmias**. *Cardiovasc. Res.* 50 2001, 224-231

Rosenberg NA: **Distruct: a program for the graphical display of population structure.** *Molecular Ecology* 2004, Notes 4: 137-138.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R: **DnaSP, DNA polymorphism analyses by the coalescent and other methods**. *Bioinformatics* 2003, **19**:2496-2497.

Russell JC, Towns DR, Anderson SH, Clout MN: **Intercepting the first rat ashore.** *Nature* 2005, **437**:1107.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC: **Effects of purifying and adaptive selection on regional variation in human mtDNA.** *Science* 2004, **303**:223-226.

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., Lander, E. S., Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhao, H., Zhou, J., Gabriel, S. B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R. C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Shen, Y., Sun, W., Wang, H., Wang, Y., Wang, Y., Xiong, X., Xu, L., Waye, M. M., Tsui, S. K., Xue, H., Wong, J. T., Galver, L. M., Fan, J. B., Gunderson, K., Murray, S. S., Oliphant, A. R., Chee, M. S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J. F., Phillips, M. S., Roumy, S., Sallee, C., Verner, A., Hudson, T. J., Kwok, P. Y., Cai, D., Koboldt, D. C., Miller, R. D., Pawlikowska, L., et al. 2007 **Genome-wide detection and characterization of positive selection in human populations**. *Nature* 449, 913-8.

Sakai T, Kikkawa Y, Miura I, Inoue T, Moriwaki K, Shiroishi T, Satta Y, Takahata N, Yonekawa H: **Origins of mouse inbred strains deduced from whole-genome scanning by polymorphic microsatellite loci.** *Mammalian Genome* 2005, **16**:11-19.

Sanguinetti MC, Curran ME, Zou A, Shen J, Spector PS, Atkinson DL, Keating MT: **Coassembly of KVLQT1 and minK (IsK) proteins to form cardiac IKs potassium channel**. *Nature* 1996, **384**:80-83.

Schlötterer C.: **A microsatellite-based multilocus screen for the identification of local selective sweeps**. *Genetics* 2002, **160**:753-763.

Schlötterer C: **Hitchhiking mapping - functional genomics from the population genetics perspective**. *Trends in Genetics* 2003, **19**(1): 32-38.

Schulz M, Gales R: **Winter diet of the subantarctic skua *Catharacta Antarctica* at Macquarie Island**. *Marine Ornithology* 2004, **32**:179–181.

Searle JB, Jones CS, Gündüz I, Scascitelli M, Jones EP, Herman JS, Rambau RV, Noble LR, Berry RJ, Giménez MD, Jóhannesdóttir F: **Of mice and (Viking?) men: phylogeography of British and Irish house mice.** *Proceedings of the Royal Society* 2009a, **276**:201-207.

Searle JB, Kotlík P, Rambau RV, Marková S, Herman JS, McDevitt AD: **The Celtic fringe of Britain: insights from small mammal phylogeography.** *Proceedings of the Royal Society* 2009b, **276**:4287-4294.

Searle JB, Jamieson PM, Gunduz I, Stevens MI, Jones EP, Gemmill CEC, King CM: **The diverse origins of New Zealand house mice.** *Proceedings of the Royal Society* 2009c, **276**:209-217.

Sher A, Coffman RL: **Regulation of immunity to parasites by T cells and T cell-derived cytokines.** *Annual Review Immunology* 1992, **10**:385-409.

Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schönfisch B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C: **The proteome of Saccharomyces cerevisiae mitochondria.** *Proc. Natl. Acad. Sci. U. S. A.* 2003 **100**:13207–13212.

Singleton GR, Krebs CJ: **The secret world of wild mice**. *The house mice in biomedical research*, 2007. Vol. 1, Chap.3:25-51.

Smith VR, Avenant NL, Chown SL: **The diet and impact of house mice on a sub-Antarctic island**. *Polar Biology* 2002, **25**:703-715.

Sotnikov I, Louis NA: **CD73-Dependent regulation of interferon AlphaA and Interleukin-10 in the inflamed mucosa**. *The scientific World Journal* 2010, **10**:2167-2180.

Tamura K, Dudley J, Nei M, Kumar S: MEGA4: **Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** *Molecular Biology and Evolution* 2007, **24**:1596-1599.

Tattersall FH, Nowell F, Smith RH: **A review of the endoparasites of wild House Mice *Mus domesticus*.** *Mammal Reviews* 1994, **24**(2)61-71.

Tautz D, Renz M: **Simple sequences are ubiquitous repetitive components of eukaryotic genomes**. *Nucleic Acids Research* 1984, **12**(10)4127-4138.

Tautz D: **Hypervariability of simple sequences as a general source for polymorphic DNA markers**. *Nucleic Acids Research* 1989, **17**(16)6463-6471.

Teschke M, Mukabayire O, Wiehe T, Tautz D: **Identification of selective sweeps in closely related populations of the house mouse based on microsatellite scans**. *Genetics* 2008, **180**:1537-1545.

The 1000 Genomes Project Consortium: **A map of human genome variation from population-scale sequencing**. *Nature* 2010, **467**:1061-1073.

Thomas M: **A systematic assessment of signatures of positive selection events in natural populations of the House Mouse**. Dissertation, 2006.

Thomas M, Möller F, Wiehe T, Tautz D: **A pooling approach to detect signatures of selective sweeps in genome scans using microsatellites**. *Molecular Ecology Notes* 2007, **7**:400–403.

Thompson LF, Eltzschig HK, Ibla JC, Van De Wiele CJ, Resta R, Morote-Garcia JC, Colgan SP: **Crucial Role for Ecto-5-Nucleotidase (CD73) in Vascular Leakage during Hypoxia**. *The Journal of Experimental Medicine* 2004, **200**(11) 1395–1405.

Thornton KR, Jensen JD, Becquet C, Andolfatto P: **Progress and prospects in mapping recent selection in the genome**. *Heredity* 2007, **98**(6):340-348.

Tucker PK: **Systematics of the genus *Mus***. *The mouse in Biomedical research 2nd Edition* 2007, **1**(2):13-23.

Van Aarde RJ: **The diet and feeding behaviour of feral cats, *Felis catus* at Marion Island**. *South African Journal of Wildlife Research* 1980, **10**:123-128.

Van Aarde RJ, Ferreira SM, Wassenaar TD: **With the cats away the mice may play**. *South African Journal of Science* 1996, **92**:357-358.

Van Aarde RJ and Jackson TP: **Food, reproduction and survival in mice on Sub-Antarctic Marion Island**. *Polar Biology* 2007, **25**:703-715.

Van der Laan M, Hutu DP, Rehling P: **On the mechanism of preprotein import by the mitochondrial presequence translocase.** *Biochimica et Biophysica Acta* 2010, **1803**:732-739.

Varnum MD, Busch AE, Bond CT, Maylie J, Adelman JP: **The min K channel underlies the cardiac potassium current IKs and mediates species-specific responses to protein kinase**. *Proc. Natl. Acad. Sci.* USA 90 1993, 11528-11532.

Vetter DE, Mann JR, Wangemann P, Liu Z, McLaughlin KJ, Lesage F, Marcus DC, Lazdunski M, Heinemann SF, Barhanin J: **Inner ear defects induced by null mutation of IsK gene.** *Neuron* 1996, 17:1251-1264.

Wade CM, Kulbokas EJ, Kirby AW, Zody MC, Mullikin JC, Lander ES, Lindblad-Toh K, Daly MJ: **The mosaic structure of variation in the laboratory mouse genome**. *Nature* 2002, 420:574-578.

Wade CM, Daly MJ: **Genetic variation in laboratory mice**. *Nature Genetics* 2005, **37**(11): 1175-80.

Wangemann P: **K+ cycling and the endocochlear potential**. *Hearing Research* 2002, **165**:1-9.

Wanless RM, Angel A, Cuthbert RJ, Hilton GM, Ryan PG: **Can predation by invasive mice drive seabird extinction?** *Biology letters* 2007, **3**:241-244.

Webb PI, Ellison GTH, Skinner JD, van Aarde RJ: **Are feral house mice from the sub-Antarctic adapted to cold?** *International journal of mammalian biology* 1997, **62**:58-62.

Wiehe TH, Stephan W: **Analysis of a genetic hitchhiking model, and its application to DNA polymorphism data from *Drosophila melanogaster***. *Molecular Biology and Evolution* 1993, **10**(4):842-854

Wolfe KH, Sharp PM, Li WH: **Mutation rates differ among regions of the mammalian genome**. *Nature* 1989, **337**(6204), 283-285.

Woodhams M: **Can deleterious mutations explain the time dependency of molecular rate estimates?** *Molecular Biology and Evolution* 2006, **23**:2271-2273.

Wynn TA: IL-13 **Effector functions**. *Annu. Rev Immunol*. 2003, **21** :425-456.

Yonekawa H, Moriwaki K, Gotoh O, Miyashita N, Matsushima Y, Shi LM, Cho WS, Zhen XL, Tagashira Y: **Hybrid origin of Japanese mice *Mus musculus molossinus*: evidence from restriction analysis of mitochondrial DNA**. *Molecular Biology and Evolution* 1988, **5**:63-73.

Zhang, B., Kirov, S.A., Snoddy, J.R. (2005). **WebGestalt: an integrated system for exploring gene sets in various biological contexts**. *Nucleic Acids Res*, 33(Web Server issue), W741-748.

Affidavit

I declare that the dissertation in form and content and except for advices given by my supervisor constitutes my own work. The first chapter: House mouse colonization patterns on the sub-Antarctic Kerguelen Archipelago suggest singular primary invasions and resilience against re-invasion was accepted on the 26th of October 2010 as a research article in BMC Evolutionary Biology. This work has been undertaken in compliance with the German Research Foundation's (Deutsche Forschungsgemeinschaft, DFG) rules of good academic practice.

Plön, 6th of December 2010

Emilie Hardouin

CURRICULUM VITAE :

EMILIE HARDOUIN

Address:	Klanderstr. 18A 24306 Plön, Germany
Citizenship:	French
Birthdates:	1 September 1982
Birthplace:	Angers, France

SCHOOL:

1988 – 1993	Ecole Primaire Marie Talet, Angers, France
1993-1997	Collège Montaigne, Angers, France
1997-2000	Lycée Joachim du Bellay, Angers, France
2000	Baccalauréat option science with honor

UNIVERSITY:

2002 - 2003:	First year of medicine University of Angers, France
2003 - 2004:	DEUG of Biology with honors University of Angers, France.
2004 - 2005:	Bachelor of Biology with honors University Paris 7 Denis Diderot, France.
2005 - 2007:	Master of Biology and Magistère of genetics with honor University Paris 7 Denis Diderot, France
2007 - Present:	PhD student in the Max Planck Institute of Evolutionary Biology Kiel University, Germany. Supervisor: Prof. Dr. Tautz