FACULTÉ DE BIOLOGIE ET DE MÉDECINE

Département d'Écologie et Évolution
Laboratoire de Biologie de la Conservation

Population genetics of Capercaillie (*Tetrao urogallus*) in the Jura and the Pyrenees: a non-invasive approach to avian conservation genetics

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la
Faculté de Biologie et de Médecine
de l'Université de Lausanne

par

Sébastien Regnaut

Biologiste diplômé
de l'Université de Lund (Suède)

Jury

Prof. Jacques Hausser, Président et Rapporteur
Dr Luca Fumagalli, Directeur de thèse
Dr Pierre Taberlet, expert
Dr Isabelle Dupanloup, experte

Lausanne
2004

Ce travail est dédié à mon grand père
Kléber Lagneaux
qui m'a initié et intéressé
aux choses de l'esprit et des sciences

RÉSUMÉ

Le Grand tétras est un galliforme de montagne apparenté au faisan et au tétras lyre. Il est distribué de manière continue à travers la toundra et les montagnes de moyenne altitude en Europe de l'ouest. Toutefois, les populations d'Europe de l'ouest ont subi un déclin constant au cours des derniers siècles. Les causes de ce déclin sont probablement liées à l'activité humaine, telle que l'élevage ou le tourisme, qui ont engendré une modification et une fragmentation de l'habitat de l'espèce. Malheureusement, les populations soumises à de forts déclins démographiques peuvent subir des effets génétiques (augmentation de la consanguinité et perte de diversité génétique) pouvant diminuer leur potentiel de reproduction et conduire irrémédiablement à l'extinction.
Cette thèse présente les analyses conduites dans le but d'estimer l'impact du déclin démographique des populations de Grand tétras sur l'étendue et la distribution de leur variabilité génétique dans le Jura et dans les Pyrénées.
Du fait de la législation locale protégeant les tétraonidés en général, mais également en raison de la biologie très cryptique du Grand tétras, l'ensemble des analyses de cette étude a été réalisé à partir de matériel génétique extrait des fientes (ou échantillonnage génétique non invasif).
Dans la première partie de l'étude, je détaille les protocoles d'extraction d'ADN et d'amplification par PCR modifiés à partir des protocoles classiques utilisant des échantillons conventionnels, riches en ADN. L'utilisation d'ADN fécal impose des contraintes dues à la mauvaise qualité et à la faible quantité du matériel génétique à disposition dans les fientes. Ces contraintes ont pu être partiellement contournées en réalisant des répétitions multiples du génotypage afin d'obtenir un degré de fiabilité suffisante. J'ai également analysé les causes de la dégradation de l'ADN dans les excréments. Parmi les causes les plus communes, telles que l'activité bactérienne, l'hydrolyse spontanée et la dégradation enzymatique par les DNases libres, c'est ce dernier facteur qui apparaît comme étant la cause majeure et la plus rapide responsable de la dégradation de la qualité des échantillons. La rapidité de l'action enzymatique suggère que les plans d'échantillonnages de excréments sur le terrain pourraient être optimisés en les réalisant dans des conditions climatiques froides et sèches, favorisant ainsi l'inhibition des DNases.

La seconde partie de la thèse est une étude par simulation visant à déterminer la capacité du logiciel *Structure* à identifier les structures génétiques complexes et hiérarchiques fréquemment rencontrées dans les populations naturelles, et ce en utilisant différents types de marqueurs génétiques.

Les troisième et quatrième parties de cette thèse décrivent le statut génétique des populations résiduelles du Jura et des Pyrénées à partir de l'analyse de 11 loci microsatellites. Nous n'avons pas pu mettre en évidence dans les deux populations des effets liés à la consanguinité ou à la réduction de la diversité génétique. De plus, la différenciation génétique entre les patches d'habitats favorables reste modérée et corrélée à la distance géographique, ce qui suggère que la dispersion d'individus entre les patches a été importante au moins pendant ces dernières générations. La comparaison des paramètres de la diversité génétique avec ceux d'autres populations de Grand tétras, ou d'autres espèces proches, indique que la population du Jura a retenu une proportion importante de sa diversité originelle. Ces résultats suggèrent que le déclin récent des populations a jusqu'ici eu un impact modéré sur les facteurs génétiques et que ces populations semblent avoir conservé le potentiel génétique nécessaire à leur survie à long terme.

Finalement, en cinquième partie, l'analyse de l'apparentement entre les mâles qui participent à la parade sur les places de chant (leks) indique que ces derniers sont distribués en agrégats de manière non aléatoire, préférentiellement entre individus apparentés. De plus, la corrélation entre les distances génétique et géographique entre les leks est en accord avec les motifs d'isolement par la distance mis en évidence à d'autres niveaux hiérarchiques (entre patches d'habitat et populations), ainsi qu'avec les études menées sur d'autres espèces ayant choisi ce même système de reproduction.

En conclusion, cette première étude basée uniquement sur de l'ADN nucléaire aviaire extrait à partir de fèces a fourni des informations nouvelles qui n'auraient pas pu être obtenues par une méthode d'observation sur le terrain ou d'échantillonnage génétique classique. Aucun oiseau n'a été dérangé ou capturé, et les résultats sont comparables à d'autres études concernant des espèces proches. Néanmoins, la taille de ces populations approche des niveaux au dessous desquels la survie à long terme est fortement incertaine. La persistance de la diversité génétique pour les prochaines générations reste en conséquence liée à la survie des adultes et à une reprise du succès de la reproduction.

ABSTRACT

Capercaillie (*Tetrao urogallus*) is a large grouse that is continuously distributed across the tundra and the mid-high mountains of Western Europe. However, the populations in Western Europe have been showing a constant decline during the last decades. The causes for this decline are possibly related to human activities, such as cattle breeding and tourism that have both led to habitat modification and fragmentation. Unfortunately, populations that have undergone drastic demographic bottlenecks often go through genetic processes of inbreeding and loss of diversity that decrease their fitness and eventually lead to extinction.

This thesis presents the investigations conducted to estimate the impact of the demographic decline of capercaillie populations on the extent and distribution of their genetic variability in the Jura and in the Pyrenees mountains.

Because grouse are protected by wildlife legislation, and also because of the cryptic behaviour of capercaillie, all DNA material used in this study was extracted from faeces (non-invasive genetic sampling).

In the first part of my thesis, I detail the protocols of DNA extraction and PCR amplification adapted from classical methods using conventional DNA-rich samples. The use of faecal DNA imposes specific constraints due to the low quantity and the highly degraded genetic material available. These constraints are partially overcome by performing multiple genotyping repetitions to obtain sufficient reliability. I also investigate the causes of DNA degradation in faeces. Among the main degraders, namely bacterial activity, spontaneous hydrolysis, and free-DNase activities, the latter was pointed out as the most important according to our experiments. These enzymes degrade DNA very rapidly, and, as a consequence, faeces sampling schemes must be planned preferably in cold and dry weather conditions, allowing for enzyme activity inhibition.

The second part of the thesis is a simulation study aiming to assess the capacity of the software *Structure* to detect population structure in hierarchical models relevant to situations encountered in wild populations, using several genetic markers. The methods implemented in *Structure* appear efficient in detecting the highest hierarchical structure.

The third and fourth parts of the thesis describe the population genetics status of the remaining Jura and Pyrenees populations using 11

microsatellite loci. In either of these populations, no inbreeding nor reduced genetic diversity was detected. Furthermore, the genetic differentiation between patches defined by habitat suitability remains moderate and correlated with geographical distance, suggesting that significant dispersion between patches was at work at least until the last generations. The comparison of diversity indicators with other species or other populations of capercaillie indicate that population in the Jura has retained a large part of its original genetic diversity. These results suggest that the recent decline has had so forth a moderate impact on genetic factors and that these populations might have retained the potential for long term survival, if the decline is stopped.

Finally, in the fifth part, the analysis of relatedness between males participating in the reproduction parade, or lek, indicate that capercaillie males, like has been shown for some other grouse species, gather on leks among individuals that are more related than the average of the population. This pattern appears to be due to both population structure and kin-association.

As a conclusion, this first study relying exclusively on nuclear DNA extracted from faeces has provided novel information that was not available through field observation or classical genetic sampling. No bird has been captured or disturbed, and the results are consistent with other studies of closely related species. However, the size of these populations is approaching thresholds below which long-term survival is unlikely. The persistence of genetic diversity for the forthcoming generations remains therefore bond to adult survival and to the increase of reproduction success.

TABLE DES MATIÈRES

REMERCIEMENTS

Je souhaite remercier les membres de Jury qui ont participé à la relecture de mon manuscrit et l'ont éclairé de commentaires avisés tant sur le fond que sur la forme.

Merci à Luca Fumagalli de m'avoir accueilli au sein de son laboratoire, de m'avoir fait profiter des ses connaissances et de son expérience.

Merci au Professeur Jacques Hausser pour avoir été le rapporteur de thèse, de m'avoir accueilli comme superviseur la première année, et aussi pour avoir été son assistant à l'enseignement pendant cinq ans.

Merci à Pierre Taberlet pour avoir participé à la mise au point des méthodes d'analyse génétique, et pour avoir collaboré à ma thèse tant pour la mise en route du projet que pour l'établissement de collaborations fructueuses avec des spécialistes du Grand Tétras. Grâce à son expérience, j'ai pu faire des progrès cruciaux aux moments opportuns et ai pu suivre des repères méthodologiques clairs et efficaces.

Merci enfin à Isabelle Dupanloup pour ses commentaires éclairés en génétique des populations, et pour son temps passé sur mes différents manuscrits. Merci pour son soutien et pour son aide scientifique précieuse et encourageante.

Je remercie aussi le Laboratoire de Biologie de la Conservation, ainsi que ses membres et sa direction pour m'avoir permis de débuter et mener à bien ce projet en mettant en œuvre les moyens techniques et financiers nécessaires.

Je remercie le Département d'Ecologie et Evolution, ses membres et sa direction de m'avoir accueilli comme enseignant et comme étudiant. Cet environnement m'a fourni les conditions de travail et l'encadrement scientifique du meilleur niveau. J'ai aussi profité de la renommée et de la qualité de notre département pour accéder à toutes les compétences dont j'ai eu besoin, à toutes les ressources matérielles et bibliographiques et à toutes les collaborations nécessaires aux avancées d'un tel projet. J'espère que le DEE continuera longtemps à apporter à d'autres chercheurs ce qu'il m'a apporté.

Un grand merci aussi aux étudiants de Biologie que j'ai connu aux cours de ces cinq années. Ils m'ont donné la mesure de l'importance du travail d'enseignement pour un chercheur. Ils m'ont aussi fourni un miroir qui m'a permis de me remettre souvent en question et, je l'espère, d'évoluer dans ma façon d'appréhender la biologie, le naturalisme et les relations de groupe. Je remercie en particulier les nombreux étudiants, souvent brillants et intéressants, avec lesquels j'ai pu collaborer au cours des travaux qu'ils ont effectué pour leur second cycle.

J'exprime encore des remerciements particuliers à Emmanuel Ménoni pour sa collaboration scientifique à ce projet. Cette collaboration a été nécessaire au cours des différentes phases de ma thèse. Je le remercie d'avoir mis à ma disposition une grande quantité d'échantillons et d'informations, ainsi que son excellente connaissance de la biologie des tétraonidés et des Pyrénées. Je le remercie enfin pour sa gentillesse, sa compétence, et pour m'avoir fait découvrir ses montagnes.

Je remercie Sébastien Sachot de m'avoir initié à la biologie du grand tétras et des tétraonidés en général, de m'avoir permis de réaliser mon échantillonnage et de m'avoir

fait connaître le Jura par sa faune, sa flore et sa géographie. Je le remercie aussi pour m'avoir mis en contact avec les gestionnaires de la faune et de la forêt en Suisse.

Je suis aussi extrêmement reconnaissant à Bernard Leclercq pour m'avoir suivi au cours de ce projet, de m'avoir fourni une somme d'échantillons et d'informations qui furent nécessaires à mon étude. Je le remercie aussi pour sa gentillesse, pour son aide constante, pour m'avoir guidé et éclairé de son expérience et de ses connaissances.

Merci encore à Jérôme Goudet pour son enseignement, pour la collaboration très instructive qu'il m'a proposé, merci pour son aide technique et pour ses idées qui furent aussi parfois déterminantes.

Je remercie aussi Nelly di Marco et Nadège Duvoisin pour leur aide et pour m'avoir fait profiter de leur travail et de leur compétence. Merci aussi pour votre soutien et pour les heures passées enfermés dans les labos.

Un grand merci aussi à Mme Maret et à France Pham pour leur aide administrative et leur soutien réconfortant tout au long de ma thèse.

Je remercie tout particulièrement le Professeur Peter Vogel pour avoir été son assistant à l'enseignement. J'ai appris beaucoup de la biologie et de l'enseignement à son contact. Son support, sa gentillesse et son soutien m'ont été extrêmement profitables à de nombreuses reprises.

Je remercie aussi grandement Alain Bloc et les employés de la Réserve Naturelle de la Haute Chaîne du Jura qui m'ont fourni un échantillonnage extrêmement précieux et une aide logistique cruciale. De la même façon, je remercie les personnes qui ont participé à la constitution de mon échantillonnage :, F. Preiss et le Groupe Tétras-Vosges, ainsi que tous les agents de l'ONCFS, de la Conservation de la Faune du Canton de Vaud et les volontaires qui ont participé à l'échantillonnage : MM. Arbel, Baraillé, Berjoan, Bibal, Bousset, Caillabet, Candebat, Cazassus, Chamouton, Corcelle, Covato, De Meerler, Depraz, Devilliers, Fabien, Filliol, Francès, Freches, Gazaux, Getaz, Guisset, Lacoste, Lacroix, Lacrouts, Leblanc, Lecomte, Mallejac, Montadert, Monteil, Montoloy, Moreau, Morsheidt, Mourany, Novoa, Olive, Patthey, Petit, Pibouleau, Rachou-Langlatte, Reymond, Roland, Ronoblet, Rumeau, Sébirou, StOyand, Tudon, Vacher, Weigmuller et Willens. Leur participation a, en grande partie, permis la réalisation de mon étude, et je leur suis redevable de leur temps ainsi que de la qualité de la collaboration qu'ils m'ont fourni.

Je dois aussi beaucoup à Hélène Niculita et à Pierre Fontanillas ainsi qu'à Nicolas Juillet, à Sylvain Ursenbacher, à Françoise Lucas, à Guillaume Evanno, à Nathaniel Valière, à Gwenaël Jacob, à Lorie Lawson, à Rob Hammond, à Mélanie Glaettli, Michel Chapuisat, Philippe Christe, Max Reuter et à Christian Parisod pour le travail réalisé en commun ou en leur collaboration. Leur influence et leur travail est présent tout au long de ma thèse, qu'ils ont chacun influencé ou amélioré. Merci aussi pour l'amitié et la confiance qu'ils m'ont témoigné.

Mille mercis aussi à mes collègues de bureau Brigitte « Bribri » Reutter, Barbara « barbie » Holzer et Hervé Rosset pour ces années passées en votre compagnie agréable et stimulante. Merci pour votre gentillesse, pour les discussions, les idées et pour votre collaboration scientifique et lors des enseignements.

Section - 1 INTRODUCTION

I CAPERCAILLIE BIOLOGY AND CONSERVATION

I . 1 GROUSE PHYLOGENY

I. 1. 1 The place of capercaillie in the order Galliforms

The genus *Tetrao* is a member of the Tetraoninae subfamily, usually placed in the family of Phasianidae, order Galliforms (Dimcheff *et al.*, 2002; Drovetski, 2002), although some authors consider this group as a family (Tetraonidae), member of Galliforms alongside the Phasianidaes (Snow & Perrins, 1998). The Tetraoninae subfamily comprises 18 species of grouse and ptarmigans (Table 1), distinguishable from other galliforms by their morphological adaptation to harsh climate conditions (long intestine, feathered nostril and tarsi) and their characteristic traits associated with courtship (for example: red comb above the eyes, fanning tail) (Dimcheff *et al.*, 2002).

Their distribution is circumpolar, holarctic, and grouse are found in tundra, prairies and low-lands, mid- and high mountains (Snow & Perrins, 1998).

Table 1: The 18 species of Tetraoninae. (Source: the Grouse Specialists Group)

Genus Falcipennis : Spruce Grouse
Siberian Spruce grouse *Falcipennis falcipennis* (Hartlaub, 1855)
American Spruce grouse *Falcipennis canadensis* (Linnaeus, 1758)
Genus Dendragapus : Blue Grouse
Blue Grouse *Dendragapus obscurus* (Say, 1823)
Genus Lagopus : ptarmigans
Willow Grouse *Lagopus lagopus* (Linnaeus, 1758)
Red Grouse *Lagopus scoticus* (Latham, 1787)
Rock Ptarmigan *Lagopus mutus* (Montin, 1776)
White-tailed Ptarmigan *Lagopus leucurus* (Richardson, 1831)
Genus Bonasa: Hazel grouse
Hazel Grouse *Bonasa bonasia* (Linnaeus, 1758)
Severtzov's Grouse or Chinese Gouse *Bonasa sewerzowi* (Przhevalski, 1876)
Ruffed Grouse *Bonasa umbellus* (Linnaeus, 1766)
Genus Centrocercus: Sage Grouse
Greater Sage Grouse *Centrocercus urophasianus* (Bonaparte, 1827)
Gunnison sage Grouse *Centrocercus minimus* (Young, Braun, Oyler-McCance, Hupp & Quinn, 2000)
Genus Tympanuchus: Prairie-Phickens
Sharp-tailed Grouse *Tympanuchus phasianellus* (Linnaeus, 1758)
Greater Prairie-chicken *Tympanuchus cupido* (Linnaeus, 1758)
Lesser Prairie-chicken *Tympanuchus pallidicinctus* (Ridgway, 1873)
Genus Tetrao: Capercaillies and Black grouse
Caucasian Grouse *Tetrao mlokosiewiczi* (Taczanowski, 1875)
Black-billed Capercaillie *Tetrao parvirostris* (Middendorff, 1851)
Black Grouse *Tetrao tetrix* (Linnaeus, 1758)
Western Capercaillie *Tetrao urogallus* **(Linnaeus, 1758)**

I. 2 BIOLOGY OF CAPERCAILLIE

I. 2. 1 Morphology

The capercaillie (*Tetrao urogallus,* Linnae, 1758) is the largest grouse among the Tetraoninae. The male can weigh up to 5 kg and the female 2.5 kg. The male's distinctive traits are the dark slate-grey plumage with brown coloured back and wings, a metallic green shield on the breast, white markings on the wing and tail and a red comb over the eyes. The female plumage is a black, grey and white camouflage, with brown feathers on the belly.

I. 2. 2 Habitat and diet

The Western Europe mountain habitat for this species is dominated by old-growth mixed mid-high forests with open areas such as bogs, roads, wind-blows and fires. Habitats are characterised by coniferous trees, open structure with moderate canopy cover, and rich ground vegetation dominated by bilberry (*Vaccinium myrtillus*) and other ericaceous shrubs (Sachot, 2002).
In winter, they are arboreal and follow an energetically stringent diet of coniferous needles. Their wintering survival is facilitated by their intestine that is adapted to the time consuming and poorly-nutritional digestion of pine needles by caecal extensions of their intestine (Andreev & Linden, 1994). In spring and summer they shift to live on the ground and feed on leaves, buds, flowers, seeds, fruits and shoots (Sachot, 2002).

I. 2. 3 Reproduction

The eggs, from five to sixteen, are ca. 6cm long, yellowing white with a few reddish-brown streaks. They are laid once a year in a nest hidden on the ground, were they are sensitive to predation from foxes (Summers *et al.*, 2004), martens and boars (B. Reymond, *pers. comm.*). The synchronous hatching occurs in May-June, after 24 to 29 days of incubation. The young's are mobile and follow the mother. They feed on invertebrates for a period of 3 to 4 weeks. During this period, weather conditions acting on lepidoteran larvaes abundance is crucial for survival (Picozzi *et al.*, 1999). Renesting potentially occurs late July when the first brood is unsuccessful.

Male capercaillies have a complex courtship display that they use to attract females to mate with them in the spring. This display usually takes place at a traditional site known as a lek, which originates from the Scandinavian word 'to play'. The display song involves tapping and gurgling which accelerates to a drum roll followed by a noise that resembles a cork being pulled out of a bottle. The final song phase involves alternating gurgling and wheezing. During the song the male fans the tail held vertically and leaps rapidly with noisily flapping wings.
Leks are suitable grounds dispatched across the whole range of distribution of the population on which males, that are regularly non-territorial, gather up and become very territorial and aggressive towards their likes. The gathering comes with a ritualized courtship, display of exuberant secondary sexual ornaments, singing and territorial defence (Pagel, 2002). Females are attracted by the leks and mate preferentially with dominant males, causing a potential skew in males reproductive success (Fiske *et al.*, 1998).
Across the animal kingdom, many species have adopted this type of mating system, such as, in avian, most of the grouse (prairie-chickens, sage grouse, black grouse and capercaillie), peacocks, manakins, some shorebirds, hummingbirds. In mammals, this behaviour has also been documented in bats (*Hypsignathus* ssp., *Miniopterus* ssp.), ungulates (noble deer *Cervus elaphus*, topi antelopes *Damaliscus lunatus*), sirenidaes (dugong), pinnipeds (walrus), then also fish like cichlids, amphibians, and even invertebrates (dipterans, lepidopterans) and more (Widemo & Owens, 1995).
In 1996, a model to study maximization of lek size relative to individuals' fitness highlighted the importance of inclusive fitness in black grouse *T. tetrix* as a factor to be considered in the evolution and maintenance of the lek mating syste (Kokko & Lindström, 1996). Thereafter, since 1999, reviews, molecular and observational studies described within population structure in grouse at the lek level. Differentiation between Black grouse leks was detected using microsatellites markers (Hoglund *et al.*, 1999). Significant differentiation between leks was also found in prairie chicken (Bouzat & Johnson, 2004; Van den Bussche *et al.*, 2003) as well as correlations between genetic similarity and geographic distances in peacocks (Petrie *et al.*, 1999; Sherman, 1999).
These observational results tend to suggest that potential kin selection is occurring on leks (Hoglund *et al.*, 1999). The latter hypothesis implies

that males can associate by kin discrimination and social cooperation, assuming inclusive fitness increase acting as a positive selection pressure at the group level. The lek behaviour would increase the population's fitness favouring mate choice by females. Long-life males would access dominant positions and get direct fitness benefits. Subordinates males, or helpers, would get indirect fitness benefit according to Hamilton's rule (Hamilton, 1964). This theory implies that lekking animals have a way to identify the degree of relatedness they share with any individual they meet. Assuming ability to discriminate between kin and non-kin, we can imagine that subordinate males would stay together with their dominant male to help him obtain females preference. Evidence of this discrimination ability (possibly by "phenotype matching") has been published (Petrie *et al.*, 1999) based on geographical proximity of kin-related individuals.
However, in species showing biased sex ratios, individuals of the most phylopatric sex tend to cluster in family groups (Stacey & Lingon, 1991). Furthermore, the extent of capercaillie dispersion remains limited at shorts distances, particularly for males (Storch, 1995). Therefore, the geographical clustering of genetically similar individuals could also be caused by geographic structure of population. In the latter case, the issue of kin-clusters on leks does not involve kin selection or kin recognition.

In this thesis, I will test whether lekking capercaillie males cluster preferentially among related individuals by measuring the relatedness of lekking individuals within and among leks from samples collected on leks. If this is so, I will identify whether it is due only to population geographic structure or true kin association by comparing relatedness values with the global and local genetic structure of the population

I. 4 DISTRIBUTION OF CAPERCAILLIE IN EUROPE

Capercaillie is widely distributed across Eurasia. Its presence is continuous from Scandinavia to eastern Siberia in the boreal forest (approx. 125° east). In the opposite, the westernmost populations of Europe are fragmented, primarily due to patchy distribution of mountain coniferous forests (Klaus *et al.*, 1989) (Fig. 3). Most of these populations have shown a strong decline in the last decades.

Fig. 3: Capercaillie distribution in Eurasia. Source: (Klaus *et al.*, 1989).

In Scotland, capercaillie has disappeared in 1785 and was unsuccessfully reintroduced several times, and still remains in constant decline (Moss *et al.*, 2000). Today, the remaining population is estimated around 1000 to 2000 individuals (source: RSCB Scotland). In France, capercaillie is distributed in the Jura, the Pyrenees, the Vosges and the Cevennes. These populations have shown a strong decline for the last decades, and only a few individuals remain in the Vosges as well as in the Cevennes, where only about 50 to 100 are now estimated in spite of reintroduction attempts of more than 500 individuals since 1975.

In the Swiss, Italian, Austrian and Bavarian Alps, the populations are still large enough to insure sustainability, but decline is already documented in Germany and Switzerland (Storch, 2000). The decline is also observed in the large Carpathian populations (Saniga, 2003).

In Scandinavia, large populations present from south-west Norway to Lappland are continuous and connected with the larger Siberian populations. Although assuming demographic cycles (Lindstrom *et al.*, 1995; Selås, 2000), these populations are probably very large. Finally, the largest population is covers most part western Russia, from the Baltic sea to the limits of Asia.

I. 5 HISTORICAL DISTRIBUTION IN THE JURA AND IN THE PYRENEES

I. 5. 1 Distribution and habitat in the Jura

In the Jura capercaillie was abundant in most forests before the beginning of the 20th century, and populations were probably connected to the ones in the Alps (Fig. 4). At this time, extensive exploitation for timber had fragmented the Swiss forests. Since then, woodlands have replenished so much that the cover has become to dense to satisfy capercaillie's habitat requirements. Today, protection measures are taken in coordination with forestry departments, and guidelines followed by rangers and forest managers offer the possibility for a long term sustainability of both wildlife and timber industry (Mollet & Marti, 2001). However, the decline is still observed, and presence on leks in 2001 was down to 250 individuals in this area (Sachot *et al.*, 2002). Moreover, the recent increase of human frequentation of the Jura for leisure and ski resorts drives disturbance further into the protected areas, and in all seasons. Managers and grouse specialists estimate that the present habitat structure is not sufficient to maintain these populations in the long term (Sachot, 2002).

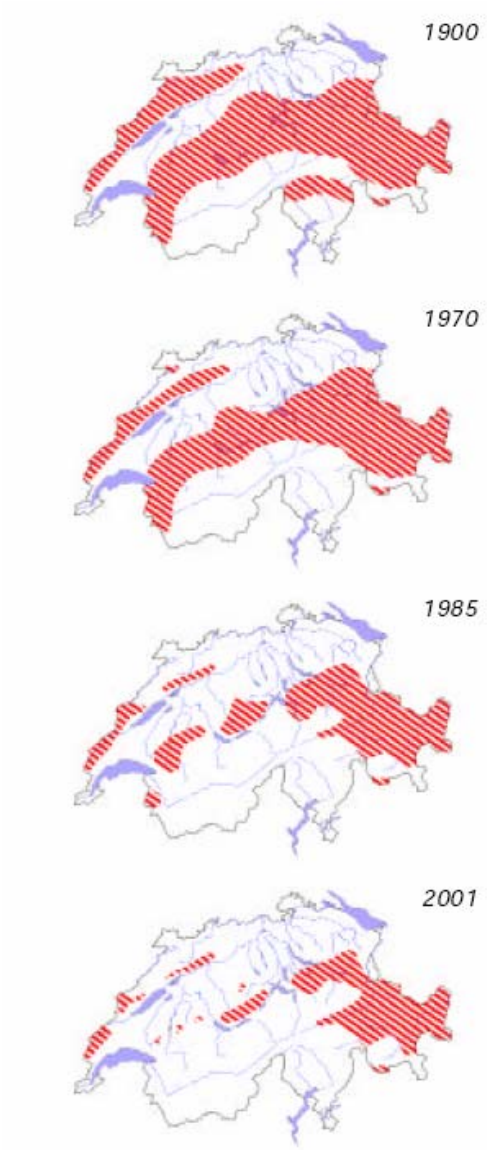


Fig. 4: Distribution of capercaillie in Switzerland from 1900 to 2001. Between 1970 and 1985, the populations in the Jura and in the pre-Alps became fragmented. Since then, isolation has been increasing. Source: (Mollet, 2002)

A recent study of population dynamics and habitat selection has found that the risk of extinction of capercaillie populations in the Jura is moderate if adequate management is insured (Sachot, 2002). However, this study has also pointed out the susceptibility of these populations to demographic stochasticity and predicts accelerated decline at low-density. Thus, the presence of capercaillie in the Jura seems highly uncertain if the decline persists.

The analyses of the population genetic status conducted in this thesis will help understanding population dynamics, but also verify whether the geographic isolation and recent decline are likely to aggravate the extinction risk.

I. 5. 2 The sub-species described in the Pyrenees

In the Pyrenees, the subspecies *Tetrao urogallus aquitanicus* (Ingram, 1915) inhabits subalpine mixed forest of *Abies pectinata* (silver fir) and *Fagus sylvatica* (beech) between 1300 and 2100m (Menoni, 1991). it has a larger ecological niche than the species standards, utilizing the niche usually attributed to *Tetrao tetrix*: higher mountains and upper forest limits (Menoni, 1991). Its distribution spans along the border between France and Spain, including Andorra (see map section 2- IV). Densities are heterogeneous and the whole population census size is thought to count ca. 6000 individuals. This form is smaller than *T. urogallus* standards, as the males weight a maximum of 4 kg and females 2 kg (Ingram, 1915).

Between 1850 and the 1920's, the population was bottlenecked because of an intensive deforestation that affected the whole mountain massif {Ingram, 1915 #559}. By the 1960's, the population had expanded to a maximum of ca. 9000 individuals and is now declining (E. Menoni, *pers.comm.*). The present distribution of genetic diversity in this population has likely kept traces and patterns of the recent demographic history that I will try to identify in order to estimate whether the observed decline is a threat to the long-term population sustainability.

II WHY PROTECT CAPERCAILLIE?

II . 1 THE CAUSES OF DECLINE

The causes of decline observed in the fragmented populations of western Europe are primarily due to habitat loss (Storch, 1991). Fragmented forest and timber industry has reduced drastically suitable habitats during the last century. The use of forests in terms of tourism industry not only degrades the habitat, but also creates disturbance that can be fatal in winter, and that can hinder reproduction in spring (e.g.: Brenot *et al.*, 1996). Other acknowledged causes of decline are predation on nest and chicks by red foxes and other mammals (Menoni *et al.*, 1991), and poaching. Finally, as population sizes decrease dramatically, genetic factors will prevent maintenance of diversity and adaptability, and increase inbreeding. Populations have already disappeared in the Cevennes and in Scotland, it is thus likely that extinction will also occur in other fragmented populations if further decline is not prevented.

II . 2 CAPERCAILLIE, A FLAGSHIP SPECIES

Capercaillie is a flagship species recognized as a symbol for wildlife fauna in Western Europe forests. Capercaillie is listed in the European Council Directive 79/409/EEC on the conservation of wild birds (the 'Birds Directive'), and is protected locally by laws on nature conservation, reserves and parks, and also in the Natura 2000 framework. However, because of the large populations in the original habitat of northern Europe and Siberia, the species is not threatened. It is therefore not on the "2003 IUCN Red List of Threatened Species" nor on the "CITES" Washington convention. In the Bern Convention (1979), only the subspecies from the Cantabric mountains in north-west Spain, *Tetrao urogallus cantabricus*, is mentioned.

II . 3 CAPERCAILLIE, AN UMBRELLA SPECIES

Capercaillie is considered as an "Umbrella species" (Pakkala *et al.*, 2003; Suter *et al.*, 2002)- species whose conservation is expected to confer protection to a large number of species naturally occurring in the same habitat. This concept has been proposed as a tool for determining the minimum size for conservation areas, for selecting sites to be included in reserve networks and for setting minimum standards for the composition, structure, and processes of ecosystems (Roberge & Angelstam, 2004).
The capercaillie requires strict ecological conditions and is very sensitive to habitat modifications and disturbance. As such, protecting the capercaillie is equivalent to protecting the whole habitat of European forests, and therefore, the biodiversity found in these forests.

II . 4 PROTECTION ACTIONS

The causes of decline are therefore well documented, and policies and actions can be endeavoured to control or stop the decline. Protecting capercaillie necessitates protecting the whole ecosystem it depends upon: the remaining forests of Europe.
Protection and reintroductions have been the lead research subject for capercaillie publications in the latest years, although until the mid-1990's habitat and ecology, or behaviour were mostly the focus of research in the Pyrenees and in the Black Forest (Menoni, 1991; Storch, 1993). A large part of direct conservation actions, such as reintroductions and management, as well as their effects, is also reported in recent literature from Scotland (Marshall & Edwards-Jones, 1998; Moss, 2000; Moss *et al.*, 2000). As game birds, grouse research has been funded by gaming departments and associations in Scotland. It is also a symbol of wildlife fauna in the Pyrenees, the Jura and the Alps. Therefore habitat choice and ecology (Helle *et al.*, 1999; Menoni, 1991; Picozzi *et al.*, 1999; Sachot, 2002), physiological requirements (Andreev & Linden, 1994) or effects of predation (Baines *et al.*, 2004) were studied to identify factors influencing life history traits and survival in threatened populations.

III CONSERVATION GENETICS

III . 1 AIMS AND SCOPE OF CONSERVATION GENETICS

The aim of conservation genetics is to describe the genetic variation found in plants and animals and to minimize the risks of population extinction due to genetic factors. Endangered organisms occur in small populations and therefore may have experienced a reduction in the level of genetic variation. Genetic variation can be lost by several different events including founder effects, inbreeding, genetic drift and bottlenecks. Populations with lower genetic diversity may be less capable of dealing with, or adapting to, changes in its environment, such as the introduction of a new disease, climate change, and arrival of invasive species or anthropogenic modification of their habitat.
The conservation geneticist uses indirect molecular methods to estimate extinction risks related to these genetic factors, and to advise population management policies. The actions and scopes of conservation genetics encompass:

- The theoretical study of the impact of genetic factors in small population in relation with demography and geography.
- Resolving taxonomic uncertainties and defining managements units within species
- The maintenance of residual populations, either in captive breeding programs or in reintroduction plans of extinct populations by minimizing deleterious genetic impact.

Other applications of this field have been found, for instance, in the biology of invasive species and in forensic and poaching control.

III . 2 MOLECULAR STUDIES OF GROUSE POPULATIONS

III. 2. 1 Molecular markers

With the increased awareness of the effects of genetic factors in population extinction (Frankham *et al.*, 2002), conservation plans and

fauna managers found interest in population genetics studies with conservation applications.
Genetic studies using allozyme techniques on grouse are scarce (Linden & Teeri, 1985; Schreiber *et al.*, 1998), as fresh tissue are needed, and shooting or sampling in the field is not adapted to threatened populations. In 1997, the University of Edimburgh developed the first microsatellite markers for ptarmigan (Piertney & Dallas, 1997), and later other markers were made available for black grouse (Caizergues *et al.*, 2001) and capercaillie (Segelbacher *et al.*, 2000).

III. 2. 2 Behavioural advances

These molecular advances have led population genetic research to investigate the distribution of genetic diversity in grouse populations and have provided valuable new understandings of behavioural observations, particularly in the lekking behaviour. For example, estimation of relatedness among breeding females around leks confirmed field observation but also highlighted the possibility of unseen components in reproductive strategies in American sage-grouse (Semple *et al.*, 2001). Theories about male aggregation by female choice (Kokko, 1997) revealed insufficient to provide full explanatory models. Furthermore, new genetic findings suggest that kin selection concurs to the maintenance of the lekking behaviour in a lekking Galliforms species (Petrie *et al.*, 1999; Sherman, 1999). High relatedness among males on leks was then confirmed by microsatellite data in black grouse (Hoglund *et al.*, 1999) and hierarchical genetic structure was described in other grouse species considering the lek as lower level of structure (Bouzat & Johnson, 2004; Gibson *et al.*, 2002; Hoglund *et al.*, 1999).

III. 2. 3 Population structure studies

Molecular methods have also allowed inferences of population dynamics parameters such as gene flow and population structure in black grouse (Caizergues & Ellison, 2002) and ptarmigans (Piertney *et al.*, 1998), or population differentiation in capercaillie across Europe revealing patterns of isolation by distance (Segelbacher *et al.*, 2003). These methods have also provided information on the loss of genetic diversity in declining populations for example in Chinese grouse (Larsson *et al.*, 2003), in black grouse (Caizergues *et al.*, 2003) and particularly in capercaillie populations in the Alps (Segelbacher & Storch, 2002).

Genetic diversity has also been well documented among the populations of capercaillie in Europe. Furthermore, at a local scale, studies based on samples collected on leks indicate significant structuring between lek places. However, the extent and the distribution of genetic diversity at the within population level remains unknown. This is why a large part of this thesis is dedicated to the study of the genetic structure within continuous populations.
In the Jura and the Pyrenees, where I will focus my study, capercaillie is present in Mountains forests habitat. This habitat is naturally fragmented by landscape features, such as high mountains and deep valleys, or urbanised areas. Nevertheless, the extent of genetic structuring among these patches of suitable habitat occupied by capercaillie has not been studied.
In this thesis, I will describe the distribution of genetic diversity at the population level. These results will provide valuable information on the dispersal ability of capercaillie as well as information about the genetic effects of their recent demographic history.

III. 3 METHODS

III. 3. 1 Microsatellite markers

The most commonly used molecular marker of genetic diversity since the 1990's are the microsatellites, and we decided on this choice because of the applicability of microsatellite to intra-population genetic analysis (Ellegren, 1992) and because laboratory techniques allow the use of non-invasive sampling strategies (Taberlet *et al.*, 1996).
Microsatellites (also referred to as SSR: simple short repeats, or STR: short tandem repeats, or again VNTR : variable number tandem repeats), are 2 to 5-mer nucleic repeats duplicated 5 to 50 times. The pattern can be any repeated arrangement of any of the 4 nucleic bases. These repeats are commonly found in all organisms, although they are less common in plants and birds than in mammals.
The mutation rate in STR is estimated at 10^{-2} to 10^{-5} mutation per generation. Three biochemical processes are suspected to be source of this very high mutation rate: slippage of DNA polymerase made instable by the repeated pattern of DNA, stochastic insertion or deletion of large parts of the STR due to non-specific annealing during replication and

unbalanced crossing-overs. This high mutation rate induces a high polymorphism and makes it a powerful tool for the genetic study of neutral genetic variation within the species level.
However, recent studies indicate that microsatellites used as molecular markers are inadequate indicators of ecologically important quantitative genetic variation (Reed *et a.l* 2001). Unfortunately, the vast majority of population genetic information available for endangered species is for neutral molecular markers, not for quantitative characters upon which selection can act (Frankham 1995a). Neutral markers are not direct indicators of a population's short evolutionary potential (Ashley *et al.*, 2003) but they provide useful pieces of information to compare diversity in different populations and make inferences on genetic fragmentations and geneflow, with the limitations inherent to the method (Balloux *et al.*, 2000).

III. 3. 2 Non-invasive sampling

The use of molecular techniques to answer biological, ecological or behavioural questions is however limited by the access to biological samples and DNA sources. Regular samples for PCR analysis include blood or tissue samples. These tissues can be obtained either by destructive sampling, which requires the destruction of the individual captured for DNA analysis, or by non-destructive methods such as biopsies or taking blood samples (Frankham *et al.*, 2002). Unfortunately, sampling in threatened species is difficult to perform because of their and because some species are elusive. Furthermore, capturing individuals can be dangerous for the populations under investigation as these manipulations can stress, hurt or disturb individuals. Additionally, capture and transport permits, often required, are difficult to obtain in protected populations.
The use of alternative DNA sources and of non-invasive sampling, such as hair, feces, urine, feathers, buccal cells on food left-overs or vomit, owl pellets (Kohn & Wayne, 1997; Taberlet & Fumagalli, 1996; Taberlet & Luikart, 1999) allows conducting studies on protected populations without having to catch or manipulate individuals.
Capercaillie DNA from feathers have been used for a recent conservation study at the scale of the European continent (Segelbacher, 2002), but feathers are too rare to provide enough samples for within population investigations. Faeces, in the opposite, are found frequently and testify of the presence of the animal. They are a useful source of information on the biology of the species to investigate diet, hormone levels, to define sex, or simply as an evidence of presence. Thus, Office National de la Chasse

(ONCFS) and Fauna professionals as well as volunteers were already experienced in the idea of collecting stool. They made a great number of samples available for our genetic study. To insure effective storage and protect our samples for degradation, several methods were proposed in the literature, such as keeping the samples in alcohol, or in freezers (Frantzen *et al.*, 1998). We chose to keep the faecal sample dry in 20ml vials, filled to the quarter with silica gel beads to insure that the vials will remain dry. This method was the most appropriate, because the composition of the stool is mostly pine needle residues and most of the samples found on the field were already dried.

However, PCR amplification of DNA from stool samples presents two major limitations: first, the presence of PCR inhibitors in the DNA extract originating from the faecal matter. Second, the low reliability of the genotyping, that is due to the low quality and quantity of DNA in the samples, causing large error rates (Taberlet et al., 1999). In the first chapter of the section 2, I detail the protocols that we used to prevent the effect of inhibitors. I also detail the multitube approach (Taberlet *et al.*, 1996), a protocol that I adopted to compensates for the low reliability of genotyping.

III. 3. 3 Molecular sexing

Identification of the sex of samples collected on the field is a fundamental information used to estimate sex ratios in populations, sex biased dispersal and sex identification of a tracked animal. Several PCR-based protocols are available to differentially amplify genes located on heterochromosomes.

In mammals, as in most taxa, the heterogametic sex is the male, being XY. Two genes located on the X and Y chromosomes have already been used for molecular sex identification. These genes are ZFY (Zfy: male or Zfx: female) and SRY. SRY (Sex-determining Region) is involved in sex determination during the development, and is only present on the Y chromosome. ZFY (Zing Finger) is located on the homologous region, hence on both X and Y chromosome.

SRY being Y specific will only amplify fragments in male samples. ZFY will amplify fragments in both sex and serves as a positive control to check the amplification reaction has worked well.

Going to the avian example, the heterochromosomes are W and Z. The female is heterogametic (WZ), the male homogametic (ZZ) (Fig. 5) as opposed to mammals where the male is heterogametic.

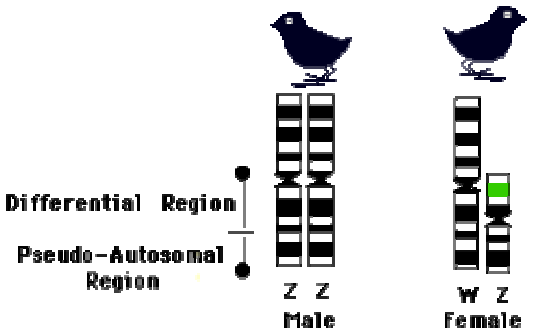


Fig. 5: Sketch of variation between avian sex chromosomes.

Primers for PCR amplification of avian sex specific genes are available (Fridolfsson & Ellegren, 1999; Griffiths *et al.*, 1998). However, limitations to general PCR protocols on non-invasive samples also apply to molecular sexing (see above) and in the course of this study, new sets of primers will be designed for use on faecal DNA extracts.

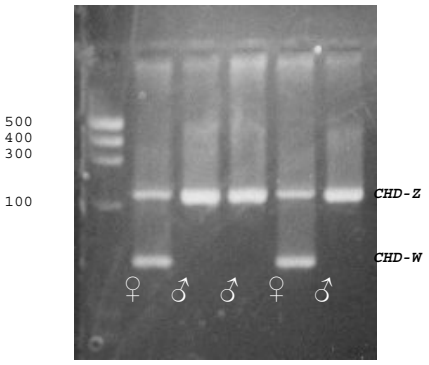


Fig. 6: Agarose gel electrophoresis is used to visualize amplified genes from capercaillie stool samples: The male bears the two copies of the same gene, because it has two Z chromosomes, the PCR will amplify the two genes but the PCR products will have the same size, so only one band will show. The female has one copy of the gene on W and one copy on Z. These two copies have different sizes, and the amplification product will show two bands on the agarose gel. On the left: size of the scale in base pairs. On the left, amplicon size standards. On the right: label of the amplified band. Below: sex of the sample.

Andreev AV and Linden H (1994) Winter energetics of the capercaillie - a methodological approach. *Ornis Fennica* **71**, 33-42.

Ashley MV, Willson MF, Pergams ORW*, et al.* (2003) Evolutionarily enlightened management. *Biology Conservation* **111**, 115-123.

Baines D, Moss R and Dugan D (2004) Capercaillie breeding success in relation to forest habitat and predator abundance. *Journal of Applied Ecology* **41**, 59-71.

Balloux F, Brünner H, Lugon-Moulin N, Hausser J and Goudet J (2000) Microsatellites can be misleading: An empirical and simulation study. *Evolution* **54**, 1414-1422.

Bouzat JL and Johnson K (2004) Genetic structure among closely spaced leks on a peripheral population of lesser prairie-chickens. *Molecular Ecology* **13**, 499-505.

Brenot JF, Catusse M and Menoni E (1996) Effet de la station de ski de fond du plateau de Beille (Ariège) sur une important population de grand tétras *Tetrao urogallus*. *Alauda* **64**, 249-260.

Caizergues A, Dubois S, Loiseau A, Mondor G and Rasplus JY (2001) Isolation and characterization of microsatellite loci in black grouse (Tetrao tetrix). *Molecular Ecology Notes* **1**, 36-38.

Caizergues A and Ellison LN (2002) Natal dispersal and its consequences in Black Grouse Tetrao tetrix. *Ibis* **144**, 478-487.

Caizergues A, Ratti O, Helle P*, et al.* (2003) Population genetic structure of male black grouse (Tetrao tetrix L.) in fragmented vs. continuous landscapes. *Molecular Ecology* **12**, 2297-2305.

Dimcheff DD, Drovetski SV and Mindell DP (2002) Molecular evolution and systematics of Tetraoninae and other Galliformes using mitochondrial 12S and ND2 genes. *Molecular Phylogenetics and Evolution* **24**, 203-215.

Drovetski SV (2002) Molecular phylogeny of grouse: individual and combined performance of W-linked, autosomal, and mitochondrial loci. *Systematic Biology* **51**, 930-945.

Ellegren H (1992) Polymerase chain reaction (PCR) analysis of avian microsatellites- a new approach to genetic studies in birds. *Auk* **109**, 886-895.

Fiske P, Rintamaki P and Karvonen E (1998) Mating success in lekking males: a meta-analysis. *Behavioural Ecology* **9**, 328-338.

Frankham R, Ballou JD and Briscoe DA (2002) *An introduction to conservation genetics* University Press, Cambridge.

Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK and Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428.

Fridolfsson AK and Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**, 116-121.

Gibson RM, Aspbury AS and McDaniel LL (2002) Active formation of mixed-species grouse leks: a role for predation in lek evolution ? *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**, 2503-2507.

Griffiths R, Double MC, Orr K and Dawson RJG (1998) A DNA test to sex most birds. *Molecular Ecology* **7**, 1071-1075.

Hamilton WD (1964) Genetical evolution of social behaviour. *Journal of Theoretical Biology* **7**, 17-52.

Helle P, Kurki S and Lindén H (1999) Change in the sex ratio of the Finnish capercaillie *Tetrao urogallus* population. *Wildlife Biology* **5**, 25-31.

Hoglund J, Alatalo RV, Lundberg A, Rintamaki PT and Lindell J (1999) Microsatellite markers reveal the potential for kin selection on black grouse leks. *Proceedings of the Royal Society of London Series B-Biological Sciences* **266**, 813-816.

Ingram C (1915) A few notes on *tetrao urogallus* and its allies. *Ibis*, 128-133.

Klaus S, Andreev AV, Bergmann H-H*, et al.* (1989) Die Auerhühner. Die Neue Brehm-Bücherei. Band 86. In: *Westarp Wissenschaften* (ed. Magdeburg), Germany.

Kohn MH and Wayne RK (1997) Facts from faeces revisited. *Trends in Ecology and Evolution* **12**, 223-227.

Kokko H (1997) The lekking game: Can female choice explain aggregated male displays ? *Journal of Theoretical Biology* **187**, 57-64.

Kokko H and Lindström J (1996) Kin selection and the evolution of leks: Whose success do young male maximize ? *Proceedings of the Royal Society of London Series B-Biological Sciences* **263**, 919-923.

Larsson JK, Sun YH, Fang Y, Segelbacher G and Hoglund J (2003) Microsatellite variation in a Chinese grouse *Bonasa sewerzowi* population: signs of genetic impoverishment? *Wildlife Biology* **9**, 261-266.

Linden H and Teeri T (1985) Genetic differentiation in the capercaillie, *Tetrao urogallus*, populations. *Hereditas* **102**, 297-299.

Lindstrom J, Ranta E, Kaitala V and Linden H (1995) The clockwork of Finnish tetraonid population dynamics. *Oikos* **74**, 185-194.

Marshall K and Edwards-Jones G (1998) Reintroducing capercaillie (Tetrao urogallus) into southern Scotland: identification of minimum viable populations at potential release sites. *Biodiversity and Conservation* **7**, 275-296.

Menoni E (1991) *PhD thesis: Ecologie et dynamique des populations du grand tétras dans les Pyrénées, avec des reférences spatiales à la biologie de la reproduction chez les poules: quelques applications à sa conservation*, Université Paul Sabatier.

Menoni E, Catusse M and Novoa C (1991) Mortalité par prédation de grand tétras (*Tetrao urogallus aquitanicus*) dans les Pyrénées. Résultats d'une enquête. *Game and Wildlife Science* **8**.

Mollet P (2002) Mesures d'urgence en faveur du Grand Tétras. In: *Fiche Info Grand Tétras*. Sation ornithologique suisse de Sempach.

Mollet P and Marti C (2001) Grand Tétras et gestion de la forêt (ed. OFEFP), Berne.

Moss R (2000) Second extinction of capercaillie (*Tetrao urogallus*) in Scotland ? *Conservation Biology* **10**, 255-257.

Moss R, Picozzi N, Summers RW and Baines D (2000) Capercaillie *Tetrao urogallus* in Scotland – demography of a declining population. *The Ibis* **142**, 259.

Pagel M (2002) *Encyclopedia of Evolution* Oxford University Press, New York.

Pakkala T, Pellikka Jo and Lindén H (2003) Capercaillie Tetrao urogallus - a good candidate for an umbrella species in taiga forest. *Wildlife Biology* **3**, 309-316.

Petrie M, Krupa A and Burke T (1999) Peacocks lek with relatives even in the absence of social and evironmental cues. *Nature* **401**, 155-157.

Picozzi N, Moss R and Kortland K (1999) Diet and survival of capercaillie *Tetrao urogallus* chicks in Scotland. *Wildlife Biology* **5**, 11-23.

Piertney SB and Dallas JF (1997) Isolation and characterisation of hypervariable microsatellites in the red grouse *Lagopus lagopus scoticus*. *Molecular Ecology* **6**, 93-95.

Piertney SB, McColl AD, Bacon PJ and Dallas JF (1998) Local genetic structure in the red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite markers. *Molecular Ecology* **7**, 1645-1654.

Roberge Jo and Angelstam P (2004) Usefulness of the umbrella species concept as a conservation tool. *Conservation Biology* **18**, 76-85.

Sachot S (2002) *PhD thesis : Viability and management of an endangered capercaillie (Tetrao urogallus) metapopulation*, Université de Lausanne.

Sachot S, Leclercq B and Montadert M (2002) Population trends of Capercaillie (*Tetrao urogallus*) in the Jura mountains between 1991 and 1999. *Game and Wildlife Science* **19**, 41–54.

Saniga M (2003) Causes of the population decline in capercaillie (*Tetrao urogallus*) in the West Carpathians. *Biologia* **58**, 265-273.

Schreiber A, Weitzel T and Strauss E (1998) Allozyme variability in Black Grouse (*Tetrao tetrix*), a tetraonid with lek behaviour. *Journal Fur Ornithologie* **139**, 55-66.

Segelbacher G (2002) *PhD thesis*, University of Tübingen, Germany.

Segelbacher G, Hoglund J and Storch I (2003) From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology* **12**, 1773-1780.

Segelbacher G, Paxton RJ, Steinbruck G, Trontelj P and Storch I (2000) Characterization of microsatellites in capercaillie Tetrao urogallus (AVES). *Molecular Ecology* **9**, 1934-1935.

Segelbacher G and Storch I (2002) Capercaillie in the Alps: genetic evidence of metapopulation structure and population decline. *Molecular Ecology* **11**, 1669-1677.

Selås V (2000) Population dynamics of capercaillie *Tetrao urogallus* in relation to bilberry *Vaccinium myrtillus* production in southern Norway. *Wildlife Biology* **6**, 1-11.

Semple K, Wayne RK and Gibson RM (2001) Microsatellite analysis of female mating behaviour in lek-breeding sage grouse. *Molecular Ecology* **10**, 2043-2048.

Sherman PW (1999) Birds of a feather lek together. *Nature* **410**, 119-120.

Snow DW and Perrins CM (1998) *The Birds of the Western Palaearctic* Oxford University Press, London.

Stacey PB and Lingon JD (1991) The benefits-of-phylopatry hypothesis for the evolution of cooperative breeding: variation in territory quality and group size effect. *American Naturalist* **137**, 831-846.

Storch I (1991) Habitat fragmentation, nest site selection, and nest predation risk in Capercaillie. *Ornis Scandanavia* **22**, 213–217.

Storch I (1993) *Habitat use and spacing of capercaillie in relation to forest fragmentation patterns*, University Ludwig-Maximilian.

Storch I (1995) Annual home ranges and spacing patterns of capercaillie in central Europe. *Journal of Wildlife Management* **59**, 392-400.

Storch I (2000) Grouse status survey and conservation action plan 2000-2004. In: *WPA-Birdlife-SSC Grouse Specialist Group*. IUCN, Switzerland.

Summers RW, Green RE, Proctor R*, et al.* (2004) An experimental study of the effects of predation on the breeding productivity of capercaillie and black grouse. *Journal of Applied Ecology* **41**, 513-525.

Suter W, Graf RF and Hess R (2002) Capercaillie (Tetrao urogallus) and avian biodiversity: Testing the umbrella-species concept. *Conservation Biology* **16**, 778-788.

Taberlet P and Fumagalli L (1996) Owl pellets as a source of DNA for genetic studies of small mammals. *Molecular Ecology* **5**, 301-305.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Taberlet P and Luikart G (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society* **68**, 41-55.

Taberlet P, Waits LP and Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution* **14**, 323-327.

Van den Bussche RA, Hoofer SR, Weidenfeld DA, Wolfe DH and Sherrod SK (2003) Genetic variation within and among frangemented populations of lesser prairie-chickens (*Tympanuchus pallidicinctus)*. *Molecular Ecology* **12**, 675-683.

Widemo F and Owens IPF (1995) Lek size, male mating skew and the evolution of lekking. *Nature* **373**, 148-151.

Section - 2 RESEARCH WORK

This section presents results of laboratory work and their statistical and biological interpretation.

I FEASIBILITY OF LARGE SCALE NON-INVASIVE GENETIC STUDIES APPLIED TO THREATENED GROUSE POPULATIONS

Sébastien Regnaut, Françoise S. Lucas
and Luca Fumagalli

I. 1 ABSTRACT

To evaluate the technical feasibility of a large scale population genetic study of an endangered capercaillie (*Tetrao urogallus*) population using only faeces as non-invasive sampling strategy, we used 11 microsatellite markers and a multitube approach. A protocol for sexing individuals using triplets of primers for short amplifications of *CHD* introns was also developed. We found 69% of amplification success on first attempt of extraction, and 99% after 4 extraction attempts. Allelic dropout was estimated at 21% and false allele rates at 3%. Furthermore, using an experimental DNA incubation design in wet conditions, we showed that free DNases constituted major degraders in grouse faecal material. Therefore, we suggest that climate conditions during sampling as well as storage condition can have a significant impact of overall sample quality. Our results suggested that a non-invasive approach based on DNA microsatellite analysis of faeces can be a powerful tool in population monitoring of capercaillie.

I. 2 INTRODUCTION

Small isolated populations or populations on the edges of the distribution range of a species are now known to potentially show inbreeding depression and loss of genetic diversity (Frankham *et al.*, 2002). Conservation genetics aim at identifying and monitoring these genetic processes at work in risk population. Such studies help focusing management policies on proper issues (Kohn & Wayne, 1997). Unfortunately, all populations are not susceptible to endure genetic sampling in the scale required for classic population genetic studies. Genetic sampling could be difficult to achieve for technical reasons or because it could threaten the survival of the population (Harrison *et al.*, 1991). The use of various non-invasive DNA sources such as hair, faeces, feathers or urine have made available DNA sources without capture, stress or injury caused to free-ranging individuals (Kohn & Wayne, 1997). In the last years, DNA extraction and amplification methods have been developed to sequence or genotype these non-invasive samples, often stemming from forensic and ancient-DNA analysis protocols (e.g. (Höss *et al.*, 1992). A wide range of mammalian species have been the subject of such analyses, in particular by using faeces and hair collected in the field as DNA source (e.g.: Hedmark *et al.*, 2004; Kohn & Wayne, 1997; Morin *et al.*, 1993; Taberlet *et al.*, 1997; Taberlet *et al.*, 1993). This approach is of great interest in the field of conservation biology, where populations are usually small and endangered, and animal capture could have dramatic effects on survival.

Among avian species, the capercaillie (*Tetrao urogallus* L.; *Aves; Galliforms; Tetraonidae*) populations in the Jura Mountains (500 km², France and Switzerland) are precisely in this precarious situation where a conservation genetic study will help assessing the genetic status of the populations. The population has been declining until the 1980's since when it fluctuates around 600 individuals (Leclercq, 2004; Sachot *et al.*, 2002). Capturing this elusive species is delicate without introducing sampling biases and without reducing population survival expectancies (Harrison *et al.*, 1991). As a consequence, traditional sampling is therefore inappropriate in the case of capercaillie in the Jura, and alternative sources of DNA had to be found in non-invasive genetic sampling. Feathers are known as a usable source of DNA in this species (Segelbacher, 2002), but they are scarce and cannot be found in sufficient

quantities to allow a fine scale genetic study. Faeces, in the opposite, are easy to find in all seasons on birds lekking sites as well as sleeping and feeding sites.

Two main impairment have been reported in the use of faecal DNA: First, excrements contain large concentrations of potential PCR inhibitors, such as complex polysaccharides linked to the vegetarian diet of capercaillie or produced by gram-negative faecal bacteria (Monteiro *et al.*, 1997). Secondly, nucleic acids left in wet and warm conditions within faeces may be degraded by physico-chemical, enzymatic or microbial factors (Matsui *et al.*, 2001). Limitations of excremental analysis by PCR applies even stronger to nuclear than to mitochondrial DNA (mtDNA), as the latter is present in many more copies than the former.

As a consequence, the DNA extracted from faecal samples often offers a poor quality and low concentration template for microsatellite screening resulting in two types of genotyping mistakes: i) allelic dropout (non-amplification of an allele, resulting in estimating a heterozygous genotype as homozygous). ii) PCR artefacts (productions of false alleles during the amplification that will estimate an homozygous genotype as heterozygous, or make ambiguous genotypes if PCR are repeated).

To date, the only published study using faecal bird material collected non-invasively concerns mtDNA sequence data (Broderick *et al.*, 2003; Idaghdour *et al.*, 2003), but no information is so far available for nuclear DNA.

In this study, we estimated the feasibility and we optimised the use of microsatellites on a large scale faeces sampling scheme, for conservation genetics of avian populations. Primarily, we improved DNA extraction and set up a new protocol for sexing the individuals. The reliability of microsatellite genotyping from capercaillie faeces was quantified using repeated time-release PCR and software simulation (Valière *et al.*, 2002) to assess the optimal number of repeats to obtain a satisfactory level of reliability. This approach was conducted following the multitube protocol proposed by Taberlet *et al* (1996). Secondly, we designed an experiment to identify the main cause of DNA degradation in stool. Using capercaillie faeces added to known quantity of fresh liver DNA, we compared the relative contribution of bacterial microorganisms, nucleases and hydrolytic activities to DNA degradation (Woegerbauer *et al.*, 2000).

I. 3 MATERIEL AND METHODS

I. 3. 1 Sample storage

Three hundred samples collected from the field were temporarily stored into individual labelled plastic bags. In the laboratory, the samples were left to dry on absorbing paper sheets for 3 days at room temperature before storage in plastic vials (50 or 100 ml.) with approx. 4cl. of silica gel beads to prevent re-hydratation (Frantzen *et al.*, 1998).

I. 3. 2 DNA extraction

All extractions took place in an isolated room only dedicated to non-invasive sample extractions and pre-PCR mixes.
We used QIAamp DNA stool Kit (Qiagen) modified to suit the requirements of our study. As we decided to apply the multitube approach (Taberlet *et al.*, 1996), several repeated amplification reactions were performed and the volume of DNA elution obtained from each extraction had to be maximized. Therefore, twice the amount of sample indicated by the manufacturer was processed in a single Qiagen column. Only 11 samples plus one internal negative control were processed at one time. Approximately 200mg of dry stool were placed in two regular 2ml transparent Eppendorf tube. 1.4 to 1.6ml of ASL buffer was added to each of the two tubes. The samples were homogenised with a flat metal rod and vortexed. When the two duplicates of each of the 12 samples were homogenised, the supernatant was transferred into a single new 2ml tube containing the adsorbing matrix (Inhibitex™) provided in the kit. After vortexing and centrifugation, the supernatant was split in two new 2ml tubes containing lysis buffer and proteinase and incubated at 70°C for 10 minutes. Then, alcohol was added and the tubes were vortexed. The solutions of the duplicated samples were then transferred into the Qiagen column by a maximum of 700µl at a time and the column was centrifuged. This step was repeated until the whole solution of the duplicated samples was passed trough the column. The columns were then washed twice with washing buffers. DNA was eluted in 375µl of heated elution buffer and recovered after centrifugation. All eleven samples and the control were subsequently checked for success by amplification, together with a positive control made of highly diluted female capercaillie liver DNA, then stored at -20°C until further processing.

I. 3. 3 *Molecular sexing*

Sex was determined by a PCR-based method and checked against stool size. Capercaillie stool size is characteristic and allows sex determination (Leclercq, 1987). However, faeces from females cannot be discriminated from juveniles of both sexes by their size. As a consequence, we propose a new PCR-based method to assess the sex of all samples. Using sex-specific universal primers (2550F and 2718R; (Fridolfsson & Ellegren, 1999), we were able to confirm the sex of liver extract of two reference samples. To obtain shorter products adapted for degraded faecal sample amplification, we sequenced for both sexes the products obtained in the previous PCR reactions and we designed two reverse primers, one located in the *W* intron (TuWR: TAATCAGAGCAACCTGAATGC) and one in the *Z* intron (TuZR: GGAATGTTAACATACTCCTTCACA), both to be used with the 2550F forward primer. Amplifications were carried out in 25 µl mix of Perkin Elmer Gold Taq, 1 X PCR reaction buffer, 2.5 mM $MgCl_2$, 0.15 $mg.ml^{-1}$ bovine serum albumin, 0.5 to 1 unit Gold Taq (Applied Biosystems), 5 µl template DNA and 2X 7.5 to 15 pmoles of each primer. The reactions were performed in PE9700 thermocycler with the program as follows: initial denaturation at 96°C for 2 minutes then, 10 touch down cycles (from 60 to50°C) with the same duration as following cycles, then 40 times the following cycle: Denaturation at 94°C, annealing at 53°C and elongation at 72°C, all of 45" long. The cycle ends by a terminal elongation of 300" at 72°C. Amplified fragments were visualized on 1.5% agarose gel stained with ethidium bromide and sex was determined based on the presence of two bands in females (W Z) and one band in males (Z Z).

I. 3. 4 *Microsatellite amplification*

PCR protocols were optimized according to the experimental procedure used when genotyping low-DNA-quantity samples (Taberlet *et al.*, 1996), using Ampli*Taq* Gold DNA polymerase (Applied Biosystems).
We used 11 microsatellite markers, using primers modified for use on ABI377 automated sequencer (5' fluorescent labelling: FAM, HEX or NED). Seven of these primers were designed for capercaillie (TUT1, TUT2, TUT3, TUT4, TUD1, TUD3, TUD5 (Segelbacher *et al.*, 2000)) and 4 for a related species (TTD1, TTD2, TTD6, TTT1 (Caizergues *et al.*, 2001)), thus allowing to test for cross species amplification. Amplifications parameters were optimized for each primer set individually, first on liver extract, and subsequently on multiple stool extracts. All amplifications were carried out in 20 µl reaction mix, containing: 5µl of DNA template,

PCR buffer, 2.5 mM $MgCl_2$, 0.15 $mg.ml^{-1}$ Bovine serum-albumin from (Roche Diagnostics GmbH, Mannheim, Germany), 0.1mM of each dNTP, 0.5 to 1 unit of Gold *Taq*) and 0.3 µM of each primer in each tube. Diluted liver DNA was used for allele size reference and PCR positive control. As negative controls, DNA template was replaced by 5µl nanopore filtered water. To avoid contaminations, PCR reactions were conducted in a different room. An Applied Biosystems 9700 thermocycler was used, following a time-release PCR protocol allowing delayed activation of DNA polymerase along the amplification cycles: After an initial denaturation at 96°C for 2 min., 45 to 55 cycles were conducted with 45 sec. steps: denaturation at 94°C, annealing at 59°C and elongation at 72°C. Amplification products were run in an ABI377 automated sequencer electrophoresis apparatus, into two sets of six multiplexed primer sets and on 64 wells, 4.5% Sequagel-4.75 (National Diagnostics, Hessle Hull, UK) mix plates.

To assess the success rate of the extraction protocol, we tested 85 stool samples randomly chosen out of a total set of capercaillie stool sampled in the field. Extraction success was subsequently checked by visualisation of amplicons on 1.5% ethidium-bromide stained agarose gels. Positive extractions were assessed after two successive PCR amplifications of the TTT1 microsatellite loci (Caizergues *et al.*, 2001). The negative PCR after two amplifications were re-extracted and tested again with PCR until positive PCR or until extraction was attempted four times.

To estimate error rates on our method, we genotyped 57 individuals with these amplifications and genotyping protocols repeated 6 times. The software Gimlet (Valière, 2002) was used to estimate allelic drop-out (ADO), false allele (FA) rates, and probability of identity (P.I.). The consensus genotypes were reported according to the multitube approach indications: heterozygote when the double genotype appeared at least twice and homozygotes when the simple genotype appeared at least 4 times out of 6.

Allele frequencies and heterozygosity rates were calculated using FSTAT (Goudet, 1995-2001). These data were subsequently used in the software Gemini (Valière *et al.*, 2002) to estimate the probability of exact genotyping when two to 10 repeats were performed.

I. 3. 5 DNA incubation experiments

To unravel the potential causes of DNA degradation in our samples (hydrolysis, enzymes or microbes), we collected faeces samples from two males and three females in Bern Zoological garden (Switzerland). The material was dried 24 hours at 65°C, then homogenised together with

mortar and pestle. Homogenized faeces (89g) were mixed with 870ml autoclaved PBS buffer (NaCl 8.56mM K2HPO41.72mM and KH2PO4 2.94mM; pH 7.2). The solid and liquid phases were separated by filtration (LS 14 ½, Schleicher & Schluer). The faecal liquid phase was subsequently filtered on glass-fibre membranes (GF/C, Whatman). Solid faecal material was autoclaved 20 min at 120°C. Capercaillie DNA was obtained from capercaillie liver using DNeasy Tissue Kit (Qiagen) following manufacturer instructions.

The faecal material was submitted to four different treatments with 8 replicates and 2 controls without DNA addition. Each replicate was prepared as faecal slurry in 15ml Falcon tubes containing 0.5g of autoclaved solid material, 1.1 µg of capercaillie DNA, and 3 ml of the following solutions prepared from the faecal liquid phase (Fig. 1).

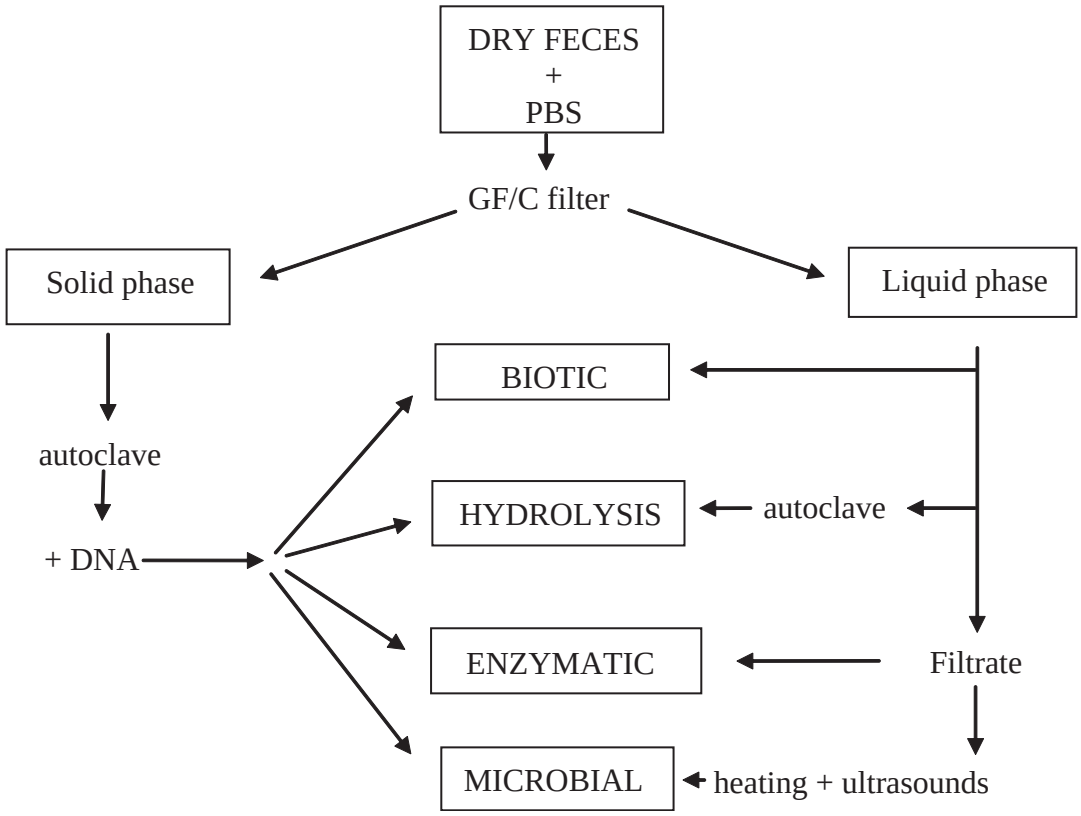


Fig. 1: Preparation of biotic, hydrolysis, enzymatic and microbial treatments
After homogenizing dry faeces in PBS buffer, solid and liquid phases were separated. The solid phase was autoclaved to inactivate biotic and enzymatic components and was separated in four part to receive the four treatments applied to the liquid phase.

a- Biotic treatment: Non-autoclaved faecal liquid. This treatment included hydrolysis, enzymatic and microbial factors.

b- Hydrolysis treatment: Autoclaved faecal liquid. Autoclaving denatured enzymes and killed microorganisms.

c- Enzymatic treatment: Faecal liquid filtered on cellulose acetate membrane (porosity 0.2µm, Whatman). Filtration removed the microorganisms.
d- Microbial treatment: Microbial suspension. After filtration of the faecal liquid on 0.2µm cellulose acetate membranes, the filter was kept and the filtrate was heated 10 min at 90°C to denature DNases. Microorganisms trapped on the filter were re-suspended in the heated filtrate with 5 min in an ultrasonic bath (Telsonic ultrasonics).

All faecal slurries were prepared under laminary flux hood to avoid lab contamination. For each treatment, four replicates and one negative control (without DNA addition and without treatment) were incubated at 12°C and at 25°C in order to evaluate temperature effect. Subsamples of 200 µl were retrieved after 1, 3 and 7 days of incubation and were stored at -20°C until estimation of DNA quality and DNases activity.

I. 3. 6 DNA Quality

DNA quality was tested as the capacity to amplify different sizes of nuclear DNA as in (Höss *et al.*, 1996). Sub-samples from the DNA incubation experiment were centrifuged 10 min at 10,000 rpm. Supernatant was kept for further DNase activity assay and DNA was extracted from the pellet using QIAamp Stool Kit (Qiagen, Chatsworth, CA, USA) as described above. Nuclear DNA fragments of different length were amplified using PCR methods described in the molecular sexing or amplification section. The following primers were used: the sexing TuWZF/TuZR primers for 350pb fragment, the microsatellite primers TTT1 for 220pb (Caizergues et al., 2001) and TTD6 for 130pb (Segelbacher et al., 2002). The size of PCR products was visualized on 1.5% agarose gels. Amplification success was ranked from 0 to 3, according to the number of positive PCR products (130, 220 and 350 bp). In negative controls without DNA addition, only the 130 bp amplified.

I. 3. 7 DNase activity assay

We estimated the DNase activity according to (Ruiz *et al.*, 2000), modified. DNA degradation reactions involved 2 µl of the incubated subsample supernatant, 2µg of herring DNA (Sigma), 0.8 µl 1 M Tris, 0.2 µl of 1 M MgCl2 (Fermentas) and 6µl nanopore filtrated H_2O. In negative controls, the supernatant sample was replaced by 2µl nanopore filtrated water. In positive control, the sample was replaced by 2µl DNaseI

solution (1 U/µl, Fermentas). After one hour at 37°C, samples were run on 1% agarose gel. Gels showing no DNA left were marked as "2", gels showing noticeable degradation (smear) were marked as "1", and gels showing no degradation were marked as "0".
Results were tested for significance with the program JMP5.01 (SAS Institute, 1995) by ordinal logistic regression for DNA quality and DNase activity.

I. 4 RESULTS

I. 4. 1 DNA extraction and amplification efficiency

At first attempt, we successfully amplified 69% of the DNA extracted from faeces samples using the TTT1 microsatellite loci (Table 1). We also found that ca. 99% of the samples tested were usable for microsatellite typing if, after failing, the extraction was repeated again up to a maximum of 4 times.

Table 1: Success rate of extractions, tested by PCR

Number and rate of positive PCR of a microsatellite loci after 1 to 4 extraction attemps

	success	rate
1st attempt	59	69.4%
2nd attempt	75	88.2%
3rd attempt	81	95.3%
4th attempt	84	98.8%

I. 4. 2 Error rates

The errors rates, averaged under 11 loci, were estimated to 21 % (+/- 6%) of allelic drop-out (ADO total, Table 2) and 3%(+/- 2%) of false alleles (FA total, Table 2), (respectively 19% +/-15% and 2% +/- 3% when averaged over the 57 samples).
Furthermore, the mean ADO, FA and success rate across loci designed specifically for *T. urogallus* (TUT1, TUT2, TUT3, TUT4, TUD1, TUD3,

TUD5) did not differ significantly better then the cross-species loci (TTD6, TTD2, TTD1, TTT1) designed for *T. tetrix* in terms of.

Table 2: *Per locus error rates*

nb PCR: total number of PCR conducted; Failed: number of PCR for which no signal was detected on the automated sequencer; success: peak detected; ADO & FA: Allelic drop-out and False allele rates; PI: probability of identity; PI(*sib*) PI if individuals of focus are sibs. PI rank: loci are ranked according to the lower PI. see (Valière N., 2002)

	TTD2	TUT2	TTD6	TTD1	TTT1	TUT4	TUD3	TUT1	TUD1	TUD5	TUT3	total	SD
nb PCR	342	342	342	342	342	342	342	342	342	342	342	3762	
failed	28	9	12	62	52	46	27	15	50	51	45	36.09	18.51
%failed	8.19	2.63	3.51	18.13	15.2	13.45	7.89	4.39	14.62	14.91	13.16	10.55	5.41
%success	91.81	97.37	96.49	81.87	84.8	86.55	92.11	95.61	85.38	85.09	86.84	89.45	
ADO	0.169	0.239	0.114	0.258	0.19	0.28	0.104	0.199	0.223	0.264	0.285	**0.21**	0.06
FA	0.086	0.045	0.033	0.015	0.007	0	0.033	0.029	0.033	0.006	0.008	**0.03**	0.02
PI	0.0671	0.0865	0.115	0.1244	0.1446	0.1916	0.1957	0.2502	0.2557	0.3876	0.5921	6.6E-09	
PI*(sib)*	0.3767	0.4068	0.4217	0.4312	0.4515	0.4937	0.4846	0.5279	0.5469	0.6158	0.7805	4.2E-04	
PI rank	1	2	3	4	5	6	7	8	9	10	11		

I. 4. 3 Estimation of exact genotype

The genotyping reliability obtained estimated using computer simulations (Valière *et al.*, 2002) towers at 98% after average of 10000 simulations when 5 genotyping repetitions are performed (Fig. 2). The success rate for more than 6 PCR repetitions remains at 98% or less (the number of ADO and FA increasing with PCR repetitions, the number of ambiguous genotypes increases as well, decreasing the success rate.)

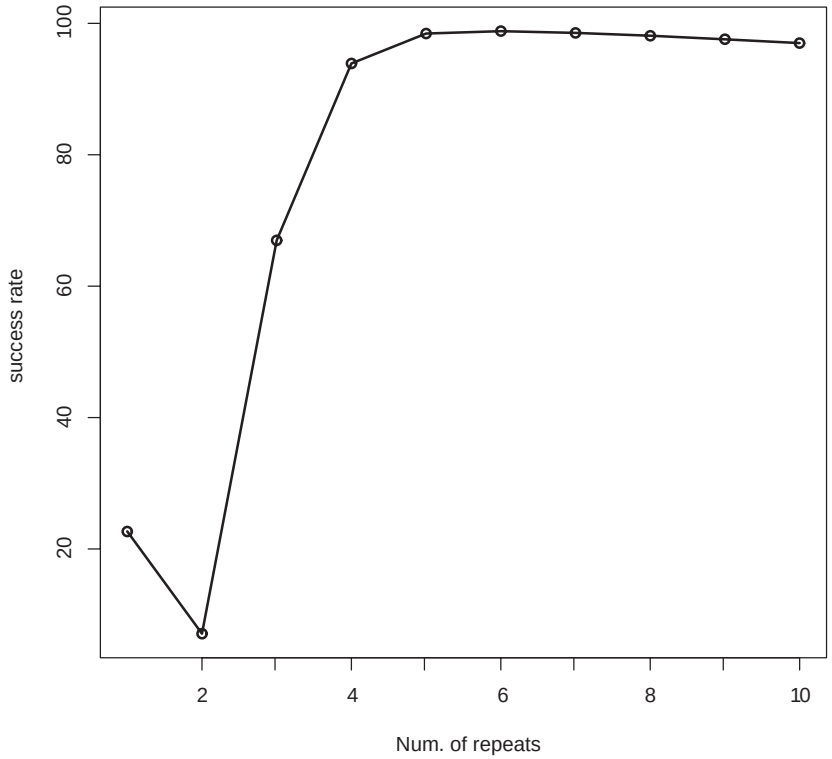


Fig. 2: Genotyping success estimated by Gemini software: probability of exact genotyping obtained after 1 to 10 PCR and genotyping repetitions

In hydrolysis and microbial treatments, DNA quality remained high (2 to 3 amplified bands per sample) after 7 days of incubation (Fig. 3 A and B) at both temperatures. In 25°C microbial treatment, DNA quality dropped from 3 to 2, suggesting microbial degradation of DNA (Fig. 3 B), while at 12°C DNA quality in the microbial treatment was stable over 7 days. In enzymatic and biotic treatment, the quality was low (1 band) after 24 hours of incubation at 12 and 25°C.

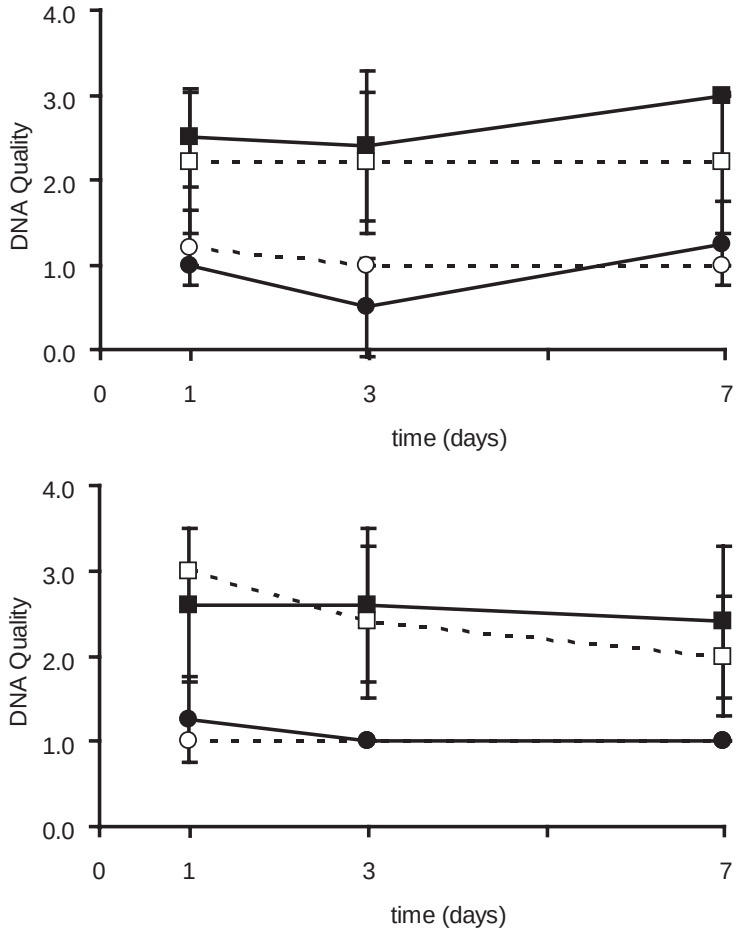


Fig. 3: DNA quality in incubated faecal slurry (average and standard deviation of PCR product number, n=4). Incubations were conducted at 12°C (A) and 25°C (B). The DNA quality is plotted against incubation time in hydrolysis treatment (■, solid line), microbial treatment (□, dashed line), enzymatic treatment (●, solid line) and biotic treatment (❍, dashed line)

DNA quality was significantly different among treatments (Tab. 3). Temperature and incubation time had no significant effect on DNA quality. However their interaction tended to be significant (Tab. 3).

Table 3: *Factors of DNA degradation*

Difference of DNA quality in incubated fecal slurries (n=91) among treatments tested using an ordinal logistic regression

Variables	dl	Chi-square	p-val
Treatment	3	35.48	0
Temperature	1	1.33	0.25
Time	2	3.75	0.15
Time * temperature	2	5.65	0.06
Final model	8	116.52	<0.0001

Enzymatic treatment (DNA quality 25%=1.0, median= 1.0, 75%=1.0) and biotic treatment (DNA quality 25%=1.0, median= 1.0, 75%=1.0) had the highest degradation effects. microbial treatment (DNA quality 25%=2.0, median= 3.0, 75%=3.0) and Hydrolysis treatment (DNA quality 25%=3.0, median= 3.0, 75%=3.0) had the lowest degradation effects. Ordinal logistic regressions (SAS institute) showed that DNA quality and DNase activities were significantly correlated at each sampling step, after 1 day ($\chi_{4,24}$=31.30, p<0.0001), 3 ($\chi_{4,24}$=21.06, p=0.0003) and 7 days ($\chi_{4,25}$=35.47, p<0.0001).

I. 5 DISCUSSION

This paper reports the first successful attempt to estimate the reliability of microsatellite genotyping with non-invasive molecular techniques in a bird species. We also describe a simple molecular sexing protocol applied to capercaillie faecal samples. The large elution volume of DNA extracted through an extraction protocol that ensures cleansing of PCR inhibitors, the time-release PCR amplification, together with a reasonable number of 5 replicates using the multitube approach (Taberlet *et al.*, 1996) for homozygote genotypes offer a theoretical average of 98% certainty on the genotype.

I. 5. 1 Comparison of reliability level with other studies

Our reliability level was similar to those reported in previous studies. For instance, Bayes, Smith *et al* (2000) and Hedmark, Flagst *et al.* (2004) assessed genotyping fidelity by comparing blood or tissues against faeces

genotypes, and found 92 and 100% accuracy, respectively. We were not able to gather enough captive capercaillies to perform such a comparison between tissue and faecal samples that could have provided us with a direct estimate of reliability on genotyping. Nevertheless, our results allow us to plan a large scale population genetic study where the errors dues to low quality of the DNA templates will produce a negligible background noise in the data to be analysed for population genetic analysis and parentage studies.

Our estimation of 21% allelic drop-out and 3% false allele rates in capercaillie stool, as the first study on nuclear DNA obtained from avian faeces is in the range of that found in genotyping conducted on mammals. For example, previous studies on primates reported the following rates of allelic drop-out: 4.2% ADO rates were found in *Pongo* faeces, (Goossens *et al.*, 2000), 12 to 36 % in *Pan troglodytes* (Morin *et al.*, 2001) and between 0 and 6% ADO in *Macaca sylvanus* (Lathuilliere *et al.*, 2001). There are examples in other mammal species also: 27% ADO rate was reported in badgers (*Meles meles*) (Frantz *et al.*, 2004), 10% in Swedish wolverine (Flagstad *et al.*, 2004) and finally 2% in deers, for which stool are covered by a cell-rich mucous (Flagstad *et al.*, 1999). Samples from small populations or social groups intensively followed by biologists, such as the ones conducted in primates, can provide larger and fresher samples. In the case of wild grouse, faecal DNA obtained from field screening is more likely scarce and in small quantities.

I. 5. 2 Effect of diet on reliability

The variation in ADO rates in literature reflects the various sampling procedures, extractions and amplification protocols used, and the physiological variations between species. Most studies so far have been done on carnivorous or omnivorous species and few herbivorous have been studies. We can suppose that diet can have a strong impact on genotyping quality when using faecal samples as most PCR inhibitors are glycolipids of vegetal origin, and as the texture of the food ingested can influence the turnover of epithelial cells that are the most likely source of DNA in faecal samples (Murphy *et al.*, 2003).

I. 5. 3 Applications for forensic studies

The level of reliability found in our study is sufficient for population genetic studies, and performing 1 or 2 more repeats would not improve the reliability. However, this result is prone to a level of bias that would

not be acceptable for other types of studies, such as forensic investigations. In the latter case, we would suggest to increase the number of markers in order to decrease the probability of identity to allow higher definition of individual identification.

I. 5. 4 Storage conditions

To stress the importance of storage condition, several studies have investigated the alternative possibilities of dry storage, mostly in alcohol (Murphy *et al.*, 2002). We submitted to our new amplification protocol to 20 samples taken from a long lasting survey in the Pyrenees conducted by the French Office National de la Chasse et de la Faune Sauvage (ONCFS) under the direction of Dr. E. Menoni dating back from as far as 1984, and that were kept in paper bags. The success rate with these samples was low (approx. 20% PCR success, data not shown) suggesting that the lack of protection from humidity. Furthermore, these storage conditions did not protect the samples from contamination by neighbouring samples.

I. 5. 5 DNA degradation

In our experiment, we demonstrated the rapid DNA degradation in moist conditions (within 24h). Our results showed clearly that enzymatic degradation was the principal cause for DNA degradation in these conditions at 25 and 12 °C. At the opposite, hydrolysis caused minor degradation in our experiment. Evidences of a strong DNase degrading effect in grouse stool found in our study conflicts with other studies reporting hydrolyse as major degrader. For instance, Lindhal (1993) showed that nucleic acids left in wet conditions were degraded by physico-chemical factors (hydrolysis and oxidation). Another factor of degradation suggested to be important was faecal bacteria (Flint & Thompson, 1990). We did not find that microorganism activity had a major effect in our samples. Although increased microbial degradation at 25°C compared to 12°C suggested that microbial DNase activity could have occurred, this increase was marginal compared to the degradation obtained in the enzymatic treatment. Soluble DNases released either from dead epithelial cells or gut microbial cells were probably responsible for this rapid degradation. Ruiz et al. (2000) showed that gut fluids can exhibit strong DNase activities.

Therefore, samples collected after rain days or left in humidity at temperature over 0°C are expected to show lower reliability that fresh samples. Stool sampled from the field should have remained in conditions

that restrain DNase activity. Frozen winter samples, stool collected on the snow (Lucchini *et al.*, 2002) or dried in the open air shall be preferred. Storage of samples, also, should involve avoidance of nucleolytic activity. Long term storage is therefore advised in dry and enclosed vials.

I. 5. 6 Conclusions

Finally, the financial cost in laboratory disposables and reactants (extraction kits, PCR, labelled primers and electrophoresis) is not comparable with classic genetic studies. If the stool extraction kits are somewhat in the same range of cost as usual Qiagen kits, the relatively low success increases the numbers of columns. The number of gels and PCR performed is to be multiplied by ca. 4 comparatively to the regular typing. On the other hand, increased time expenses in the lab are largely compensated by the gain in sampling effort compared to usual invasive or destructive sampling surveys conducted in grouse.
The estimations of genotyping reliability exposed here indicate that non-invasive sampling of avian populations over large scale and large populations can be performed with success. By choosing an appropriate approach, we were able to provide optimized extraction and amplification protocols for reliable genotyping in spite of the hindrances uncovered in previous studies (inhibition, contaminations, allelic drop-out and false alleles). By demonstrating the effect of free DNases as major DNA degrader in our samples, we highlighted the role of sampling scheme and storage conditions in loss of DNA quality.

I. 6 ACKNOWLEDGMENTS

We are indebt to P. Taberlet for discussing and help improving the protocols and to N. Valière for valuable comments on the manuscript. We thank the students who participated in this study: I. Py, C. Salamin, C. Decoster. Further, we thank M. Rosset, from Berne Zoological Garden, and Dr M. Menoni who provided the samples.

Bayes MK, Smith KL, Alberts SC, Altmann Jo and Bruford MW (2000) Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. *Conservation Biology* **1**, 173-176.

Broderick D, Idaghdour Y, Korrida A and Hellmich J (2003) Gene flow in great bustard populations across the Strait of Gibraltar as elucidated from excremental PCR and mtDNA sequencing. *Conservation Genetics* **4**, 793-800.

Caizergues A, Dubois S, Loiseau A, Mondor G and Rasplus JY (2001) Isolation and characterization of microsatellite loci in black grouse (Tetrao tetrix). *Molecular Ecology Notes* **1**, 36-38.

Flagstad O, Hedmark E, Landa A*, et al.* (2004) Colonization history and noninvasive monitoring of a reestablished wolverine population. *Conservation Biology* **13**, 676-688.

Flagstad O, Roed K, Stacy JE and Jakobsen KS (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. *Molecular Ecology* **8**, 879-883.

Flint HJo and Thompson AM (1990) Desoxyribonucase activity in rumen bacteria. *Applied Microbiology Letters* **11**, 18-21.

Frankham R, Ballou JD and Briscoe DA (2002) *An introduction to conservation genetics* University Press, Cambridge.

Frantz AC, Pope LC, Carpenter PJ*, et al.* (2004) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. *Molecular Ecology* **12**, 1649-1661.

Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK and Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428.

Fridolfsson AK and Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**, 116-121.

Goossens B, Chikhi L, Utami SS, de Ruiter Jo and Bruford MW (2000) A multi-sample, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. *Conservation Genetics* **1**, 157-162.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3), available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet

J, 1995. FSTAT (vers. 1.2): a computer program to calculate F-statistics. Journal of Heredity **86**: 485-486.

Harrison S, Quinn JF, Baaughman JF and Murphy DD (1991) Estimating the effect of scientific study on 2 butterfly populations. *American Naturalist* **137**, 227-243.

Hedmark E, Flagstad O, Segerstrom P*, et al.* (2004) DNA-based individual and sex identification from wolverine (Gulo gulo) faeces and urine. *Conservation Genetics* **5**, 405-410.

Höss M, Jagura P, Zastawny TH, Dizdaroglu M and Pääbo S (1996) DNA damages and DNA sequences retrival from ancient tissues. *Nucleic Acids Research* **24**, 1304-1307.

Höss M, Kohn MH, Pääbo S, Knauer F and Schröder W (1992) Excrement analysis by PCR. *Nature* **359**, 199.

Idaghdour Y, Broderick D and Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustard. *Conservation Genetics* **4**, 789-792.

Kohn MH and Wayne RK (1997) Facts from faeces revisited. *Trends in Ecology and Evolution* **12**, 223-227.

Lathuilliere M, Menard N, Gautier-Hion A and Crouau-Roy B (2001) Testing the reliability of noninvasive genetic sampling by comparing analyses of blood and fecal samples in barbary macaques (*Macaca sylvanus*). *American Journal of Primatology* **55**, 151-159.

Leclercq B (1987) *Ecologie et dynamique des populations du grand tétras dans le Jura francais* PhD thesis, Université de Bourgogne.

Leclercq B (2004) Evolution des populations du grand tétras dans le jura Francais.

Lindhal T (1993) Instability and decay of the primary structure of the DNA. *Nature* **362**, 709-715.

Lucchini V, Fabbri E, Marucco F*, et al.* (2002) Noninvasive molecular tracking of colonizing wolf (Canis lupus) packs in the western Italian Alps. *Molecular Ecology* **11**, 857-868.

Matsui K, Honjo M and Kawabata Z (2001) Estimation of the fate of dissolved DNA on thermally stratified lake water from the stability of exogenous plamid DNA. *Aquatic Microbiology and Ecology* **26**, 95-102.

Monteiro L, Bonnemaison D, Vekris A*, et al.* (1997) Complex polysaccharides as PCR inhibitors in feces: *Helicobacter pylori* model. *Journal of Clinical Microbiology* **35**, 995-998.

Morin PA, Chambers KE, Boesch C and Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology* **10**, 1835–1844.

Morin PA, Wallis J, Moore JJ, Chakraborty R and Woodruff DS (1993) Noninvasive Sampling and DNA Amplification for Paternity Exclusion, Community Structure, and Phylogeography in Wild Chimpanzees. *Primates* **34**, 347-356.

Murphy MA, Waits LP and Kendall KC (2003) The influence of diet on faecal DNA amplification and sex identification in brown bears (Ursus arctos). *Molecular Ecology* **12**, 2261-2265.

Murphy MA, Waits LP, Kendall KC*, et al.* (2002) An evaluation of long-term preservation methods for brown bear (Ursus arctos) faecal DNA samples. *Conservation Genetics* **3**, 435-440.

Ruiz TR, Andrews S and Smith GB (2000) Identification and characterisation of nuclease activity un anaerobic environmental samples. *Canadian Journal of Microbiology* **46**, 736-740.

Sachot S (2002) *PhD thesis : Viability and management of an endangered capercaillie (Tetrao urogallus) metapopulation*, Université de Lausanne.

Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. *Molecular Ecology Notes* **2**, 367-369.

Segelbacher G, Paxton RJ, Steinbruck G, Trontelj P and Storch I (2000) Characterization of microsatellites in capercaillie Tetrao urogallus (AVES). *Molecular Ecology* **9**, 1934-1935.

Taberlet P, Camarra JJ, Griffin S*, et al.* (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology* **6**, 869-876.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Taberlet P, Mattock H, Duboispaganon C and Bouvet J (1993) Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. *Molecular Ecology* **2**, 399-403.

Valière N (2002) GIMLET: A computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**, 377-379.

Valière N, Berthier P, Mouchiroud D and Pontier D (2002) Gemini: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83-86.

Woegerbauer M, Burgmann H, Davies J and Graninger W (2000) DNase I induced DNA degradation is inhibited by Neomycin. *The Journal of Antibiotics* **53**, 276.

II DETECTING THE NUMBER OF CLUSTERS OF INDIVIDUALS USING THE SOFTWARE STRUCTURE: A SIMULATION STUDY

Sébastien Regnaut, Guillaume Evanno & Jérôme Goudet

II . 1 ABSTRACT

The identification of genetically homogeneous groups of individuals is a long standing issue in population genetics. A recent Bayesian algorithm implemented in the software STRUCTURE (Pritchard *et al.*, 2000) allows the identification of such groups. However, the ability of this algorithm to detect the true number of clusters in a sample of individuals when patterns of dispersal among populations are not homogeneous has not yet been tested. The goal of this study is to test the ability of this algorithm to detect the true number of clusters under various dispersal scenarios using data generated from an individual based computer model. Using a simple graphical method, we found that STRUCTURE always detects the highest level of structure. As might be expected, the results are sensitive to the type of genetic marker used (AFLP vs. microsatellite), the number of loci scored, the number of populations sampled, and the number of individuals typed in each sample. Interestingly, it seems that dominant AFLPs like markers give a stronger signal than microsatellites, as long as there is five times more polymorphic AFLP bands than microsatellite loci.

Population genetics deals with the variations of allele frequencies between and within populations. The most widely used measures of population structure are the Wright's F statistics (Wright, 1931). To calculate these indices, one needs first to define groups of individuals and then to use their genotypes to compute variance in allele frequencies. Thus, a fundamental prerequisite of any inference on the genetic structure of populations is the definition of populations themselves. Population determination is usually based upon geographic origin of samples or phenotypes. However, the genetic structure of populations is not always reflected in the geographical proximity of individuals. Populations that are not discretely distributed can nevertheless be genetically structured, due to unidentified barriers to gene flow. In addition, groups of individuals with different geographic locations, behavioural patterns or phenotypes are not necessarily genetically differentiated (for instance, migratory bats from the same breeding roost could be sampled thousands of kilometres apart in winter, see e.g. (Petit *et al.*, 2001).

Among the methods not assuming pre-defined structure, tree-based methods use genetic distance between individuals and tree construction algorithms such as UPGMA or Neighbour-joining to group them in clusters (Saitou & Nei, 1987). Similarly, multivariate analyses such as multidimensional scaling can help in identifying clusters of individuals. However, these graphical methods are only loosely connected to statistical procedures allowing the identification of homogeneous clusters of individuals.

An alternative method developed recently by Pritchard *et al.* (2000) and implemented in the software STRUCTURE aims at delineating clusters of individuals on the basis of their genotypes at multiple loci using a Bayesian approach. Given a number of clusters K, the method works by finding K groups of individuals within which heterozygote deficiency and linkage disequilibrium are minimized. The likelihood for each value of K is given, allowing estimating the more likely number of clusters. A quantification of how likely each individual is to belong to each group is also given, information that can be then used to assign individuals to populations. The program has been widely used for detection of genetic structure in sample populations for medical purposes (Pritchard &

Donnelly, 2001; Satten *et al.*, 2001), assignment studies (Rosenberg *et al.*, 2001), population admixture and hybridization analysis (Beaumont *et al.*, 2001; Goossens *et al.*, 2002; Randi & Lucchini, 2002), migration and dispersal analysis (Arnaud *et al.*, 2003; Berry *et al.*, 2004; Cegelski *et al.*, 2003) and also to detect, with or without success, cryptic genetic structure of natural populations (Caizergues, Bernard-Laurent *et al.*, 2003; Ciofi *et al.*, 2002; Hampton *et al.*, 2004; Rosenberg *et al.*, 2002; Vernesi *et al.*, 2003; Zeisset & Beebee, 2003). Among the Bayesian clustering methods, STRUCTURE is the most widely used. While other methods have been developed (Banks & Eichert, 2000; Corander *et al.*, 2003; Dawson & Belkhir, 2001) and still other methods for the assignment of individuals to populations exist but imply the a priori knowledge of source populations (Cornuet *et al.*, 1999; Paetkau *et al.*, 1995; Rannala & Mountain, 1997), we will focus here exclusively on the software STRUCTURE.

Tests and comparative studies using empirical datasets have been performed to assess STRUCTURE's ability in assigning individuals to their known cluster of origin (Pritchard *et al.*, 2000; Rosenberg *et al.*, 2001; Turakulov & Easteal, 2003). Most of these studies have proven the software to be efficient in assigning individuals to their populations of origin (albeit most are based on simulations with limited number of populations and absence of dispersal between them). However, little is known on the crucial ability of STRUCTURE to detect the real number of clusters (K) which composes a data set. Pritchard *et al.* (2000) showed that STRUCTURE easily detects 2 to 4 highly differentiated populations but studies in molecular ecology usually include many more populations and very often these populations are not spatially evenly distributed. Many studies have described migration patterns departing from Wright's island model and including several hierarchical levels and/or isolation by distance. For instance, Giles *et al.* (1998), Chapuisat *et al.* (1997), Bouzat *et al.* (2004) or Trouvé *et al.* (2004) have documented situations with a hierarchical pattern of population structure, as groups are themselves clusters of differentiated populations. Another pattern frequently described is a contact zone between otherwise isolated populations. This situation implies a relative genetic isolation between the two groups of populations and sometimes also a pattern of isolation by distance within each group. Such a migration scheme was found for instance by Lugon-Moulin *et al.* (1997) who describe two longitudinal geographical patterns of isolated shrew populations separated by a zone through which dispersal is strongly reduced.

Many of these studies have been conducted using microsatellite markers to assess polymorphism. These DNA markers are widely used because

they are both codominant and highly polymorphic (Jarne & Lagoda, 1996). However, their development is relatively expensive, time-consuming and can be difficult. An alternative family of markers also commonly used in populations studies are the Amplified Fragment Length Polymorphism (AFLPs) (Vos *et al.*, 1995). AFLPs generate hundreds of polymorphic bands and are easier to develop than microsatellites, but they have the potential inconvenient of being dominant (a DNA band is either present or absent). These two types of markers have different properties. For instance, Gaudeul *et al* (2004) reported very different levels of population structuring inferred from AFLPs and microsatellite markers. Both AFLP and microsatellites can be used for assignment studies but their respective ability to delineate clusters of individuals have not been compared so far.

The goal of this study is to test the ability of the algorithm underlying the software STRUCTURE to detect the number of clusters in situations including more than two populations. While the program is increasingly used, it is unknown whether it can efficiently detect the real number of clusters in hierarchical systems where migration between populations is uneven. We present an evaluation of the performances of the method under three models of population structure: the island model, a contact zone, and a hierarchical island model. For each model, we simulated AFLP and microsatellite genotypic datasets that were subsequently run in STRUCTURE, and analyzed the output. We find that the maximum of the second derivative of the likelihood function with respect to the number of clusters is a good predictor of the real number of clusters. STRUCTURE identifies groups of individuals corresponding to the highest hierarchical level, and performs well with both dominant and codominant markers.

II . 3 MATERIAL AND METHODS

II. 3. 1 Simulation of the three migration models

We used the software EASYPOP (Balloux, 2001) to generate genotypic data from 3 different models of population structure: an island model, a hierarchical island model and a contact-zone model. For all simulations and model of population structure, mutation process followed the K allele model (equal probability of mutations to any allelic state) at a rate of $\mu = 10^{-3}$. The modelled organisms are diploid, hermaphroditic and randomly mating (excluding selfing). Each simulation was run for 10 000

generations to obtain populations at drift, migration and mutation equilibrium. For each model, we generated 10 replicates where each individual genotype was made of 100 microsatellite loci each with 10 possible allelic states.

The parameters that were varied for the simulations are the number of populations, the number of individuals per population, and the migration rates. These parameters are summarized in Table 1. For the finite island model, 5 populations of 100 individuals each are exchanging migrants at a rate 0.01. The expected value of F_{ST} for these simulations is 0.15.

The hierarchical island model (Slatkin & Voelm, 1991) consists in 5 sets of 4 populations, each made of 50 individuals. Migration occurs at a rate 0.02 within archipelago and 0.001 between archipelagoes, (Table 1). The expected value of F_{ST} is 0.30 between archipelagos ($F_{ARCHIPELAGO-TOTAL}$), 0.16 between islands within archipelagos ($F_{ISLAND-ARCHIPELAGO}$), and 0.41 overall ($F_{ISLAND-TOTAL}$).

The contact zone model is characterized by 2 sets of 5 populations, which are organized in a one dimension stepping-stone scheme (Kimura & Weiss, 1964). Migration between the 2 sets occurs through the two central populations at a rate 10 times lower than within each set (Table 1). The expected value of F_{ST} for this model cannot be easily analytically resolved, but global F_{ST} estimated over the 10 replicates (ten times 100 microsatellite loci) is 0.33 and pairwise F_{ST} range from 0.16 to 0.43. The observed value of F_{ST} is 0.17 between the two sets ($F_{SET-TOTAL}$), 0.25 between populations within sets ($F_{POPULATION-SET}$), and 0.38 overall ($F_{POPULATION-TOTAL}$).

EASYPOP generates codominant, microsatellite -like genotypic data. In order to simulate dominant AFLP data, the genotypes generated by EASYPOP were re-coded as biallelic loci, in a manner similar to Mariette *et al.* (2002): A randomly chosen half of the microsatellite alleles were coded as "1" and considered dominant while the second half was coded as "2" and considered recessive ("2"). Because with dominant data, one cannot distinguish between a dominant homozygote and a heterozygote, dominant phenotypes (obtained from genotypes 1-1 and 1-2) were recoded as 1-0, where 0 indicates a missing datum. Thus, AFLP datasets bear a proportion of missing data that microsatellite sets do not. This is different from what is recommended in the user's manual of STRUCTURE (Pritchard & Wen, 2003), which suggests that dominant markers can be dealt with by coding each phenotype (absence or presence of a band) by a single allele and a missing datum (1-0 for dominant and 2-0 for recessive). We did not use this method because this implies adding a missing value also for recessive homozygous, which seems unnecessary.

Microsatellite datasets given to STRUCTURE were made of 10 loci as this is a number commonly found in molecular ecology studies. AFLP datasets were made of 100 loci, which seems conservative as AFLP-based studies often include hundreds of markers (Luikart *et al.*, 2003). A further reason for this 1:10 ratio of microsatellite loci to AFLP bands comes from a recent simulation-based study (Mariette *et al.*, 2002) showing that at least 10 times more AFLP than microsatellite loci are necessary to reach a similar accuracy in the estimation of genetic diversity.

II. 3. 2 Sampling Scheme

To assess the effects of sampling strategies on the method's accuracy, analyses were also carried out on partial datasets. We investigated first the effect of the number of typed loci by sampling only 5 microsatellites or 50 AFLP bands (Table 2). We also looked at the effect of sampling a subset of individuals from each population (Table 2). Last, for the hierarchical island model, we also looked at the effect of sampling a subset of the populations by randomly omitting one island per archipelago (Table 2). We tested whether partial sampling affected the detection of the true K by comparing results between full and partial datasets.

II. 3. 3 Structure runs

We set most of parameters to their default values as advised in the user's manual of STRUCTURE 2.0 (Pritchard & Wen, 2003). Specifically, we chose the admixture model and the option of correlated allele frequencies between populations, as this configuration is considered best by Falush *et al.* (2003) in cases of subtle population structure. Similarly, we let the degree of admixture Alpha be inferred from the data. When Alpha is close to zero, most individuals are essentially from one population or another, while Alpha > 1 means that most individuals are admixed (Falush *et al.*, 2003). Lambda, the parameter of the distribution of allelic frequencies, was set to one, as the manual advices. From a pilot study, we found that a length of the burnin and MCMC (Monte Carlo Markov Chain) of 10'000 each was sufficient. Longer burnin or MCMC did not change significantly the results. As we found that different runs could produce different likelihood values (even with much longer chains- e.g. 1'000'000), for each data set 20 runs were carried out in order to quantify the amount of variation of the likelihood for each K. The range of possible Ks we tested was from 1 or 2 to the true number of populations plus 3.

II. 3. 4 Statistics used to select K

True number of populations (K) is often identified using the maximal likelihood value returned by STRUCTURE (Ciofi *et al.*, 2002; Hampton *et al.*, 2004; Vernesi *et al.*, 2003; Zeisset & Beebee, 2003). However, we observed in our simulations that in most cases, once the real K is reached, the value for the likelihood at larger Ks plateaus (this phenomenon is mentioned in the Structure's manual (Pritchard & Wen, 2003) and the variance between runs increases. The likelihood distribution for K did not show a clear mode for the true K, but we found that the second derivative of the likelihood function with respect to K did show a clear peak at the real value of K. The rational for this is best explained graphically, as is shown on Figure 1. First, we plotted the mean likelihood over 20 runs for each K (Fig. 1A). Second, we plotted the mean difference between successive likelihood values of K (Fig. 1B). These differences correspond loosely to the first derivative of the likelihood function with respect to K, L(K)'. In a third step we plotted the (absolute value of the) difference between successive values of L(K)' (Fig. 1C). This corresponds loosely to the second derivative of L(K) with respect to K, L(K)''. We found the modal value of this distribution to be located at the real K. We used the height of this modal value as an indicator of the strength of the signal detected by STRUCTURE.

Figure 1

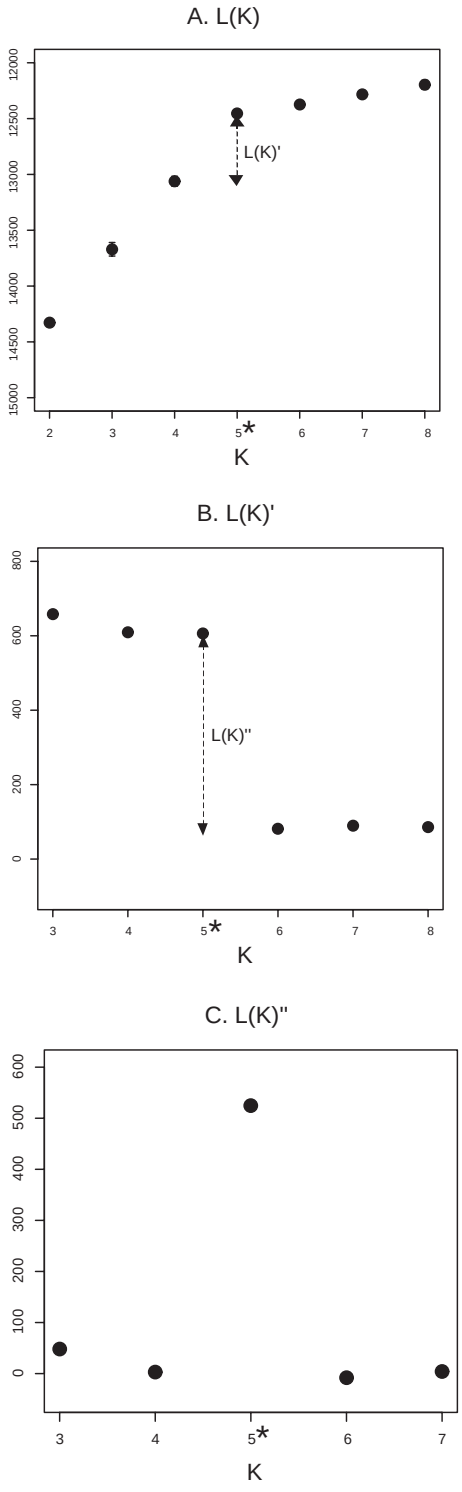


Figure 1. Description of the three steps for the graphical method allowing detection of the true number of groups K*. (A) Mean likelihood (± s.d.) over 20 runs for each K value. The model considered here is an island model using all 100 individuals per population and 50 AFLP loci. (B) First derivative of the likelihood distribution calculated as L(K)' = L(K+1) – L(K). (C) Absolute values of the second derivative of the likelihood distribution calculated according to the formula: |L(K)''| = |L(K+1)' - L(K)'|. The modal value of this distribution is the true K(*), here 5 populations.

Overall simulation scenarios, we seldom found a mode of the likelihood distribution at the real K. In most cases, the likelihood increased until the real K was reached, and then levelled off (often increasing slightly after the real K). On the other hand, the distribution of the second derivative of the likelihood as a function of K almost always showed a mode at the real K (Fig. 2).

For all three models, and both in full or partial configurations, STRUCTURE identified a number of groups corresponding to the highest hierarchical level of genetic partitioning between populations. STRUCTURE primarily highlights the between-sets of populations level for the hierarchical island model and the contact zone, and the between populations level for the island model. Importantly, these results were obtained by using the modal value of L(K)'' rather than the maximum value of L(K) (Fig. 1a, c). In Figure 2, the magnitude of the modal value of L(K)'' is plotted for each model, which allows the comparison of results obtained with different parameters sets. Overall, there was some variance among likelihood values L(K) for the different replicates of the same parameter set, but for all models except one (detailed below), all replicates had the same modal value for L(K)''.

Figure 2

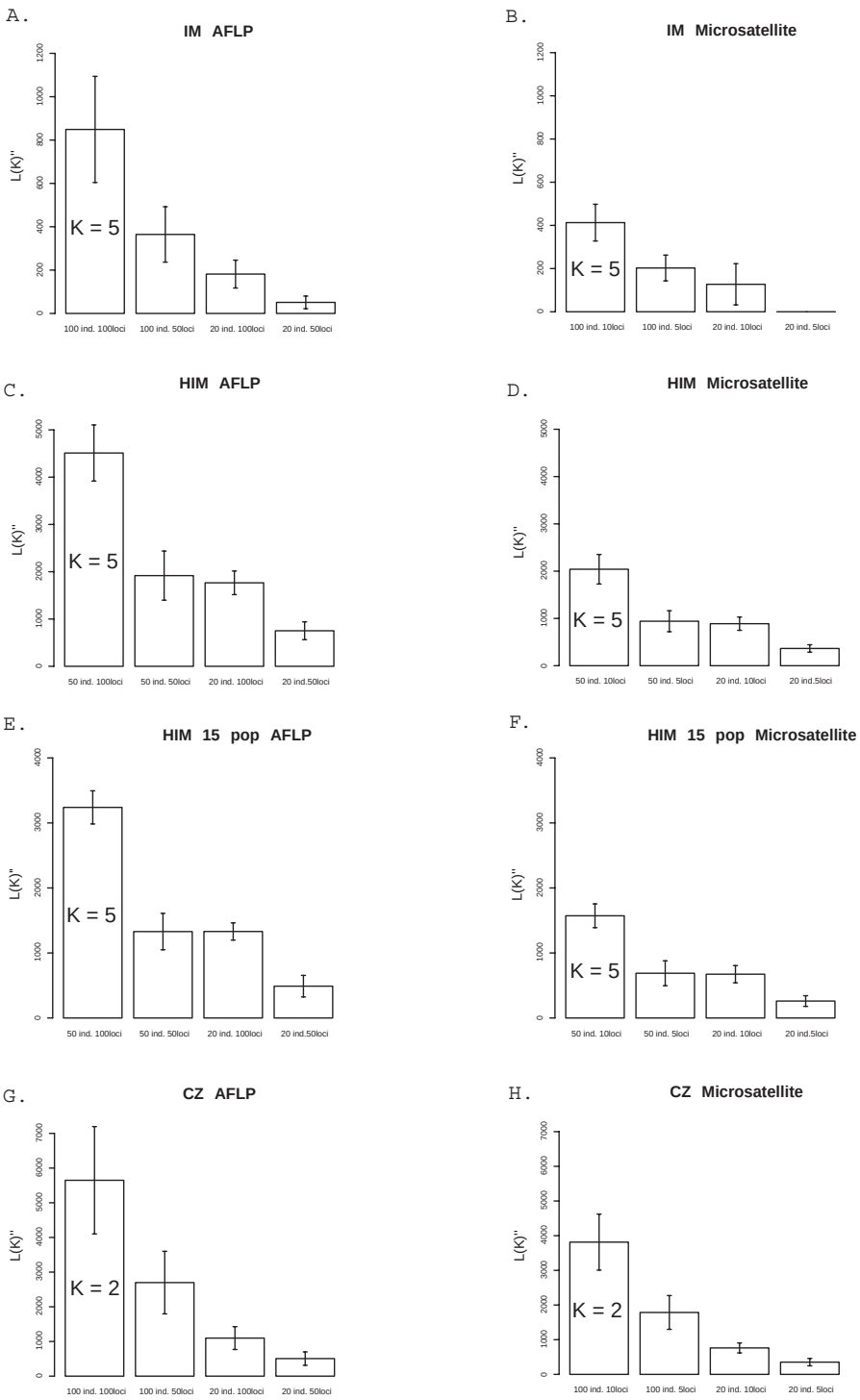


Figure 2. Histograms showing the magnitude of the modal value of the second derivative of the likelihood distribution (L(K)'') (mean of the absolute values ± s.d. over 10 replicates), calculated for each model using the procedure illustrated in Fig. 1. For the Hierarchical island model (HIM), the highest value of L(K)''was detected at K = 5; it was detected at K = 2 for the Contact zone model (CZM) and at K = 5 for the Island model (IM). (A) IM with AFLP loci, (B) IM with microsatellite loci, (C) HIM with AFLP loci, (D) HIM with microsatellite loci, (E) HIM with AFLP loci and 15 populations sampled out of 20, (F) HIM with microsatellite loci and 15 populations sampled out of 20, (G) CZM with AFLP loci, (H) CZM with microsatellite loci.

II. 4. 1 *Island model*

For the full data set, as well as for the partial samplings, the modal value of L(K)'' was K = 5, the real number of populations(Fig. 2a and 2b). The only situation in which STRUCTURE failed to detect the real K was the partial sampling of 20 individuals and 5 microsatellite markers (Fig. 2b). In this case, we did not see any plateau nor clear maximum in the likelihood distribution of K for any of the 10 replicates, and the software found a maximal likelihood value at K = 5 in 2 replicates, at K = 2 twice, at K = 3 four times and at K = 4 twice.
There is a strong effect of the partial sampling of individuals and loci for both microsatellite and AFLP markers (Fig. 2a and 2b). However, the partial sampling of 20 individuals per population affected more strongly the results than the sub-sampling of markers. The results were better with AFLPs than microsatellites since the average magnitude of modal value of L(K)'' is twice lower for the latter (Fig. 2a and 2b).

II. 4. 2 *Hierarchical island model*

For this model, the highest likelihood was observed between K = 13 and K = 17 but the modal value of L(K)'' was at K = 5, which corresponds to the number of archipelagoes. Using L(K)'', we observed that STRUCTURE always found the modal value to be K = 5 even when we omitted one island in each of the archipelagoes (Fig. 2c, d, e and f).
The height of the modal value increased with the intensity of sampling and the number of loci typed, as expected. When comparing AFLPs and microsatellites datasets for the same sampling intensity, we found the height of the modal value higher for AFLPs than for microsatellites, an indication that the signal was stronger in the former.
In order to detect sub-structuring within archipelagoes, we used the best assignment of individuals to one of the five groups to define five subgroups. Each of this sub-group was subsequently analyzed with STRUCTURE to detect number of sub-groups in each cluster. We did not apply this method to all the sub-sets of data but for the 3 sub-sets we tested, we always found the modal value of L(K)'' to be K = 4, which corresponds to the number of populations within each sub-set.

II. 4. 3 *Contact zone*

Across replicates of the different scenarios, the maximum likelihood for the number of populations oscillated between 8 and 11. However, the modal value of L(K)'' was K = 2 for all replicates, using either full or

partial datasets with 20 individuals out of 100 and 5 microsatellite loci or 50 AFLP loci only (Fig. 2g and 2h). K = 2 corresponds to the highest level of structuring in the model, as the 10 demes were partitioned into 2 sets of 5 populations by a "contact zone" of restricted gene flow. Similarly to the hierarchical island model, a division of the data set in two groups corresponding to the best assignment of individual to groups made by STRUCTURE and a subsequent analysis of each sub-set detected five populations in each subset.
Sub sampling of individuals or loci reduced the height of the modal value of L(K)'' (Fig. 2g and 2h). And ten AFLPs produced a stronger signal than one microsatellite since the average magnitude of the height of the modal value of L(K)'' is 1.5 higher for the former.

II . 5 DISCUSSION

Our goal in these simulations was to confront the algorithm underlying the program STRUCTURE with populations organized less simply than the standard island model. We emphasize here that our purpose was not to test the quality of the assignment of individuals to groups, as this has been done (for simpler population structure) by others (e.g.: Manel *et al.*, 2002; Rosenberg *et al.*, 2001). We showed that while the maximum likelihood estimate for the number of groups given by STRUCTURE often does not correspond to the real number, the second derivative of the likelihood function with respect to K has a mode at the true K for all the situations investigated. We further found that the algorithm underlying STRUCTURE detects the highest level of population structure, and also that subgroups created by the best individual assignment produced by STRUCTURE permits to identify sub-levels of structuring. We restricted our simulations to cases of moderate to strong structure at different hierarchical levels because our goal was to test the ability of the algorithm to detect the number of groups of individuals in situations when different layers of population structure exist, as is often the case in real situations. Limited simulations for the hierarchical island model with a higher migration rate equal to 0.004 between archipelagoes ($F_{ARCHIPELAGO-TOTAL}$ = 0.17) and 0.02 within archipelagoes ($F_{ISLAND-ARCHIPELAGO}$ = 0.14) still detected the (correct) number of archipelagoes. The correct number of archipelagoes was also detected with a migration rate of 0.08 between archipelagoes ($F_{ARCHIPELAGO-TOTAL}$= 0.038) and 0.02 within ($F_{ISLAND-ARCHIPELAGO}$= 0.035), but only with the genetic information from 100 AFLPs.

As might be expected, we found that the intensity of sampling both of individuals and markers plays a role in the correct detection of the number of groups. Among the types of markers commonly used for population structure detection, it seems that AFLPs perform at least as well as microsatellites. This is unexpected as AFLPs coding implies a number of missing observations, absent from microsatellite datasets. Using the height of the modal value of L(K)'' as an indicator of the ability of STRUCTURE to detect the real number of groups, it seems that roughly five AFLP bands give the same result as one microsatellite locus. This ratio is lower than the one advocated by Mariette *et al.* (2002) for analysis of gene diversity.

STRUCTURE was still able to detect the real number of groups with 5 microsatellites or 50 AFLPs. However, for the three models we simulated, the intensity of the signal detected with 5 microsatellites or 50 AFLP loci was always at least twice lower than when the full set of loci was considered. For the microsatellites datasets with 5 loci, the signal was the weakest and thus we suggest a minimum of ten loci is necessary to insure the detection of the correct number of groups by STRUCTURE.
We also tested the ability of STRUCTURE to detect the correct number of groups with a realistic sample size of 20 individuals per deme and it performed quite well in this way. For the hierarchical island model, the signal was decreased by the same order of magnitude as when we divided the number of loci by two. For the contact zone and the island model, the effect of the reduction in the number of individuals was stronger than the effect caused by fewer loci. This can be explained by the intensity of sampling: in these models only 20% of the individuals in each population were sampled, while 40% were sampled in the hierarchical island model. In the case of the partial sampling including 15 demes out of 20 in the hierarchical island model (3 out of 4 demes on each island) STRUCTURE still detected a strong signal at K = 5. The height of the modal value of L(K)'' only decreased by about 25% in comparison with the full hierarchical island model. While the exhaustive sampling of all potential sources of migrants is crucial if one wants to investigate the comprehensive pattern of migration and structure in an area, our results indicate that the program still works with missing sources, given the level of structure we simulated.
Finally, it must be emphasized that while our simulations provide some indications as to how the STRUCTURE's algorithm reacts to limited sampling, a much more thorough investigation remains to be done. Similarly, the ability of STRUCTURE to detect clusters of individuals at different levels when dispersal among the clusters is more intense is not

clear. However, Rosenberg et al (2002) showed empirically on a very large microsatellite data set (377 loci) encompassing 1'026 individuals from the 5 continents that humans cluster in 5 groups, loosely corresponding to the 5 continents. They obtain these results despite the notoriously weak genetic differentiation among humans populations (F_{ST} among continents around 5%, and lower between populations within continents). Obviously, few non-human species could be genotyped with such intensity, but this study indicates that detection of the correct number of clusters can still be found when differentiation is weaker than in our simulations.

In conclusion, we showed that STRUCTURE is not only able to detect the structure of data-sets simulated according to an island model but performs also very well when confronted with more complex hierarchical migration schemes. In such situations, the highest hierarchical level of population structure is detected. Subsequent analyses of subsets defined by the best assignment of individuals to groups provided by the program allow finding the hidden within-group structure. Importantly, we showed that the real number of groups is best detected by the modal value of the second derivative with respect to K of the likelihood function. Using only the maximum likelihood is generally misleading. While STRUCTURE is not profiled to analyze data from dominant markers, our simulations show that AFLPs can give results at least as accurate as microsatellites.

II . 6 ACKNOWLEDGEMENTS

We thank M. Schupbach for his help with the computer partitioning of simulations, and N. Juillet, S. Trouvé, J. Jaquiéry and S. Manel for comments on the manuscript. J.G. was supported by the Swiss NSF (grants no: 31-59326.99 (PhD of G.E.), 31-55945.98 and 31-068325.02).

II . 7 REFERENCES

Arnaud Jo, Viard F, Delescluse M and Cuguen Jo (2003) Evidence for gene flow via seed dispersal from crop to wild relatives in *Beta vulgaris* (Chenopodiaceae): consequences for the release of genetically

modified crop species with weedy lineages. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 1565-1571.

Balloux F (2001) EASYPOP (version 1.7): A computer program for population genetics simulations. *Journal of Heredity* **92**, 301-302.

Banks MA and Eichert W (2000) Computer note. WHICHRUN (version 3.2): a computer program for population assignment of individuals based on multilocus genotype data. *Journal of Heredity* **91**, 87-89.

Beaumont M, Barratt EM, Gottelli D*, et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology* **10**, 319-336.

Berry O, Tocher MD and Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology* **13**, 551-561.

Bouzat JL and Johnson K (2004) Genetic structure among closely spaced leks on a peripheral population of lesser prairie-chickens. *Molecular Ecology* **13**, 499-505.

Caizergues A, Bernard-Laurent A, Brenot JF, Ellison L and Rasplus JY (2003) Population genetic structure of rock ptarmigan Lagopus mutus in Northern and Western Europe. *Molecular Ecology* **12**, 2267-2274.

Cegelski CC, Waits LP and Anderson NJ (2003) Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Molecular Ecology* **12**, 2907-2918.

Chapuisat M, Goudet J and Keller L (1997) Microsatellites reveal high population viscosity and limited dispersal in the ant *Formica paralugubris*. *Evolution* **51**, 475-482.

Ciofi C, Milinkovitch MC, Gibbs JP, Caccone A and Powell JR (2002) Microsatellite analysis of genetic divergence among populations of giant Galapagos tortoises. *Molecular Ecology* **11**, 2265-2283.

Corander J, Waldmann P and Sillanpaa MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics* **163**, 367-374.

Cornuet JM, Piry S, Luikart G, Estoup A and Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989-2000.

Dawson KJo and Belkhir K (2001) A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genetical Research* **78**, 59-77.

Falush D, Stephens M and Pritchard JK (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Gaudeul M, Till-Bottraud I, Barjon F and Manel S (2004) Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae):

comparison of AFLP and microsatellite markers. *Heredity* **92**, 508-518.

Giles BE, Lundqvist E and Goudet Jo (1998) Restricted gene flow and subpopulation differentiation in *Silene dioica*. *Heredity* **80**, 715-723.

Goossens B, Funk SM, Vidal C*, et al.* (2002) Measuring genetic diversity in translocation programmes: principles and application to a chimpanzee release project. *Animal Conservation* **5**, 225-236.

Hampton JO, Spencer PBS, Alpers DL*, et al.* (2004) Molecular techniques, wildlife management and the importance of genetic population structure and dispersal: a case study with feral pigs. *Journal of Applied Ecology* **41**, 735-743.

Jarne P and Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* **11**.

Kimura M and Weiss GH (1964) Stepping stone model of population structure + decrease of genetic correlation with distance. *Genetics* **49**, 561.

Lugon Moulin N, Wyttenbach A, Brünner H, Goudet J and Hausser J (1997) Study of gene flow through a hybrid zone in the common shrew (*Sorex araneus*) using microsatellites. *Hereditas* **125**, 159-168.

Luikart G, England PR, Tallmon D, Jordan S and Taberlet P (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics* **4**, 981-994.

Manel S, Berthier P and Luikart G (2002) Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. *Conservation Biology* **16**, 650-659.

Mariette S, Cottrell Jo, Csaikl UM*, et al.* (2002) Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed Q-petraea (MATT.)LIEBL. and Q-robur L. stands. *Silvae Genetica* **51**, 72-79.

Mariette S, Le Corre V, Austerlitz F and Kremer A (2002) Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Molecular Ecology* **11**, 1145-1156.

Paetkau D, Calvert W, Stirling I and Strobeck C (1995) Microsatellite analysis of population structure in Canadian polars bears. *Molecular Ecology* **4**, 347-354.

Petit RJ, Bialozyt R, Brewer S, Cheddadi R and Comps B (2001) From spatial patterns of genetic diversity to postglacial migration processes in forest trees. In: *Integrating Ecology and Evolution in a Spatial Context. 14th Special Symposium of the British Ecological Society* (eds. Silvertown J, Antonovics J), p. 295. Blackwell Publishing, Oxford.

Pritchard JK and Donnelly P (2001) Case-control studies of association in structured or admixed populations. *Theoretical Population Biology* **60**, 227-237.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Pritchard JK and Wen W (2003) Documentation for *structure* software: Version 2. Available from http://prtich.bsd.uchicago.edu.

Randi E and Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics* **3**, 31-45.

Rannala B and Mountain Jo (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9197-9201.

Rosenberg NA, Burke T, Elo K, *et al.* (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* **159**, 699-713.

Rosenberg NA, Pritchard JK, Weber JL, *et al.* (2002) Genetic structure of human populations. *Science* **298**, 2381-2385.

Saitou N and Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology Evolution* **4**, 406-425.

Satten GA, Flanders WD and Yang QH (2001) Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model. *American Journal of Human Genetics* **68**, 466-477.

Slatkin M and Voelm L (1991) Fst in hierarchical island model. *Genetics* **127**, 627-629.

Trouve S, Degen L and Goudet J (2004) Ecological components and evolution of selfing in the freshwater snail *Galba truncatula*. *Journal of Evolutionary Biology* **17**, In press.

Turakulov R and Easteal S (2003) Number of SNPS loci needed to detect population structure. *Human Heredity* **55**, 37-45.

Vernesi C, Crestanello B, Pecchioli E, *et al.* (2003) The genetic impact of demographic decline and reintroduction in the wild boar (*Sus scrofa*): A microsatellite analysis. *Molecular Ecology* **12**, 585-595.

Vos P, Hogers R, Bleeker M, *et al.* (1995) Aflp - a New Technique for DNA-Fingerprinting. *Nucleic Acids Research* **23**, 4407-4414.

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.

Zeisset I and Beebee TJC (2003) Population genetics of a successful invader: the marsh frog Rana ridibunda in Britain. *Molecular Ecology* **12**, 639-646.

III ABSENCE OF DEME DIFFERENTIATION OF CAPERCAILLIE POPULATIONS IN THE SWISS AND FRENCH JURA, IN SPITE OF HABITAT FRAGMENTATION

Sébastien Regnaut, Sébastien Sachot* and Luca Fumagalli

III . 1 ABSTRACT

Capercaillie populations in Europe have been steadily declining since the 1980's. The causes of this decline are mainly related to habitat factors. To asses the genetic status of the remaining populations in the French and Swiss Jura, we have adapted recent methods of non-invasive sampling and conservation genetics. Hundreds of stools samples have been collected over the area and a subset of these have been genotyped, using microsatellites, to perform a large-scale population genetic study. The populations in the Jura, as originally described into landscape-defined patches, are heterogeneous in density and the forest is fragmented by lakes and humanized valleys. These geographical patches show moderate and non-significant genetic differentiation (F_{ST} between 0.003 and 0.28, N.S.). All patches are individually in mutation-drift equilibrium, as is the population taken as a whole (F_{IS}=-0.029, N.S.). In the light of these findings, it appears that the population decline has not left genetic traces as yet. We suggest considering this population as a continuum of individuals dispersing in a viscous environment, viscosity due to partial natal phylopatry and limited dispersal distance.

Management policies should therefore focus on habitat quality as potential immigrants from the Alps are susceptible to produce gene flow, and as the local Jura population has not attained threshold population size that is critical for population survival and do not show immediate risk of demographic decline caused by mutational meltdown or inbreeding depression.

III . 2 INTRODUCTION

Genetic processes in declining populations can induce the vortex of sudden population extinction: via mutational meltdown, the accumulation of deleterious mutations induces population size reduction, which in turn causes inbreeding depression and increases genetic load by drift (Lynch *et al.* 1990). Small isolated populations, or populations on the edges of the distribution range of a species, potentially show loss of genetic adaptability (Frankham *et al.* 2002). In principle, conservation genetics aims to identify and monitor these genetic processes in threatened populations to reinforce ongoing demographic surveys and help population management policies.

Capercaillie (*Tetrao urogallus* L.; Aves; Galliforms; Tetraoninae) is a flagship species of the European Mountain wildlife. Due to the declining status of all the populations in Europe (Moss *et al.*, 2000; Storch, 2000), it has also become a flagship species of biodiversity conservation. A recent genetic study demonstrated a gradual increase in genetic isolation with loss of habitat connectivity between 14 capercaillie populations throughout Europe (Segelbacher *et al.*, 2003). In another study, it was also suggested that, at local scale, habitat fragmentation related to human land use has increased loss of genetic diversity in isolated Alpine populations (Segelbacher & Storch, 2002).

Similarly to these "edge" Alpine populations, the populations in the Jura Mountains in France and Switzerland are presumably reproductively isolated from the closest populations (Alps or Black Forest). The census size is thought to be close to 600 individuals, with a balanced sex-ratio, over an area of ca. 5000 km² (Sachot, 2002). The population has been declining until the 1980's since when it fluctuates according to direct counts in France or shows a declining pattern based on reproductive success estimations and on counts on leks in Switzerland (Leclercq, 2004; Sachot *et al.*, 2002).

Suitable habitat is not continuous, as the forest is fragmented by lakes and urbanised valleys, leading to potential genetic fragmentation effects resembling a metapopulation (Sachot, 2002). To evaluate the extinction probabilities of the Jura population, a Population Viability Analysis was conducted on the Swiss side of the border that runs along the mountain range (Sachot, 2002). This study produced a fine-scale map of habitat suitability, and concluded that the connectivity between suitable patches

was important for population conservation. To complement the latter study, we carried out a population genetic analysis to evaluate whether any barriers to gene flow exist among patches of suitable habitat. Information regarding the extent of gene flow, levels of genetic diversity and the potential for inbreeding are critical aspects in identifying whether genetic issues could affect the population's long-term survival. If the causes of capercaillie population decline have genetic basis, decisions concerning genetic restoration (captive breeding programs, reintroduction plans, translocation from other populations) will be advisable as much as habitat restoration policies.

Lekking species, such as the capercaillie, are sensitive to disturbance caused by invasive sampling. Non-invasive sampling has therefore provided tools to study these cases (Kohn & Wayne, 1997). To date, hair (Taberlet *et al.*, 1993), feather (Leeton *et al.*, 1993) and even urine (Hayakawa & Takenaka, 1999) or stool (Kohn *et al.*, 1995) have been used as DNA source for PCR based molecular genetic analysis and surveys. Faecal DNA is a source that has proven useful in other types of genetic studies when appropriate methods and protocols adapted to low quantity DNA are applied (Idaghdour *et al.*, 2003; Kohn *et al.*, 1995; Taberlet *et al.*, 1997). We therefore collected faeces left on lekking and feeding sites across the studied area (French and Swiss Jura).

This study has three principal aims: first, to describe the population genetic structure of capercaillie in the Jura, and investigate whether the habitat patches lead to genetic isolation at a local scale. Second, to estimate whether some patches are isolated from the whole population and if recent barriers to gene flow are defined in the landscape. Third, if we detect genetic isolation, we will assess whether it represents a threat to the population's survival in the middle or short range.

III. 3. 1 Sampling and study area

The study covers an area between Chamfromier in France and Grandson in Switzerland (VD) that is 100 km long and up to 20 km wide (Fig. 1). The area has been surveyed for over 20 years by the *Groupe Tétras Jura*, the *Réserve naturelle de la Haute Chaîne du Jura* and the *Service de la conservation de la faune du canton de Vaud*. Several hundreds of stool samples were collected from the field in 1999 and 2000 with the help of many volunteers throughout the Moutain ranges of Grand Risoux, Mont Tendre, Mont Salâ, Crêt de la Neuve in Switzerland, and the ranges of Massacre, Risoux, Haute-Chaîne du Jura and Chamfromier in France.
The Jura is a mid-to-high mountain range covered with a mixed forest of fir, spruce, beech and Mountain ash. Preferred habitats selected are open and grazed areas of spruce, with few fir and very low understorey cover. The highest point in the study area is the *Crêt de la Neige* (1,723 m). All samples were collected between 1200 and 1650m. We sampled exhaustively in 11 patches described by Sachot *et al* (2002).

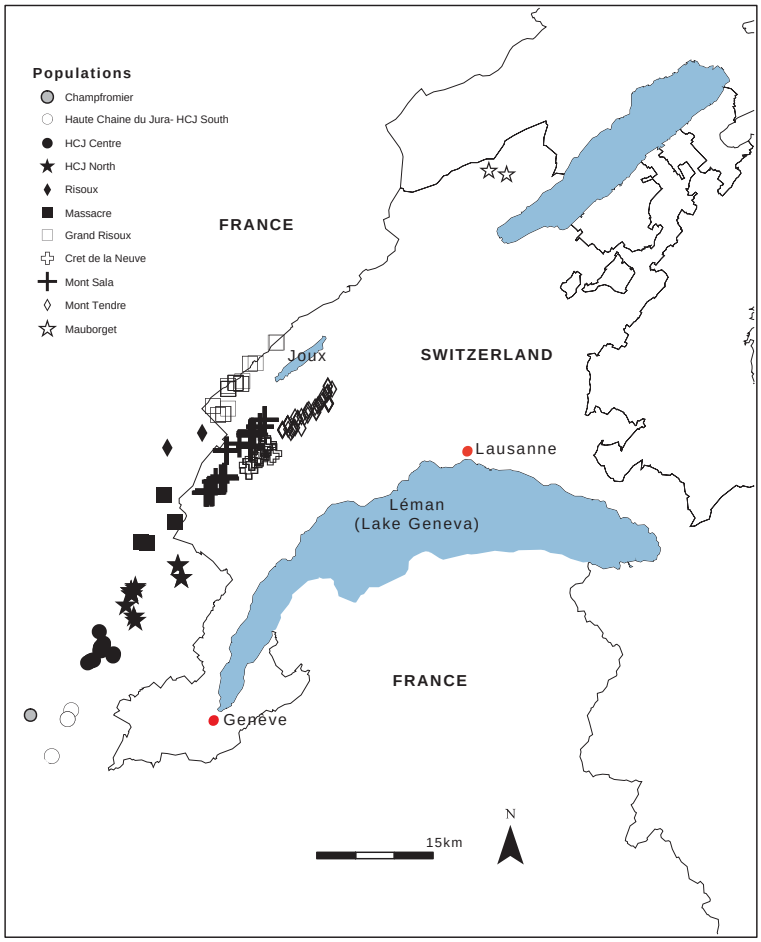


Fig. 1: Sampling area and populations according to Sachot (2002)

III. 3. 2 Extraction and amplification

Stool samples were dried at room temperature for 3 days and then stored in individual vials together with silica-gel beads to maintain dryness (Frantzen *et al.*, 1998). We extracted DNA from ca. 200 mg of dried stool using the commercial kit specifically designed for stool samples (Qiagen) in a special room dedicated to the preparation of scat samples. For these extractions, we followed the protocols provided by the company, except that the elution required 375µl of elution buffer.

Sex was primarily estimated using stool diameter (Leclercq, 1987), however, this method does not distinguish females and juveniles (females or males). Therefore, sex was subsequently confirmed by PCR (*see Chap. I Feasibility*) of *CHD* gene introns located on the heterochromosomes (W & Z, see (Fridolfsson & Ellegren, 1999)). Sex determination amplifications were carried out in 25 µl mix of Perkin Elmer Gold Taq, 1 X PCR reaction buffer, 2.5 mM $MgCl_2$, 0.15 $mg.ml^{-1}$ bovine serum albumin, 0.5 to 1 unit Gold Taq (Applied Biosystems), 5 µl template DNA and 2X 7.5 to 15 pmoles of each primer. The reactions were performed in PE9700

thermocycler with the program as follows: initial denaturation at 96°C for 2 minutes then, 10 touch down cycles (from 60 to50°C) with the same duration as following cycles, then 40 times the following cycle: Denaturation at 94°C, annealing at 53°C and elongation at 72°C, all of 45" long. The cycle ends by a terminal elongation of 300" at 72°C. Amplified fragments were visualized on 1.5% agarose gel stained with ethidium bromide and sex was determined based on the presence of two bands in females (W Z) and one band in males (Z Z)

We used 12 microsatellite markers, 8 of which were designed for capercaillie: TUT1, TUT2, TUT3, TUT4, TUD1, TUD3, TUD5, TUD6 (Segelbacher *et al.*, 2000) and 4 designed for black grouse (*T. tetrix*) TTD1, TTD2, TTD6, TTT1 (Caizergues *et al.*, 2001). All amplifications were carried out with the same reagent concentrations as for sex determination without touchdown, at an annealing temperature of 59°C and with 50 PCR cycles rather than 40. Amplification products were combined into 2 sets of 6 multiplexed primer sets and on 4.5% Sequagel-4.75 (national instruments) mix plates, and electrophorised in an ABI377 automated sequencer.

To avoid allelic dropout and false allele effects due to poor DNA quality, a multitube approach genotyping (Taberlet *et al.*, 1996) was performed over the whole sample set *(see Chap. I Feasability)*. All genotypes were firstly repeated twice. When the two repeats were not identical or were homozygous, we performed three further repeats. Necessary repeatability assessment was conducted in the course of this study by testing multiple repeats per samples to measure allelic drop out (ADO) and false allele (FA) rates (Taberlet *et al.*, 1996). These measurements were used to estimate genotyping reliability with the software Gemini (Valière *et al.*, 2002). The probability of identity between genotypes was estimated using the methods implemented in the software Gimlet (Valière, 2002). Furthermore, two references of liver DNA extracts were run on each electrophoresis gel to insure concordance between runs.

III. 3. 3 Statistical analyses

The microsatellite genotypes were tested for Hardy-Weinberg (H.W.) equilibrium within all patches at each locus using the randomization procedures available in FSTAT (Goudet, 1995-2001). Presence of null alleles in the microsatellite loci was tested using the software Micro-Checker (Van-Oosterhout & Shipley, 2003). We used the Software FSTAT to estimate allele frequencies and to test genetic substructure according to the F statistics and pairwise F_{ST} methods of Weir & Cockerham (Weir & Cockerham, 1984). We investigated the geographical distribution of

microsatellite diversity at two levels: within the Jura Mountains, then among favourable habitat patches.

Genetic structure within the Jura Moutains

First, we conducted an *a-posteriori* analysis using the software *Structure* (Pritchard *et al.*, 2000). We investigated substructures from K=2 to Kmax depending on situation, with 20 replicates under the Admixture model. Second, we described the distribution of microsatellite genotypes in relation to favourable habitat patches using a principal component analysis (PCA) on individuals genotypes. This PCA was conducted with the software PCAgen (available from http://www2.unil.ch/izea/softwares/pcagen.html) and the spatial arrangement of the individuals was compared with the geographical origin of the sample. As an outgroup, we also included in this analysis 33 samples taken from a distant population (Pyrenees, France) and genotyped the same way as all the samples from the Jura.

Genetic structure among patches:

We investigated the genetic structure among patches by testing the relation between geographical distance and pairwise F_{ST} between the 11 patches (see map Fig. 1) using the Mantel tests procedure in FSTAT. To calculate the distance between two patches, we considered the geographical centre of the area covered by the samples in each patch. The centroïd of a patch was defined as the average of the coordinates of all samples in their patch. In continuous populations with limited dispersal ranges, as it is possibly the case in the capercaillie populations in the Jura, we expect the genetic relatedness between individuals to decrease with geographical distance.

Testing recent demographic bottlenecks

As the populations of interest are known for undergoing significant demographic declines, we investigated whether a recent bottleneck had affected the pattern of genetic diversity. This test was performed using *Bottleneck* (Cornuet & Luikart, 1996). Observed distribution of allele frequencies were tested for their expected L-shape under mutation-drift evolution.

Sexed biased dispersal

To asses the effect of sex biased dispersal on our data, we tested whether structure of the two sexes was different across the patches where data was available (2 patches were sampled only for males and were therefore removed from this analysis). *T. urogallus* populations contain mixed age

cohorts, but the low reproduction success and the great number of adults sampled allows us to assume that the samples contains a large majority of post-dispersal adults. We expect differentiation (F_{ST}) and assignment index (A.I.) to be lower in the most dispersing sex (females), whereas inbreeding index (F_{IS}) is expected to be higher in females (Goudet *et al.*, 2002).

Finally, geographical distance to the Alps and demographic parameters (patch size, carrying capacity) were compared to genetic diversity indexes using Pearson or Spearman rank correlations in the statistical package R (Vers. 1.6.1).

III . 4 RESULTS

III. 4. 1 Genotyping

Twelve loci were successfully amplified. Out of 283 samples processed, we found 238 unique genotypes. Since the probability of identity between siblings genotypes was low (P.I.(sibs)= 5.10^{-13}), we considered each unique genotype as a distinct individual. Most of the duplicated genotypes (41) were collected at the same location, at the same date and by the same collector and were therefore discarded from the data set, assuming they were samples of the same individual. The remaining two pairs of duplicated samples were collected 10 and 20 km from each other in 1999 and 2000, indicating possible long-range dispersal. Furthermore, 109 samples had from 1 to 6 missing genotypes (45% of samples and for a total of 8.5% missing values). The sex composition of the data is unbalanced: 70 (30%) females, 154 (65%) males, and 14 (5%) samples for which sex remained unknown.

III. 4. 2 Global Structure at the population level

Across the 238 genotypes, the observed heterozygosity (0.559) is lower than expected under Hardy-Weinberg's equilibrium expectation (0.594). An excess of homozygotes can be expected under several circumstances: Sub-structure in the data set (Wahlund effect), inbreeding and non-random mating, but also if some alleles are not amplified due to technical reasons (null alleles). The first two reasons would effect equally all loci, leading to a trend consistent in the Fis values across loci. The latter reason (null alleles) would only effect the Fis for the loci that are concerned by null alleles, leading to inconsistency across loci. Four loci (TTD2, TUT1, TUD5, and TUD6) diverge significantly from H.W. expectation (see Table 1). Three of these exhibited null alleles according to the method implemented in the software Micro-Checker (Van-Oosterhout & Shipley, 2003) (see Table 1: TTD2, TUT1, TUD6). Excluding these three loci, assuming they contained null allelels, the allele frequencies did not diverge from H.W. expectancies (F_{IS}=-0.03, N.S.).

Table 1: Diversity indexes and Null alleles detection.
A=number of alleles; Ho: Observed het. under H.W. equilibrium; Hs: mean expected het. Over patches; Ht: global expected het.; F_{IS} = 1-(Ho/Hs) inbreeding index. *refers to

significance at the 5% level; R: Allelic richness; H: Gene diversity (Hs corrected for sample size); Null alleles: presence of null alleles detected according to Van-Oosterhout. Null allele frequency estimator disregarding non-amplified data (Chakraborty) or considering missing null homozygous (Brookfield 2)

Locus	A	Ho	Hs	Ht	Fis	R	H	Null Allele	Chakraborty	Brookfield 2
TTTD6	13	0.73	0.74	0.78	0.04	12.69	0.76	no	0.02	0.11
TTD2	17	0.63	0.68	0.71	0.12*	17.00	0.73	yes	0.06	0.32
TUT1	11	0.42	0.62	0.66	0.28*	10.43	0.65	yes	0.16	0.26
TUT2	5	0.45	0.50	0.51	0.07	4.69	0.51	no	0.04	0.11
TUD3	8	0.28	0.28	0.29	0.03	7.65	0.31	no	0.01	0.20
TUD5	13	0.72	0.61	0.62	-0.13	12.91	0.63	no	-0.06	0.24
TTTD1	11	0.49	0.48	0.51	0.01	10.85	0.57	no	0.00	0.24
TUT4	9	0.60	0.54	0.55	-0.09	8.67	0.51	no	-0.04	0.10
TUD1	7	0.55	0.56	0.56	0.05	6.95	0.56	no	0.02	0.28
TTTT1	3	0.31	0.36	0.40	0.10	3.00	0.36	no	0.05	0.25
TUT3	6	0.70	0.63	0.70	0.00	5.72	0.67	no	0.00	0.09
TUD6	13	0.71	0.79	0.86	0.16*	13.00	0.86	yes	0.09	0.31
All	9.7	0.55	0.57	0.59	0.06	3.10				

The approach using the program *Structure* (Pritchard *et al.*, 2000), did not indicate a clear structuring of the genetic data comparable to the geographical distribution of the Jura populations. Testing from K=2 to K=13, where K is the number of substructures, all assignment matrices showed that almost all individuals are admixed. Even if the geographical patches was the not the best descriptor of the genetic diversity distribution, *Structure* would indicate the highest hierarchical structure pattern at work (see chap II).

III. 4. 3 PCA analysis of individual genotypes

The figure 2 shows the spatial organisation of the 238 samples from the Jura with the 33 samples taken from the Pyrenees. It appears that the first component discriminates perfectly the two groups of samples from the Jura and the Pyrenees (12.5% inertia. Pval=0.334). However, the samples from the Jura do not appear to be separated from each other by any of the principal components of this analysis. The gene pool of these samples seems homogenous, as all patches (indicated by the coloured ellipses) are mingled.

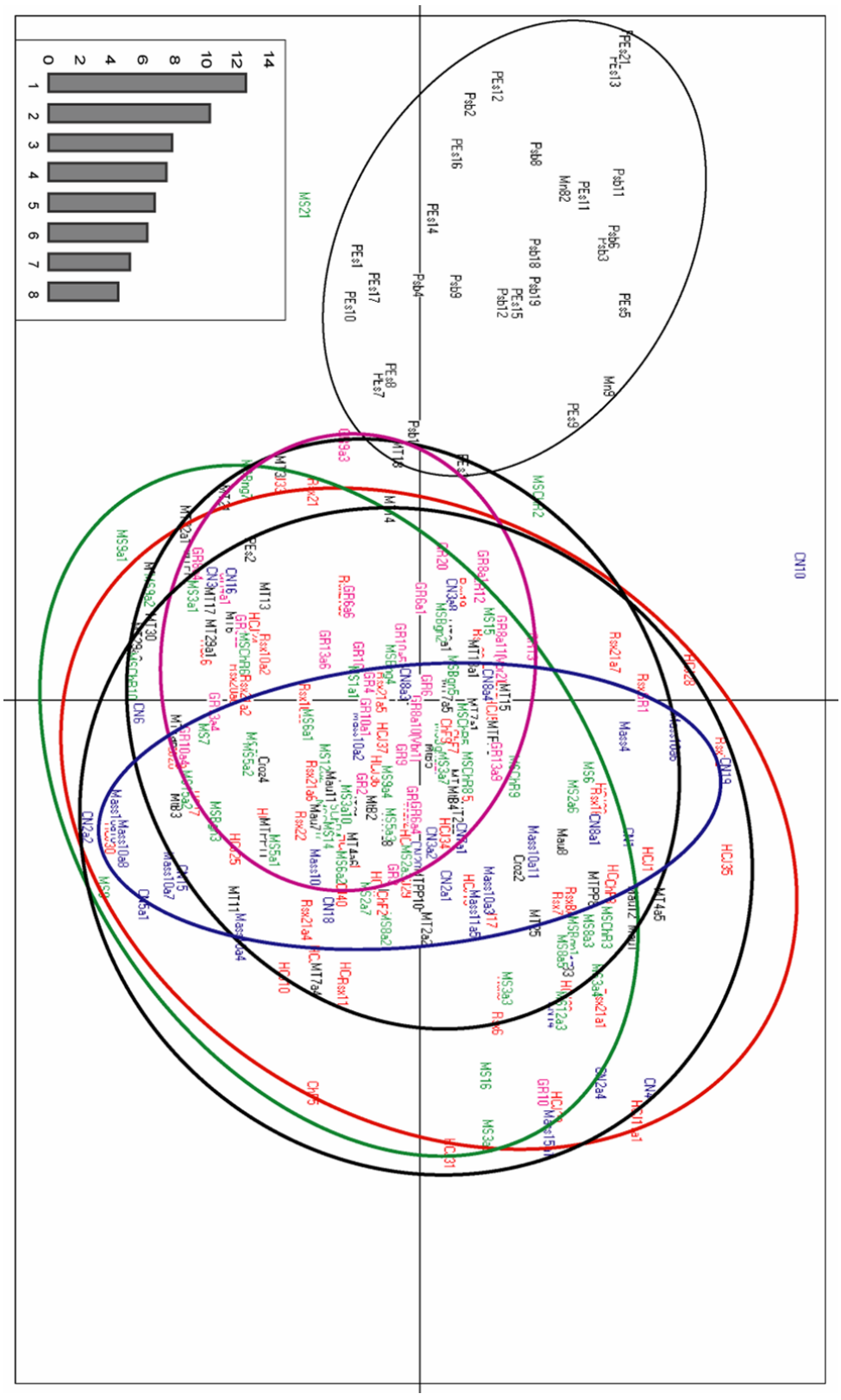


Fig. 2: Principal component analysis (PCA) based on genetic data (nine microsatellite markers) of capercaillie showingnseparation (differentiation) of the population of the Pyrenees (left ellipse) and the Jura (six ellipses on the right, patches are grouped by geographical proximity for display convenience). Lower left corner: histogram of the inertia of the first 8 axis.

III. 4. 4 Genetic Structure between patches

Global F_{ST}, the genetic differentiation among the 11 patches with 11 loci is 0.037. After removing the three loci that have homozygote excess with suspected null alleles, the global F_{ST} remains moderate (F_{ST}=0.036, SE= 0.012 obtained after jackknifing over loci). All patches are in H.W. equilibrium except Mont Sâla (F_{IS}=-0.147, Table 2).

Table 2: Diversity indices by population

F_{IS}: deviation from H.W. expected allele frequencies (significance indicated as *). n: number of samples; K: carrying capacity (data from Sachot 2002); N: initial female abundance of the patch (in Sachot S. 2002). DistAlp: relative distance to the nearest population (Western Alps) as to the closest distance=0; H: Gene diversity averaged over loci; R: Allelic richness averaged over loci.

Patch	Fis	n	K	N(female)	DistAlp	H	sd	R	sd
HCJCen	0.08	14	22	30	74	0.50	(0.16)	2.03	(0.36)
HCJNorth	0.01	17	14	12	85	0.49	(0.17)	2.01	(0.41)
HCJSouth	0.29	7	30	13	98	0.50	(0.31)	1.86	(0.51)
Risoux	0.04	24	14	12	57	0.56	(0.17)	2.21	(0.47)
Massacre	0.01	14	8	7	67	0.48	(0.16)	1.99	(0.41)
Mont Tendre	0.02	45	40	12	42	0.51	(0.19)	2.10	(0.48)
Grand risoux	-0.04	31	117	64	46	0.59	(0.13)	2.28	(0.39)
Crêt de la neuve	-0.03	25	23	3	50	0.54	(0.13)	2.15	(0.34)
Mont Sâla	-0.15*	48	70	25	52	0.54	(0.12)	2.10	(0.33)
Champfromier	-0.14	6	17	1	98	0.55	(0.23)	2.16	(0.51)
Mauborget	-0.04	7	17	11	0	0.43	(0.26)	1.91	(0.65)

Pairwise F_{ST} values between patches range from 0.003 to 0.28 with the highest values found between HCJSouth and other patches, however these values are not significant because of low sample size (n=7) within the HCJSouth patch. Only 16 remain significant after correction for multiple testing and most of these (14) concern two patches, HCJNorth and ChF, which appear differentiated from most other patches (Table 3).

Table 3: 9 loci analysis of pairwise F_{ST}
Pairwise genetic distances between 11 demes inferred from landscape and habitat quality maps (F_{ST}, lower half) and differentiation test (upper half). Stars indicate the level of significance of a G test by genotype permutation. NS: non significant. NA: available data insufficient to perform test.

	HCJCent	HCJNorth	HCJSouth	Risoux	MTendr	Mass	GRisou	CNeuve	MSala	ChF	Maubor
HCJCent		NS	NA	NS	NS	NS	NS	NS	NS	NS	NS
HCJNorth	0.006		NA	**	**	**	**	**	**	**	NS
HCJSouth	0.173	0.220		NA	NA	NA	NA	NA	NA	NA	NA
Risoux	0.015	0.057	0.177		NS	NS	NS	NS	NS	**	NS
MTendr	0.018	0.062	0.172	0.007		**	**	NS	*	NS	NS
Mass	0.038	0.112	0.186	0.014	0.053		NS	NS	NS	**	NS
GRisou	0.033	0.088	0.164	0.002	0.043	0.003		NS	NS	**	NS
CNeuve	0.014	0.074	0.161	0.001	0.035	0.015	0.009		NS	**	NS
MSala	0.031	0.083	0.118	0.010	0.027	0.015	0.010	0.010		*	*
ChF	0.049	0.064	0.111	0.073	0.061	0.122	0.095	0.071	0.078		NS
Maubor	0.024	0.042	0.282	0.023	0.009	0.078	0.056	0.061	0.075	0.089	

III. 4. 5 Correlation between geographic and genetic distances

Geographic distance explains a large proportion of the genetic divergence only within core populations. Geographical distances and pairwise Fst/(1-Fst) are not correlated over the whole sample set, however, removing Mauborget, the most northerly and geographically isolated patch, the relation become significant (Mantel test, $p<0.01$, Fig. 3b). The correlation is still significant when the most southerly populations (HCJ and ChF) are also removed ($p<0.05$, Fig. 3c). The correlation is obscured at long geographical distances because the genetic signal looses in accuracy. Stochastic fluctuation in the correlation at long distances can be due to genealogic history and to the small population sizes.

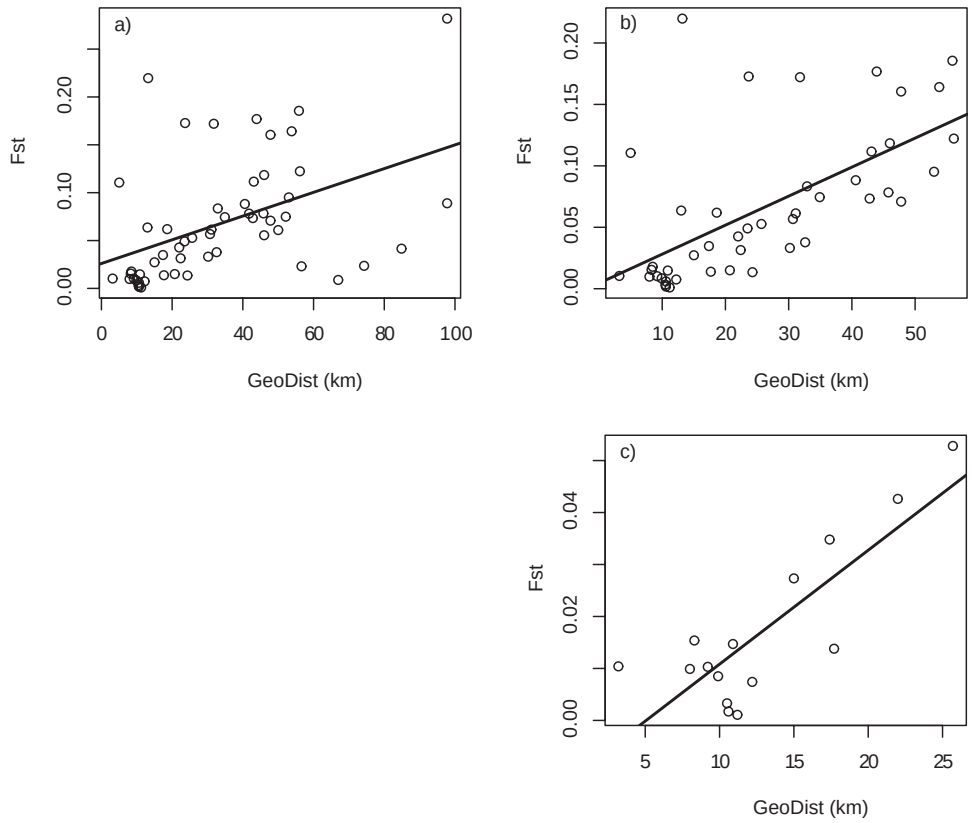


Fig. 3: Correlation of genetic and geographical distances in capercaillie populations in the Jura.
Pairwise genetic differentiation between suitable habitat patches are plotted against geographical distance (in kilometres) and regression line computed from these estimates. (R^2 are given for information as regression statistics normally do not apply to non-indendent sets of data, such as pairwise distances) a) All pops included, over 100 km. R^2=0.21, $p < 0.07$. b) Most remote pop excluded (Mau). R^2=0.39, $p < 0.01$. c) Core pops only: HCJ, ChF and Mau excluded. R^2=0.68, $p < 0.05$. P-values are the significances of the observed correlations as estimated by a Mantel test.

III. 4. 6 Test Male-Female Structure

Mean AI are significantly lower for the females than for the males (one-tailed $p<0.051$), as are the estimates of Fis and F_{ST} but the difference is not significant. Furthermore, variance of AI is significantly higher in females ($p<0.05$; Table 4). These results indicate that among population genetic structuring is higher in males than it is in females. These indirect measurements tend to abide field observations, if only partially, that female disperse more than males.

Table 4: Sex-biased dispersal in 5 of the 8 available populations
N: number of genotyped individuals (total). P-val: one-sided p-value from sex-biased dispersal option in FSTAT. Fis: inbreeding index; Theta: Weir & Cockerham (1984) Fst; Relat: average relatedness within patches compared to the whole population; A.I.: assignment index; Var(A.I.): variance in assignment index.

	Fis	Theta	Relat	A.I.	Var(A.I.)
Males					
N (128)	-0.06	0.02	0.05	0.27	8.08
Females					
N (67)	-0.02	0.01	0.03	-0.52	12.30
p- val	N.S.	N.S.	N.S.	>0.049	>0.022

III. 4. 7 Between seasons differentiation

To determine whether the population structure remained stable during the breeding and non-breeding seasons, we compared structure in 5 patches where samples were available, Risoux, Mont Tendre, Massacre, Grand Risoux, Crêt de la Neuve and Mont Salâ, including 76 samples from the reproduction period and 93 from winter period, and found that there was no difference among the two sub samples (comparison of F_{IS}, Ho and allelic richness non significant between groups. Option in FSTAT).

III. 4. 8 Diversity indexes

None of diversity indexes, gene diversity or allelic richness, nor F_{IS} correlates with any structural characteristic of the patches (distance to closest population, size or density, Table 2).

III. 5 DISCUSSION

In the course of our study, we did not find evidence that genetic factors (geographical genetic structure or inbreeding) has affected the diversity of this population and hinder the result of the efforts of conservation and population management strategies aiming at improving habitat quality.

III. 5. 1 HW, barriers to gene flow and continuum

Our investigations of the distribution of genetic diversity in the Jura population showed that the global structure inferred from F-statistics is moderate, yet significant. However, most pairwise Fst values are not significant. Furthermore, Allele frequencies within patches and within the Jura population as a whole are in H.W. equilibrium. This suggests they are all drawn from a single population. From the well-known biology of the species, we do not suggest that individuals mate randomly over the whole region. Dispersal data gathered by I. Storch (1995) indicate that capercaillies do not fly far from their birth place, and also that they show strong lek fidelity, trend which is also documented in other grouse species (Dunn & Braun, 1985; Schroeder & Robb, 2003).
Furthermore, significant differentiation and patterns of isolation by distance were found at various scales. Isolation by distance has also been documented in other capercaillie populations, both at the European scale, across several hundred kilometres (Segelbacher *et al.*, 2003) and at more local scale, between isolated and differentiated patches within a same mountain range (Segelbacher & Storch, 2002). Furthermore, related grouse species also show such trends (Caizergues, Bernard-Laurent *et al.*, 2003; Caizergues, Ratti *et al.*, 2003). In the Jura population, the pairwise F_{ST} values between patches of suitable habitat are significantly correlated with geographical distance when patches are separated by less than 40 km. This trend was detected in spite of low and non-significant pairwise F_{ST} values. The pairwise F_{ST} values appear very low because few migrants are necessary to homogenize allele frequencies over large distances in a continuous habitat (Hedrick, 2000) and if we look at these patches as a two dimensional stepping stone model, two patches that never exchange migrants do not have to show high F_{ST} values (Castric & Bernatchez, 2004).

Therefore, the homogenous distribution of microsatellite diversity and the trends of isolation by distance detected are consistent with a model of continuous population with close-to-close dispersal of the individuals, with natal fidelity of capercaillie acting as an agent of viscosity. This pattern of dispersion homogenizes allele frequencies at short distance, with the observed increase of pairwise Fst values with geographical distance.

This homogenizing effect of short distance dispersal might be primarily due to female dispersal as females might travel over 20 km within one season, whereas males remain within 1 km around their lek place (Storch 1995). Our microsatellite data shows a trend of biased sex dispersal (Goudet *et al.*, 2002), although not significant, in the expected direction with greater differentiation in males than in females. The low power to detect a significant effect was likely due to the small number of females in our sample, however there is no information on how this test performs when the assumptions of non-overlapping generations are violated.

III. 5. 2 Genetic diversity of capercaillie in the Jura

The Jura population shows a genetic diversity that is in the range of those observed in other capercaillie populations and other grouse species (Table 6). Values of expected heterozygosity in the Jura population are in the same range than the capercaillie populations described in Europe and in the Alps by Segelbacher (2002 and 2003). The latter values are also influenced by the larger population sizes and the differences in mating systems. However, the same microsatellite loci used in this study are also used in Segelbacher and Caizergues studies and values are comparable.

Table 5: Comparison with other grouse studies using microsatellites
The list on the left of the table is the bibliographic source and the species. Nb loci: number of loci used; Ho observed heterozygosity, Hs: mean expected heterozygosity over loci and demes; R: allelic richness; A: mean number of alleles over loci (Allelic diversity). The diversity indexes presented in this study do not diverge from the range of those found in other studies.

Species	nb loci	Ho (SD)	Hs	R	A	Study
capercaillie	12	0.55	0.57	3.1	9.7	Jura (this study)
	9	0.515	0.515	2.15	8.3	
capercaillie	11	0.486	0.549	3.6	7.8	Pyrenees (this thesis)
capercaillie	10	0.56 - 0.78	0.49 - 0.66	2.38	2.7 - 5.7	(Segelbacher et al. 2002)
capercaillie	10	0.44 - 0.72	0.53 - 0.71	2.33	4.43	(Segelbacher et al. 2003)
black grouse (*T. Tetrix*)	13	0.74 - 0.76	0.75 - 0.78		5.4 - 6.3	(Caizergues et al. 2003b)
red grouse (*Lagopus scoticus*)	7	0.81	0.94		18.9	(Piertney et al. 1998)
ptarmigan (*Lagopus mutus)*	6	0.68 - .092	0.77 - 0.91		5.9 - 14.6	(Caizergues et al. 2003a)
chinese grouse (*Bonasa sewerzowi*)	5	0.56 - 0.71	0.52 - 0.74	2.15 - 8.11	3.5 - 8.8	(Larsson et al. 2003)
greater prairie chicken (*Tympanuchus cupido*)	6	0.54 (0.05)	0.68 - 0.76		6.17 (1.4)	(Bouzat et al. 2004)
lesser prairie chicken (*Tympanuchus pallidicinctus*)	6	0.22-0.75	0.55 - 0.86	1.8-5.8	4.33 - 5.83	(Van den Bussche et al. 2003)

III. 5. 3 Loss of diversity with time

Theoretical expectation of loss of diversity in small isolated populations depends on the inverse of the effective population size (Hedrick, 2000). We shall expect effective size (Ne) of ca. 10% of the census size in natural populations (Frankham, 1995), a value also found in a closely related species (Johnson *et al.*, 2004). In the Jura population, the census size is ca. 600, thus Ne would be of ca. 60 individuals.

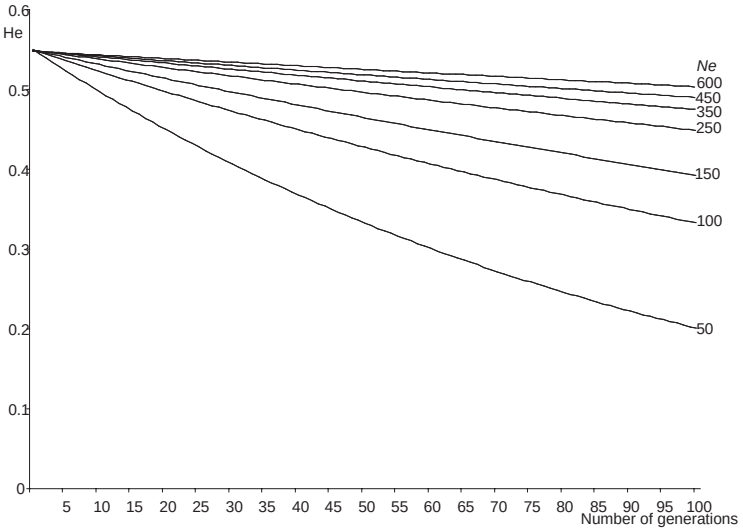


Fig. 4: Expected loss of diversity in a small population.
Expected decrease in heterozygosity in closed, small population, in absence of mutation, generation overlap and migration. He depends on Ne, indicated on the right.

Thus, if the habitat connectivity does not degrade further and if the decline in population is stabilized, the heterozygosity would drop to 0.4 in 35 to 65 generations (which represents between 130 and 260 years)(Fig. 3), disregarding mutation rate, reproduction success skews, assuming no immigration of new alleles, and non-overlapping generations. Nevertheless, this impressive rate of loss has to be moderated by two other features of the biology of capercaillie: i) Overlapping generation allows individuals to participate in several generations of reproduction. ii) The maximal dispersal ability of capercaillie almost certainly exceeds the average dispersal distance (Menoni E. pers. comm.; (Holzinger & Rosler, 1990). Therefore, potentially, immigration from the Alps might lead to gene flow into the Jura Mountains that would help maintain genetic diversity and the "adaptability" of these populations in the long term. Unfortunately, the extent of gene flow between the Alps and the Jura is unknown. Furthermore, populations in the Alps could be undergoing a strong decline and isolation (Segelbacher, 2002). Nevertheless, the gain in gene flow from long-term isolated populations may not always be beneficial. Genetic load theories state that immigration from small isolated population may lead to the introduction of deleterious alleles that have accumulated or fixed (Lynch & Gabriel, 1990). Gene flow of this type would be detrimental to the recipient population by increasing the number of deleterious alleles present.

The data presented in this study provides very little information about dispersal. Only two pairs of identical genotypes were sampled at distant localities and in different seasons. These two individuals, one female, one male, would have travelled respectively 20 km in one year, and 10 km within one year. Furthermore, the Figure 2 shows two genotypes that are clearly isolated from all the other clusters. These two genotypes display more than one allele each that do not appear in any other genotype in the data set. After several re-checks, it appears that these new and discrete alleles are not due to PCR artefacts. The principal component analysis (fig. 2) suggests that these two individuals originate from a distant population, possibly by long distance dispersal from the Alps, the Black forest or the Vosges. Alternatively, they may represent undocumented releases in the Jura.

Finally, the value of effective population size we considered in this study (60) is reported in a PVA of the Scottish population (Marshall & Edwards-Jones, 1998) as a population having 95% probability of surviving over 50 years with an available habitat ten fold smaller than the distribution area of capercaillie in the Jura.

In summary, the patches are not yet differentiated and evidence of strong gene flow within the mountain range is highlighted. This population has retained a large amount of genetic diversity, and still represents a healthy and viable population in regard to genetic factors. The recent observed decline has not left genetic signatures within the Jura population as yet. Many other factors may also cause this decline, for example: human disturbance in winter, poor habitat quality, predation on chicks and eggs by mammals (Baines *et al.*, 2004), and suboptimal weather condition at rearing period (Picozzi *et al.*, 1999). These factors are currently under investigation in the Jura as well and in the Pyrenees and in Scotland (Marshall & Edwards-Jones, 1998; Moss, 2000; Moss *et al.*, 2000). The capercaillie populations in the Jura do not show the genetic symptoms expected previously to imminent mutational meltdown or rapid extinction due to genetic factors.

III . 7 ACKNOWLEDGEMENTS

Many professional and volunteers have participated in this study by gathering and collecting data and samples. We are in debt to the Group Tétras Jura and Dr. B. Leclercq for the help in biology of the species and all the data and samples. To Parc régional de la Haute Chaîne du Jura and A. Block, to the Conservation de la Faune du Canton de Vaud : Dr. C. Neet and ranger B. Reymond. Many thanks also to R. Hammond, L. Handley, P. Fontanillas, C. Parisod, N. Juillet for corrections and discussions.

III . 8 REFERENCES

Baines D, Moss R and Dugan D (2004) Capercaillie breeding success in relation to forest habitat and predator abundance. *Journal of Applied Ecology* **41**, 59-71.

Bouzat JL, Cheng HH, Lewin HA*, et al.* (1998) Genetic evaluation of a demographic bottleneck in the Greater Prairie Chicken. *Conservation Biology* **12**, 836-843.

Caizergues A, Bernard-Laurent A, Brenot JF, Ellison L and Rasplus JY (2003) Population genetic structure of rock ptarmigan Lagopus mutus in Northern and Western Europe. *Molecular Ecology* **12**, 2267-2274.

Caizergues A, Dubois S, Loiseau A, Mondor G and Rasplus JY (2001) Isolation and characterization of microsatellite loci in black grouse (Tetrao tetrix). *Molecular Ecology Notes* **1**, 36-38.

Caizergues A, Ratti O, Helle P*, et al.* (2003) Population genetic structure of male black grouse (*Tetrao tetrix* L.) in fragmented vs. continuous landscapes. *Molecular Ecology* **12**, 2297-2305.

Castric V and Bernatchez L (2004) Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase of genetic differences with distance. *Molecular Ecology* **13**, 1299.

Cornuet JM and Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Dunn PO and Braun CE (1985) Natal dispersal and lek fidelity of Sage Grouse. *Auk* **102**, 621–627.

Frankham R (1995) Effective Population-Size Adult-Population Size Ratios in Wildlife - a Review. *Genetical Research* **66**, 95-107.

Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK and Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428.

Fridolfsson AK and Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**, 116-121.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3), available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet J, 1995. FSTAT (vers. 1.2): a computer program to calculate F-statistics. Journal of Heredity **86**: 485-486.

Goudet J, Perrin N and Wasser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11**, 1103-1114.

Hayakawa S and Takenaka O (1999) Urine as another potential source for template DNA in polymerase chain reaction (PCR). *American Journal of Primatology* **48**, 299-304.

Hedrick PW (2000) *Genetics of populations, 2nd edition* Jones & Bartlett, London.

Holzinger Jo and Rosler M (1990) The occurence of the capercaillie (*Tetrao urogallus- major* CL Brehn, 1831) at the mount Athos (Greece). *Journal of Ornithology* **131**, 95-99.

Idaghdour Y, Broderick D and Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustard. *Conservation Genetics* **4**, 789-792.

Johnson JA, Bellinger MR, Toepfer J and Dunn P (2004) Temporal changes in allele frequencies and low effective population size in greater prairie-chickens. *Molecular Ecology* **13**, 2617-2630.

Kohn MH, Knauer F, Stoffella A, Schröder W and Pääbo S (1995) Conservation genetics of the European brown bear- A study using excremental PCR of nuclear and mitochondrial sequences. *Molecular Ecology* **4**, 95-103.

Kohn MH and Wayne RK (1997) Facts from faeces revisited. *Trends in Ecology and Evolution* **12**, 223-227.

Larsson JoK, Sun YH, Fang Y, Segelbacher G and Höglund Jo (2003) Microsatellite variation in a Chinese grouse *Bonasa sewerzowi* population: signs of genetic impoverishment? *Wildlife Biology* **9**, 261-266.

Leclercq B (1987) *Ecologie et dynamique des populations du grand tétras dans le Jura francais* PhD thesis, Université de Bourgogne.

Leclercq B (2004) Evolution des populations du grand tétras dans le jura Francais. In: *Actes des "4èmes Rencontres Jurassiennes"* (ed. J. B), Mijoux.

Leeton P, Christidis L and Westerman M (1993) Feathers from museum bird skins - A good source of DNA for phylogenetic studies. *Condor* **95**, 465–466.

Lynch M and Gabriel W (1990) Mutation load and the survival of small populations. *Evolution* **44**, 1725-1737.

Marshall K and Edwards-Jones G (1998) Reintroducing capercaillie (Tetrao urogallus) into southern Scotland: identification of minimum viable populations at potential release sites. *Biodiversity and Conservation* **7**, 275-296.

Moss R (2000) Second extinction of capercaillie (*Tetrao urogallus*) in Scotland ? *Conservation Biology* **10**, 255-257.

Moss R, Picozzi N, Summers RW and Baines D (2000) Capercaillie *Tetrao urogallus* in Scotland – demography of a declining population. *The Ibis* **142**, 259.

Picozzi N, Moss R and Kortland K (1999) Diet and survival of capercaillie *Tetrao urogallus* chicks in Scotland. *Wildlife Biology* **5**, 11-23.

Piertney SB, McColl AD, Bacon PJ and Dallas JF (1998) Local genetic structure in the red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite markers. *Molecular Ecology* **7**, 1645-1654.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Sachot S (2002) *PhD thesis : Viability and management of an endangered capercaillie (Tetrao urogallus) metapopulation*, Université de Lausanne.

Sachot S, Leclercq B and Montadert M (2002) Population trends of Capercaillie (*Tetrao urogallus*) in the Jura mountains between 1991 and 1999. *Game and Wildlife Science* **19**, 41–54.

Schroeder MA and Robb LA (2003) Fidelity of greater sage-grouse *Centrocercus urophasianus* to breeding areas in a fragmented landscape. *Wildlife Biology* **9**, 291-299.

Segelbacher G (2002) *PhD thesis*, University of Tübingen, Germany.

Segelbacher G, Hoglund J and Storch I (2003) From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology* **12**, 1773-1780.

Segelbacher G, Paxton RJ, Steinbruck G, Trontelj P and Storch I (2000) Characterization of microsatellites in capercaillie Tetrao urogallus (AVES). *Molecular Ecology* **9**, 1934-1935.

Segelbacher G and Storch I (2002) Capercaillie in the Alps: genetic evidence of metapopulation structure and population decline. *Molecular Ecology* **11**, 1669-1677.

Storch I (1995) Annual home ranges and spacing patterns of capercaillie in central Europe. *Journal of Wildlife Management* **59**, 392-400.

Storch I (2000) Grouse status survey and conservation action plan 2000-2004. In: *WPA-Birdlife-SSC Grouse Specialist Group*. IUCN, Switzerland.

Taberlet P, Camarra JJ, Griffin S*, et al.* (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology* **6**, 869-876.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Taberlet P, Mattock H, Duboispaganon C and Bouvet J (1993) Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. *Molecular Ecology* **2**, 399-403.

Valière N (2002) GIMLET: A computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**, 377-379.

Valière N, Berthier P, Mouchiroud D and Pontier D (2002) Gemini: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83-86.

Van den Bussche RA, Hoofer SR, Weidenfeld DA, Wolfe DH and Sherrod SK (2003) Genetic variation within and among frangemented populations of lesser prairie-chickens (*Tympanuchus pallidicinctus)*. *Molecular Ecology* **12**, 675-683.

Van-Oosterhout C and Shipley P (2003) Micro-checker (ed. http://www.microchecker.hull.ac.uk/MicroCheckForm.jsp).

Weir BS and Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

IV EFFECTS OF BOTTLENECK AND DISTURBANCE ON CAPERCAILLIE (*TETRAO UROGALLUS*) POPULATIONS GENETICS IN THE PYRENEES MOUNTAINS

Sébastien Regnaut, Emmanuel Menoni[1]
and Luca Fumagalli

[1] Office National de la Chasse et de la Faune Sauvage. DER - CNERA Faune de Montagne Station des Pyrénées - R.N. 117 31800 Villeneuve de Rivière- France.

IV . 1 ABSTRACT

The capercaillie populations in Europe have been declining for the last few years. We investigated the potential genetic factors involved in this decline at a local scale within the Pyrenees Mountains by estimating inbreeding and genetic fragmentation using microsatellite data and non-invasive sampling. This study, based uniquely on faeces, showed that the samples within the Pyrenees showed high Fis (0.06) and genetic structure between geographical units is significant (Fst = 0.07).

Genetic diversities differed significantly among localities: places where the forest was re-grown more recently and more recently disturbed showed higher diversity than older forests, and places at the edges of the distribution area showed a decrease of genetic diversity compared to core populations. Furthermore, the distribution of genetic diversity suggests both the presence of barriers to gene flow and metapopulation dynamics due to human activities and forests harvesting. Therefore, we concluded that drift-mutation-dispersion has not been attained within the Pyrenees capercaillie populations after the large bottlenecked that occurred at the beginning of the 20th century. We suggest that the high Fis values are due to Wahlund effects rather than inbreeding and that gene flow is important between localities but that landscape items must hinder dispersion between some of the units. Capercaillie shows a great ability to recover from large bottlenecks, which can be due to its great dispersal capacity. However, we hypothesize that the absence of black grouse in the Pyrenees might have reduced the competition pressure for ecological niche and therefore facilitated capercaillies' recovery from a recent near-to-extinction situation. Finally, we suggest that the long-term protection of numerous and large reserves in the Pyrenees would protect capercaillie and the local flora and fauna from the successive trends of disturbance and destruction of the forest by the cyclic human activity increases.

The Pyrenees Mountains range across approximately 270km east to west, along the border between France and Spain. It is an area rich in flora and fauna and many species are endemic (Maurin, 2003). Great efforts of nature conservation are conducted in this area and it is protected by law. To restore the past diversity that has been eroded by human activities (gaming, agriculture and urbanisation), ongoing plans of reintroduction of threatened species are organised, such as brown bears, marmots or vultures. These plans aim to remind of the importance of conserving the remaining biodiversity, and that the factors causing population declines can be studied, understood and modified.

One of the avian species present in the Pyrenees, the capercaillie (*Tetrao urogallus aquitanicus* L.; Aves; Galliformes; Tetraoninae) was defined as a subspecies by Ingram (1915) on the basis of morphological criteria. Size and colour varied greatly between *T.u. aquitanicus* and all other subspecies of capercaillie (Ingram, 1915). Moreover, its ecological niche is less specific than the species standards as it also occupies ecological niches characteristic of *Tetrao tetrix*, a closely related species with which capercaillie lives in sympatry on all its distribution range but in the Pyrenees (Menoni, 1991). Furthermore, the specificity of the Pyrenean capercaillie was confirmed recently by Segelbacher *et al.* (2003) in a genetic study based on microsatellites. In this European-wide study, the Pyrenees population showed the highest pairwise Fst values with all other populations, confirming its genetic isolation. In the Pyrenees, they also showed the lowest diversity indexes of the European capercaillie populations. The closest populations are located in the Spanish Cantabric mountains, where resides the subspecies *T.u. cantabricus*, and in the Jura mountains, where a remnant population of ca. 600 isolated individuals are found. These two populations, distant of, respectively 300 and 800 km, are very unlikely to provide immigrants into the Pyrenees since the populations became isolated probably during the Middle Age for the Jura, and several hundred years ago for the Cantabric massif.

In this context of ecological and genetic specificities, the population in the Pyrenees appear to be a unique and differentiated gene pool that needs to be protected as part of the Pyrenean patrimony. However, the recent history of capercaillie in the Pyrenees Mountains has been highly disturbed. A very strong decline due to intensive deforestation between

the 17th and 20th century has led to a situation of near extinction (Ingram, 1915). Later, after World War I, spontaneous and active reforestation allowed a rapid expansion of the capercaillie populations, probably from relictual refuges. The census size reached a climax of about 9000 individuals in the 1960's (Menoni, 1991), as the aging forest provided more suitable habitats. Since this time, the population has been showing a steady decline due to increasing human activity (roads, tourism, gaming and poaching, sylviculture) down to ca. 6000 individuals in 2002 (*pers. obs.*). However, the decline observed in the Pyrenees follows a trend described in other populations in Scotland (Moss, 2000), in Switzerland (Sachot *et al.*, 2002) and eastern Europe (Saniga, 2003).
In the context of demographic decline, small populations become threatened by genetic factors such as loss of diversity and increased inbreeding (Nunney & Campbell, 1993). The harsh and heterogeneous landscape of the Pyrenees could contribute to habitat fragmentation, reducing gene flow and increasing local genetic differentiation and drift. Fragmentation is here relevant to take into account especially because recent data indicate that capercaillie dispersal distances are short (Hjeljord *et al.*, 2000; Marshall & Edwards-Jones, 1998) and Menoni, *pers. obs.*).

The present study aims at estimating the present degree of fragmentation within the Pyrenees and at identifying the source of the low diversity. We will also investigate whether the present distribution of genetic diversity is still influenced by the bottleneck that occurred between 1850 and 1915 due to forest degradation. Furthermore, we will assess whether the low microsatellite diversity found in previous studies (Segelbacher *et al.*, 2003) was due to sampling effect or whether this low diversity affects all areas along the distribution of capercaillie in the Alps. Since these animals cannot be captured without affecting the chances of survival, we will use faecal samples as DNA source.

IV . 3 MATERIAL & METHODS

IV. 3. 1 Sampling and study area

Several hundred of stool samples were collected since the 1980's during population surveys on lek sites. The area surveyed spans along the Pyrenees mountain range and the Andorra valley, along ca. 230 km (Fig.

1). The range is divided into Natural Units (N.U.) defined as areas of continuous suitable habitat. N.U.'s are separated from each other by landscape features.
Density of capercaillie individuals and habitat structure are heterogeneous amongst N.U. The population has been monitored for over 30 by local rangers of the Office National de la Chasse (ONCFS) (Catusse, 1993). For the genetic analysis, we selected faecal samples based on their geographic repartition in order to optimise the samples size within each locality.

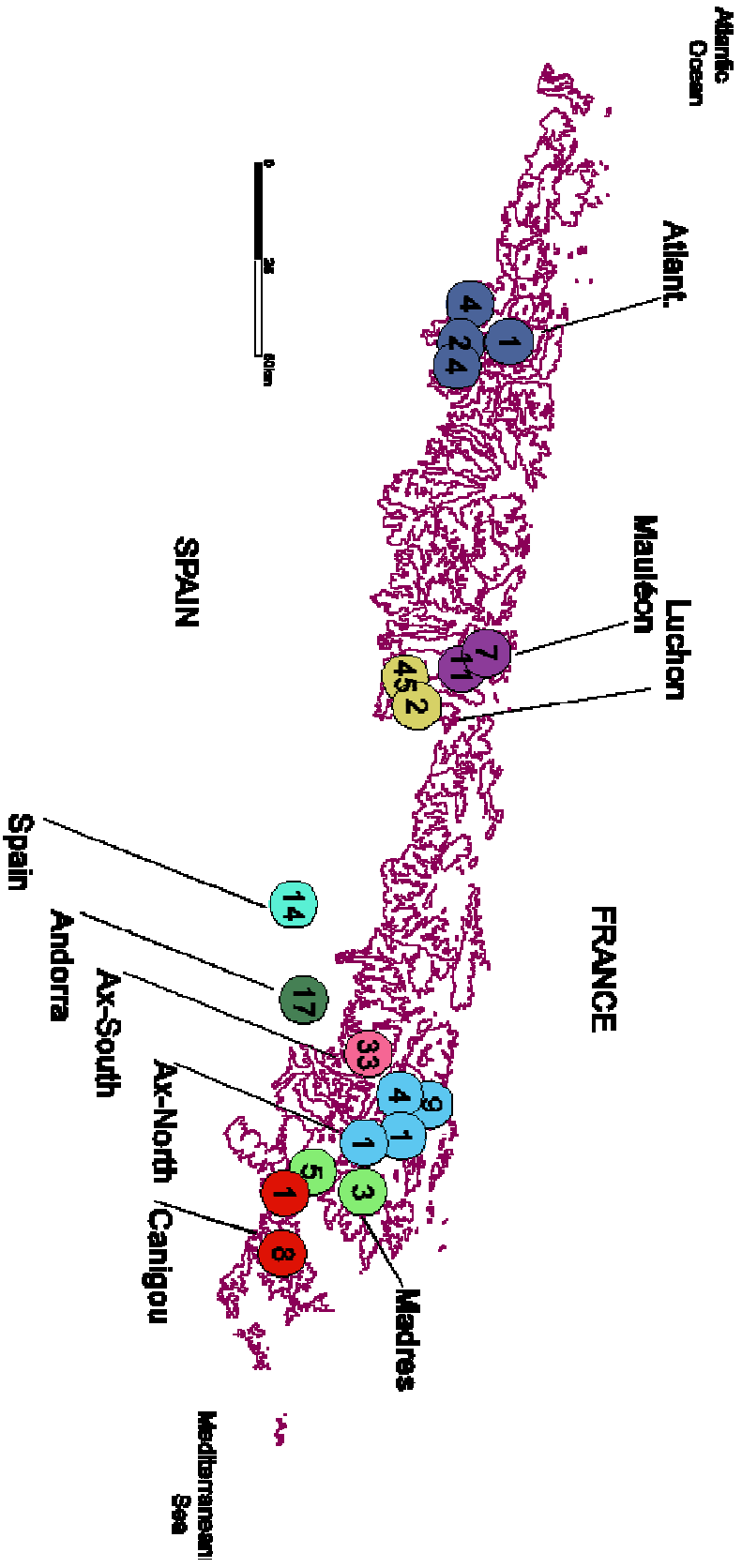


Fig.1: Sampling area
Map of Natural Units division (Source: ONCFS). Circles: N.U. sampled during this study. Numbers: Number of samples per N.U. Sample sites are grouped by geographical proximity.

Among the 202 faecal samples used in this study, 170 were sampled in 17 Natural Units located in France, 12 stool samples and 5 blood samples originated from Andorra, whereas 15 samples came from Spain, in a spread-out valley between Luchon and Andorra (Fig 1). On the French side of the border, most N.U. did not add up enough samples to be considered individually, so that they were grouped together on a geographical basis into 7 demes with sufficient sample size. Furthermore, two of the demes (Luchon and Ax-South) provided enough samples and enough local resolution to be divided into three distinct areas of lek in Luchon (Bareille, Esbas and Superbagnère) or two in Ax-South (plateaus of Beille and Bourbourou). Whereas Bareille and Superbagnère forested areas were relatively conserved during the deforestation period (1850-1920) to protect the city and the local therms, the forest surrounding Esbas was completely razed and only restored between 1900 and 1950. In the Ax-South demes, the case is similar as in Esbas, except that, since the forest recovery, the site of Bourbourou has been particularly protected and Beille was left to large-scale tourism. We suspect that the time lag between the recoveries of different forest patches might have had an effect on the local population dynamics.

IV. 3. 2 Extraction and amplification

All faeces extraction took place in a room only dedicated to low-copy DNA samples, using Qiagen QiaAmp DNA stool kit (Qiagen GmbH, Hilden, Germany) modified to suit our requirements. Amplification protocols are detailed in a previous study (see Section 2 - I). Blood samples were extracted using a QIAamp Tissue Kit (Qiagen) and the manufacturer's recommended protocols. We typed 11 microsatellite loci, seven of which were designed for capercaillie: TUT1, TUT2, TUT3, TUT4, TUD1, TUD3, TUD5 (Segelbacher *et al.*, 2000) and four for a related species TTD1, TTD2, TTD6, TTT1 (Caizergues *et al.*, 2001). All amplification protocols and extraction procedures were adapted to non-invasive samples, performing 2 to 5 repeats for each sample on each primer, according to the guidelines of the Multitube approach (Taberlet *et al.*, 1996). All samples were sexed using a PCR-based method using the CHD gene. We adapted a universal method to sex non ratite bird already published (Fridolfsson & Ellegren, 1999) to our non-invasive samples (see Section2-I).

IV. 3. 3 Statistical analyses

Maps were created using ArcGIS software (ESRI Inc., 231 Redlands, California) with Natural Units data from ONCFS. The geographic distances considered between demes are the geographic centre of the Natural Unit. When demes were composed of several N.U., we considered the average coordinates over all N.U. as centre of the demes. To investigate the genetic structure between demes, heterozygosities, Weir & Cockerham (1984) fixation indexes, and population differentiation tests (G-test) were estimated using the methods implemented in the software FSTAT (Goudet, 1995-2001). We further tested the relation between genetic distance and geographic distance using the Isolation by Distance approach explained by F. Rousset (1997): exploring whether Fst/(1-Fst) correlates or not with the distance or with the logarithm of the distance. The significance tests for these correlations (Mantel tests) were done with the appropriate function in FSTAT.

We analyzed the geographical arrangement of the microsatellite diversity of the patches by a PCA analysis on allele frequencies that we performed with the software PCAgen (available from the web : http://www.unil.ch/dee/page6767_en.html#4).

Concerning the demographic history of the population of interest, we had reasons to suspect that some of the demes could bear detectable traces of more recent bottlenecks than the other demes. We therefore used the approach provided by Cornuet and Luikart in the software Bottleneck (1996) to test for heterozygosity excess (under T.P.M. expectation) and deviation from the expected L-shape distribution of alleles frequencies.

We also tested diversity indexes between these more recently bottlenecked patches versus the others, using the option “comparison among groups” implemented in FSTAT. By this mean, we also tested source-sink patterns between the marginal demes (Atlant., Madres and Canigou) against the central demes.

Finally, the diversity indexes were also used to compare the populations of the Jura with the one sampled in the Pyrenees to estimate the divergence between the Jura and the Pyrenees and to compare it with the results found in Segelbacher *et al.* study (2003).

We successfully amplified 11 loci for 202 samples. Out of these 202 samples, 14 had duplicates and 15 had more than 5 loci that didn't amplify. These 23 samples were therefore removed from further analyses. The 172 unique genotypes had 5% missing data and were composed of 117 males and 56 females. The sex of six of the samples remained unidentified.
The biased sex ratio might be due to the sampling scheme that involved sample collection mostly around leks. Leks aggregate a large portion of males permanently, whereas females do not always remain for a long time on the site, so that their faeces are less frequent.

IV. 4. 1 Differentiation between Jura and Pyrenees

The two populations of the Jura and Pyrenees show a significant Fst of 0.17 (0.09 – 0.26, 95%C.I. bootstrap over loci) and the G-test indicates significant differentiation, using the nine loci common to both studies. Only nine loci are considered in this analysis because three loci were removed from the Jura data set as they were strongly suspected of being impaired with null alleles (see Section 2 - III).
Although allelic richness and expected heterozygosity (diversity indexes) are similar between the two distant populations, observed heterozygosity (Ho) and Fis are significantly higher in the Pyrenees than in the Jura.

Table 1: Differentiation between Jura and the Pyrenees using nine loci

	Jura(n=11)	Pyrenees (n=8)	one-sided p-val
A (Allelic Richness)	2.071	2.084	NS
Ho (Obs. Heterozygotie)	0.542	0.476	0.045*
Hs (Mean Exp. Het.)	0.527	0.504	NS
Fis ((Hs-Ho)/Hs)	-0.029	0.056	0.038*
Fst	0.036	0.071	NS
Relatedness	0.072	0.127	NS

IV. 4. 2 Genetic Structure between patches in the Pyrenees

The inbreeding index, Fis, is significantly different from zero (11 loci, Fis= 0.069; S.E.=0.026), which is probably due to the Wahlund effect as the global Fst=0.07 ($p < 0.001$; S.E.=0.01) is moderate yet significant (Weir & Cockerham, 1984).

Graphically, the principal component analysis of allele frequencies reveals that 57% of the inertia is distributed on the first two axes (Fig. 2). These axes tend to group together the two Ax populations as well as Mauléon and Luchon that are neighbouring demes. The first axis separates the pops accordingly to the geographical arrangement, apart from the demes of Madres and Canigou that, following the geographical arrangement, should be placed on the left border of the graph, beyond the Ax demes. The Atlant. population is isolated and Mauléon and Luchon are neighbours and are placed between Atlant. and Ax, closer to Andor and Spain, that are themselves located between Luchon and Ax on the map (Fig 2).

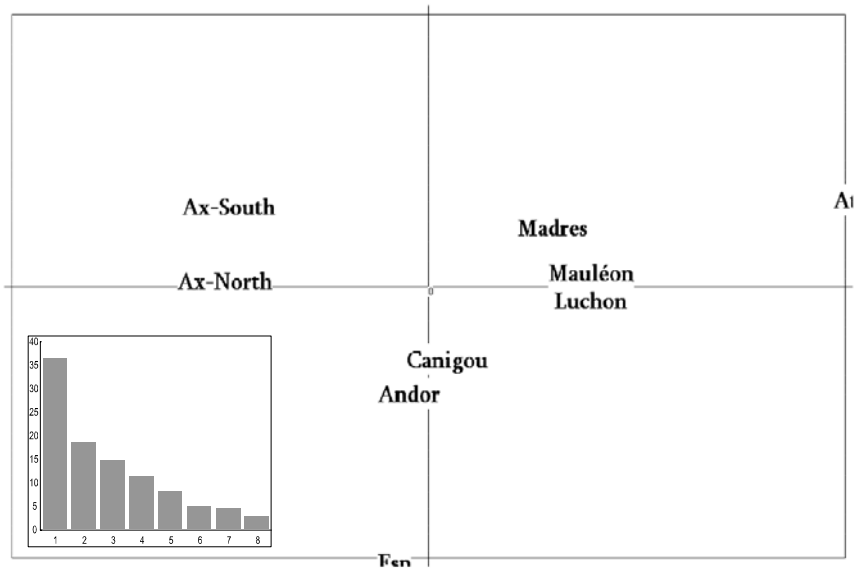


Fig. 2: PCA of gene frequencies. The first axis is significant (broken-stick) and the first 2 axis (resp. horizontal and vertical) sum up 57% of the total inertia. Lower left corner: graph of inertia for the first 8 each components.

The effect of geographical distance on the distribution of genetic diversity in the population was tested using pairwise Fst values (Table 2).
The relation between Fst/(1-Fst) and the logarithm of the geographic distance is significant (R^2=0.36; pval=0.0005, Mantel test, 2000 randomizations). However, the deme of Atlant. is remote and shows the largest geographical distance with all other demes, and we suspected that this large geographical distance could account for a large part of the significance of the model. Indeed, removing the latter deme from the data to illustrate this relation at short distance, the relation between Fst or Fst/(1-Fst) and geographical distances is not significant (R^2=0.02; pval=0.94).

	Atlant.	Mauléon	Luchon	Esp	Andor	Ax-Sud	Ax-Nord	Madres	Canigou
Atlant.		85	95	156	180	190	207	225	238
Mauléon	**0.082**		18	80	100	107	123	142	157
Luchon	**0.078**	0.010		64	86	96	113	131	144
Esp	**0.177**	**0.073**	**0.072**		25	44	63	74	84
Andor	**0.130**	**0.045**	**0.029**	**0.046**		22	40	49	59
Ax-Sud	**0.185**	**0.071**	**0.066**	**0.107**	**0.054**		19	35	50
Ax-Nord	**0.199**	**0.075**	**0.082**	**0.096**	**0.049**	**0.041**		23	41
Madres	**0.073**	0.030	0.047	0.089	0.066	0.068	0.100		18
Canigou	**0.131**	**0.075**	**0.053**	**0.052**	**0.053**	0.060	**0.107**	0.025	

Table 2: Pairwise Fst between 7 populations
The populations from left to right are placed East to West. Upper half-matrix: Geographic distances in Km. Lower half: Fst (bold: significant at the 5% level).

As no robust pattern of isolation by distance was detected geographically, we investigated other potential patterns. Geographical distribution of genetic variability could be placed at another geographical scale than the one tested, or could be linked to biological parameters. The geographic distribution of genetic variation between the demes is visualized by a graphic representation of variance components of alleles using individuals (Fig 3).

The first two axis add up to 36.7% of the total variance (Fig. 3). The neighbouring demes Mauléon-Luchon, Ax-North and South, Spain and Andorra, cannot be differentiated. Atlant. individuals are grouped and the samples from the demes of Madres and Canigou, located at the westernmost edge of the distribution area (Fig. 1) appeared spread across the other demes.

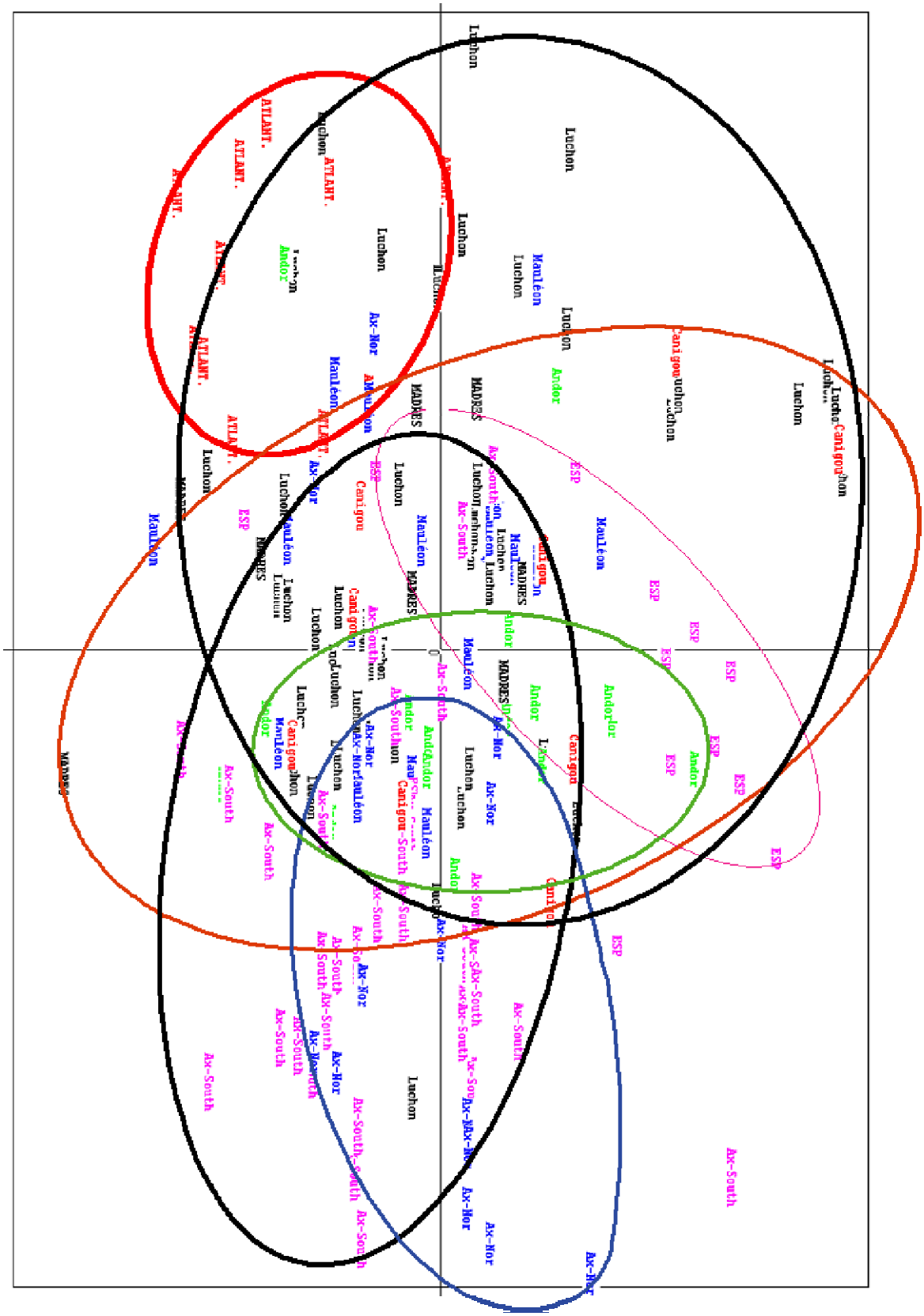


Fig. 3: Graphic principal component analysis of genotypes using PCAgen. Ellipse from left to right: Mauléon-Luchon, Atlant., Canigou-Madres, Esp, Ax-South, Luchon, Andorra, Ax-North.

IV. 4. 3

Genetic Structure within patches

The number of samples per deme is unbalanced, ranging from 11 in Atlant., the furthermost deme, to 47 in Luchon. Fis are not consistent across demes, ranging from 0 to 0.16 (Table 3).

Table 3: List of demes sampled in the Pyrenees, downwards from west to east; F: number of females; M: number of males, Total: total number of samples (males plus females plus samples for which sex was not confirmed); C.S.: estimated min. and max. census size of patch; d.: estimated min. and max. density of patch; N.U.: number of Natural Unit comprised in the deme. A: mean number of alleles over loci; *rank A*: rank of A value among patches; R: allelic richness; *rank R*: rank of R value among patches; Fis: inbreeding index. * indicates significance at the 5% level.

Patch	subdivision	F	M	Total	N.U.	C.S.	d.	A	*rank A*	R	*rank R*	Fis
Atlant.		2	8	11	4	171~399	1,5~3,6	2.72	*13*	3.27	*10*	-0.04
Mauléon		6	11	18	2	88~88	0,6	3.19	*6*	4.36	*2*	0.01
Luchon	*total*	19	28	47	2	164~293	2,4~4,2	6.27	*1*	3.52	*8*	0.11*
	Bareille	3	3	6	-	32~44	1,6~2,2	3.00	*8*	3.09	*14*	0.08
	Esbas	12	15	27	-	22~62	2,2~6,2	3.38	*3*	5.82	*1*	0.17*
	SuperBagn	4	10	14	-	12~28	2,4~5,6	3.05	*7*	4.00	*5*	-0.05
Andor		2	14	17	-	350~500	-	3.20	*5*	4.09	*4*	-0.05
Ax-Sud	*total*	16	15	33	1	48~112	0,7~1,8	4.64	*2*	3.11	*13*	0.15*
	Beille	6	5	12	-	15~20	-	3.34	*4*	4.18	*3*	0.36*
	Bourbourou	10	10	21	-	18~42	-	2.55	*14*	3.27	*11*	-0.01
Ax-Nord		2	13	15	4	64~64	0,6	2.80	*12*	3.64	*7*	0.00
Madres		1	6	8	2	71~205	0,6~1,8	2.84	*11*	3.18	*12*	0.11
Canigou		1	8	9	2	83~169	0,8~1,6	2.91	*10*	3.45	*9*	0.00
Esp		5	9	14	-	-	-	2.97	*9*	4.00	*6*	0.10

Only two Fis values are significantly different from zero: Luchon and Ax-South, the largest demes in terms of sample size. This significant excess of homozygotes indicated by the microsatellite data in the two patches containing the biggest number of samples suggest sub structuring due to Wahlund effect (Hedrick, 2000). We investigated the possibility of sub structuring within these two demes by subdividing them into known aggregates of lek or wintering sites: we first grouped genotypes in the deme of Luchon based on the two leks from which they were sampled. The latter three sampling site (Bareille, Esbas and Superbagnère) are 5 km apart from each other. The three subsamples Bareille, Esbas and Superbagnère do not appear to be differentiated (Gtest $p < 0.005$) and only Esbas is still deviating significantly from H.W. expectation (Table 3). We also divided the Ax-South genotypes into two groups, as samples were collected on two different wintering forests (plateaus of Beille and Bourbourou) distant of a few kilometres from each other and about 100km from the Luchon population. The wintering site of Bourbourou (21

samples, Table 3) does not diverge from H.W. expectations, whereas the patch of Beille (12 samples, Table 3) still shows significant Fis (0.36*, Table 3).
The allelic richness (R) in the two patches that are not in H.W. equilibrium is significantly higher than in all other patches, excluding the deme Esp (pval= 0.008, one-tailed test, 2000 permutations). The allelic richness between the marginal demes (Madres, Canigou and Atlant.) is lower than the one measured in the central demes (pval= 0.043, one-tailed test 2000 permutations).
Finally, no demes showed evidence for recent bottleneck, apart from the deme of Luchon, showing a slight mode-shift compared to the expected distribution of allele frequency distribution.

IV. 5 DISCUSSION

Our general results indicate that the capercaillie populations in the Pyrenees have not attained mutation-migration-drift equilibrium and that several local parameters affect the demographic dynamics within this population. Furthermore, the genetic differentiation appears significant, even between neighbouring demes. However, dispersal does not seem to occur according to a pattern of isolation by distance.

IV. 5. 1 Within population differentiation

The capercaillie population in the Pyrenees is composed of genetically differentiated demes as Fst between demes is moderate, yet significant (Fst = 0.07) and most pairwise differentiation test were significant, with pairwise Fst values ranging from zero to 0.2.
According to the principal component analysis we performed, the distribution of genetic variability fits the geographical arrangement of the patches. In spite of this geographical arrangement, our data did not reveal a significant pattern of isolation by distance or correlation between genetic and geographic distance at a local scale. The only case for which genetic structure is explained by geographical distance concerns the furthermost patch (Atlant.). This patch is in H.W. equilibrium and shows higher pairwise Fst. We can suspect that the western-most deme of Atlant. is differentiated from the rest of the demes sampled. These results hint at isolation by distance at long range.

Patterns of isolation by distance were reported in other capercaillie populations across Europe (Segelbacher *et al.*, 2003; Segelbacher & Storch, 2002) or in the "edge" populations in the Alps (Segelbacher & Storch, 2002). Also in black grouse and ptarmigan, the same pattern was documented (Caizergues & Ellison, 2002; Caizergues, Ratti *et al.*, 2003).
However, at short distances, our data indicate either that the level of gene flow among patches in the Pyrenees is heterogeneous, either that the recent demographic history is dissimilar among patches, which could be explained by metapopulation dynamics or recent expansion.

The first possibility, that gene flow is heterogeneous among patches, could be due to the presence of natural barriers. This possibility is likely, as the Pyrenees are a mountain massif, with very accidented relief. A deeper study comparing landscape details and distribution of individuals or of genetic diversity could shed light on this situation.

However, dissimilar demographic history between patches is also possible, as the Pyrenees population as recently recovered from a bottleneck (Ingram, 1915) and is presently suffering from demographic decline. To investigate further this scenario, we explored the distribution of microsatellite diversity within patches, when data was available.
The way we defined the demes and grouped N.U. into patches was uniquely based on geographical parameters, but the significance of the geographical division remains ambiguous because some demes are not in H.W. equilibrium and two demes are not significantly differentiated.
First, the deme located in Spain (Esp: Fis=0.10, N.S., Table 3) has been sampled across a long valley that would correspond to an equivalent area covered by several N.U. on the other side of the border with France. It is therefore likely that the high Fis value found in the Esp deme is due to Wahlund effect rather than inbreeding. This deme was therefore removed from the following comparisons of gene diversities.
Second, the allele frequencies in other patches (Luchon, Ax-south and Madres) deviate from H.W. expectations. Among these three patches, Luchon and Ax-south are known for having experienced the latest forest recovery among the patches considered here. In these two areas, forests were totally eradicated at the beginning of the century, whereas all the other patches only suffered forest degradation. It is probable that these patches have not attained drift-migration equilibrium since recolonisation (after W.W.I). The allelic richness (R) in the three localities (patch sub-division) known for having had the latest forest re-growth, namely Esbas, Bourbourou and Beille, when grouped together, is not significantly higher than in the other patches (excluding the Spanish pool Esp). However, if

we consider Esbas and Beille grouped, these two localities together show higher allelic richness and Hs (gene diversity) than in all the other patches pooled. This discrepancy between levels of diversity could originate from the recent disturbances affecting the demographics of these patches compared to other elder patches of forest: colonisation being more recent, the gene diversity would remain higher than in other patches because the samples taken on these sites have multiple origins. The multiple origins of individuals in Esbas and Beille would also explain the high Fis observed by Wahlund effect.
Third, the diversity in the central patches is higher than in the three patches at the edge of the distribution. This pattern of relative genetic impoverishment is typical of population edge effects (Vucetich & Waite, 2003).
In the latter two situations (most disturbed patches and "edge impoverishment" effect), it seems that habitat quality and disturbance due to human activity had a significant effect on the local demography dynamics, often modifying the distribution range and the dispersal patterns of capercaillie in the Pyrenees.

IV. 5. 2 Jura and Pyrenees: differentiated sub-species

The populations in the Jura and the Pyrenees appear significantly differentiated, as confirmed by the divergence in allele frequencies and the significant population differentiation test. The two populations are considered as sub-species because of morphological differences (Snow & Perrins, 1998). We could certainly expect higher Fst between the two distant populations (800km) as we suppose there has not been consistent gene flow between them for at least the last 800 years, time of human expansion and fragmentation of the forests in the south of France. The only potential stepping-stone between Jura and Pyrenees is the massif of Cevennes where individuals have been re-introduced several time since 1979. Nevertheless, the morphological difference is strong and the one measured with microsatellites is moderate. We suppose that the microsatellite loci have evolved differentially by drift and mutation and that mutation has caused homoplasy, thus underestimating the Fst measures. This pattern is somehow also indicated in Segelbacher et al (2003), as an Fst of 0.16 was measured between the Pyrenees and the Jaroslawl population, that is located far into Russia.
The low diversity measured in the Pyrenean populations in the latter study (Segelbacher *et al.*, 2003) (mean number of allele per loci A=3.0; allelic richness R=1.99; observed heterozygosity Ho=0.44) is the lowest amongst all samples included in this study. Our analysis confirms the

values of heterozygosity, and also that A and R are lower in the Pyrenees than in the Jura, an isolated population of ca. 600 individuals. We suspect that this relatively low microsatellite diversity is a result of the strong drift that follows the bottleneck that occurred in the early 20th century.
This bottleneck was due to intensive forest harvesting for cattle breeding, ore processing and construction needs. It has maintained the capercaillie populations in the Pyrenees in a situation of near extinction for over 50 years, which corresponds to ca. 15 generations. During this time, capercaillie must have survived to the benefit of scattered refuges of probably sub-optimal habitat and genetic drift must have increased drastically, hence, the low diversity compared to other European populations (Segelbacher *et al.*, 2003). At the same period, the French Alps were also totally deforested but the local capercaillie population never recovered. We suggest that the bottleneck in the Alps might have lasted longer than in the Pyrenees, but it is also likely that the Pyrenees populations benefited from the absence of black grouse. The competition between these two closely related species has never been studied but a previous study demonstrated clearly that capercaillie in the Pyrenees occupies currently the ecological niche elsewhere held by black grouse (Menoni, 1991). This factor might have been crucial in the survival of the population during intensive forest destructions and might have permitted its recovery in the Pyrenees.

In summary, in spite of significant genetic differentiation indicated by the F statistics, molecular data analysis suggests that neighbouring demes are connected by a substantial level of gene flow, possibly locally impaired by natural barriers. Furthermore, these populations have not attained mutation-drift-migration equilibrium, likely because of the recent bottleneck and disturbances caused by human activities. Therefore, if human activity in the area is bond to a new cycle due to increased tourism, as it was caused in the 1850's by industries, there is little hope of stopping the present decline. Capercaillie population dynamics would benefit from the long-term establishment of many large and protected forest patches. Suitable habitats (Menoni, 1991) protected over long periods could facilitate capercaillie populations' recovery more rapidly than it has recovered after the last bottleneck and they would therefore suffer less from the successive trends of human activities regularly deteriorating its habitat.

IV . 6 ACKNOWLEDGEMENT

Many thanks to M. Catusse and M. Monta, for supporting this work and providing the GIS maps, to M. Mossoll Torres, who provided the samples from Andorra. N. di Marco and C. Decoster helped with the extractions. We are also in debt to M. Glaettli, C. Parisod, N. Juillet for providing helpful comments on the manuscript.

IV . 7 REFERENCES

Caizergues A, Dubois S, Loiseau A, Mondor G and Rasplus JY (2001) Isolation and characterization of microsatellite loci in black grouse (Tetrao tetrix). *Molecular Ecology Notes* **1**, 36-38.

Caizergues A and Ellison LN (2002) Natal dispersal and its consequences in Black Grouse Tetrao tetrix. *Ibis* **144**, 478-487.

Caizergues A, Ratti O, Helle P*, et al.* (2003) Population genetic structure of male black grouse (Tetrao tetrix L.) in fragmented vs. continuous landscapes. *Molecular Ecology* **12**, 2297-2305.

Catusse M (1993) Spatial and Temporal Plasticity of a Capercaillie (Tetrao-Urogallus) Arena in the French Pyrenees. *Ethology Ecology & Evolution* **5**, 145-156.

Cornuet JM and Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Fridolfsson AK and Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**, 116-121.

Goudet J (1995-2001) FSTAT (Version 2.9.3): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Hedrick PW (2000) *Genetics of populations, 2nd edition* Jones & Bartlett, London.

Hjeljord O, Wegge P, Rolstad Jo, Ivanova M and Beshkarev AB (2000) Spring-summer movements of male capercaillie *Tetrao urogallus*: A test of the 'landscape mosaic' hypothesis. *Wildlife Biology* **6**, 251-256.

Ingram C (1915) A few notes on *tetrao urogallus* and its allies. *Ibis*, 128-133.

Marshall K and Edwards-Jones G (1998) Reintroducing capercaillie (Tetrao urogallus) into southern Scotland: identification of minimum viable populations at potential release sites. *Biodiversity and Conservation* **7**, 275-296.

Maurin K (2003) *Faune et Flore des Pyrénées* Arthaud.

Menoni E (1991) PhD thesis: Ecologie et dynamique des populations du grand tétras dans les Pyrénées, avec des reférences spatiales à la biologie de la reproduction chez les poules: quelques applications à sa conservation, Université Paul Sabatier.

Moss R (2000) Second extinction of capercaillie (*Tetrao urogallus*) in Scotland ? *Conservation Biology* **10**, 255-257.

Nunney L and Campbell KA (1993) Assessing minimum viable population size: Demography meets population genetics. *Trends in Ecology and Evolution* **8**, 234-239.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Sachot S, Leclercq B and Montadert M (2002) Population trends of Capercaillie (*Tetrao urogallus*) in the Jura mountains between 1991 and 1999. *Game and Wildlife Science* **19**, 41–54.

Saniga M (2003) Causes of the population decline in capercaillie (*Tetrao urogallus*) in the West Carpathians. *Biologia* **58**, 265-273.

Segelbacher G, Hoglund J and Storch I (2003) From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology* **12**, 1773-1780.

Segelbacher G, Paxton RJ, Steinbruck G, Trontelj P and Storch I (2000) Characterization of microsatellites in capercaillie Tetrao urogallus (AVES). *Molecular Ecology* **9**, 1934-1935.

Segelbacher G and Storch I (2002) Capercaillie in the Alps: genetic evidence of metapopulation structure and population decline. *Molecular Ecology* **11**, 1669-1677.

Snow DW and Perrins CM (1998) *The Birds of the Western Palaearctic* Oxford University Press, London.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Vucetich JA and Waite TA (2003) Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation Genetics* **4**, 639-645.

Weir BS and Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

V CAPERCAILLIE MATING SYSTEM: RELATEDNESS ON LEKS

Sébastien Regnaut, Philippe Christe,
Michel Chapuisat, and Luca Fumagalli

V. 1 ABSTRACT

Recent theories, supported by molecular data, suggest that males displaying on grouse leks are related to each-other. This study aims at measuring relatedness on leks on new molecular data from non-invasive samples of capercaillie, another lekking grouse. We find that, although relatedness is higher on lek than in the average population (R_{global}=0.2), this value decreases when relatedness is measured with respect to the local allelic frequencies (R_{local}=0.12). These results, interpreted in the light of the isolation by distance pattern found in our data among leks and among individuals indicate that both geographical structure of the population and kin association account for genetic similarity between males on leks.

Mating among relatives is a threat to population survival as it tends to increase inbreeding. However, our results demonstrate that, even if males do display preferentially among relatives, the lek mating system concurs to the establishment of microsatellite diversity in capercaillie populations and that leks represent a lower level genetic structure that should be considered in conservation plans.

The evolution of lekking as a mating system remains a puzzling issue and a debated question in sexual selection theories (Andersson, 1995; Hoglund, 2003). Until recently, direct fitness benefits mediated through male-male competition have been considered to be the main factor driving the evolution of lek mating. However, because females usually preferred to mate in larger male aggregations rather than in smaller lek, it has been proposed that low-rank males could enhance their inclusive fitness by increasing lek size of related high-rank males (Kokko & Lindström, 1996). Indeed, the idea that kin selection could play an important role in the evolution and in the maintenance of lekking has received empirical supports from at least four bird species. Peacocks (*Pavo cristatus*) were more closely related to males within the same lek than to males at other leks (Petrie *et al.*, 1999). Similarly molecular data on lekking black grouse (*Tetrao tetrix*) (Hoglund *et al.*, 1999), lesser prairie-chicken (Bouzat & Johnson, 2004) and white-bearded manakin (*Manacus manacus*) (Shorey *et al.*, 2000) have revealed that relatedness was high among males sharing a same lek. However, Genetic similarity between males on leks may also result from limited dispersal by males from natal sites (phylopatry) and from their particular mating system that involves a large degree of lek fidelity (Hoglund, 2003; Watson *et al.*, 1994).

Recently, Møller (2003) emphasized the importance of including sex when theoretically studying extinction risk in conservation biology and Kokko & Brooks (2003) investigated whether sexual selection and inbreeding influenced extinction risks. Indeed reproductive skew is predicted to be particularly high in a lek mating system (Fiske *et al.*, 1998). Therefore if mating success is highly skewed towards small number of males within lek, effective population size (Ne) is expected to decrease. Moreover if males within leks are kin, the decrease of effective population size will be more pronounced and the level of inbreeding will increase. The resulting loss of genetic diversity in small and fragmented populations of threatened species is then expected to reduce the ability to evolve and thus to increase extinction risk due to environmental changes (Frankham *et al.*, 2002).

The aim of the present study was to investigate the genetic structure within and between leks of a declining population of a tetraonid species,

the western capercaillie (*Tetrao urogallus*). This will enable us to investigate if males cluster on leks preferentially by kin in this species. Because this species is particularly sensitive to human disturbance and because the population of the Swiss and French Jura mountains has drastically declined since 20 years (Sachot *et al.*, 2002), we used a non-invasive method to collect biological samples (faeces) on different leks.

V. 3 SAMPLING & ANALYSIS

V. 3. 1 Sampling

The study was conducted across the south-western portion of the Jura mountains, between Nantua (France) and Yverdon (Switzerland; Fig. 1). We collected several hundred capercaillie faecal samples either under deciduous and fir trees, where individuals spend the night or feed, or on lek sites. We also sampled faeces on 15 lek sites during the breeding season (average lek distance is 30km with S.D.=22.3). Faeces were first individually dried in vials containing silica gel beads, and then processed in a separated laboratory only dedicated to low-content DNA samples.

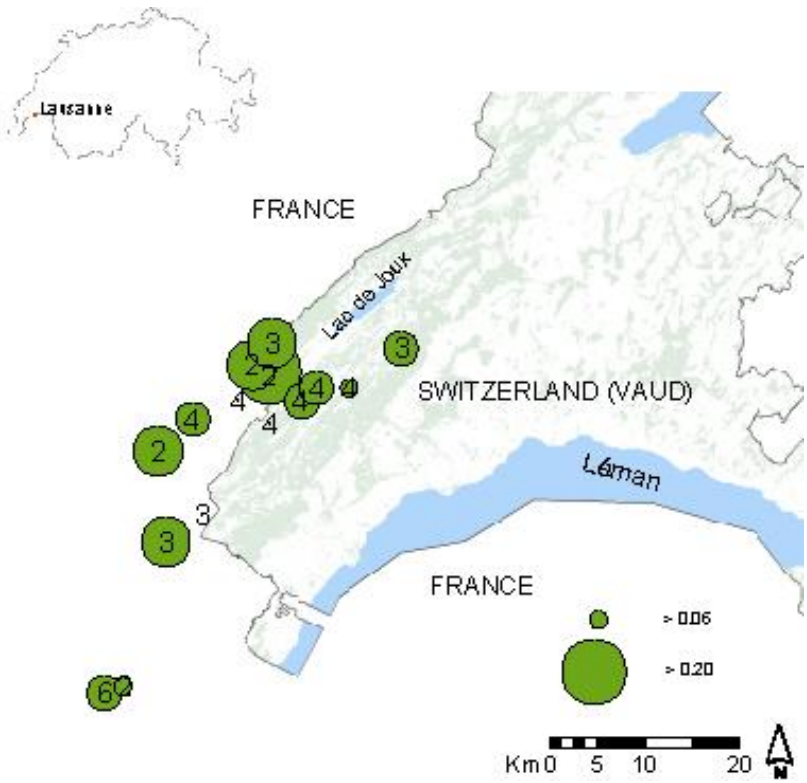


Fig. 1: Sampling area, map of lek positions (black numbers correspond to sample size) and relatedness within leks (circle sizes correspond to *r* values from Goodnight & Queller 1999)

V. 3. 2 Faecal samples DNA extraction

DNA extraction was performed with the QIAamp Stool kit (QIAGEN) optimized for convenience (see section 2-Chap II). We amplified 11 microsatellite loci, 7 of which were specifically designed for capercaillie (Segelbacher *et al.*, 2000), and 4 from a related species, *Tetrao tetrix* (Caizergues *et al.*, 2001). We monitored DNA contamination by performing controls for both extractions and amplifications. Amplification products were visualized using an ABI377 automated sequencer (Applied Biosystems). In order to minimize genotyping errors due to the analysis of noninvasive samples (i.e. allelic dropout, false alleles), we adopted a multitube approach (Taberlet *et al.*, 1996). Heterozygotes were repeated twice and homozygotes five times. We estimated the reliability of the latter genotyping procedure to be as high as 98% using computer simulations with the software Gemini (Valière *et al.*, 2002) based on measures of allelic dropout and false alleles rates (Taberlet *et al.*, 1996). Sex was determined by a PCR-based method adapted from (Fridolfsson & Ellegren, 1999).

V. 3. 3 Genetic data analysis

Population structure was explored with the help of descriptive F-statistics using the FSTAT software (Goudet, 1995). Nei's distances D (1978) were computed with the GENETIX software (Belkhir *et al.*, 1996-2002). Pairwise genetic distances between lekking males (D) were plotted against geographical distances and the matrix correlations were tested with a Mantel test using the software FSTAT. Spatial autocorrelation of individual genotypes was investigated using Rousset's a_r with the method implemented in the software SPAGeDi (Hardy & Vekemans, 2002).
We measured the Relatedness among males lekking together using the software *Relatedness* (Goodnight & Queller, 1999) first, with respect to the whole reference population, and second with respect to the individuals sampled in the neighbouring population (patch) defined according to habitat fragmentation (Sachot 2002).
To estimate the expected Relatedness values obtained in our data set using our microsatellite markers and analyse the frequency distribution of pairwise relatedness values among lekking males, we used the software KINSHIP to simulate pairwise *r* value distributions for pairs of

unrelated, of half-sibs and of full-sibs drawn from a population with the same allelic frequencies as the reference population.

V. 4 RESULTS

V. 4. 1 Population genetic structure

Two hundred and thirty eight unique genotypes were scored throughout the whole sample set, of which 30% were genotyped as females, 65% as males and 5% remained of sex unknown. The sex ratio is strongly male biased (2:1) in the data set. We suspect this an effect of sampling because other counting methods did not reveal unbalanced sex ratio (Leclercq, 2004). Sampling on leks offers indeed more opportunities to collect males faeces.

Among the 9 patches included in our study, overall F_{ST} = 0.04 (S.E.= 0.02; jackknife over loci), showing a significant level of differentiation among leks (G-test (Goudet, 1995-2001) was significant over all loci (p < 0.0001) at this small spatial scale. Among these 9 patches, only one showed significant departure from Hardy-Weinberg expectations by homozygotes excess.

Genetic structure among lekking males was based on a dataset of 59 samples distributed among 17 leks sites confirmed by field observations (Fig. 1).

V. 4. 2 Isolation by Distance

A significant pattern of isolation by distance is detected among leks. As shown by the significant correlation between geographic distances to Nei's distances between pairs of leks within 40 km maximum distances (Fig. 2), it appears that the leks within 40km show the most explanatory model (r^2=0.43, Mantel p <0.001). A genetic spatial autocorrelation of individual genotypes reveals that genotypes are, on average of the 11 loci, significantly more correlated than average for pairs of males distant of less than ca. 15km (Fig. 3).

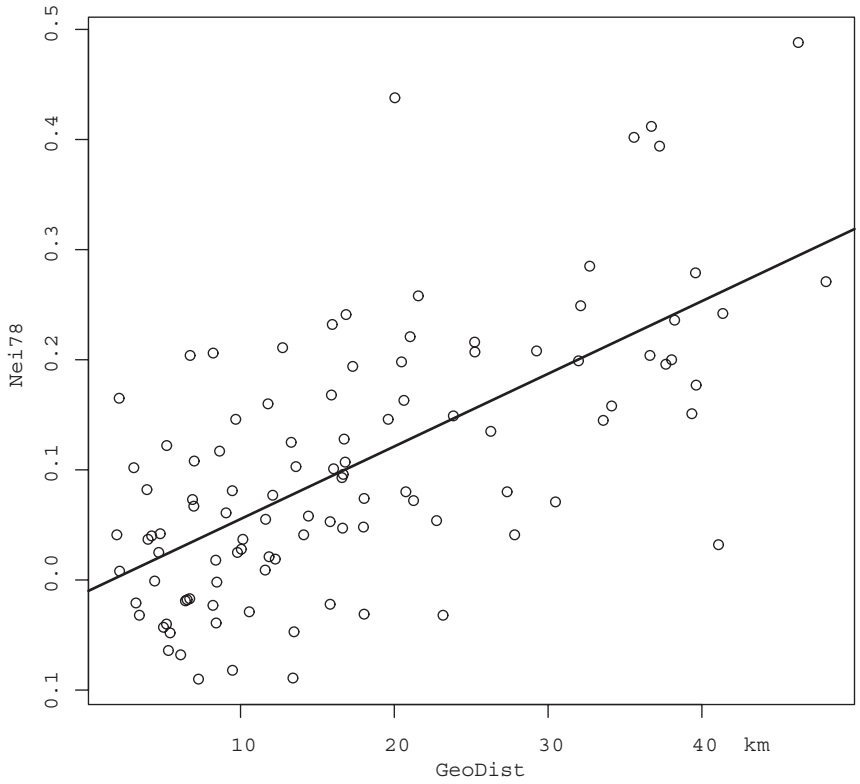


Fig. 2: Correlation between geographic and Nei's distances over 40km.

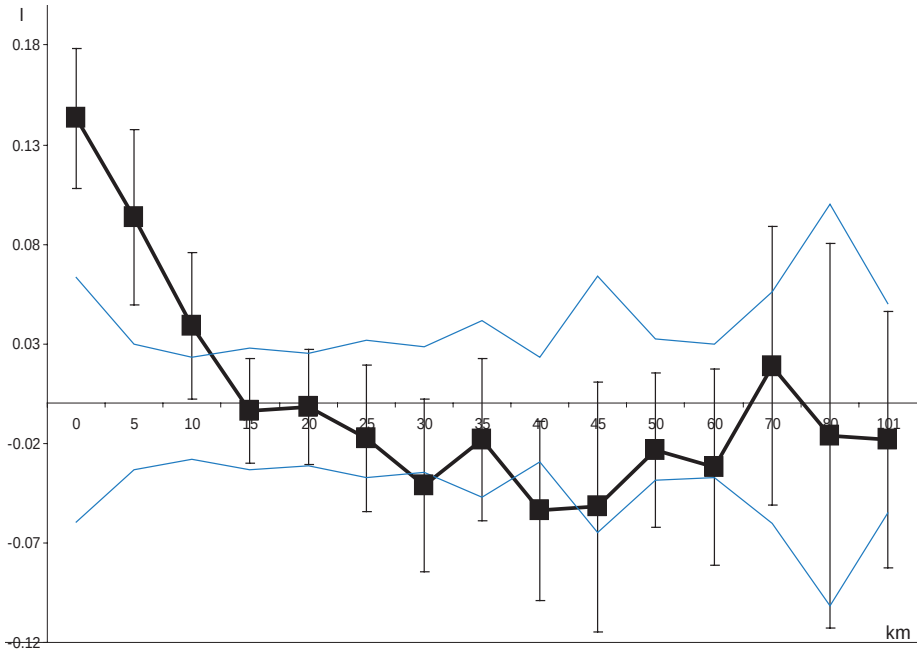


Fig. 3: Consensus autocorrelogram of within lek individual genotypes, using Moran's index (I) averaged over all loci. The lines represent the 95% envelope of average I distribution after 1000 permutations of individual genotypes, the vertical bars represent the S.D. over loci.

First, relatedness (r) among males sampled on the same lek is higher than the population's average (r = 0.22, Fig. 4) showing that males don't cluster randomly at lek sites. Second, we found that the relatedness measured between lekking males at a local scale, with respect to the neighbouring population, is lower than the former (r = 0.12, Fig. 4).

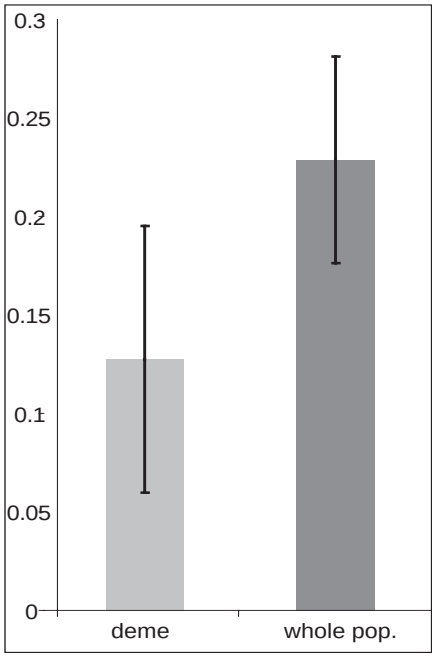


Fig. 4: Average pairwise relatedness (R) among pairs of lekking cocks within leks. In light gray: with local deme as reference sample (with 95% C.I. jacknifed over 11 loci). In dark gray: with the whole population as reference (with 95% C.I. jacknifed over 11 loci).

An r value of 0.25 characterizes half-brothers and uncle-nephew, on average, but the within-leks pairwise r values found range from -0.34 to 0.77 (Fig. 5). To investigate the biological significance of the averaged value of R=0.2, we compared the frequency distribution of pairwise relatedness values (Fig. 5 – empty circles) with frequency distribution of r values for three levels of relatedness, simulated from the microsatellite data in our Jura reference population.

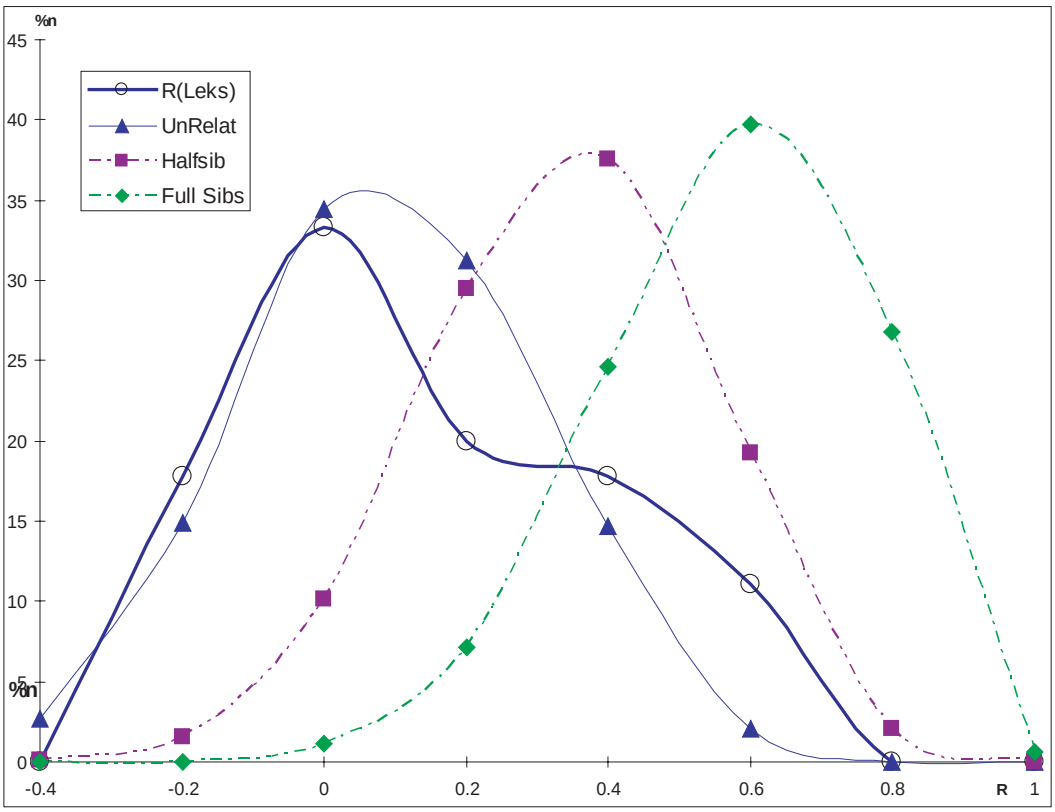


Fig. 5: Histogram of observed and simulated *r* values densities
Observed distribution of paired relatedness values among cocks lekking on the same arena with local deme as reference sample (full circles, *r*(Leks)) - Simulated R values distribution (as in Queller, 2000) distributions of brothers (FullSibs, true *r*= 0.5), half-brothers (HalfSibs, true *r*= 0.25) and unrelated individuals (UnRelat, true *r*= 0) based on allele frequencies of the total Jura population (238 unique genotypes).

V. 5 DISCUSSION

The measure of relatedness (*r*) among males indicated that within-leks pairwise *r* is higher (0.12) than the average *r* of the local population, showing that males don't cluster randomly at lek sites.
The comparison of this value of local relatedness with the higher global average *r* (0.22) suggests that geographical proximity accounts for a large part of the relatedness measured among males on leks. In geographically structured populations, genetic similarity among individuals is higher within than among populations. Therefore, the structure found in our data set (Fst = 0.04) accounts for a part of genetic similarity measure by the global *r*.

Indeed, the study area that spreads over a distance beyond the dispersal ability of the species as field data suggest that males' capercaillie

dispersing distance remains within a few kilometres (Storch, 1993), whereas the sampling spread over 100 km. Furthermore, the patterns isolation by distance among leks and of spatial autocorrelation of individual genotypes indicate that, within a certain range, geographical proximity explains a large part of genetic similarity between individuals. Such patterns of isolation by distance is also described in other grouse populations (Caizergues, Bernard-Laurent *et al.*, 2003; Caizergues, Ratti *et al.*, 2003; Piertney *et al.*, 1998) and documented in capercaillie (Segelbacher *et al.*, 2003; Segelbacher & Storch, 2002).
Nevertheless, the high local relatedness among males when measured with respect to the local population (local $r = 0.12$) is consistent with results from studies conducted in related species showing that grouse and their allies modify their social behaviour accordingly to the degree of relatedness of their neighbours (Watson *et al.*, 1994).
Therefore, our results suggest that clustering of related males on leks is partly due to geographic population structure and partly to kin association.

Furthermore, the pairwise values of relatedness between males lekking together are distributed across a broad range. We find that 60% of the pairs of individuals are more related than the average ($r>0$), even possibly inbred (high r value), but we also find that many pairs of cocks are unrelated. Actually, 12% of the pairs lekking together have a r value inferior to the low value of -0.25, meaning that they are less related to each other than the average of all pairs of males drawn randomly in the whole Jura population genotyped. This broad distribution of values can possibly originate from flaws in the methods. The noise in microsatellite data due to pedigree history of the population, but also the large variance of relatedness estimations from genotyping data could lead to such approximations and produce this large range of values, even for full brothers for instance (Milligan, 2003). The other possibility explaining the variance in pairwise relatedness between lekking males is that many unrelated individuals do display together with related individuals.
The possibility of unrelated males on leks suggests that dispersion among leks occurs among males. This suggestion is consistent with the pattern of correlation of genetic and geographical distances between leks found in our data.

Indirect measures such as distribution of microsatellite diversity can only bring partial insight into the understanding of the factors that concur to the establishment and to the maintenance of lekking behaviour. Nevertheless, our results of local relatedness among males on leks

demonstrate that the geographical structure and the patterns of dispersion in not the only parameter that accounts for kin association on leks, as suggested in previous studies (Hoglund, 2003; Kokko & Lindström, 1996).

In conclusion, it appears that the mating system of capercaillie has a strong effect on the distribution of genetic diversity within populations. The lek is therefore possibly the lower level of population structure, and the geographical distribution of leks should be taken into account when defining conservation units and population management policies.

V. 6 ACKNOWLEDGEMENTS

For analysis: P. Fontanillas, H. Niculita, J. Goudet, I. Dupanloup, and P. Taberlet.
For samples: B. Leclercq (Groupe Tétras Jura), A. Bloc (Parc Naturel Regional de la Haute Chaine du Jura), S. Sachot (Conservation de la Faune du Canton de Vaud)

V. 7 REFERENCES

Andersson M (1995) *Sexual Selection* Princeton University Press, Princeton.

Belkhir K, Borsa P, Chikhi L, Raufaste N and Bonhomme F (1996-2002) GENETIX 4.04, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions. *available on the www:* *http://www.univ-montp2.fr/~genetix/genetix/genetix.htm*.

Bouzat JL and Johnson K (2004) Genetic structure among closely spaced leks on a peripheral population of lesser prairie-chickens. *Molecular Ecology* **13**, 499-505.

Caizergues A, Bernard-Laurent A, Brenot JF, Ellison L and Rasplus JY (2003) Population genetic structure of rock ptarmigan *Lagopus mutus* in Northern and Western Europe. *Molecular Ecology* **12**, 2267-2274.

Caizergues A, Dubois S, Loiseau A, Mondor G and Rasplus JY (2001) Isolation and characterization of microsatellite loci in black grouse (Tetrao tetrix). *Molecular Ecology Notes* **1**, 36-38.

Caizergues A, Ratti O, Helle P*, et al.* (2003) Population genetic structure of male black grouse (Tetrao tetrix L.) in fragmented vs. continuous landscapes. *Molecular Ecology* **12**, 2297-2305.

Fiske P, Rintamaki P and Karvonen E (1998) Mating success in lekking males: a meta-analysis. *Behavioural Ecology* **9**, 328-338.

Frankham R, Ballou JD and Briscoe DA (2002) *An introduction to conservation genetics* University Press, Cambridge.

Fridolfsson AK and Ellegren H (1999) A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**, 116-121.

Goodnight KF and Queller DC (1999) Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Molecular Ecology* **8**, 1231-1234.

Goudet J (1995-2001) FSTAT (Version 2.9.3): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Goudet Jo (1995) FSTAT (vers. 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Hardy OJ and Vekemans X (2002) SPAGeDi : a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.

Hoglund J (2003) Lek-kin in birds - provoking theory and surprising new results. *Annales Zoologici Fennici* **40**, 249-253.

Hoglund J, Alatalo RV, Lundberg A, Rintamaki PT and Lindell J (1999) Microsatellite markers reveal the potential for kin selection on black grouse leks. *Proceedings of the Royal Society of London Series B-Biological Sciences* **266**, 813-816.

Kokko H and Brooks R (2003) Sexy to die for ? Sexual selection and the risk of extinction. *Annales Zoologici Fennici* **40**, 207-219.

Kokko H and Lindström J (1996) Kin selection and the evolution of leks: Whose success do young male maximize ? *Proceedings of the Royal Society of London Series B-Biological Sciences* **263**, 919-923.

Leclercq B (2004) Evolution des populations du grand tétras dans le jura Francais. In: *Actes des "4èmes Rencontres Jurassiennes"* (ed. J. B), Mijoux.

Milligan BG (2003) Maximum-likelihood estimation of relatedness. *Genetics* **163**, 1153-1167.

Møller AP (2003) Sexual selection and extinction: Why sex matters and why asexual models are insufficent. *Annales Zoologici Fennici* **40**, 221-230.

Petrie M, Krupa A and Burke T (1999) Peacocks lek with relatives even in the absence of social and evironmental cues. *Nature* **401**, 155-157.

Piertney SB, McColl AD, Bacon PJ and Dallas JF (1998) Local genetic structure in the red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite markers. *Molecular Ecology* **7**, 1645-1654.

Sachot S, Leclercq B and Montadert M (2002) Population trends of Capercaillie (*Tetrao urogallus*) in the Jura mountains between 1991 and 1999. *Game and Wildlife Science* **19**, 41–54.

Segelbacher G, Hoglund J and Storch I (2003) From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology* **12**, 1773-1780.

Segelbacher G, Paxton RJ, Steinbruck G, Trontelj P and Storch I (2000) Characterization of microsatellites in capercaillie *Tetrao urogallus* (AVES). *Molecular Ecology* **9**, 1934-1935.

Segelbacher G and Storch I (2002) Capercaillie in the Alps: genetic evidence of metapopulation structure and population decline. *Molecular Ecology* **11**, 1669-1677.

Shorey L, Piertney S, Stone J and Höglund J (2000) Fine-scale genetic structuring on *Manacus manacus* leks. *Nature* **408**, 352-253.

Storch I (1993) *Habitat use and spacing of capercaillie in relation to forest fragmentation patterns*, University Ludwig-Maximilian.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Valière N, Berthier P, Mouchiroud D and Pontier D (2002) Gemini: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83-86.

Watson A, Moss R, Parr R, Mountford MD and Rothery P (1994) Kin Landownership, Differential Aggression between Kin and Nonkin, and Population Fluctuations in Red Grouse. *Journal of Animal Ecology* **63**, 39-50.

Section - 3 CONCLUSIONS

I RESULTS OF THE FEASIBILITY STUDY

The first studies in capercaillie population genetics used feathers as DNA source (Segelbacher, 2002). The first studies of avian species to use faecal sample were done in the great bustard (Idaghdour *et al.*, 2003) and capercaillie (this thesis), suggesting that these methods can also be applied to other avian species.

Faecal samples have proven very efficient in past studies to monitor small populations where regular sampling methods (blood, muscles or organs) were not available. However, this type of non-invasive methods such as the use of faecal samples suffers from technical limitations due to low genotyping reliability. Therefore, the first part of this thesis aimed at adapting laboratory protocols to capercaillie faecal samples and at assessing the reliability of the genotyping procedure.

I. 1 GENOTYPING

I. 1. 1 Choice of faecal samples as source of DNA

Faecal samples collected from the field came in large quantities. Therefore, these samples have the advantage of potentially providing samples from all individuals present on the field of study without increasing disturbance or capture. Furthermore, they provided accurate geographical information on the activity of the individual sampled: while collecting faeces from the field, the scientist knows if the subject was passing through the area or feeding or sleeping. Other type of non-invasive DNA sources, such as urine or hair need to be trapped and are difficult to detect and feathers are scarce and they are available only during moulting season.

I. 1. 2 Protocol optimisations

Because of the low quality and quantity of template DNA available from stool, the laboratory processing of the samples included in this study have shown to be a lot more expensive and time consuming than usual tissue and DNA-rich samples. In spite of these constraints, I was able to genotype ca. 500 stool samples in the course of this study. I performed many comparative tests of alternative extraction procedures and of amplification protocols before processing a large amount of samples. First, I optimised the commercial extraction protocols to obtain enough extract to perform ca. 60 PCR on each sample. Second, I optimised the PCR conditions (Type of DNA polymerase, concentrations of reactants, annealing temperature) and managed to amplify and obtain reliable results from several microsatellite loci primarily developed for both capercaillie and black grouse. Third, I also tested alternative protocols to increase genotyping reliability or reduce processing time. For example, with the help of the LBC, we tested a two-step amplification procedure (Piggott *et al.*, 2004) that, however, did not reveal significantly better than the procedure we were using (Appendix I, Table 4-5).

I. 1. 3 Reliability

I further used these routines of extraction and amplification to estimate the reliability of the genotyping: First, to avoid genotyping error due to electrophoresis across gels, I used two cross reference samples (Appendix I, Table 1) on all gels. Second, to estimate allelic drop out and false alleles rates (Taberlet *et al.*, 1996), I performed systematic tests on sub-sets of samples using two alternative procedures (8 independent repeats, or 3 repeats then 4 further repeats for the homozygous samples; Appendix I, Table 2). Then, I tested these parameters more precisely using 56 samples (section 2, part I) and estimated the reliability of the genotypes at 98% using the softwares Gimlet and Gemini (Valière, 2002, 2002).
In the future, new technologies and new technical optimisations will likely increase the range of applicability and reduce the labour and the funding required.

I . 2 POPULATION GENETIC METHODS

The microsatellite data set was interpreted in the light of population genetics models and frameworks. With the advances in computer science in the recent years, population geneticists have made their methods for population parameters estimation available in software packages. In the course of my study, I have collected exclusive information about levels of gene flow at local scales (section 2, part III), genetic effect of recent demographic history (part IV) and behavioural ecology (parts V). These pieces of information were not available from field observation, and provided new insights in the biology of capercaillie.

I. 2. 1 The software Structure

The previous studies conducted on capercaillie in my model populations gave fundamental insights in capercaillie's ecology (Ménoni, 1991; Sachot, 2002). Both studies considered the population structures and its conservation based on ecological and habitat requirement of the species. However, there was no evidence that the genetic structure of the capercaillie populations was reflected in their habitat structure.

Therefore, a large part of this thesis is dedicated to a population genetic software, Structure (Pritchard *et al.*, 2000), developed to investigate and infer population genetic structure from microsatellite data. However, the ability of the software Structure to identify genetically differentiated clusters of individuals has not helped my investigations conducted with the microsatellite data sets from the Jura or the Pyrenees.
In the Jura, the software Structure did not detect any particular genetic structure. Using further genetic analysis, with the help of other methods, such as PCA analysis and F statistics, I concluded that the individuals living in the habitat patches were not significantly differentiated at the genetic level, confirming the findings of the software Structure.
In the Pyrenees, Structure did not highlight a clear genetic clustering that I could relate to a geographical pattern. Using F-statistics methods based on geographical clusters of individuals, it appeared that the genetic structure was higher in the Pyrenees than in the Jura. However, the allelic frequencies in some patches were not in Hardy-Weinberg equilibrium. I concluded for this study that this population might not have attained drift-migration-migration equilibrium. This situation causes deviations from Structure method's pre-requisites, making it unable to unravel the clusters.

The use of this software in my study did not provide major information, but merely confirmed those provided by other methods. The use of clustering methods in Conservation genetic studies, such as my study, was limited by the fact that data sets obtained from threatened populations often include populations that are not in equilibrium: small populations often suffer from inbreeding and fragmented populations often display metapopulation dynamics. The use of this software can therefore be misleading and will require other models and methods of investigation that include the possibility of genetic disequilibrium within demes.

I. 2. 2 Other softwares

I have used several softwares to describe the basic statistics of the populations, such as allele frequencies, heterozygosity, deviation from Hardy-Weinberg expectations, and also to produce estimates of F statistics following Weir and Cockerham estimates (Weir & Cockerham, 1984). These first steps were performed using the methods implemented in FSTAT (Goudet, 1995-2001), and Genepop (Raymond & Rousset, 1995).
I used Easypop (Balloux, 2001) to simulate data, SPaGeDi to calculate spatial autocorrelations (Hardy & Vekemans, 2002), Gimlet and Gemini

(Valière, 2002; Valière *et al.*, 2002) to estimate genotyping error rates and reliability, Bottleneck (Cornuet & Luikart, 1996) to measure the impact of recent demographic decline on the genetic diversity in the populations I studied.
Often, these methods developed for genetic analysis purpose could not be used with the required statistical power (Kokko *et al.*, 1998; Ryman & Jorde, 2001) for the following reasons: genetic stochasticity is increased in small demes and Hardy-Weinberg expectations are rarely met in declining populations. Furthermore, sampling is not balanced between demes because of sampling limitations, but also because decline in threatened populations do not occur at the same rate across the whole distribution area.

I. 2. 3 Estimation of effective population size

The inference of estimates of population size have revealed infeasible in the course of my study, even using the recent developments of Bayesian methods. The limitations encountered with MCMC and Bayesian procedures have also been documented in a study on cat populations, for which, in spite of great availability of samples, the authors have failed to produce usable samples (Kaeuffer *et al.*, 2004). From discussion and help provided by M. Beaumont (O'Ryan *et al.*, 1998), from Readings University and P. Berthier (Manel *et al.*, 2002), from the University of Bern, I concluded that none of these methods were adapted to my data sets because of the violation of the prerequisites of the respective models: deviation from H.W. and overlapping generations. Furthermore, another method to estimated effective population sizes based on a gene identity, Estim (Vitalis & Couvet, 2001) did not provide reliable estimates either. Hopefully, in the future, new models and methods will be developed to concur to more accurate estimations of the crucial parameters, such as geneflow and effective population sizes.

I. 2. 4 Limitations

The interpretation of this data in the light of Evolutionary theories have given inclusive information on the population structure and behaviour of capercaillie that were not available trough other techniques, such as field observation of radio-tracking. However, the statistical power that population genetics models required could not always be obtained from sampling in threatened populations. Therefore, making biological and

conservation inferences using population genetic models in declining populations have to be considered with respect to these limitations.

II RESULTS OF POPULATION GENETIC STUDIES

Two populations of capercaillie, the one in the Jura and the one in the Pyrenees, were included in this thesis. These two populations differ in census size, in demographic history, and in ecological requirements. However, the populations in the Jura and the Pyrenees have shown a consistent decline since the 1980's (Sachot *et al.*, 2002). The population genetic study presented in this thesis aimed at describing the genetic structure of these two populations and assess whether genetic factors, such as genetic fragmentation or inbreeding, could contribute to increase their extinction probabilities

II . 1 THE JURA

The capercaillie in the Jura Mountains has seen its habitat reduced during the 20th century by human activity. Consequently, the distance between this population and the populations of the Alps has increased (Mollet, 2002). Furthermore, urbanised areas and ski resorts have recently increased, which potentially creates barriers to gene flow among patches of suitable habitats. This apparent fragmentation has raised the issue of the possibility that genetic factors could increase extinction risk. In this part of the thesis, we investigated the genetic connectivity between habitat patches within the Jura, and whether these populations suffer from increased inbreeding.

II. 1. 1 Inbreeding

No genetic inbreeding was detected and the genetic relatedness observed on leks does not seem to cause preferential mating among related individuals. In the opposite, both the mating system and the short distance dispersal, creating geographic allelic frequencies correlations among neighbouring habitat patches concur to long distance homogenisation of the population.

II. 1. 2 Genetic structure and gene flow within the Jura

Fragmented habitat potentially causes genetic differentiation between demes, as shown in many conservation studies conducted in many species, for example in invertebrates, reptile and amphibians (Clark *et al.*, 1999; Hitchings & Beebee, 1996; Keller & Largiader, 2003; Rowe *et al.*, 2000). Fragmentation causing genetic differentiation and inbreeding is documented in avian populations in a past study (Lens *et al.*, 2000). However, other studies describe fragmented and threatened avian population for which no genetic differentiation was found (Galbusera *et al.*, 2004).

In my study, molecular data suggest that the Jura population is a continuum of individuals dispersing at short distance but at a large rate. The flight ability of capercaillie is certainly sufficient to maintain connectivity across valleys and urbanised areas present in the Jura.

II. 1. 3 Gene flow between the Jura and other populations

The great dispersal ability of capercaillie found in the present study suggests that gene flow might still exist between the Jura and other disconnected populations (The Alps, The Vosges, The Black Forest...). Furthermore, two samples did stand out of the clusters defined by the principal component analysis of allelic frequencies. However, I was not able to suggest whether these two originated from undocumented release or if they were long distance migrants. Therefore, the connection between the Jura and other European populations remains at question. An independent federal research centre in Switzerland (WSL, Birminsdorf) is presently conducting a study similar to mine in the Swiss Alps. Gernot Segelbacher (who conducted the past study on capercaillie genetics) and I expect a lot from the WSL study and we hope to be able to merge out data in a future study. This future study will be necessary to estimate whether gene flow still occurs between the large isolated populations. If such gene flow still occurs, I suggest that capercaillie will be able to recolonize depleted populations, such as the Alpes Vaudoises et the French Alps once the reasons for this decline has been understood and modified.

II . 2 THE PYRENEES

The Pyrenees Mountains range over a larger geographical scale than the Jura. A strong bottleneck has depleted the population between 1850 and 1920. The bottleneck was due to massive deforestation of the Pyrenees for

cattle breeding, mining industry and ore processing or again for construction. The situation leading to deforestation has been modified by human demographic factors after W.W.I. Since then, the forest had re-grown and the capercaillie population have expanded to a climax of 9000 individuals in the 1960's. The population is declining again, as well as the other European populations, since the 1980's, certainly due to another cycle of human activity in the area, increased tourism and poaching.

II. 2. 1 Effects of the bottleneck, of the subsequent expansion and of the disturbance

The low genetic diversity found within the Pyrenees compared to the Jura population suggests that the bottleneck that occurred until the 1920's has had a great impact on the genetic variation within this population. Furthermore, my analysis of the genetic data indicates that the Pyrenees population is not in mutation-drift equilibrium. Some patches show high Fis and high diversity, suggesting recent recolonisation. It appeared indeed that the latter patches have had the latest forest recovery of all patches, and that they are still undergoing high disturbance due to tourism. One the most recently re-grown forest patch, however, has been very protected by local legislation and did not show this pattern of genetic disequilibrium. In addition, the population patches located at the edges of the population's distribution showed lower levels of diversity.
These findings of heterogeneous distribution of genetic variation across the population's distribution suggest that the population has not fully recovered from the bottleneck and its demographics and dispersal patterns are still strongly influenced by human activity in the area.

II. 2. 2 Conservation and protection of the habitat

Seeing how the population has recovered between 1920 and 1960, from almost complete depletion up to 9000 individuals, I assume that capercaillie in the Pyrenees have a strong demographic potential. To avoid the future depletions that follow the irregular trends of human activities, I would propose the idea of protecting large and numerous patches of forests that would be used as sanctuaries for capercaillie, but also for the whole fauna and flora of the Pyrenees. These sanctuaries would become the core of future population expansions in case forest destructions happen again, as it has in the past.

II . 3 LEVELS OF STRUCTURE

II. 3. 1 Leks: the lower level

Significant genetic structuring was found between leks. The lek is best described as clusters of kin-related individuals, males sharing the parades with their sibs, paternal half-sibs or cousins and some migrants. Many other studies report differentiation between leks in grouse species such as black grouse (Hoglund *et al.*, 1999) and American grouse (Bouzat & Johnson, 2004; Van den Bussche *et al.*, 2003). This pattern is temporal and does not represent the permanent distribution of genetic diversity within the populations as individuals move seasonally to wintering, summering of breeding sites (Hjeljord *et al.*, 2000; Menoni, 1991; Storch, 1993).

II. 3. 2 Within population continuums

The habitat structure in the Jura did not reflect the actual genetic structure. The Fst values found among habitat patches is due isolation by distance and the gene pools of these patches are strongly mingled. In the Pyrenees, the Fst values were higher than in the Jura. This result is certainly due to larger geographic distances between patches. However, no pattern of isolation by distance was found in the Pyrenees. This can be explained by the presence of possible barriers to gene flow in this area where the relief is much more contrasted than in the Jura.
Globally, patches where the capercaillie is present appear linked by gene flow, and their flight ability seems beyond habitat structure, or at least, it was such in the recent past. Anecdotal observations of individuals performing very long distances flight to Greece (Holzinger & Rosler, 1990), to Belgium (Snow & Perrins, 1998) or observed at more than 3000m high (E. Menoni, pers. comm.). Furthermore, theories among grouse specialists suggest that the song of males can be heard from females at very long distance, therefore providing information about habitat occupation to surrounding individuals. These findings and observations suggest that geneflow between suitable patches within mountain ranges are not impaired by forest fragmentation and only partially impaired by relief structures.

II. 3. 3 Among populations: perspectives

Between separated mountain ranges, however, the genetic data suggest that genetic isolation is strong. F_{ST} measures between Jura and Pyrenees are moderate but measures performed in a recent study (Segelbacher *et al.*, 2003) suggest that the microsatellite markers used are saturated with homoplasy as the maximum F_{ST} found between distant is 0.28, and F_{ST} as high as 0.22 was found between patches within the Jura. Therefore, the F_{ST} of 0.17 found in this thesis between Jura and Pyrenees can be considered as high, and geneflow between the Jura and the Pyrenees is not likely.

In summary, we can distinguish two levels of genetic differentiation among capercaillie populations that are biologically significant: first, a geographical pattern of genetic differentiation between mountain ranges due to long distance. Second, a temporal pattern of structure among leks.

II . 4 MATING BEHAVIOR AND MAINTENANCE OF GENETIC DIVERSITY

The lek mating system groups related individuals at the time of reproduction. I suggested that inbreeding could be increased by this pattern if it induces assortative mating between relatives. Furthermore, the lek favours dominant males in access to females, which could potentially reduce effective population and diversity by reducing the number of males participating in reproduction (Kokko & Brooks, 2003).
The mating system could therefore bear a part of responsibility in the general decline by increasing effects of habitat deterioration, fragmentation and demographic stochasticity documented in grouse (Lindstrom *et al.*, 1995).
However, our results did not support this hypothesis. Measures of diversity in the Jura and the occurrence of many unrelated males present at lek sites suggest that the lek mating system has been efficient in maintaining genetic diversity in the Jura so far.
If male relatedness on leks is documented in lekking grouse species, one study relates the absence of relatedness among females at this stage (Bouzat & Johnson, 2004). The dispersion of female can suffice in homogenizing the population.

III CONSERVATION

III . 1 CAUSES FOR DECLINE

The present decline observed in most European populations does not find a consensual origin among the authors. Although no controversy exists, no precise cause to this decline has been found. This decline seems to occur in most habitats, even in large populations of Scandinavia and in the Alps (Moss, 2000; Sachot, 2002; Storch, 2000). Apart from the known demographic cycles (Lindstrom *et al.*, 1995), this decline probably finds it source in a large range of human related causes, from habitat loss due to sylviculture, to human disturbance with tourism (Saniga, 2003), and is made worse by various causes of mortality, such as predation (Helle *et al.*, 1999), ski runs (Brenot *et al.*, 1996) and fences.
Other potential causes, such as poaching, increased predation or climate change, are also related to human activity.
Whereas no evidence was found in the molecular data presented here that the populations I studied suffer from sound genetic impoverishment, the on-going decline will probably create isolation within population or increased homozygotie in the near future.
In the Jura, the demographic decline will certainly deplete the most sensitive areas, thus creating some fragmentation in the distribution unless this decline is stopped or gene flow other populations is restored.
In the Pyrenees, the population size is much larger than in the Jura, however, the genetic diversity seems to suffer from intense cycles of human activity. It would be probably beneficial for the long-term survival of this population to have access to large and suitable refuges that would remain untouched by events such as the latest deforestation (between 1850 and 1920) or such as the present increase of tourism and ski resorts.

III . 2 PROTECTION AND ACTIONS AGAINST THE DECLINE

Policies to stop the decline have already been initiated. For example, new habitats are provided and protected through collaborations with the forestry services (Mollet & Marti, 2001). Forest management plans following guideline for grouse conservation are already at work in the

Jura Vaudois (Mollet, 2002). Poaching can be also reduced by organising collaborations with hunting societies, as I saw in the Pyrenees with E. Ménoni.
Thus, once new habitats will be available, how can we estimate the chance of population expansion? According to the present study, it appears that capercaillie is able to disperse in spite of the habitat fragmentation and apparent barriers to migration. Furthermore, the possibility that the concept of metapopulation is a good descriptor of capercaillie population genetic structure is discussed by I. Storch and G. Segelbacher (2002). The authors assume the possibility of an extinction and recolonisation of habitat patches. However, whereas long distance communications mediated through the ultra-sonic component of their song (Moss & Lockie. I, 1979) can insure that even isolated males can find partners (pers. obs.) and form new leks to recolonized newly available habitats, a recent study reports suitable patches that are not colonized in the Jura (Sachot, 2002).

The solutions generally proposed by the specialists often appear idealistic: to protect the populations and the habitat, access to large areas of forest should be restricted or banished and forestry should be managed accordingly to capercaillie's preferences. The reality is however way beyond such ecological considerations. Social events, such as motor, bicycle or sleigh races are organised within protected areas and tourists, including myself, regularly wander far off the tracks. Furthermore, choices and policies regarding forest management do not include grouse populations in their priorities. Grouse populations in the Jura and the Pyrenees are indeed in the way of the large economical prospects that their habitat offers to the human population. Furthermore, local and personal initiatives led by motivated individuals, such as the grouse specialists who have offered their help in this study, are seldom supported by the authorities or the government and they sometimes cease their activities because of lack of funding.
I believe that occidental culture is acquiring an ecological consciousness, and that dramatic events such as the massive extermination of species in Europe are less likely to occur now than it has in the past. The evidence for these ecological concerns can be found in the reintroduction policies led for other flagship species such as brown bears or bearded vultures, in the large research efforts in conservation biology funded through universities, or in the monitoring of the return of the wolf in the French and Swiss Alps. However, Capercaillie does not attract the mediatic attention it would take to preserve its natural habitat in Western Europe.

In the past, several capercaillie populations have gone extinct or nearly extinct and were replenished by re-introduction plans:
-The Scottish population went to extinction in the 1785 (Moss, 2000) and was subsequently replenished by whole families of capercaillie imported from Scandinavia. However, this population still undergoing strong decline.
-The Population in the Massif Central in central France was also re-introduced by more than 500 captive bred individuals imported from all over Europe. The objective of this massive importation was to maximise adaptability of the individuals. This plan also failed as the population is now down to below 100. The causes of this failure were not identified, as possible outbreeding depression was not considered in this plan (Nappée, *pers.comm.*).
-The Black Forest population was also replenished by imported individuals, and the status of this population is still at decline.
The Pyrenees also went to near extinction in the early 20th century, but the population's demographics had fully recovered by the 1950's. The population expansion that occurred from the 1920's was probably due to two reasons, as was suggested to me: First, capercaillie has survived during the bottleneck probably using relictual patches of forest that were not valuable enough to harvest. Second, their survival and subsequent expansion might have been facilitated by the absence of black grouse in the area, releasing some of the competition pressure upon capercaillie (Ménoni, pers. comm.).
Therefore, the dynamics of recolonization appears obscure, and more experiments should focus on the causes of the failure of the past reintroductions and on methods to facilitate captive bred individuals to new habitats.

IV PERSPECTIVES ON KIN SELECTION

This thesis has brought more results to support the hypothesis that kin selection and inclusive fitness act on males' aggregations on leks by demonstrating that capercaillie males display among related individuals. However, my results also suggest that the observed distribution of relatedness value within lek is partly due to natal phylopatry. To clearly

demonstrate whether group or kin selection can act on both natal phylopatry and kin recognition, experimental manipulation on leks will be necessary.

Brood could be manipulated and eggs moved among nests to measure the ability of males to find leks composed of kindred. This experiment would reproduce the one conducted in peacocks (Petrie *et al.*, 1999).

Measuring the rate of aggressive behaviour between local males and immigrants, and finally assessing the reproductive success of phylopatric and dispersing males would also provide some insight into this process.

Finally, estimating male mating success or number of offspring as a function of relatedness within leks would demonstrate the possibility of kin selection on leks. However, definite understanding of this behaviour would ultimately come from the identification of the genes involved. Nevertheless, because of its threatened status and of its long generation time, capercaillie is not a good model for genetic studies, and results of such studies would be obtained after many years of hard labour and would probably also suffer from lack of statistical power...

V REFERENCES

Balloux F (2001) EASYPOP (version 1.7): A computer program for population genetics simulations. *Journal of Heredity* **92**, 301-302.

Bouzat JL and Johnson K (2004) Genetic structure among closely spaced leks on a peripheral population of lesser prairie-chickens. *Molecular Ecology* **13**, 499-505.

Brenot JF, Catusse M and Menoni E (1996) Effet de la station de ski de fond du plateau de Beille (Ariège) sur une important population de grand tétras *Tetrao urogallus*. *Alauda* **64**, 249-260.

Clark AM, Bowen BW and Branch LC (1999) Effects of natural habitat fragmentation on an endemic scrub lizard (*Sceloporus woodi*): an historical perspective based on a mitochondrial DNA gene genealogy. *Molecular Ecology* **8**, 1093-1104.

Cornuet JM and Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001-2014.

Galbusera P, Githiru M, Lens L and Matthysen E (2004) Genetic equilibrium despite habitat fragmentation in an Afrotropical bird. *Molecular Ecology* **13**, 1409-1421.

Goudet J (1995-2001) FSTAT (Version 2.9.3): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Hardy OJ and Vekemans X (2002) SPAGeDi : a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.

Helle P, Kurki S and Lindén H (1999) Change in the sex ratio of the Finnish capercaillie *Tetrao urogallus* population. *Wildlife Biology* **5**, 25-31.

Hitchings SP and Beebee TCJ (1996) Persistence of British natterjack toad bufo *Calamita laurenti* (Anura:Bufonidae) populations despite low genetic diversity. *Biological Journal of the Linnean Society* **57**, 69-80.

Hjeljord O, Wegge P, Rolstad Jo, Ivanova M and Beshkarev AB (2000) Spring-summer movements of male capercaillie *Tetrao urogallus*: A test of the 'landscape mosaic' hypothesis. *Wildlife Biology* **6**, 251-256.

Hoglund J, Alatalo RV, Lundberg A, Rintamaki PT and Lindell J (1999) Microsatellite markers reveal the potential for kin selection on black grouse leks. *Proceedings of the Royal Society of London Series B-Biological Sciences* **266**, 813-816.

Holzinger Jo and Rosler M (1990) The occurence of the capercaillie (*Tetrao urogallus- major* CL Brehn, 1831) at the mount Athos (Greece). *Journal of Ornithology* **131**, 95-99.

Idaghdour Y, Broderick D and Korrida A (2003) Faeces as a source of DNA for molecular studies in a threatened population of great bustard. *Conservation Genetics* **4**, 789-792.

Kaeuffer R, Pontier D, Devillard S and Perrin N (2004) Effective size of two feral domestic cat populations (*Felis catus* L.): effect of the mating system. *Molecular Ecology* **13**, 483-490.

Keller I and Largiader CR (2003) Recent habitat fragmentation caused by major roads leads to reduction of gene flow and loss of genetic variability in ground beetles. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 417-423.

Kokko H and Brooks R (2003) Sexy to die for ? Sexual selection and the risk of extinction. *Annales Zoologici Fennici* **40**, 207-219.

Kokko H, Sutherland WJo, Lindström Jo, Reynolds JoD and Mackenzie A (1998) Individual mating success, lek stability, and the neglected limitation of statistical power. *Animal Behaviour* **56**, 755-762.

Lens L, Van Dongen S, Galbusera P*, et al.* (2000) Developmental instability and inbreeding in natural bird populations exposed to different levels of habitat disturbance. *Journal of Evolution Biology* **13**, 889-896.

Lindstrom J, Ranta E, Kaitala V and Linden H (1995) The clockwork of Finnish tetraonid population dynamics. *Oikos* **74**, 185-194.

Manel S, Berthier P and Luikart G (2002) Detecting wildlife poaching: Identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. *Conservation Biology* **16**, 650-659.

Menoni E (1991) *PhD thesis: Ecologie et dynamique des populations du grand tétras dans les Pyrénées, avec des reférences spatiales à la biologie de la reproduction chez les poules: quelques applications à sa conservation*, Université Paul Sabatier.

Mollet P (2002) Mesures d'urgence en faveur du Grand Tétras. In: *Fiche Info Grand Tétras*. Sation ornithologique suisse de Sempach.

Mollet P and Marti C (2001) Grand Tétras et gestion de la forêt (ed. OFEFP), Berne.

Moss R (2000) Second extinction of capercaillie (*Tetrao urogallus*) in Scotland ? *Conservation Biology* **10**, 255-257.

Moss R and Lockie. I (1979) Infrasonic components in the song of the Capercaille Tetrao urogallus. *Ibis* **121**, 95–97.

O'Ryan C, Harley EH, Bruford MW*, et al.* (1998) Microsatellite analysis of genetic diversity in fragemented South African buffalo populations. *Animal Conservation* **1**, 85-94.

Petrie M, Krupa A and Burke T (1999) Peacocks lek with relatives even in the absence of social and evironmental cues. *Nature* **401**, 155-157.

Piggott MP, Bellemain E, Taberlet P and Taylor AC (2004) A multiplex pre-amplification method that significantly improves microsatellite amplification and error rates for faecal DNA in limiting conditions. *Conservation Genetics* **5**, 417-420.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Raymond M and Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rowe G, Beebee TJC and Burke T (2000) A microsatellite analysis of natterjack toad, *Bufo calamita,* metapopulations. *Oikos* **88**, 641-651.

Ryman N and Jorde PE (2001) Statistical power when testing for genetic differentiation. *Molecular Ecology* **10**, 2361-2373.

Sachot S (2002) *PhD thesis : Viability and management of an endangered capercaillie (Tetrao urogallus) metapopulation*, Université de Lausanne.

Sachot S, Leclercq B and Montadert M (2002) Population trends of Capercaillie (*Tetrao urogallus*) in the Jura mountains between 1991 and 1999. *Game and Wildlife Science* **19**, 41–54.

Saniga M (2003) Causes of the population decline in capercaillie (*Tetrao urogallus*) in the West Carpathians. *Biologia* **58**, 265-273.

Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. *Molecular Ecology Notes* **2**, 367-369.

Segelbacher G, Hoglund J and Storch I (2003) From connectivity to isolation: genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology* **12**, 1773-1780.

Segelbacher G and Storch I (2002) Capercaillie in the Alps: genetic evidence of metapopulation structure and population decline. *Molecular Ecology* **11**, 1669-1677.

Snow DW and Perrins CM (1998) *The Birds of the Western Palaearctic* Oxford University Press, London.

Storch I (1993) *Habitat use and spacing of capercaillie in relation to forest fragmentation patterns*, University Ludwig-Maximilian.

Storch I (2000) Grouse status survey and conservation action plan 2000-2004. In: *WPA-Birdlife-SSC Grouse Specialist Group*. IUCN, Switzerland.

Taberlet P, Griffin S, Goossens B*, et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Valière N (2002) Gemini- Manual.

Valière N (2002) GIMLET: A computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**, 377-379.

Valière N, Berthier P, Mouchiroud D and Pontier D (2002) Gemini: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83-86.

Van den Bussche RA, Hoofer SR, Weidenfeld DA, Wolfe DH and Sherrod SK (2003) Genetic variation within and among frangemented populations of lesser prairie-chickens (*Tympanuchus pallidicinctus). Molecular Ecology* **12**, 675-683.

Vitalis R and Couvet D (2001) Estim 1.0: a computer program to infer population parameters from one- and two-locus gene identity probabilities. *Molecular Ecology Notes* **1**, 354-356.

Watson A, Moss R, Parr R, Mountford MD and Rothery P (1994) Kin Landownership, Differential Aggression between Kin and Nonkin, and Population Fluctuations in Red Grouse. *Journal of Animal Ecology* **63**, 39-50.

Weir BS and Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Section - 4 APPENDIX

III APPENDIX I : PROTOCOLS

III . 1 PROTOCOL FOR FAECAL DNA EXTRACTION

Qiagen Stool Kit – Qiavac24
July 17, 2001

III. 1. 1 -Stool processing

Turn on eating block : 70°C
For 11 samples to process in one session:
Two 2ml tubes per sample, plus un tube for negative control =(23 tubes)
Use the tube opener! Do not handle open sample tubes

1. fill up the tubes up to 2/3 with sample material. Break the structure of the stool with forceps
2. Add 1.5ml de ASL buffer
3. Vortex twice vigorously until homogenization
4. Wash the outside of the tubes with javel water in a beaker
5. Incubate 10min at 70°C
6. Change gloves, wash tube-racks, vortex and vortex
7. Vortex again samples twice
8. Centrifuge for 2 min at 10 000 rpm
9. Turn the tubes ½ a tour in the centrifuge
10. Centrifuge for 2 min at 10 000 rpm

Prepare: one 2 ml tube per sample, plus one for neg. control
Transfer ½ Inhibitex ® tablet in each tube

III. 1. 2 -Inhibitors adsorption

11. Transfer the surnatant (ca. 1.4 ml) into the tubes with Inhibitex. Discard the pellet
12. Vortex 10 sec.
13. Incubate 1 min at room temp.
14. Vortex all tubes twice
15. Centrifuge for 2 min at 10 000 rpm
16. Turn the tubes ½ a tour in the centrifuge
17. Centrifuge for 2 min at 10 000 rpm

Prepare: two 2 ml tubes per sample, plus one for neg. control =(23 tubes), each containing :

- 25 µl Prot K
- 600 µl AL buffer (lysis buffer)

III. 1. 3 -Lysis and Precipitation

18. Transfer the surnatant (ca. 600 ml) into the tubes with 25 µl Prot. K & 600 µl de AL buffer
19. Vortex
20. Incubate 10 min. at 70

Prepare: Quiavac
Pre-heat AE buffer on the heating block
Prepare twelve 1.5 ml tubes
Prepare labels for the extract tubes.

21. After incubation, add 600 µl Ethanol (stored at -20°C)
22. Vortex
23. Centrifuge 10 sec.

III. 1. 4 -Extraction

24. Transfer the whole supernatant into one Qiagen column per sample (several times necessary)
25. Add 600 µl AW1 buffer to the column
26. Centrifuge 30 sec. at 8 000 rpm.
27. Add 600 µl AW2 buffer to the column
28. Centrifuge 30 sec. at 8 000 rpm.
29. (repeat steps 27 and 28 if necessary for dark samples)
30. discard the collector tubes then transfer the column into new collector tubes
31. Centrifuge 3 min. at max. speed to dry the alcohol out

III. 1. 5 -Elution

32. Change gloves
33. Transfer the columns into 1.5 ml tubes
34. Add 200 µl AE buffer
35. Incubate 5 min. at room temp
36. Centrifuge 1.5 min at 10 000 rpm.
37. Add 175 µl AE buffer
38. Centrifuge 1.5 min at 10 000 rpm.
39. Discard the column (CAREFUL: DO NOT DISCARD THE TUBES!!)
40. Close the tubes and apply label
41. Store extracts in the refrigerator until PCR check

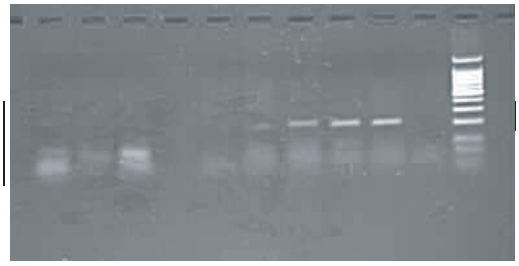
III . 2 AMPLIFICATION

All calculations required to process the numerous PCR protocols were handled automatically in an Excel sheet by a series of short scripts that I created in "Visual Basic for Application" or "VBA macros" (Microsoft®Excel 2002 SP3). VBA offers functionalities to handle specifically automated calculations that I could apply to my work. Furthermore, I used the same program Excel as data base for my genotypes and other applications. This choice facilitates the compatibility between the softwares used in the first steps of the analysis of data.
The PCR protocols were managed as follows:
One worksheet contains the list of PCR primers available in the lab stock, with their respective concentrations. The second worksheet displays all reactant to the PCR, with "listbox" offering appropriate concentrations to choose by the user, the choice of the Taq, the quantity of Taq to use, various stocks of BSA, dNTP's, and buffers available in the lab. A choice of "clickbox" are linked to the primers data base on the first worksheet and to calculation parameters for various amplification protocole: Touch-down, long (Time Release) PCR, sexing and microsatellites. Finally, the major part of this second worksheet is dedicated to the calculation of reactant concentrations to prepare master-mixes for amplification, to which I added the designated primer.
Finally, the last worksheet is linked to the second and displays a print version that is clear, double checked, and contains all information required to move directly to the lab, as follows:

TYPAGE PL8-1-2

Organisme: : stool
Objectif de la PCR: : Typage
Date : 18-Jul-04
Taq : **Perkin Elmer Gold**
Primers :

Nombre de tube	:	**1**			
vol. par tube	:	**25**			
Concentrations	:	Stock	0		tube1
MgCl2 (mM)	:	25	2.5	mM	tube2
BSA(mg.ml-1)	:	20	0.15	mg.ml-1	tube3
dNTP (mM)	:	2.5	0.1	mM	tube4
Primer#1 (M)	:	10	0.3	M	tube5
Primer#2 (M)	:	10	0.3	M	tube6
Taq.(U/l)	:	5	0.5	U/l	tube7
Mix					tube8
H2O	:	20.81875	l		tube9
PCR Buffer	:	3.75	l		tube10
Qsol	:	0	l		
MgCl2	:	3.75	l		
BSA	:	0.28125	l		
dNTP	:	1.5	l		
Primer #1	:	1.125	l		
Primer #2	:	1.125	l		
T.aq.	:	0.15	l		
vol. mix:	:	32.5	l		
Post-mix					
dispatch	:	20	l		
template	:	5	l		
PCR conditions.					
hot start: 95°C 5 mn		touchdown	*non*		
nb. De cycles			50		
denat.		94	°C	45sec	
bind.		**59**	°C	45sec	
elong.		72	°C	45sec	
ending: 72°C 5 mn			*oui*		



IV APPENDIX II : ASSESSING THE RELIABILITY : GENOTYPING REPEATS

IV. 1. 1 Cross reference among gels

All the amplification products were run on an electrophoresis apparatus (ABI377 Automated Sequencer). As we measure the size polymorphism of the amplification products, all gels run on the ABI377 include a size reference in each lane. However, the electrophoresis process is highly sensitive to buffer quality and to the room temperature and humidity. Furthermore, different gel batches can produce discrepancies among gels. It is therefore common procedure, when genotyping a large number of sample, to use a reference sample that will be loaded on all gels and will be the size standard.
To this purpose, I used two DNA samples extracted from liver (Table 1), diluted 1/1000, that I included into each PCR plate and subsequently, into each gel.

Table 1: Reference genotypes
Two liver samples used as cross reference between gels. The DNA template of these samples was diluted at 1/1000 to insure sufficent quantity all through the project and also to avoid artefacts due to the large number of amplification cycles

loci	TTD6	TD2	TUT1	TUT2	TUD3	TUD5	TTD1	TUT4	TUD1	TTT1	TUT3	TUD6
Female	117121	164175	216216	159175	192192	189191	160168	175179	202208	212228	155159	194202
Male	119134	168168	212212	159163	192192	189189	160168	159175	202202	212212	155159	212212

IV. 1. 2 Preliminary assessment of reliability

Using DNA from faecal source bears several impairments and constraints that necessitate the use of particular procedures.
First, the DNA contained in faecal samples is highly degraded and contains PCR inhibitors. Thus, the extraction procedure must preserve and not degrade further the DNA, and also must remove the PCR inhibitors.
Second, the low quantity of DNA contained in faecal samples generates genotyping mistakes. These mistakes are allelic drop out and false alleles.

The allelic drop out (ADO) occurs when allele copies are not amplified because they were lost during the liquid handling process or by PCR competition with another allele.

False alleles (FA) occur randomly during the amplification cycles as a PCR artefact. During the amplification cycles, a large part of the amplicons bears mutations that modify their size, giving raise to the "tuttle" effect. When one of these mutations occurs at very early stage of the amplification procedure, it can be over-amplified and appear as an allele during gel reading.

To assess the reliability of the genotypes, I proceeded to a multiple repeats genotyping test (Table 2). I first genotyped 8 samples eight independent times each to estimate ADO and FA rates.

I also genotyped 13 further samples using the multitube approach (detailed in the second chapter of the second section of this study). This approach implies to accept heterozygous genotypes as certain when they are genotyped identically two or three times.
When a genotype appears ambiguous or homozygous after the first two or three repeats, it is repeated again three or four further times, according to the following scheme (from Taberlet *et al.* 1996, *Nucl. Ac. Res.*).

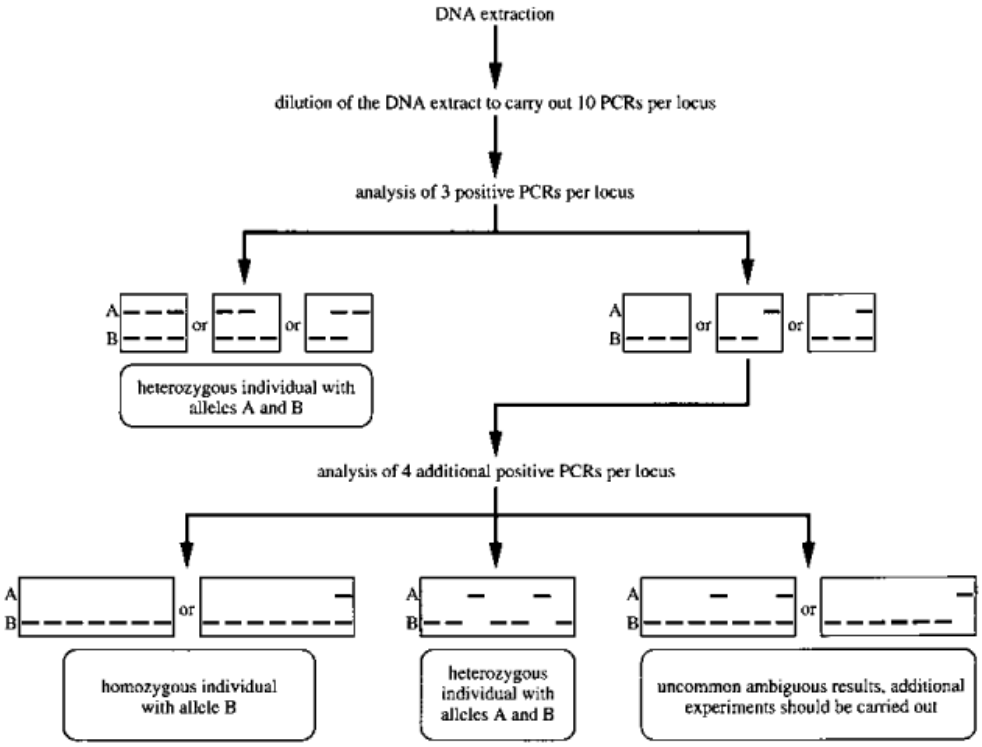


Table 2: Assessment of reliability: testing the multitube approach
To estimate the rates of allelic drop out, false alleles and PCR failure, I conducted the following experiment: I genotyped 8 samples in 8 independent repeated PCR procudures. Furthermore, I tested this approach by sorting out all homozygote scores after the third repeat. Do and FA rates were comparable and low compared to other studies.

	8 repeats Grand Total Nb of PCR : 583 8 individuals			3 Repeats then 4 for uncertain typings Grand Total Nb of PCR : 604 13 individuals		
	False Alleles	Drop Out	Failed	False Alleles	Drop Out	Failed
min:	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
maxi:	9.5%	20.8%	33.8%	18.9%	12.0%	21.6%
average across individuals:	2.5%	0.6%	11.6%	7.3%	5.9%	6.5%
SD:	3.3%	7.9%	10.2%	6.4%	4.5%	6.3%
min across loci:	0.0%	0.9%	0.0%	2.2%	0.0%	1.4%
max across loci:	11.7%	24.4%	34.3%	15.7%	13.5%	13.5%
average across loci:	3.0%	9.9%	11.4%	6.7%	5.9%	6.0%
SD:	3.6%	8.4%	10.9%	4.7%	3.9%	4.5%
total:	2.8%	9.9%	11.5%	6.8%	5.4%	6.3%

IV. 1. 3 Blind test.

The reliability of the genotyping was also tested by parallel genotyping of 6 samples over 8 loci in two independent laboratories (ours and the WSL, in Birmensdorf, conducted by G. Jacob) following similar procedures (Table 3).
Out of the 96 alleles genotyped, only one showed discrepancy due to ADO that had not been detected (underlined, Table 3). One locus in one individual was not genotyped in either lab, suggesting the presence of null alleles. Finally, one genotype failed.

Table 3: Comparative test with the WSL
This test was conducted jointly at the LBC and at the WSL in Birminsdorf in collabaration with G. Jacob. A set of 8 samples was genotyped for the same loci in Lausanne and in Birminsdorf and the scores were compared for consistency. Only one allele out of the whole set showed discrepancy

loci	TuD1		TuT4		TuD3		TuT1		TuD6		TuT3		TuD5		TuT2	
sample	WSL	LBC	WSL	LBC	WSL	LBC	WSL	LBC	WSL	LBC	WSL	LBC	WSL	LBC	WSL	LBC
Mau5	153153	202202	133137	171175	8989	192192	128128	216216	163169	200206	96100	159163	126136	181189	147147	163163
Mau7	153153	202202	121137	159175	8991	192194	128128	216216	143169	178206	96100	159163	136136	189189	143147	159163
Mau7.1	153153	202202	121137	159175	-	192194	-	-	143169	178206	96100	159163	136136	189189	143147	159163
Mau9	153163	202212	137137	175175	8989	192192	128128	216216	163163	200200	92100	155163	126136	179189	147147	163163
Mau10	153163	202212	137137	175175	8989	192192	128128	216216	163163	200200	92100	155163	126136	181189	147147	163163
Mau11	153255	202202	137145	175183	8991	192194	132132	220220	163175	200212	96100	159163	136136	189189	143147	159163

IV. 1. 4 Two-step amplification protocol

Finally, in order to improve the genotyping quality, we also tested various protocols, among which the two step-amplification procedure proposed by Taberlet and coauthors (Piggott *et al.*, 2004) that we tested in parallel with *Canis sp.* (Tab le 4 & 5). This protocol did not significantly increase the efficiency of our protocols (and even sometimes decreased it).

Table 4: Pre-amplification protocole
We tested the possibility of increasing genotyping quality by an alternative protocole proposed by P. Taberlet. This protocole consisted in a two steps amplification of the template by 2 successive reactions.

All [primers]fin=0.01mM, 50°	
nb primers	6
number of tubes:	1
vol. final (ml):	50
ml DNA (per tube):	12.0
[Buffer] initial (xX):	10
[Buffer] final (xX):	1
[primers] initial (mM):	10
[primers] final (mM):	0.01
[dNTPs] initial (mM each):	2.5
[dNTPs] final (mM each):	0.2
[BSA] initial (mg/ml)	10
[BSA] final (mg/ml)	0.2
[MgCl2] initial (mM):	25
[MgCl2] final (mM):	2.5
[Taq Gold] initial (U/ml)	5
[Taq Gold] final (U/ml)	1
dH2O	22.20
Buffer	5.00
MgCl2	5.00
dNTPs	4.00
Primer F !each primer!	0.05
Primer R !each primer!	0.05
BSA	1.00
Taq	0.20
	38.00
Total (ml Mix per tube):	38.00

Cycles (pre-amp)	
initial denaturation	5' 95°
denaturation	30" 95°
annealing,	30" 50°
elongation	1' 72°
final extension	5' 72°

Cycles (2cd amp)	
initial denaturation	5' 95°
denaturation	30" 95°
annealing,	30" TA
elongation	1' 72°
final extension	5' 72°

Table 5: Numeric results of the Pre-amplification protocole
This protocol did not increase (or sometimes decreased) the efficiency of genotyping

Comparison multitube approach / PreAmpli-ReAmpli Canis lupus & Canis fam.									Tetrao urogallus								
summary of results																	
	Dropout		False allele		Double error		Complete Error		paired t	Dropout		False allele		Double error		Complete Error		paired t :
Across loci	0.236	0.192	0.006	0.014	0	0	0	0.003	0.5421	0.161	0.174	0.019	0.033	0.031	0.009	0.013	0.004	0.9166
Across sples	0.209	0.17	0.005	0.013	0	0	0	0.003	0.5629	0.137	0.124	0.009	0.027	0.031	0.009	0.013	0.004	0.505
Across PCR	0.236	0.172	0.004	0.005	0	0	0	0.002		0.156	0.133	0.01	0.031	0.025	0.007	0.011	0.005	0.5559
Total nr	51	40	1	1	0	0	0	1		40	32	2	6	11	3	5	2	
missing data	74	92								35	40							

% positive PCR				
Locus				
FH2140	0.95	0.94	0.94	0.94
PEZ17	0.92	0.9	0.98	0.9
FH2161	0.95	0.97	0.98	0.97
FH2096	0.99	0.95	0.96	0.99
FH2137	0.98	0.97	0.99	0.99
FH2054	0.93	0.93	0.99	0.99
Mean	0.95	0.94	0.97	0.97
paired t-test, Pval=		0.1975		

V APPENDIX III : GENOTYPING DATA

V . 1 THE JURA

Together with the help of Sébastien Sachot, of Dr. Bernard Leclercq and the group Tétras Jura, of Alain Bloc and the employees of the Reserves Naturelle de la Haute Chaîne du Jura and finally of the rangers of the Conservation de la faune du Canton de Vaud under the direction of Dr. Cornelis Neet, we collected approximately 400 samples of grouse stool along the Jura Mountains massif.

Among these samples, about 20 where identified by sight as being hazel grouse stool and were removed from the set of samples. The hazel grouse, *Bonasa banasia*, is a close relative to the capercaillie and another study demonstrated that the microsatellite markers used in this study can also be used for the genus *Bonasa* (Larsson *et al.*, 2003).

Following the protocols described in the Appendix I, I genotyped 283 samples. I selected the samples for extraction based on their location of origin and also on the date of sampling to avoid duplicating the experiments for the same individuals. However, among the 283 genotypes, I found 41 duplicates that were removed from the data set. Among these duplicates, only two pairs (four genotypes) were sampled in different seasons, suggesting migrations of 10 and 20 km respectively. These two pairs are the only evidence for long range dispersal in my study. All other duplicates had been sampled at the same place and date.
The final sample size was 238 samples (Table 1), containing 8.5% missing allele data, due to either PCR failure or ambiguous genotypes.

The way I selected the samples for extraction aimed at saving lab work. It has however also prevented me from performing an estimation of population size as in Kohn *et al.* (Proc. R. Soc. London B, 1999). I believe this type of study cannot be performed when census population size is estimated at over 600 individuals, as it is the case in the Jura.

V . 2 THE PYRENEES

Dr. Emmanuel Ménoni and the rangers of the Office National de la Chasse et de la Faune Sauvage in the Pyrenees provided me with approximately 300 stool samples of capercaillie mostly collected around leks, during population counts.
Among these samples, I genotyped 189 samples. After removing potential duplicates and genotypes with more than 50% missing data, the set was composed of 173 containing 5% missing data (Table 2)

Among the 238 samples available in the Jura, 121 males were sampled during the breeding season. Among these 121, 69 samples could be assigned precisely and with certainty to their lek place (Table 3). These assignments were performed by Sébastien Sachot, Bernard Leclercq and Alain Bloc. I used these 69 samples for the study on the capercaillie mating system to study the relatedness among males lekking together.