Characterisation of the cattle, buffalo and chicken populations in the northern Vietnamese province of Ha Giang

Cécile Berthouly

Thesis

to obtain the degree

DOCTEUR D'AGROPARISTECH

Field: Animal Genetics

presented and defended by

Cécile BERTHOULY

on May 23rd, 2008

Characterisation of the cattle, buffalo and chicken populations in the Northern Vietnamese province of Ha Giang

Supervisors: Jean-Charles MAILLARD and Etienne VERRIER

Committee

Steffen WEIGEND	Senior scientist, Federal Agricultural Research Centre Neustadt (Germany)	Referee
Bernard FAYE	Senior scientist, CIRAD, Montpellier (France)	Referee
Thierry ROBERT	Associate Professor, University Paris-VI (France)	Examiner
Bernard BIBÉ	Senior scientist, INRA, Toulouse (France)	Examiner
Jean-Charles MAILLARD	Senior scientist, CIRAD, Hanoi (Vietnam)	Examiner
Etienne VERRIER	Professor, AgroParisTech (France)	President

Remerciements

Je voudrais remercier en premier lieu ma famille: mes parents qui m'ont donné l'envie de faire ce métier et qui m'ont toujours encouragée, à Anne pour ces longues discussions animées de statistiques depuis l'autre bout du monde, à Matthieu et Camille, pour leur soutien et leur patience, et à Lucien pour sa gentillesse.

A mes amis qui pour certains m'accompagnent depuis le lycée:

A Perrine pour ses blagues sur la prof de math et pour avoir trouvé un mari qui est presque aussi maladroit que moi (merci Julien);

A Aurélia pour ces crises de rire que rien ne perturbe;

A Marie pour les aprems Roland –Garros entre deux révisions de partiels et son calme imperturbable ;

A David, quien sabe como hacer que un burro se vuelva burra ;

A Adeline et Natalia pour les apéro-stats ;

A Josué, à qui je dois beaucoup ;

A Tim pour ses discussions parapente-escalade, pour ses oublis perpétuels, sa naïveté (bien qu'il pense que je le sois plus que lui!), à notre projet

A Gab pour ses discussions philosophiques du petit déjeuner, nos chorégraphies, nos interprétations musicales de variétés internationales, et nos longues soirées passées sur skype pendant que chacune travaillait à l'autre bout du monde ;

A Oliv pour sa gentillesse, et pour nous avoir supportées Gab et moi ;

A Sophie pour sa spontanéité, son humanité et sa joie ;

A Cédric, pour son accueil, les débats philosophiques, l'œuf de canard embryonné et ses repositionnements ostéologiques!

A Miguel pour les cochons d'Inde entre midi et deux, et ses longues descriptions pâtissières ;

A Irène avec qui je l'espère nous réussirons a faire une 6b+ avec une bière à la main ;

A mes collègues de l'unité qui font que c'est un plaisir de faire une pause café, particulièrement à Grégoire et sa diversité de blagues dont la rapidité d'imagination restera pour moi la plus grande énigme à résoudre et Marianne ma colocataire de l'INAP-G qui a rendu les longues soirées d'hivers moins monotones.

All this work was only possible thanks to the field team. They were not only coworkers but also roommates and friends. For western people we feel as a chance to work in a remote mountainous area, for Vietnamese people it may be the worst thing it could happened to them. After the first month on the field, the director of NIAH required a meeting with Mr Thu, he was really surprised about the extensive amount expenses, Mr Thu answered : "This is because we had worked 27 days in the month". I will never stop to be full of gratitude to my Vietnamese field coworkers for working so hard without complaining under the steady rhythm and hard conditions I asked them:

To Thu, who since the beginning of this adventure always considered me as a individual and not as a foreigner, who always try to teach and trust in me and so many other things;

A Hai, qui m'a appris le congolais, a mangé mes peaux de poulets et n'a cessé de me faire rire pendant les deux années de terrain ;

To Cong, the only driver which can take blood sample as well as he drives;

To Binh who initiate me to the underground life in a remote area;

To Hue who negotiate hard the price of my furniture I tried to sell to her;

I would like also to thanks Mr Vo Van Su, Tran Thi Chau, Dinh Thi Dung, the team from NIAH laboratory and all the students who help in the accomplishment of this project.

To the farmers that I want to thanks through this work and the projects I hope it will create.

To the children who ran across paddy rice fields in order to catch all the chickens we need and their welcoming smiles.

I would like to thanks Mss Pabiou for linguistic revisions.

Merci à Michèle Tixier-Boichard, Bertrand Bed'hom, Denis Laloë, Renaud Lancelot et Stéphane Robin pour leurs précieux conseils.

Je souhaiterais remercier Xavier Rognon pour son aide et sa disponibilité.

J'ai eu la chance d'avoir des directeurs de thèse, Prof. Etienne Verrier et Dr Jean-Charles Maillard, qui m'ont toujours soutenue dans mes choix, et grâce auxquels ce travail a pu aboutir, merci.

I would like to thank the direction of the DARD who allow me to work in good conditions and everywhere in the province, but also the NIAH direction who receive me.

The French Ambassy, the Vietnamese ministry of MARD, the CIRAD and INRA institutes which financed the project and this PhD thesis.

Table of Contents

List of Tables

List of Figures

CHAPTER 6

List of Boxes, Pictures* and Annexes

*Pictures used in this thesis were taken by the field team of BIODIVA project and Dr. J-C Maillard.

INTRODUCTION

1 INTRODUCTION

The majority of the world's livestock is from Asian origin (Bruford *et al.,* 2003). The Near East is considered as the most important cradle of animal domestication, where sheep, goats, taurine (humpless) cattle and pig were domesticated (this last species having also been domesticated in the Far East). The earliest evidence for zebu (humped) cattle derives from an archaeological site located in Baluchistan.. While these four species – sheep, goats, cattle, pig - dispersed throughout the world, other species, such as buffalo, yak and Bactrian (two-humped) camel, have retained narrower adaptations and remain typically "Asian" livestock. Minor cattle species exist in both wild and domesticated forms: the banteng (Bali cattle) and the gaur (mithun or gayal). Asia has also given rise to most of our poultry. The chicken is derived from the Red Jungle fowl in India and South-East Asia (Liu *et al.* 2007, Oka *et al.* 2006).

As a consequence, the many livestock breeds from Asia arose of the continent's enormous cultural and ecological diversity. In the thousands of years since animals were first domesticated, they have been introduced to different environments and subjected to breeding practices of different communities and social groups. A large variety of breeds are adapted to various ecological niches and to the differing needs and preferences of their breeders. The earliest evidence for the existence of distinct populations within a given species in Asia dates back to the 3rd millennium BC. Mesopotamian archives refer to several different breeds of sheep. Pictorial records from this time also show sheep with different horn forms. In China, a large number of different pig breeds are recorded for the Ming period (AD 1500-1644).

1.1 What is a breed?

"*One of the big problems facing genetic resource specialists, especially those working on livestock, is that of determining whether animals or populations belong to different breeds or whether they represent variations within a single breed. In Africa, in particular, livestock breeds tend to occur across countries or even regions – the East African Zebu, the Boran cattle in Ethiopia, Kenya and Somalis; Nungi or Nkone cattle in Swaziland, Zimbabwe and Botswana; and Djallonke sheep in West Africa, to name but a few. These animals are known by the same name in different places, but often look quite different from one place to another. Conversely there are breeds that look alike but have different names in different places*" (Source: ILRI, 1996)."

Scientists usually define a breed as "a group of animals with definable and identifiable external characteristics that distinguish it from other groups within the same species". The term "breed" is often associated with breeders associations and herd (or flock) books. Usually groups of breeders get together to formulate a breeding goal, and then select animals that fulfil the desired criteria. These animals are regarded as the founder population. Usually only the offspring of the founder animals are eligible for inclusion in the herd book. This process was pioneered in England with the studbook for the Thoroughbred Horse in 1791, and has since been repeated many times. Most breeders associations specify the standard of the breed, generally based on traits easy to record and with a simple determinism, i.e. coat colour, presence/absence of horns, etc. The definition of the standard allows the animals to look similar. Breed societies and written records such as herd books are rare or absent in developing countries. So scientists often have problems deciding whether a given animal population should be regarded as a "breed".

BOX 1 Some definition of a breed

" Either a subspecific group of domestic livestock with definable and identifiable external characteristics that enable it to be separated by visual appraisal from other similarly defined groups within the same species, or a group for which geographical and/or cultural separation from phenotypically similar groups has led to acceptance of its separate identity.
Note: Breeds have been developed according to geographic and cultural differences, and to meet human food and agricultural requirements. In this sense, breed is not a technical term. The differences, both visual and otherwise between breeds account for much of the diversity associated with each domestic animal species. Breed is often accepted as a cultural rather than a technical term"
Scherf (2000)

"In the strictest sense, a breed designates a closed or partially closed population. Mating pairs are drawn only from within the population and relationships among individuals are documented. Members of the breed have developed under the same selection pressures and share common ancestry"
Rege (2003)

"A breed is a population of animals sharing enough traits to be considered homogenous by one or several groups of breeders who agree on the organization of the choice of reproducing animals and of the required exchanges of animals, including international exchanges"
French Regulation on Animal Breeding (2006)

But if we look at how indigenous communities manage their animal genetic resources, we see that they may actually function much like a breeder association. Instead of written rules or bylaws, they have culturally embedded breeding goals. Individual cultures or communities have very divergent views of what a desirable animal should look like. Breeding stock is passed on from one generation to the next, so its genetic composition remains very stable. Female animals are usually not sold, but are borrowed or exchanged at the time of marriage celebrations or events. They stay in the community, resulting in a closed gene pool.

Thus, the genetic pattern of local breeds reflects the social rules of the community for exchanging animals.

A concise definition of what a breed is (or should be) is given by Köhler-Rollefson (1997):

A breed is a population that belongs to the same species and is
- *kept by a particular community*
- *in a specific environment*
- *subjected to the same utilisation pattern*
- *regarded as distinct by the community that keeps it*

1.2 Flows of animal genetic resources

"Gene flow" (movement and exchange of animal breeds and germplasm) in livestock species has probably occurred since prehistoric times, and has been driven by a range of factors. On a global scale the most significant gene flows have involved the "big six" livestock species: cattle, buffalo, sheep, goats, pigs and chickens. Focusing on mainly these six species, FAO created a global databank DAD-IS where breed information is recorded. However the information available is often sketchy and incomplete. This may be due in part because:

- there are no systematic records of breed population sizes

- breeds from temperate zones are often better defined and documented than breeds from tropical regions

Starting from prehistoric times, genes spread as a result of the dispersal of domestic animals by means of gradual diffusion, migration, exploration, colonisation and trade. Since about the middle of the XXth century, a series of technological advances has facilitated gene flow. Commercial use of semen started in the 1960s, of embryos in the 1980s, and of sexed embryos in the mid-1990s. Towards the end of the XXth century, gene flows to the South (i.e. developing countries) began to be fuelled by a growing number of consumers with a taste for higher standards. The resulting expansion of intensive livestock production systems in developing countries has been termed the "livestock revolution". Gene flows is driven by the desire of producers and breeders to obtain genotypes that perform optimally in a given production environment. Exports generate profits whereas motives for importing genetics can vary. Countries such as China and Brazil are in the process of building up their own intensive production systems and breeding programmes whereas African and South-East Asian countries traditionally import because of the high costs associated with developing their own breeding programmes. One of the best examples of this process involves dairy cattle. Cattle genetics resources are exchanged in the form of live breeding animals, semen and embryos.

By far the most widespread breed is the Holstein-Friesian, which is present in at least 128 countries, and in all regions. However, despite a very large female population size, due to the unbalanced use of a small number of sires at the World level, the genetic basis of the Holstein breed seem to be narrow; as an example, the French Holstein population comprises about 1.9 Millions of recorded cows and its realized effective size was estimated, *via* pedigree analyses, to 42 only (Mattalia *et al.,* 2006).

Gene flows can both enhance and reduce diversity. The type of impact depends on a number of factors, which include the environment suitability in the receiving country, and the organisational structures on both the receiving and providing sides (Mathias & Mundy, 2005). During the period from the beginning of animal husbandry in prehistory to the mid-twentieth century, gene flow generally enhanced diversity.

BOX 2 Comparison of livestock systems

High-performance breeds   &   Locally adapted

Specialised, provide only one product	Multipurpose, provide wide range of product
Require high feed input, including green fodder and concentrate	Subsist on natural vegetation
Require expensive housing and stabilized climate	No, or only, protection necessary
Susceptible to disease	Disease resistant
Need high level of care and time	Need little care
If kept in landless system negative effects on the environment through the accumulation of waste	Positive ecological effects by being integrated into the farm cycle, contributing manure and sometimes draft power
Often compete with human for grain	Utilize vegetation and areas that often cannot be exploited otherwise
Adoption requires large amount of capital	Traditional occupation inherited by forefathers
Lead to wealth differentials and offer benefits for only few	Kept in social contexts which maintain sharing mechanisms

Source: ILRI (1996)

However, during the past four to five decades the development and expansion of intensive livestock production and the export of entire production systems have led to a reduction in diversity through the large-scale replacement of local breeds with a small number of globally successful breeds. In Vietnam, the percentage of indigenous sows declined from 72 % of the total population in 1994 to only 26 % in 2002 (Huyen *et al.*, 2006). Of its 14 local breeds, five breeds are vulnerable, two are in a critical state, and three facing extinction. Extinction of a breed could occur by dilution through indiscriminate crossbreeding with imported stock, often without significant gains in production levels or desirable characteristics. This could lead to the total disintegration of local breeds such as the pure Taiwan Yellow cattle breed, which has been totally replaced by the Taiwan zebu, which results from a crossing between the Red Sindhi, and the Kankrej breeds.

In the context of the ongoing "livestock revolution", it seems likely that the transfer of livestock breeding systems will continue and even increase in pace in the rapidly developing

countries of the South. The crowding out of local breeds is, thus, set to accelerate in many developing countries, unless special provisions are made for their conservation by providing livestock keepers with appropriate support.

BOX 3 Threats to indigenous livestock breeds

- **Crossbreeding with exotic breeds**: breeding programmes have favored the use of exotic breeds for crossbreeding, upgrading or replacement. Although such programmes often fail to achieve their objectives (unless a stable supply of high inputs can be arranged) they have let to the dilution of indigenous breeds.
- **Agricultural intensification**: changes in cropping patterns are a major factor leading to elimination of indigenous farm animals. The switch to certain cash crops (such as tobacco) eliminates crop residues that used to be important component of fodder. Mechanisation for rice cultivation led to an abandon and decrease in swamp buffalo in Thailand.
- **Change in demands**: changes in producer preference, usually in response to changes in socio-economic factors
- **Political conflict and boundaries**: Warfare had contributed to the disappearance of indigenous breeds. Moreover, new boundaries, such as the Indo-Pakistan border, also disrupt migration patterns and undermine animal-dependent livelihoods strategies.

1.3 Importance and conservation of livestock indigenous breed

The term animal genetic resources (AnGR) is used to include all animal species, breeds and strains that are of economic, scientific and cultural interest to humans in terms of food and agricultural production for the present or in the future (Tisdell, 2003). Some 40 domesticated mammalian and avian species are now used to meet our demands for food, clothing, draught power, and manure and to satisfy various cultural or religious purposes. They also facilitate the use of marginal lands of little or no value for crop agriculture. The AnGR may represent the only durable form of storing wealth, which is particularly important where there is no financial system to perform this function. In view of the environmental and disease stresses in the tropics, only locally adapted livestock can serve these purposes, especially in low-input smallholder systems. Among the estimated 6 400 breeds of farm animal species worldwide, about 30 % are endangered and 1 % becomes extinct every year (Scherf, 2000). Over the past 15 years, about 300 breeds of farm animals identified by the FAO became extinct and 1 350 face extinction in the near future (Scherf, 2000). From a global perspective, conservation is not only about endangered breeds but also about those that are not being utilised efficiently.

While, in principle, maintaining a large random mating pool without reference to breed could preserve valuable genes, this practice is not favoured mainly because the genes would be difficult to evaluate and to retrieve. In addition, specific breeds may possess unique allele combinations at different loci as the result of adaptation to different environmental challenges, which would be difficult to recreate. Livestock possessing important quantitative traits resulting from such combinations will only be amenable to assessment and exploitation when managed as discrete breeding entities. Livestock genetic conservation measures are therefore likely focused on the maintenance of breeds. Since the possible changes in the future, such as different market needs, new diseases, climatic changes are unknown, it is essential to conserve diversity in the whole genome, and not only known functional genes. Obviously, activities to preserve between breed variation have to be accompanied by activities to conserve within-breed variation.

Conventional approaches to preserve indigenous animal breeds include *ex situ* and *in situ* conservation. *Ex situ* approaches could be of two types: (i) *in vivo*, i.e. conserving living animals outside the environment where the breed is usually found such as national centers, zoos, etc; (ii) *in vitro*, i.e. conserving genetic material (semen, oocytes,...) in an artificial environment such as for cryoconservation.

Ex situ conservation generated much criticism, because in these systems animals do not coevolve with the environment. The breed of interest conserved at the present type may not be adapted to the future ecological or market environment, and so its survival will be critical. Therefore, *in situ* conservation should be preferred as far as possible, and *ex situ* conservation should be considered as both a complementary way and a very long term insurance. The best way to conserve breeds *in situ* is by maintaining them as part of functional production systems and in the social and ecological contexts in which they were developed . Then, a prerequisite for *in situ* conservation is to have a global view about the breed it will be established for. Understanding the management practices which have led to the current genetic pattern, the needs and constraints on its husbandry, the current genetic status and the characteristics desired by breeders are just a few of the aspects which should be taken into account before undertaking any conservation project.

The context of the research work presented in this thesis was, on one hand, the assumed status of Vietnam as a "hot spot" for animal biodiversity, and on the other hand, the recent development livestock production in Vietnam. In order to supply the recent increase of the demand of animal products such as milk or meat, highly productive breeds have been imported. Therefore, there is an urgent need to better characterise the Vietnamese AnGR and to examine the needs and possibilities to improve but also preserve the indigenous breeds. This is all the more true for areas with natural and economics constraints, where the use of adapted animals to these constraints is of high value. For all these reasons, a French-Vietnamese research program, called BIODIVA (www.biodiva.org.vn), including the present work, was devoted to the survey of livestock populations in a mountainous area of Vietnam, namely the Ha Giang province. The Ha Giang province was chosen for several reasons: as it is one of the poorest provinces of Vietnam and as so a political wish to develop agriculture, presence of various ethnic groups including the H'mong group which is assumed to rear several indigenous breeds culturally identified.

The work described in this thesis is a synthesis of data obtained from three different indigenous livestock species, namely cattle, buffalo and chicken from the Ha Giang province. Each species is reared to fulfil different purposes and with specific management systems by different communities. The goal of this work is twofold: (i) to assess the level of genetic and phenotypic variation within populations from the Ha Giang province in order to provide a framework for rational genetic conservation and breed description and (ii) to correlate the genetic data obtained with the farming systems and management practices. In this study, we will firstly present an analysis of farming systems and breeding practices that mould the genetic characteristics of livestock. Afterwards, for each species we will summarise their role in the local soio-economy , i.e. Asia and Vietnam, then we will present the genetic structure of the species in the Ha Giang province. Last chapter is a synthesis of the major common results between species.

CHAPTER 2

PRESENTATION OF THE REGION UNDER STUDY, MATERIALS & METHODS

2 REGION, MATERIALS & METHODS

2.1 The Ha Giang province

2.1.1 Geography of the Ha Giang province

Vietnam, meaning "Country of Viet people", is situated in the heart of South-East Asia (8°34'N-23°22'N), and stretches along the oriental coast of the Indochinese peninsula over 1700 km from north to south. Its width does not exceed 300 km (50 km for the narrowest zone). With a total land surface area of 331 000 km^2, Vietnam is bordered by the Chinese sea to the east (3260 km of coastline), by China to the north, and by Laos and Cambodia to the west and southwest. With a population reaching 83 million inhabitants, the demographic density average is 254 inhab/km^2. Nevertheless, the distribution of the population over the country is unequal, because the population is mainly concentrated in the plains of the deltas of the main two rivers of the country, namely the Red River in the North and the Mekong River in the South. Therefore among the 59 Vietnamese provinces, a few are densely populated whereas the remaining mountainous provinces are neglected. In Vietnam, there are 54 ethnic groups, among them the Kinh ethnic group, i.e. Viet, represents 80 % of the population. At least 49 of the ethnic minorities live in the mountainous provinces, therefore mainly in the North and the center highlands.

The research area of this study was the Ha Giang province. The Ha Giang province is one of the northern provinces of Vietnam (22°08'-23°19'N; 104°33'-105°33'E). With a total surface area of 7884 km^2, 274 km border with China. In the South-East, there is the Cao Bang province and in the West, the province of Lao Cai which is popular with tourists. The Ha Giang province's terrain is strongly partitioned, which results in reduced communication between villages (see, Castella *et al.* 2005, for an analysis of such reduction in a neighboring Vietnamese province). Ha Giang's topography has sharp variation in altitude. The lowest level is 10 m above sea level in the Ha Giang prefecture, and the highest about 2400 m. The Ha Giang province has a tropical climate with an average temperature of 15°C-16°C. Temperature is unlikely below 5°C. The average rainfall is about 2 000- 2 500 mm/year and the rainfall is different from month to month. The rainy season lasts from April to September.

Figure 2.1 Administrative map of the Ha Giang province.
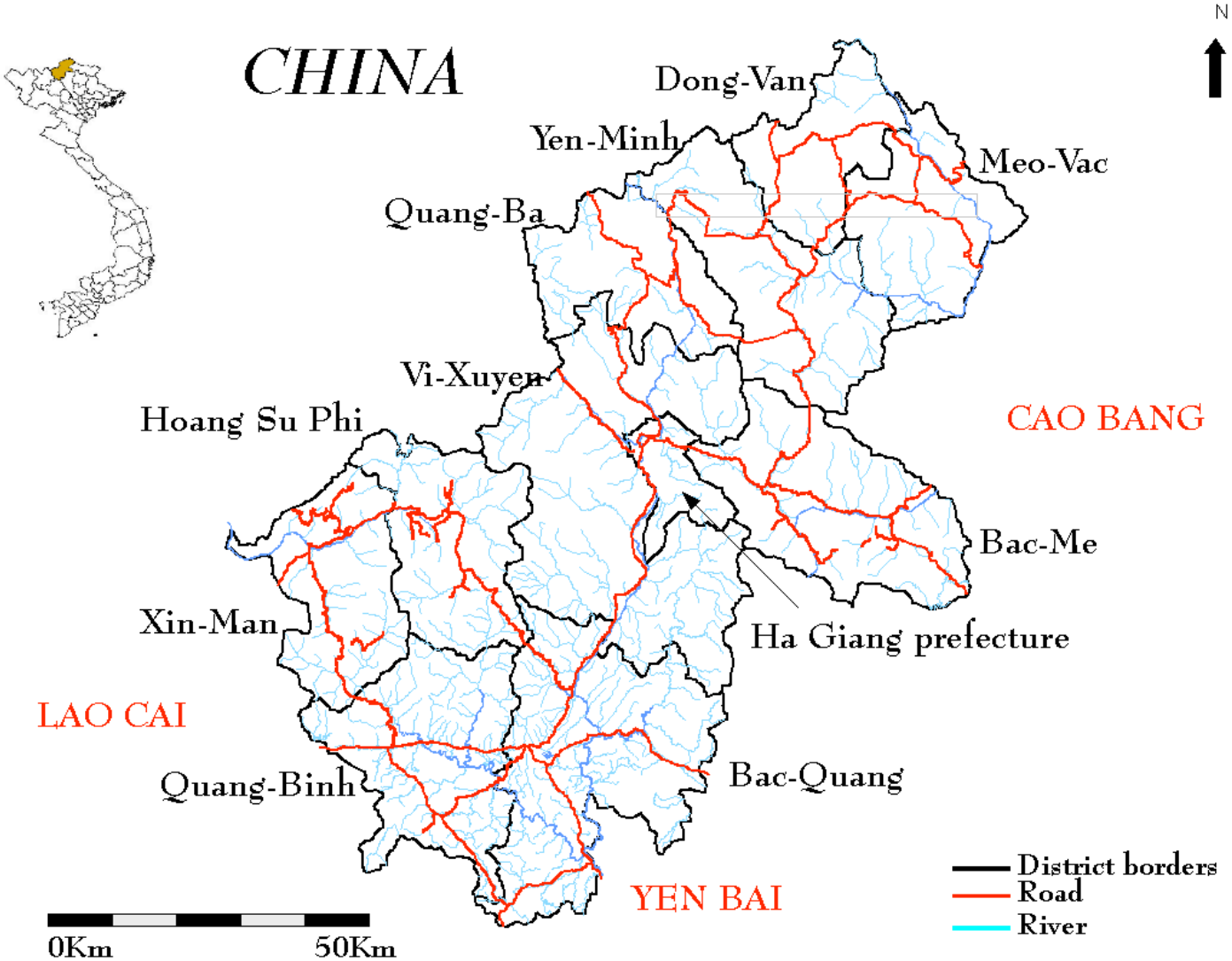
The dry season starts in October and continues until March, during that time, some communes or districts of the province suffer from a lack of water. From an administrative point of view, the Ha Giang province is subdivided into 11 districts that are composed of a total of 193 communes (Figure 2.1).

Road infrastructures allowing connections between districts were developed recently. Until the 1960's, the road network possessed only one national road: called N2, linking Ha Giang city to Hanoi. Otherwise people and animals had only tracks to travel on. In 1965, the road linking Meo-Vac district to Ha Giang city through Dong-Van district was finished. On the west side, the road from Bac-Quang – Hoang Su Phi – Xin-Man was completed in 1972, and Ha Giang –Bac Me in 1976.

2.1.2 Bref ethno-history of the Ha Giang province

Few bibliography about history of the ethnic groups and their migration is available, as a consequence this section is mainly drawn from the book titled "Histoire des Miaos" written by Savina (1924). The Ha Giang province is, with the Lao Cai province, are known to be the inhaiting area of the H'mong ethnic groups.

According to Savina (1924), H'mong people have a legend that says they come from a country where the night lasts half of the year and the day the other half. They might have migrated from as far as the valley of the Huanh He (Yellow) and the Hoai River. During the XIIIth century BC, conflict between Han people and Miao "barbarian" people (name given by Han people referring to no-Chinese people and so to H'mong people) started. In the XIIIIth century BC, all Miao people were pushed away over the Shaanxi and Guangzhou province. The Miao, from the Hoai river valley, crossed the Blue river and finally settled in the Guizhou province where they are still living today. From the Guizhou Mountain, they migrated toward neighbouring provinces such as the Guangxi and south Guangdong province, the north of Sichuan province and west of the Yunnan province.

Maize was introduced from America to China around the XVIth century. Easy to cultivate on poor lands, maize allowed the Han people to extend and to reach mountainous land previously inhabited by ethnic minorities. The tax following the arrival of the Han people was a source of conflict between Han and ethnic minorities. Some of them fought whereas others decided to migrate towards the south as far as the Tonkin area. It was at that time (1600) that the first arrival of 80 H'mong families in the district today called Dong-Van, was observed. One century later, a second wave of migrants arrived with around one hundred families. A few decided to stay in Dong-Van, and others continued until Hoang Su Phi, while some arrived in Lao Cai and more precisely in the Bac-Ha district and the commune of Si Ma Cai.

The H'mong ethnic group is in fact subdivided into four main groups which are defined by the women's clothes: the White H'mong, the Flower H'mong, the Black H'mong and the Green H'mong. Historical records never specify which kind of H'mong ethnic group migrated first or toward which direction. However, taking into account the current distribution of H'mong people, it is highly probable that it was the White H'mong people who settled in Dong-Van and the Flower H'mong in Hoan Su Phi and Lao Cai province. The last one, but the most important migration wave, with about 10 000 H'mong people, was in 1850. Most of these H'mong people came from the Chinese Guizhou province and a few from neighbouring Chinese provinces of Yunnan and Guangxi. They migrated simultaneously to Lao Cai or Ha Giang province as far as the southern province of Yen Bai (Figure 2.2). Descriptions of migration always state that entire families arrived with all their livestock, which would mean that the current H'mong cattle may be related to the H'mong cattle from China.

During their migration towards the south, H'mong people interacted with other ethnic groups such as the Thai ethnic group. This group, from the Thai-Kadai ethno-linguistic family, was also subdivided into many sub-groups such as the Black Thai, the Tho now called the Tay, the Nung and the Giay ethnic group. Savina (1924), considered that the Thai originated from the Yunnan, Guizhou, Guangxi and Guangdong Chinese provinces. The territory or province of "Nam Viet" from the Chinese Empire was in fact the current North Vietnam, and was designated in the Chinese official document from that time as the "Thai territory".

26

Figure 2.2 Map of approximate migration routes of Ethnic groups.

Figure 2.3 Map of Ethnic group's distribution in 1970.

Source: EFEO (1949)

27
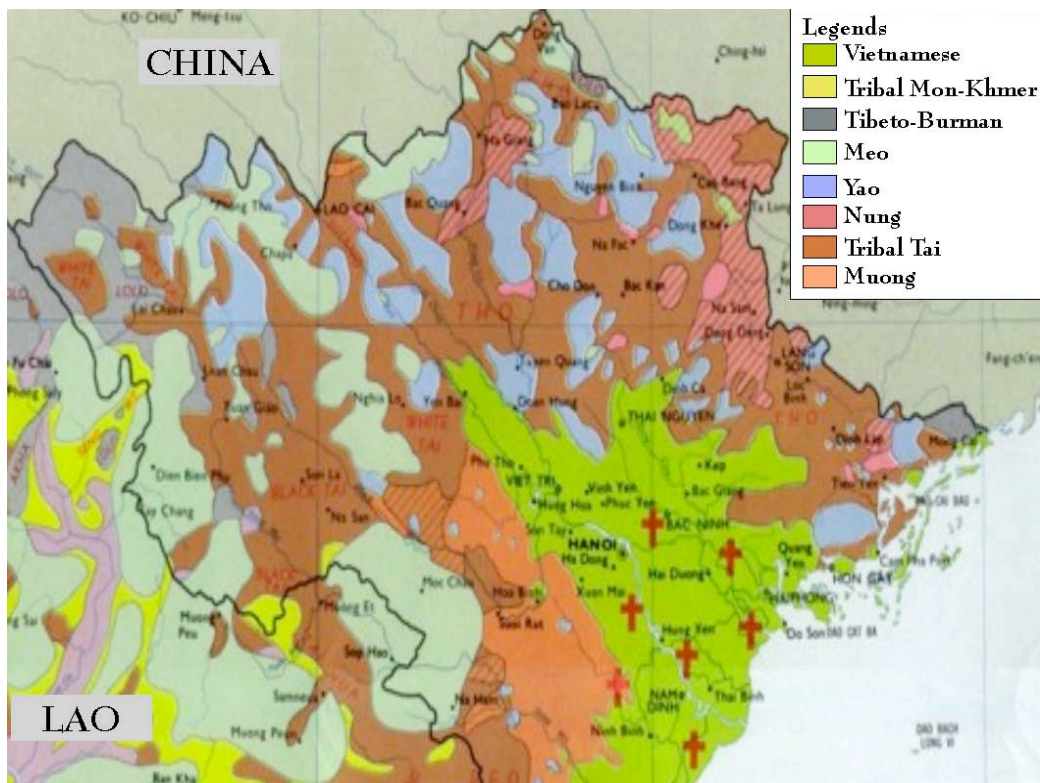
In 1503, the famous person called Cao Tri Nung, from the Nung ethnic group, as his last name indicates, rallied the sub-groups or Thai tribes from the border between China and the former "*Tonkin*" to fight against the Chinese yoke. In fact, Cao Tri Nung was successful and made the Chinese retreat as far as Canton. Then, he proclaimed himself the Emperor of the Big South: "Ta Nam", but not for very long due to a Chinese counter-attack.

The first inhabitants of North Vietnam were the Tay people (different ethnic group than the Thaï from Thailand), and after the Nung people arrived they negotiated with Tay ethnic groups. The latter awarded some middle-hill lands to Nung. Some Nung families still have the paper contract established with the Tay people (Abadie, 1924). This clearly showed that Tay and Nung people were previously inhabiting the North Vietnam (i.e. the Ha Giang province) and followed by the H'mong or Dao ethnic groups (Figure 2.3). Moreover, the power of the Tay group in the region and its political influence was well established. Therefore, it was not surprising that when the migration of H'mong people occurred, conflict between Tay and H'mong started. During the last migratory wave (1850), the Tay ethnic group appealed to Vietnamese mandarins in order to fight against the H'mong "invaders".

The H'mong ethnic groups belong to the Miao-Yao ethno-linguistic family, as do the Yao group, previously called Man by annamites and currently generally called Dao ["Zao"]. In fact, their true name is "Mien" which means "human being" in Dao dialect. There are proofs of their presence since 2000 BC in the Chinese provinces of Hunan and Guangxi and, for many centuries they have been migrating from east to west. The legend says (Abadie, 1924):

"*Tous les Mans (Yao) seraient les descendants d'un chien nommé P'an Hou. L'empereur de Chine P'an Hoang, était depuis de longues années en guerre avec le roi Kao-Wong. Ayant déclaré un jour qu'il donnerait sa fille en mariage à celui qui lui apporterait la tête de son ennemi, ce propos fut entendu par un chien nommé P'an-Hou qui tua le roi Kao-Wong et rapporta la tête à l'empereur. Celui-ci obligé de tenir sa promesse donna la jeune princesse au chien qui en eut 6 fils et 6 filles, dont la nombreuse descendance fonda la race "Yao". La princesse devait apporter en dot la moitié de l'empire, mais pour rendre ce sacrifice moins lourd, l'empereur partagea son empire […] en deux moitiés mais dans le sens de la hauteur,*

laissant à P'an-Hou et à ses descendants le sommet des collines et toutes les montagnes qui n'étaient d'aucune utilité pour les chinois".

This story is very interesting because it explains two main characteristics of Dao (i.e. Yao) people: 1) they do not eat dog meat and 2) they live at high altitudes. Moreover, as for the other ethnic groups, the Dao people are subdivided into four main groups: the Man Lan Tien: the Indigo Dao; the Man-Ta-Pan: *"Dao aux grandes planches;"* the Man Tien: *"Dao aux Sapèques";* and the Dao with White Pants. The Ha Giang province is inhabited by the Man Ta Pan in the east and the Man Tien in the west. They arrived around the year 1600 in Ha Giang province possibly for the same reasons as the H'mong. Contrary to the H'mong people, they arrived peacefully and they negotiated with the Tay ethnic groups. According to Abadie (1924), at the beginning the Dao people used to mainly cultivate maize or rice (i.e. mountain rice); but in contact with Tay people they learned how to cultivate paddy rice that finally replaced maize.

Through their migration , the ethnic groups, which originally came from distinct regions, finally mixed and shared the same territory from Chinese Yunnan province to North Vietnam. However, cohabitation did not necessarily imply good relationships, and this even within an ethnic group. As an example, the White H'mong tell to childrens that if they get near the Black H'mong they will be eaten for dinner (Morechand, 1968). Savina (1924) told that during trekking with H'mong people, even if they were short of water, they preferred to find a water source rather than stop in a Tay village and ask for water. In such a context, commercial exchanges or contacts between groups could have been limited. Obviously there were some exceptions, as Red Dao people who always used to exchange with other people without ethnic discrimination. However, Savina (1924) also reported that the H'mong people annually went to the market to buy salt and other staples for the rest of the year. Because the journey took several days and many H'mong people did not have any money, they used to bring along pigs, poultry, buffaloes or cattle to sell to the Chinese, Annamites or Tay people in order to pay for their expenses. A similar observation was made by Culas & Michaud (1997), which hypothesises that H'mong migration followed the Chinese Muslim nomad "Haw" from whom H'mong people always bought perishable staples. According to Savina (1924), the H'mong people could rear 6 cattle heads per family. For H'mong funerals, two cattle at least could be sacrificed, and even twelve at minimum for an H'mong chief. Savina (1924) described farming systems as follows:

"*Les buffles ou les bœufs vivent en liberté dans les montagnes [...] Comme ils restent parfois plusieurs semaines sans rentrer, les propriétaires connaissent rarement le nombre exact de leurs bêtes, [..] par contre ils vendent des porcs, et en très grande quantité. Ils doivent être rares les peuples qui élèvent autant de porcs que les Miaos. On en voit dans tous les coins, fouillant partout, cherchant toujours quelque chose à grignoter, et dévorant tout ce qu'ils trouvent. Si vous éprouvez le besoin de vous écarter dans la brousse, dans les environs d'un village Miao, prenez bien soin de vous munir d'un bâton pour tenir en respect tous ces vendangeurs qui vous guettent*"

This description of farming practices showes that ethnic groups did not give much importance to mating selection. They used to let all the animals roam free: cattle, buffaloes, and pigs. Also, rearing pigs seems to be a major activity in the farming system.

2.2 Sampling Strategies

In an intraspecific study of molecular variation, there are two factors, which influence the scope, intensity and thoroughness of specimen collection. These are: i) the resources required to carry out the field and laboratory work and ii) the statistical accuracy necessary to determine the genetic status and history of the population considered.

There is a large body of theoretical work in the literature examining the question of sample size and the optimal number of loci required to derive meaningful estimates for the most simple genetic parameters. The most commonly derived statistics from genotypic data are estimates of alleles frequencies, heterozygosities and genetic distances. When estimating the frequency of a given allele at a given locus, it is well known that the higher the sample size, the smaller the standard error; this standard error tends asymptotically to zero for samples sizes larger than 30 (MacHugh *et al.,* 1997). As there was no *a priori* about the genetic structure of the animal populations in the province, for a given species, a target of 30 to 60 animals sampled per district was chosen.

In order to obtain representative statistics about descriptive characteristic of farming systems, and also morphometric data, it has been showed that a stratified sampling procedure gives the best estimates of descriptors (Sokal & Rohlf, 1969). Moreover, as described previously there is a stratigraphic distribution of ethnic groups, thus only one ethnic group inhabited villages or

communes. One of the aims of the study was to characterise farming systems and management practices and highlight differences between ethnic groups. Therefore we applied a 3 strata sampling procedure: district-commune-village. We decided to sample a minimum of 4 communes per district, when possible. Because access to villages can be very difficult (2hrs by foot, or only by motorbike), once in the village, it was therefore necessary to maximise the sampling and so spend all day in the village. The minimum time needed for a farmer interview and sampling was 1 hour, in one day, the maximum number of farmers interviewed per village was 8.

During the data collection, it was necessary to stay some days in the district or the commune without coming back to the laboratory in Ha Giang city. However, in order to have good DNA quality, it was not possible to conserve biological samples without treatment for more than four days. Because two teams were organised with each one working in a different village per day, it was decided to sample 8 villages per commune before coming back to Ha Giang city.

In conclusion, sampling procedure was 8 farmers per village, 8 villages per commune and 4 communes per district across all districts. After 14 months of survey, 246 villages distributed in 37 communes from the 11 districts were surveyed (Figure 2.4, Figure 2.5).

In an idealised sampling procedure communes, villages and farmers should be randomly chosen. However it was not possible to strictly apply a random procedure, for many reasons, of practical or administrative nature. With the concern of involving the districts' agricultural departments in the project, the administrative staff of each district was asked to propose a list of communes where it was believed that animals with specificities were reared. However there was an evidence that agricultural departments mainly proposed well developed communes with good road infrastructures and easy access. Choosing only such communes would have led to a biased view of farming systems and animal populations It was then decided that among the four

Figure 2.4 Map of communes and villages surveyed.
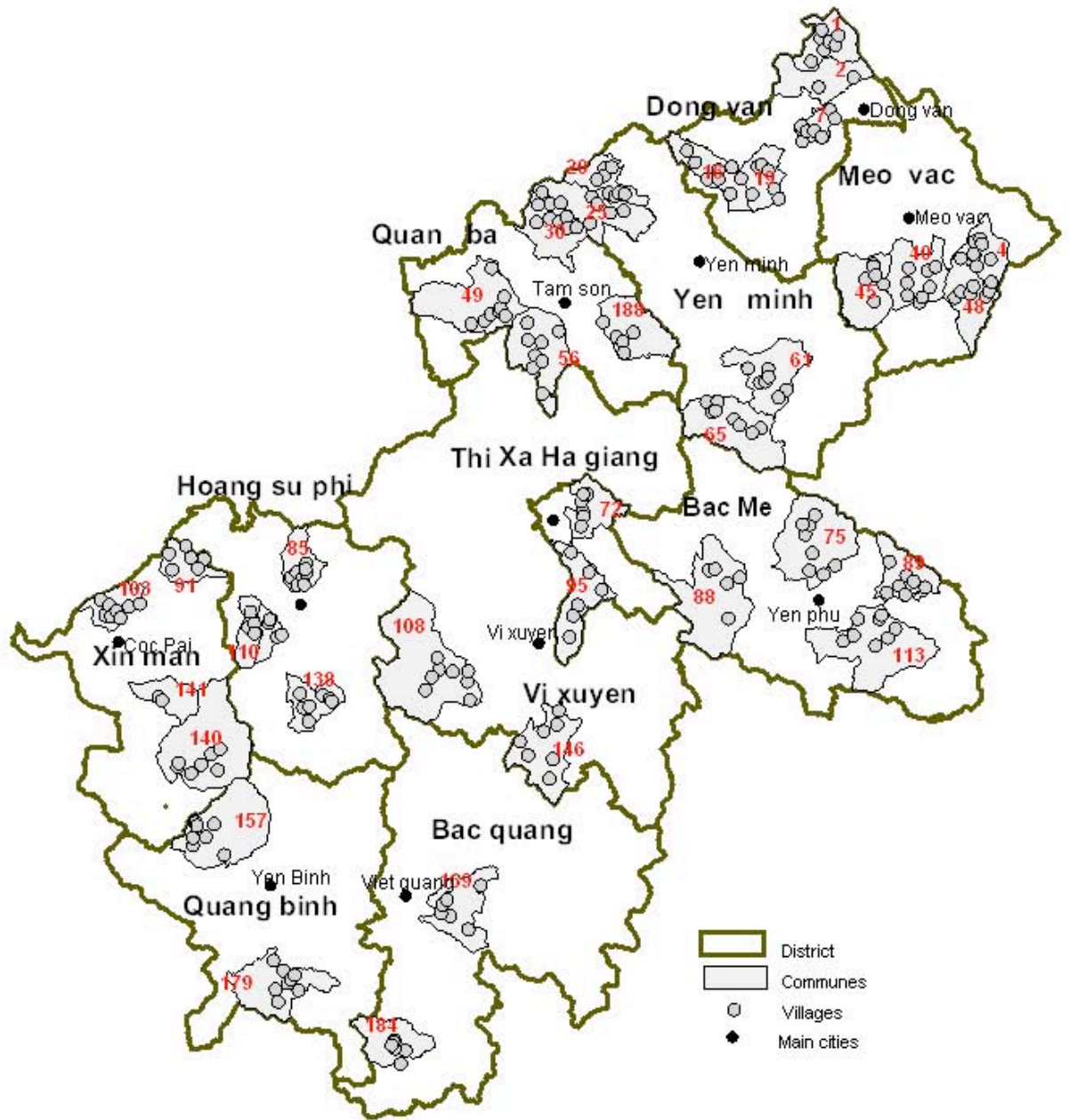
communes per district two would be chosen from the ones proposed by the agricultural department, and two by us. Based on official census data available, we chose isolated communes with the maximum number of species present (cattle, buffaloes, goats, pigs and chickens).

For a given commune, official data about the number of farmers, ethnic groups representation, number of animals, were available. We also inquired about accessibility of each village. Therefore, villages were chosen according to these information in order to have a representative representation of the ethnic groups and of environmental conditions in the commune.

Farmers use to let the animals out to graze freely early in the morning. Therefore, in some villages with difficult access, it was not possible to arrive before the animals were let out. As a consequence, it was necessary to ask the chief of the village the day before to select 8 farmers that would keep their animals at the farm.

Most of the ethnic groups in this region believed in spirits that inhabit animals and plants, and they do not trust veterinarian services. These characteristics combined to make it difficult to convince farmers to keep animals inside in order to take blood or tissue samples. As a consequence, in many cases, the chief of the village could only convince relatives, and it is known that during cultural events, or marriage, they inherit animals from each other. Caution was taken to avoid sampling related animals but missing information meant that we could not be sure, and therefore it should not be forgotten that inbreeding could be overestimated.

More than 20 % of animals of the visited communes were sampled for mammal species and nearly 10 % of the chickens. Considering the province level, the percentage of livestock sampled ranged from 0.24 % (chickens) to 1.04 % (cattle). Geographic distribution of samples area summarised in Table 2.1.

Figure 2.5 Geographic distribution of villages sampled and the ethnic groups.
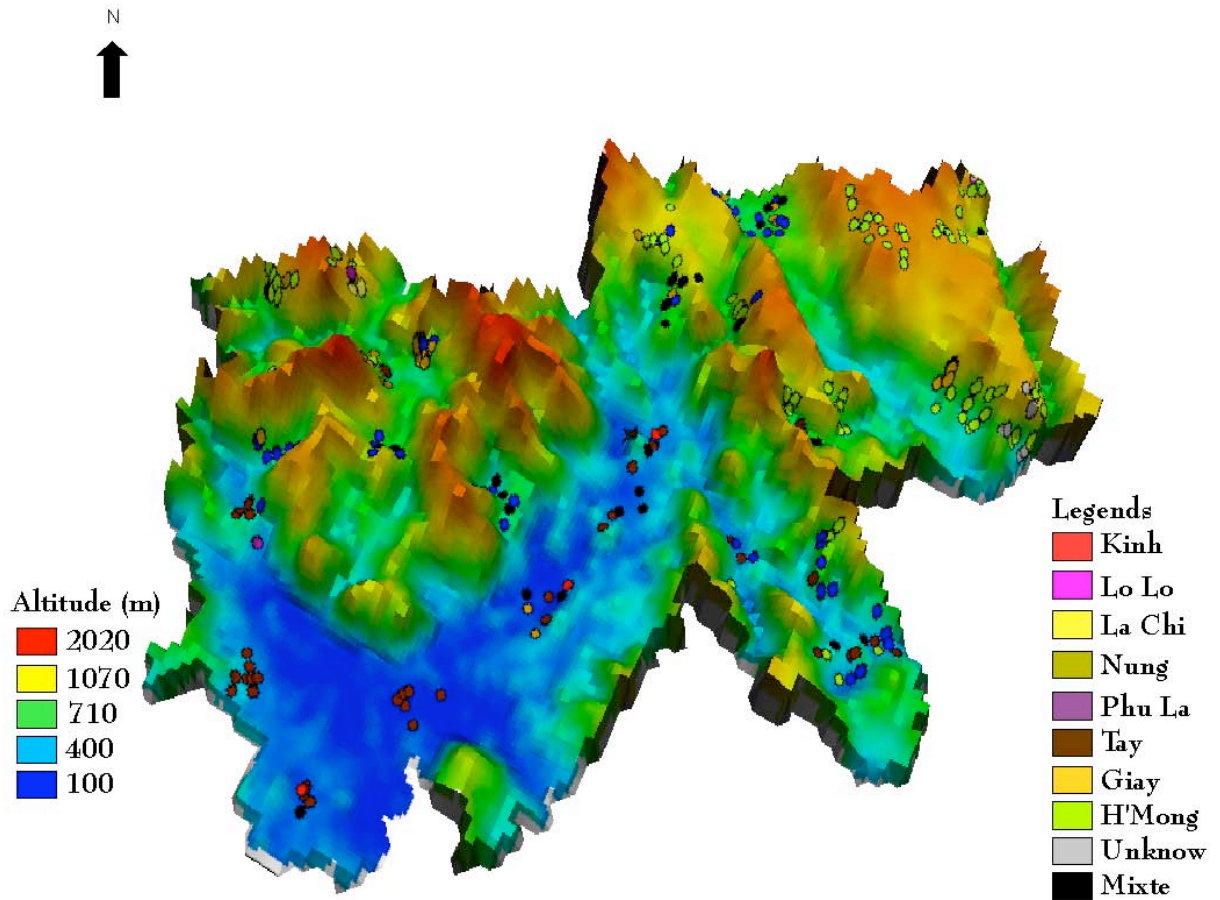
Picture 2.1 Typical landscape in Dong-Vang and Meo-Vac districts.

Table 2.1 List of samples described and genotyped.

District	Commune (code)	Nv	Nf	Cattle	Buffalo	Bk	Goat	Pig	Chicken
Bac-Quang	Quang-Minh (169)	6	50	0	50		12	44	98(15)
	Vinh-Phuc(184)	6	47	0	47(17)		0	47	96(17)
Quang-Binh	Tan-Nam(157)	7	57	6(5)	67(48)		5	49	111(52)
	Xuan-Giang(179)	8	64	18(4)	63(39)		17	56	123(88)
Hoang Su Phi	Po-Lo(85)	8	63	15	55(20)	13	50	49	122(90)
	Po Ly Ngai(110)	4	32	24(13)	13(9)		18	28	64(49)
	San Sa Ho(114)	4	34	13(10)	33(30)		18	34	60(44)
	Nam-Son(138)	7	53	2	56(46)	4	69	46	88
Xin Man	Nan-Xin(91)	6	43	6(6)	37(16)		25	35	73(64)
	Chi-Ca(103)	8	63	56(45)	35(25)		27	56	103(88)
	Quang-Nguyen(140)	7	53	4	56(44)		43	47	95(32)
	Che-La(141)	2	9	7	3(3)		3	7	17
Vi-Xuyen	Thuong-Son(108)	8	58	0	76(3)		46	47	102(16)
	Trung-Thanh(146)	7	53	1	61(42)		15	49	110(14)
HG prefecture	Ngoc-Duong(72)	6	44	3	49(7)		7	39	62(16)
	Phu-Linh(95)	6	49	0	44(26)		5	38	44(32)
Bac-Me	Giap-Trung(75)	8	47	7(4)	57(20)	5	26	51	105(52)
	Minh-Ngoc(88)	5	55	6	55(38)	5	18	44	73(32)
	Yen-Phong(89)	7	49	37(19)	34(16)	2	18	43	79
	Yen-Cuong(113)	8	54	17(9)	48(40)	5	19	42	93(113)
Quan-Ba	Bat Dai Son(30)	8	47	25(19)	27(23)	2	11	52	59
	Tung-Vai(49)	8	44	13(11)	33(17)		2	54	73(52)
	Quyet-Tien(56)	8	54	20(12)	58(26)		6	58	98(24)
	Lung-Tam(188)	5	34	19(12)	20(10)		11	28	50
Yen-Minh	Bach-Dich(20)	8	49	1	55(37)	1	1	64	72(23)
	Na-Khe(25)	7	35	19(14)	37(28)	5	5	39	54(22)
	Lung-Ho(61)	7	38	42(23)	2(2)		24	37	65(37)
	Du-Gia(65)	8	42	23(15)	30(6)		16	42	79(39)
Dong-Van	Lung-Cu(1)	7	45	44(25)	7(4)		35	40	72(36)
	Ma-Le(2)	2	8	4	12(9)	3	3	7	8(8)
	Thai Phin Tung(7)	7	49	47(24)	1		33	28	86(6)
	Pho-Cao(16)	8	37	34(18)	7(4)	2	20	45	58(39)
	Sang-Tung(19)	6	34	35(21)	1		55	31	68
Meo-Vac	Lung-Pu(4)	7	44	46(28)	0		33	42	76(44)
	Tat-Nga(40)	8	47	24(14)	19		10	39	64(26)
	Nam-Ban(45)	6	33	10(8)	23(23)		27	38	62(3)
	Khau-Vai(48)	8	48	50(32)	0		32	46	86(47)
Total		37	246	1665	678(407)	1271(744)	47	765	1541	2848(1085)

Nv: Number of villages; Nf: number of farmers; Bk: Number of buffalos for karyotyping. Numbers within brackets are the numbers of animals genotyped.

2.3 Farming systems

2.3.1 Data collected

In order to characterise the farming system and management practices, interviews with farmers were carried out. A preliminary form was written and checked during a one-week survey. Then, the form was modified and completed (Annexe 1). It was noticed that farmers were not able to provide some informations such as the interval between two parturitions, age at first parturition, etc. Therefore, "a genealogical tree" of each animal was realised and allowed to estimate such information. Moreover, because the origin of each animal was asked, we were able to know if animals came from the farm where it was sampled or from another farm reducing the probability of sampling related animals. In addition, it enabled us to check if the number of animals which should be present at the farm was similar to the first estimation given by the farmer when asked the number of animals for each species present on the farm, and so confirmed the validity of the livestock inventory of the farm.

Farmers did not know the exact surface area of their land, but they could estimate the quantity (kg) of seeds or rhizomes sown; and the number of trees planted when they had forest land for wood production. Farmers also knew how many kilograms they harvested. Then, information such as Plantation/Production were recorded in kg units. Data such as kg of seeds, different weights of pigs, or number of eggs, chicks, etc. are estimations made by farmers and therefore, this is only approximate data.

Farmers were questioned about their knowledge of the origin of each animal through the pedigree interview. Eight categories of origin were then considered: Farm (i.e. the animal is born on the farm where the dam and grand-dame were raised, bulls are rarely known); Village (another farm within the village); Commune; District; District Market (the animal was bought at the district market, with no information about its farm origin); Other district; Outside the Province; Project (many poverty alleviation projects have given buffaloes or cattle to farmers, therefore a class was done for these "gifts" because their origin was not possible to know). For the "Becoming" of animals, same categories where used plus one category: Buyers because in some cases buyers go directly to the village or the commune in order to buy animals that will be brought to slaughter houses or to big market for sale at higher price. All the data collected from interviews was recorded in an ACCESS Database with 11 thematic tables, which are summarised in Table 2.2.

Table 2.2 List of fields recorded per table in the farmers database from interviews.

Village	Farmer	Crops	Livestock (nb)	Maternity (per female in the farm)
code	code	code	Cattle (M/F)	Origin
Latitude	Name	Plantation (kg seeds)	Buffalo(M/F)	Age
Longitude	Code Village	Production (kg seeds)	Goat(M/F)	Nb Litter
Altitude	Ethnic groups	Purchase (kg/year)	Pig(M/F)	Nb ind/Litter
Nb Farmers	Family size	Sale (kg/year)	Horses(M/F)	Interval-parturitionr
Ethnic Groups	Nb children		Chicken(H/C/chicks)	Age 1st parturition
	Children education		Duck(H/C)	Nb sale/Litter
	Pig breeder (Y/N)			
	Pig fattening (Y/N)			

Pig Husbandry	Chicken Husbandry	Origin (per reproductive animal)	Becoming (per animal outgoing)
Weight at Purchase	Nb laying/year	Family	Death
Weight at Sale	Nb eggs/Laying	Village	Village
Current Weight	% Brooding	Commune	Commune
Husbandry time	% Hatching	District	District
Purchase Price	% Adult	District Market	District Market
Sale Price	Year of the last epidemic event	Other district	Other district
Nb Litter	Nb survivors	Project Gift	Other Province
Nb sale/Litter	Nb cocks introduced after epidemic	Other Province	Buyers
Disease(Y/N)	Nb hens introduced after epidemic		
Disease Frequency	Origin of introduced animals		

Nutrition (per species)	Paternity (per parturition/species)
Free Grazing (Y/N)	Selected(Y/N)
Communal Grazing (Y/N)	Grazing(Y/N)
Stalling (Y/N)	Free on purpose (Y/N)
Crops Fodder (Y/N)	Unknown
Forest Fodder (Y/N)	Related (Parent-Offspring, full-sib)
Elephant Grass (Y/N)	Farm (not related)
Canna sp. / Cassava (Y/N)	Village
Bong rio (Y/N)	Commune
Cam (Y/N)	District
Starches (kg/day/animal)	Project Gift
Vegetable (kg/day/animal)	Other Province
Total (kg/day/animal)	Nb bulls used
	Nb parturitions

Y/N: Yes/No; M/F: Male/Female; Nb: number; Bong gio: waste from rice alcohol production; Cam: mush of rice bran and leaves

2.3.2 Statistical analysis

Three types of analyses were used in order to highlight differences between ethnic groups in management practices. Two of these analyses, Factorial Discriminant Analysis FDA and the Multivariate Coinertia Analysis will be described in following pages as they were also used for analysing genetic and phenotypic data. Only the Hierarchical Agglomerative Clustering Analysis (HACA) was specifically used on the farming system data. The HACA will uses coordinates for the three first discriminant factors obtained by FDA of each ethnic group in order to build a dissimilarity matrix. The distance between two clusters will be the average of the dissimilarities between the points in one cluster and the points in the other cluster. It will build an agglomerative tree for which the vertical coordinate of the junction of two branches is the dissimilarity between the corresponding clusters. The agglomerative coefficient Ac (Rousseeuw, 1986) is a quality index for the agglomerative coefficient of the data. In other words, the discriminant analysis will perform the best synthetic variables for description of farming systems, while cluster analysis will allow to see which groups were similar using the synthetic variables.

2.4 Genetic data

2.4.1 Buffalo Karyotypes

In order to assess if crossbreeding between water buffalo types occurred in the Ha Giang province, forty-seven karyotypes were done. Indeed, Swamp buffalo is *2N*=48 and River and Wild Buffalo are *2N*=50 (Ulbich & Fishcer, 1967, F1 and F2 offspring's have *2N*=49. Realisation of karyotypes needs cells culturing and so fresh blood. Then, only animals older than 15 years old and sample the last day of the week field were used for karyotyping. The protocol was provided by the joint research unit INRA/ National Veterinary School of Toulouse "Cytogénétique des populations animals". After 72h of cell culture, cell multiplication was stopped with colchicine. Conventional GIEMSA colouration of chromosomes was used and cells were staggered on slides in order to count the number of chromosomes. Three cells per animal was counted to ensure the exact number of chromosomes.

2.4.2 DNA markers

During recent decades, a large number of markers have been developed for genetic population analyses. For our study on Ha Giang livestock populations, microsatellite markers have been chosen in each species. Microsatellites are simple sequence repeats in genomic DNA, such as poly-dinucleotide repeats (e.g., TGTGTG…). For a given marker, alleles are defined by the number of repeats and may be distinguished after migration with a capillary sequencer. Microstallite markers have been found to be common in all eukaryotic genomes and, as they are highly polymorphic, they provide extremely useful markers. Microsatellite markers are assumed to be neutral in regard to selection. If, in addition, one assumes that populations are in drift-migration equilibrium, the breed/population similarities can be interpreted as only reflecting common ancestry. Microsatellites have three primary uses in population genetics analyses for livestock studies. First, they can be used to quantify genetic variation within and among livestock populations or breeds. Second they allow to document admixture among populations. Third, microsatellite data can be used to assign individuals to genetically similar groups at the population, breed or species levels. Therefore, FAO provides guidelines of microsatellites that have been found to be informative and so should be used for population genetics studies (Hoffmann, 2004). For genetic analysis of the Ha Giang province livestock species, we chose microsatellite markers proposed by FAO in most cases.

For mammals, both blood samples (5 ml) and tissue (i.e. a small piece of ear) were taken to ensure DNA quality. For poultry, only blood (2ml) was sampled.

The Hanoi laboratory being 8 hours travel time away from the Ha Giang, the samples needed to be treated for long-term conservation before DNA extraction:

- tissues were conserved in 2 ml of pure ethanol at –20°C until DNA extraction.

- mammal's blood was separated (serum / leucocytes / erythrocytes). Serum and erythrocytes were conserved at –20°C and leucocytes at +4°C in 2 ml of Hank's media.

- for chicken blood, serum and erythrocytes were separated; erythrocytes were mixed (50/50) with a PBS-Sucrose solution. Serum and plasma were kept at –20°C.

In the laboratory in Hanoi, DNA extractions (blood, tissues) were done with the QIAamp kit from QIAGEN. The PCR products were fluorescent labeled, and genotyped using a capillary sequencer (Beckman & Coulter CQ8000).

2.4.3 Analysis of markers data

Genetic diversity has been defined as the variety of alleles and genotypes present in a population and that is reflected in morphological, physiological and behavioural differences between individuals and populations.(Frankham *et al.,* 2002)

1   Classical measures of genetic diversity:

At the molecular level, genetic diversity has been usually measured by the following parameters: 1) Frequencies of genotypes and alleles; 2) Proportion of polymorphic loci; 3) Observed and Expected heterozygosity; 4) Allelic diversity and 5) Allelic richness

a) Heterozygosity

For a given locus, the observed heterozygosity H_{Obs} is defined as the observed proportion of heterozygous individuals. Within a given population, for a marker with co-dominant alleles (which is the case for microsatellite markers), this parameter is estimated by simple counting. Considering several loci, H_{Obs} is the mean value of H_{Obs}over loci.

The expected heterozygosity H_{Exp}, or gene diversity, was defined by Nei (1973) as the probability that two alleles chosen at random at a given locus and within a population are different alleles. Let p_i be the frequency of allele i in the population, H_{Exp} is computed as the proportion of heterozygotes according to the famous Hardy-Weinberg proportions:

$$H_{Exp} = 1 - \sum_i p_i^2$$

In field studies, the number of animals with a known genotype is limited. To avoid biases in the estimation, the unbiased expected heterozygosity H_{nb} has been defined as follows (Nei, 1978):

$$H_{nb} = 1 - \frac{(n\sum (p_i)^2 - 1)}{n - 1}$$

where n is the sample size. In this study, we will always perform the unbiased expected heterozygosity which we will denote H_{Exp} and in practice, heterozygosity measures were estimated using GENETIX (Belkhir *et al.*, 2000).

b) F-statistics

In actual populations, the genotype frequencies in each defined subpopulation do not necessarily follow the Hardy-Weinberg proportions. Wright (1951) proposed that deviations of genotypes frequencies in a subdivided population be measured in terms of three parameters, F_{IS}, F_{IT} and F_{ST}. F_{IS} and F_{IT} are the correlations between the two uniting gametes relative to the subpopulation and relative to the total population, respectively, whereas F_{ST} is the correlation between two gametes drawn at random from each subpopulation:

$$(1-F_{IT})=(1-F_{IS})(1-F_{ST})$$

Weir & Cockerham (1984) have developed a variance based method for estimation of *F-statistics*. They describe a measure θ_{ST} which estimate the correlation of pairs alleles between individuals within a subpopulation through an analysis of the partitioning of variance of allele frequencies:

$$\theta = \frac{\sum a_i}{\sum (a_i + b_i + c_i)}$$

where a_i is the variance partition of allele frequencies between population, b_i the partition within population and c_i within individuals. In this thesis, *F-statistics* were calculated according to Weir & Cockerham (1984) as implemented in GENEPOP (Rousset & Raymond, 1995).

Genetic distance means the extent of the genetic differences between populations as measured by a function of allele frequencies. There is a wide range of genetic distances

available. Each distance is based on different assumptions of selection, genetic drift, etc. According to Laval *et al.* (2002), D_R (Reynolds, 1983) seemed to be the most efficient when comparing subpopulations weakly differentiated, as may be the case with current livetostck populations:

$$D_R = -\ln(1 - F_{ST})$$

2 Bayesian approaches

Bayesian approaches have been widely developed since the last decade. One of the most important uses is to determine the number of distinct genetic populations K in the sample, and to try to assign each individual to the correct population. The STRUCTURE program developed by Pritchard *et al.* (2001) is one of the standard programs for such an approach. The assumptions in STRUCTURE are of two types: i) there is a prior distribution for unobserved quantities (number of population K, and their allele frequencies) and ii) there is a likelihood function relating these unknown parameters to the observed genotypes.

There are two likelihood function. The first one was developed by Pritchard *et al.* (2001) applying Dirichlet model v.s. non-admixture model: each individual have ancestors from only one population. The second livelihood function was developed by Falush *et al.* (2003), namely the admixture model which allows each individual to have ancestors in more than one population.

When the likelihood model has been selected, the programme will run several Monte-Carlo Markov Chains (MCMC) in order to estimate probability of each K to be the true number of populations. These runs will result in a posterior Log Likelihood probability for each K. Pritchard *et al.* (2001) had suggested using an exponential method to choose the best K. However, this method did not consider the variability of Log Likelihood obtained for a given K. Then Evanno *et al* (2005) proposed a second derivative method that takes into account this variability. One weakness of this approach is that the second derivative cannot be calculated for K=1. So a more complete approach would be to first use exponential formula from Pritchard *et al.* (2001) to ensure that there is more than K=1 population in the sample and then apply the method proposed by Evanno *et al.* (2005).

3   Multiple Co-inertia analysis (MCOA)

The MCOA is a multivariate analysis which allows to simultaneously find ordinations from the loci tables that are the most congruent. For that purposes, successive axes from each table of allelic frequencies, which maximize a covariance function are searched. This method allows to extract the common information from separate analyses, in the setting up of a reference typology, and to compare each separate typology. The efficiency of a marker is assessed by its typological value (*Tv*), i.e. the contribution of the marker to the construction of the reference typology, contribution which is equal to the product of the variance (genetic diversity within a locus table) multiplied by the congruence with the consensus Cos² (i.e. the correlation between the scores of individual locus tables and the synthetic variable of the same rank) (Laloë *et al*. 2007).

4   Prioritisation of breeds/populations for conservation

To conserve maximum global diversity, Barker (1999) indicated the need for a measure of the contribution of a breed to the overall genetic diversity of a species so that breeds with a substantial contribution should be given higher priority. Simianer *et al.* (2003) and Reist-Marti *et al.* (2003) noted that the importance of a breed to the overall diversity depends not only on its own risk of extinction. Therefore the marginal diversity of each breed, which takes into account the extinction probabilities of related breeds, offers a more sophisticated tool.

Weitzman (1993) has suggested a method that uses genetic and non-genetic information to calculate a maximum-likelihood tree and the current diversity of a group of species. Thaon d'Arnoldi *et al.* (1998) proposed to set conservation priorities through the analysis of genetic distances by Weitzman's approach to measure the global diversity and the marginal loss of diversity attached to each breed.

BOX 4 Weitzman's approach

The approach uses a diversity function *V*. The intuitive property of this function is the "monotonicity in species": the gain of one j element increases the diversity of a group Q by at least *d(j,Q)*
:

$$V(Q \cup j) \geq V(Q) + d(j,Q)$$

Then the diversity of a set S is "the maximum, over all members in the set, of the distance of that member from its closest relative in the set plus the diversity of the set without that member":

$$V(S) = \max_{i \in S}[V(S \setminus i) + d(i, S \setminus i)]$$

Therefore with this function, Weitzman proposed to take into account the risk of extinction, and so it is possible to build a tree that maximises the survival probability of all the element of the set S.

One of the main reproaches to this method, is that it is based on genetic distances and consequently it only takes into account between-breed diversity and not within-breed diversity. To counter this weakness, Ollivier & Foulley (2005) had proposed supplementary measures based on the Weitzman's approach:

$$D = \alpha\ CB + (1-\alpha)\ CW$$

Weitzman values represented contribution to between-breeds diversity (*CB*). Within-breeds (*CW*) contributions to diversity are measured using heterozygosity:

$$H_k = 1 - H(S \setminus K)/H(S)$$

where *H(S\K)* is the heterozygosity of the set without element K. The aggregate diversity *D* is the sum of both elements weighted by a constant that can be set to be equal to F_{ST}. Then, the new measure *D*, will allow the prioritisation of breeds taking into account both within- and between-breed diversity.

Instead of using genetic distances which refer to dissimilarity within breeds, one could use an index of similarity. Coefficients of kinships between and within populations were proposed by Eding and Meuwissen (2001) as a tool to assess genetic diversity for conservation purposes.

Coefficients of kinship are directly related to the definition and measurement of genetic diversity. They represent a method for estimating average kinships between populations from similarities of genetic marker alleles. In addition, kinships can assess between and within breeds simultaneously. When the within breed diversity is weighted against the between breed diversity, inbred populations will receive smaller contributions (Eding & Meuwissen, 2003). In contrast, Weitzman's method gives higher priority for inbred populations. Similarly, Caballero & Toro (2002) proposed a molecular coancestry f applying the Malécot definition, but considering genes identical by state instead of by descendant, for subdivided populations. Morevover, this methodology also allows to estimate the contribution of each subdivided population to an hypothetical "conservation stock" in order to maximize genetic diversity in the next generation.

However, in addition to their genetic uniqueness, it is necessary to choose breeds for conservation based on sensible criteria, such as degree of endangerment, adaptation to specific environments, traits of economic importance, availability of unique traits, cultural and ecological values, etc. (Ruane 2002). Then, prioritizing breeds for conservation in order to develop optimum conservation strategies needs assessment of genetic and non-genetic factors. Breeds can be characterized based on genotypic and phenotypic information.

BOX 5 Molecular coancestry approach

Let f_{ij} be the average pairwise coancestry (Malécot, 1948) between individuals of subpopulation i and j, and s_i the average selfcoancestry (within an individual) of the N_i individuals of subpopulation i, then the average distance between individuals of subpopulations i and j is $D_{ij}=[(s_i + s_j)/2]-f_{ij}$

The total genetic diversity GD_T is partitioned into three component:

- The genetic diversity within individuals: $GD_{WI} = (1 - \tilde{s})$

- The genetic diversity between individuals: $GD_{BI} = (\tilde{s} - \tilde{f})$

$$GD_{WS} = GD_{WI} + GD_{BI}$$

- The genetic diversity between subpopulations: $GD_{BS} = (\tilde{f} - \bar{f})$

2.5 Phenotypic data

Because the characterization of AnGR cannot be based on neutral polymorphisms only, phenotypic descriptions and morphological measurements were done on the animals sampled. As an example of the interest of such a characterization (if necessary), one may refer to the significant morphometric differences between two cattle populations in Kenya, namely the Kamba and Massai populations, observed by Mwacharo & Drucker (2005). The Massai zebu had a tall and leggy conformation, which can also be attributed to its suitability to the almost free and wide-ranging mode of grazing in search of pasture and water. On the contrary, the Kamba ethnic minority being agropastoralists, who rely heavily on animal draught power from cattle, prefer short, small and compact animals, as it is believed that such types of animals are better at pulling ploughs for longer hours than larger, taller animals. Skeletal measurement such as body height, length and chest depth are less affected by nutrition and thus indicate inherent size better than measures related to girth measurements (Kamalzadeh *et al.* 1998; Hall 1991).

2.5.1 Measurements

1   Body Measurements for mammals

The exact age of the animals was generally unknown, although farmers were generally able to say if animals were older or not than two years old. Therefore, only animals assumed to be older than 2 years old were measured. Measurements were taken with a cloth tape:

- the height at withers ***HW*** (measured on the dorsal midline at the highest point on the wither);

- the thorax depth ***VSS*** (measured from the floor to the sternum);

- the body length ***BL*** (measurement of the distance between the point of the shoulder and the pin bone);

- the heart girth ***HG*** (body circumference immediately posterior to the front leg measured);

- the neck girth ***NG*** (for buffaloes only: neck circumference at the base);

- the ear length ***EL***;

Those measures were combined into 5 indices:

- the index of slenderness IGs = VSS/(HW-VSS) (Lauvergne & Souvenir Zafindrajaona 1992),

- the length index I_{BL}= BL/HW,

- the heart girth index I_{HG} = HG/HW

- the neck girth index I_{NG} = NG/HW

- the ear index I_{EL} = EL/HW

2 Buffalo horn measurements

Buffaloes have big horns, and it is often used as a criteria for differentiation between breeds, therefore we carried out some measurements proposed by Souvenir-Zafindrajaona (1999) for the description of the Malagasy zebu, in Madagascar. Six measurements were taken on only one horn (Figure 2.7):

- *TT* (length between the two tips)

- *ExL* (external length of the horn)

- *IL* (internal length of the horn)

- *TB* (chord distance between tip and base of the horn)

- *BG* (Horn base girth)

Figure 2.6 Description of buffalo horn measurements.
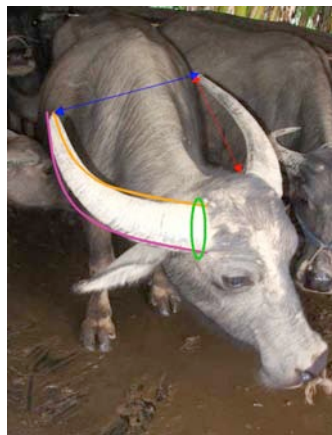
TT (in blue); *ExL* (in pink); *IL* (in orange); *TB* (in red); *BG* (in green).

For buffalo, horns still grow during adult life, so it was more appropriate to compare ratio than exact measurements. Moreover, ratios are believed to translate in a better way the general shape of the horns. Three ratios were then calculated:

- TT/ExL

- IL/ExL

- TB/IL

3   Weight estimates

Field conditions did not allow us to bring weight instruments when doing surveys. Therefore, measurements of weight were carried out on a few animals during contests organised in the province, and a regression with *HG* values was done in order to allow estimation of weight from animals sampled and summarised in Table 2.3.

Table 2.3 Linear regression for weight estimates of cattle and buffaloes from heart girth (*HG*)

Species	Sexe	No	Equation	R^2
Cattle	Bulls	44	-493.41 + 5.0185 x *HG*	0.911
	Cows	17	-273.43 + 3.5527 x *HG*	0.776
Buffalo	Bulls	19	-686.11 + 6.2080 x *HG*	0.876
	Cows	14	-464.69 + 4.7288 x *HG*	0.802

No: number of individuals sampled

2.5.2 Phenotypic descriptions

1   On cattle

Adding to morphometric measurements, we described the coat colour and horn position of the cattle:

- Coat: uniform (one colour) / not uniform (various colours)

- Coat colour (yellow / black / brown-reddish)

- Horn position (frontward / horizontal / vertical)

2   On swamp buffaloes

During the preliminary survey, it was noticed that there were two different colours in the swamp buffalo: grey and white. According to theliterature, white buffaloes are not albinos because they have some black spots on the muzzle. Moreover, we have observed the presence of pigments on the eyes. Also, presence/absence of white markings (Figure 2.8) on grey buffaloes is used to differentiate some breeds, even wild and feral buffaloes (Heinen, 1992). Then 6 phenotypic characteristics were recorded:

- Colour: Grey / White

- Presence of black spots (if white)

- White marks around the eyes (Yes / No)

- White marks on the cheeks (Yes / No)

- Number of white chevrons on the neck (0 / 1 / 2)

- White socks: absent / diffuse / very clear

Picture 2.2 Colour and white marks from swamp buffalo.

3 Poultry: genes with visible effects

For chicken, many well identified genes have a very clear phenotype, so that by just describing the phenotype, it is possible to know the genotype of the animal. Such genes also affect the colour or pattern of feathers, comb, etc (Coquerelle, 2000). In order to allow non-specialists to describe phenotypic traits and to reduce errors as far as possible, we only selected easily identifiable phenotypes. For traits known to be determined by a single biallelic gene with one recessive allele, the frequency of this allele was estimated as the square root of the frequency of animals (Table 2.4, Picture 2.3, Picture 2.4).

Table 2.4 List of chicken genes with visible effects.

Single Locus traits	Name	Alleles	Recessive Genotype	Recessive Phenotype
c	Recessive white	C+; c	[cc]	no coloration, white chicken
I	Dominant white	I; i+	[i+i+]	presence of black
Pg	Pattern feather	Pg; pg+	[pg+pg+]	no pattern
H	Silkiness	H+; h	[hh]	silkiness
F	Frizzled plumage	F; f+	[f+f+]	frizzled plumage
Cr	Crest	Cr; cr+	[cr+cr+]	without crest
vh	Vulture hocks	Vh+; vh	[vhvh]	without vulture hocks
Mb	Muffs and beards	Mb; mb+	[mb+mb+]	no muffs and beards
Double Loci traits	**Trait**	**Loci**	**Genotypes**	**Phenotypes**
R & P	Comb type	R; r+	[r+r+,p+p+]	Simple Comb
		P; p+	[R-,p+p+]	Rose Comb
			[r+r+,P-]	Pea Comb
			[R-,P-]	Walnut Comb
W & Id	Shank colour	W+; w	[id+id+,ww]	green
		Id; id+	[id+id+,W+-]	blue grey
			[Id-,W+-]	white
			[Id-,ww]	yellow

Some other traits easy to identify but for which it is not possible to know the exact genotype because of multiple genes or interactions were also studied, but only for phenotype frequencies (Table 2.5).

a)

Picture 2.3 Pictures of few phenotypic traits described on chickens.
a) Comb types from top left to bottom right: Simple comb, Rose comb, Walnut comb, Pea Comb;
b) Presence of vulture hocks;
c) Presence of beards;
d) Presence of crest.

b)

c)

d)

Table 2.5 List of complexes chickens phenotypes.

Complex traits	Code	Phenotypes
Black shank	BlackT	
Eye colour	BlackE, BrownE, OranE	Black / Brown / Orange
Skin colour	BlackS, YellowS, WhiteS	Black / Yellow / White
Comb colour	BlackCmb, CmbRB, CmbRv, CmbRc	Black / Red-Black/ Dark Red/ Light Red
Ear Colour	BlackO, RedO, WhiteO, WRO	Black / Red / White / White and Red
Shank without feathers	TnF	
Polydactily	TnP	

4 Statistical Analysis

Each measurement could be considered alone or coupled with others ones in order to highlight morphometric differentiation. The objective was to identify morphometric differences among districts or following ethnic groups. So, for this analysis, we considered all the measured animals and not only genotyped animals. Then, we transformed (Ln, x^2) measurements in order to obtain normality of residuals and homoskedasticity.

a) MANOVA

We first performed a MANOVA using the R programme. The MANOVA is used to test the significance of the effect of one or more predictor variables on a set of two or more outcome variables:

$$Y_1 Y_2 \ldots = X_1 (X_2 \ldots)$$

where X_s are discrete and Y_s are continuous. This method enabled us to create a multivariate composite variable. Then, in our cases, all the morphometric measurements (Y_s) in order to differentiate the districts (X_s). Using morphometric measurements, measurements could be naturally correlated, then we must take this correlation into account when performing the significance test. Therefore, we used Wilk's test for significance because this test is based on a comparison of the error variance/covariance matrix and the effect of variance/covariance matrix.

Picture 2.4 Pictures on chicken plumage types.
a) frizzled plumage; b) silky plumage

a)   **b)**

Picture 2.5 Sale of chickens in the Lung Phin Market (Dong-Van district)

b) ANOVA: Hierarchical mixed linear regression models

As breeders/farmers, we are interested in determining which characters precisely define the population and allow a differentiation from others. Therefore we performed ANOVA using single trait linear hierarchical mixed regression models with the SAS software. For a given trait, the model was:

$$y_{ijkl} = \mu + b \times alt_{ijkl} + s_i + d_j + c_{ijk} + e_{ijkl}$$

where y is the observation, μ the overall mean, *alt* the altitude where the animal is raised and b the coefficient of regression of performance over altitude, s_i the fixed effect of sex i ($i = 1$ or 2), d_j the fixed effect of district j, c_{ijk} the random effect nested into district and e_{ijkl} a random error The P-values of pairwise mean adjusted comparisons were corrected with Bonferroni correction.

The level of the random effect, which could be considered as repeated measurements within the district, needs to have at least three replicates. Not all species were present in every sample site:

- Not always 3 villages sampled per commune for cattle

- Not always 3 communes sampled per district for buffaloes or chicken

Therefore the level for the random effect depends on the species considered: for cattle we considered communes randomly chosen within a district, for buffalo we considered villages randomly chosen within a district.

Many distinct ethnic groups rear buffaloes. Then it was possible, in the same way as the effect of the village or the commune where the animal was sampled, that different management practices, depending on the ethnic group, could influence the conformation of the animal. Thus we performed two tests:

- ethnic groups randomly nested within a district

- district randomly nested within an ethnic group (i.e. one breed=one ethnic group=fixed effect)

c) Principal Component Analysis PCA and Factorial Discriminant Analysis FDA

The PCA is a non-parametric analysis and is an orthogonal linear transformation that transforms the data to a new coordinate system, such that the greatest variance by any projection data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. The PCA is theoretically the optimum transformation for a given set of data in least square terms. It can be used for dimensionality reduction in a data set by retaining those characteristics of the data set that contribute most to its variance. The PCA uses the eigenvectors of the covariance matrix and it finds only the independent axes of the data under Gaussian assumption. Then, groups are created by forming composite axes that maximise the overall distance between data without *a priori,* because it does not consider class separately as it does not take into account the class label of the feature vector. The PCA is done on continuous variables; the symmetric analysis for discrete variables is the Factorial Correspondence Analysis (FCA).

When an *a priori* grouping needs to be taken into account, then a class label vector is incorporated into the PCA analysis which is therefore called a Factorial Discriminant Analysis FDA. The FDA is used to determine which variables best discriminate the groups and to do so it considers within- and between-groups variability. The FDA is multivariate analysis of variance (MANOVA) reversed. In MANOVA, the independent variables are the groups and the dependant variables are the predictors. In the FDA, the independent variables are the predictors and the dependent variables are the groups. In summary, MANOVA will tell us if significant differences between groups occurred and FDA because of which variable this difference occurred. These approaches are based on functions available in the ADE4 package (Chessel *et al.,* 2004; Dray & Dufour, 2007) of the R software (R development core team, 2007).

d) *Hill & Smith Analysis*

A breed or population should be characterised by both genetic and phenotypic data. Therefore we used an approach similar to that of Hill & Smith (1976). It was a combination of an internal correspondence analysis on markers (Cazes *et al.,* 1998; Laloë *et al.,* 2002), and of principal component analysis on quantitative measurements. This approach was also applied to combine morphometric measurements and phenotypic descriptions such as coat colour.

CHAPTER 3

FARMING SYSTEMS

TYPOLOGICAL ANALYSIS OF FARMING SYSTEMS IN A NORTHERN PROVINCE OF VIETNAM IN REGARDS TO ETHNIC COMMUNITIES AND THEIR GEOGRAPHICAL LOCATION

C. Berthouly, X. Rognon, T. Nhu Van, A. Guillouzo, H. Thanh Hoang, E. Verrier, J-C Maillard

3 THE FARMING SYSTEMS
3.1 Indigenous knowledge :

Indigenous knowledge of farmers is :

"the body of knowledge acquired by a community in any given area and relating to agriculture, livestock rearing, food preparation, education, institutional management, natural resource management, heath care and other pertinent subjects. It is regarded as a valuable resource for development activities and should therefore be considered and applied in development projects wherever suitable" (Mathias, 1995)

Indigenous knowledge on animal husbandry consists of several components such as ethnoveterinary medicine, or housing, herd management, feeding, fodder management, breeding animals and processing of animal products. Indigenous knowledge about animal breeds and breeding includes strategies and social mechanisms that influence the gene pool such as passing animals to the next generation at certain life-cycle events such as birth, weddings, sharing mechanisms (lending or gifting animals to poorer relatives). Also farmers, may have preferences on colour, shape that will influence the breeding objectives and practices.

Analysing AnGR management are important to identify and evaluate aspects of traditional breeding strategies and to achieve appropriate conservation policies. Steglich & Peters (2002) showed that the agropastoralists' trait preference in cattle in Namibia was disease resistance, size and milk yield. Therefore, as N'Dama taurine cattle is trypanotolerant, farmers preferred to rear this breed, but they crossbred it with the Gobra zebu cattle that is bigger and had higher milk yield. As a consequence, N'Dama genetic pool is being diluted. So conservation policies will focus on improving size and milk yield on N'Dama cattle in order to avoid crossbreeding.

In northwestern Nigeria, Hoffman et al. (2003) showed that the breeds distribution in the landscape varies spatially and seasonally and that different breeds are kept by different producer groups depending on their ecological niche and production systems.

3.2 The farming systems in the Ha Ginag province

TYPOLOGICAL ANALYSIS OF FARMING SYSTEMS IN A NORTHERN PROVINCE OF VIETNAM IN REGARDS TO ETHNIC COMMUNITIES AND THEIR GEOGRAPHICAL LOCATION

Submitted to Agricultural Systems

Abstract

During the last decades, the decrease of the number of reared breeds has led to a loss of domestic animal biodiversity. In the context of a genetic livestock resources conservation project, a multidisciplinary approach was performed in the northern mountainous province of Ha Giang in Vietnam. Interviews of 1,570 farmers were carried out in order to understand their management practices. This study will deal with four livestock species of major interest in the province, namely cattle, buffaloes, goats and pigs, and will focus on the comparison of breeding and husbandry practices between the main five ethnic communities of the province. Multivariate analyses such as Factorial Discriminant analysis (FDA) and Hierarchical Agglomerative Clustering Analysis (HACA) were used in order to characterise farming systems and husbandry practices. It was found that ethnic communities were distributed according the altitude. Two communities: H'mong and Giay inhabited the top of mountains. The Dao community inhabited middle heights, and Tay and Nung communities were found in the plain area. Environmental conditions due to altitude will influence the type of cultivated crops: maize was observed in the heights and paddy rice in the lowlands. Similarly, crops will influence the choice of livestock for draught power as cattle was more suitable in rocky mountains and swamp buffalo in paddy rice fields. As a consequence, H'mong ethnic communities cultivated significantly more maize and raised cattle, while other ethnic communities preferentially cultivated rice and raised swamp buffaloes. All ethnic communities rarely practiced selection of reproductive animals for mating. All farmers practiced free grazing for livestock but the surrounding forest seemed to be an important source for supplementary fodder. For the H'mong ethnic community, district markets were found to have a key role for commercial livestock exchanges, whereas others ethnic communities seemed to exchange within the village or commune.

Introduction

During the last decades, the decrease of the number of reared breeds has led to a loss of domestic animal biodiversity (Scherf 2000, Taberlet *et al.* 2007). In tropical countries, some local breeds have been replaced by breeds highly selected in countries with a temperate climate. These "exotic" breeds hardly support the harsh conditions of tropical husbandry. As a consequence, there is a need to valorise and improve local breeds. Breed descriptions have mainly been focussed on phenotypic and genetic characteristics and emphasised quantitative data. Such information is of interest since the best conditions to preserve a breed are met when its use allows economic benefits by direct goods (meat, milk, …) or services (draught, fertiliser, …). Then, the increasing capacities of a breed will ensure its future use. However, this kind of data will give a picture of the breed at a given time, but it will not elucidate which mechanisms led to this picture and how this picture will be after some generations if similar mechanisms continue. Since breeders influence the genetic structure of their breed, the analysis of management practices can help to understand the genetic make-up of local breeds. Moreover, it has been shown that animal genetic resources managed in a traditional way might retain important genetic characteristics that could be of interest for the future (Granevitze *et al.*, 2007).

Mwacharo & Drucker (2005) showed significant morphometric differences between the zebu population of the Kamba and Massai ethnic communities. Since Kamba and Massai do not inhabit similar zones (plain v.s. mountain) they do not share the same preferences for the body conformation of their livestock. Kumar *et al.* (2006), found out that the river buffalo reared by the Toda ethnic community is genetically more different than other Indian breeds. This was in part due to the fact that the buffalo in this ethnic community occupied a central part in their social life since the buffalo is a sacred animal for this ethnic community and so the Toda people reared their own buffalo without exchanging with other ethnic communities.

The Ha Giang province is a mountainous province in the north of Vietnam. A total of 24 different ethnic communities inhabit this province. Because of a difficult accessibility, this province has remained semi-isolated and only a few exchanges with other provinces do happen. It is usually assumed that livestock from this province may be pure indigenous and might carry genes or traits of interest. In the context of a genetic livestock resources

conservation project, a multidisciplinary approach was performed. Farmers were interviewed in order to understand the management practices of ethnic communities who will be the primary actors of the conservation projects. The present study is the first one in the Ha Giang province. It will deal with four livestock species of major interest in the province, namely cattle, buffaloes, goats and pigs, and will focus on the comparison of breeding and husbandry practices between the main five ethnic communities of the province.

Figure 1. Map of villages sampled in the Ha Giang province.
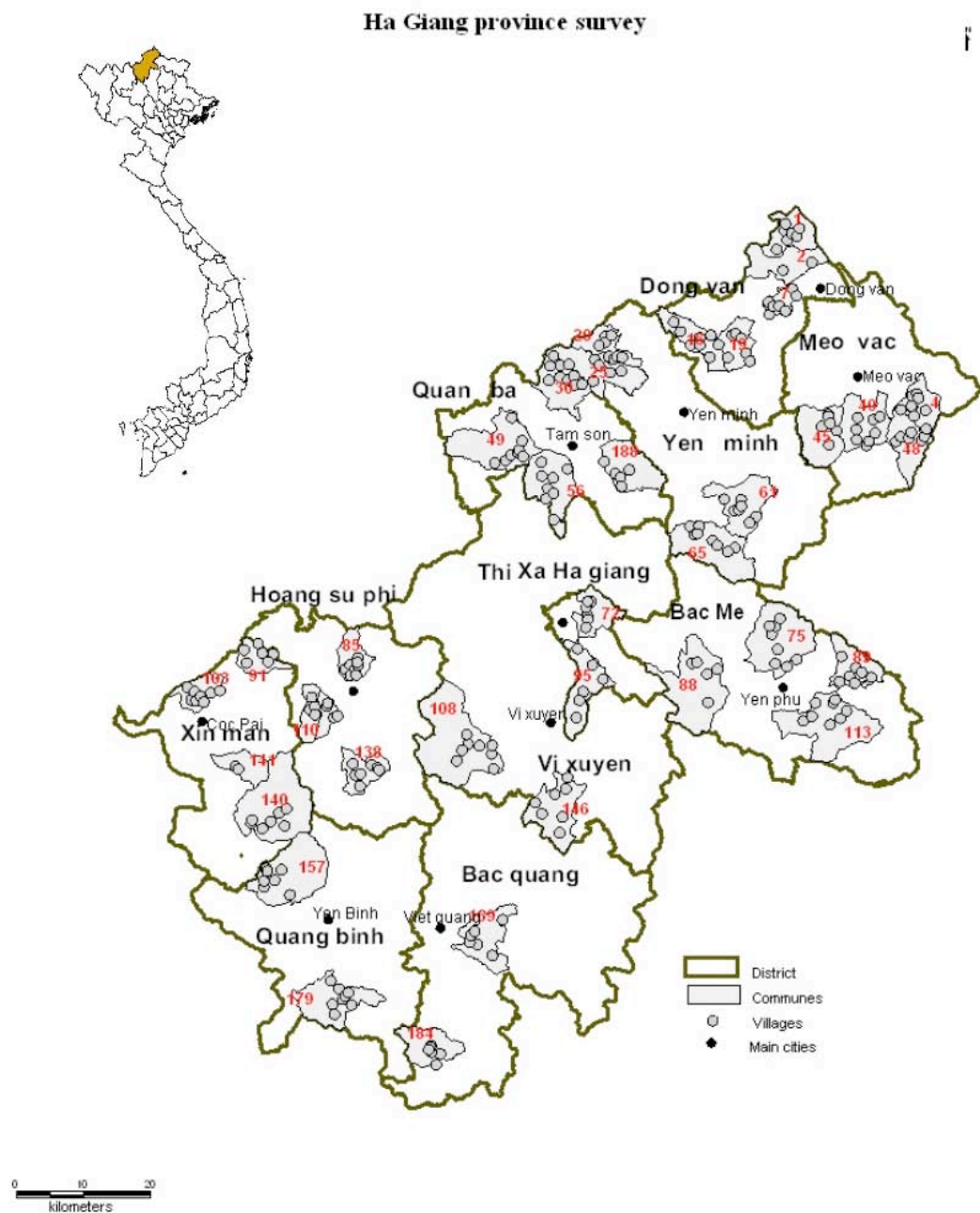
Materials and Methods

The Ha Giang province

The Ha Giang province is one of the northern provinces of Vietnam (22°08'-23°19'N; 104°33'-105°33'E). With a total area of 7884 km^2, it possesses 274 km of border with China. At the southeast, there is the Cao Bang province and at the west, the touristy province of Lao Cai. The Ha Giang province's terrain is strongly partitioned, this partition generates various big and small valleys that result in strong communication difficulties (see, Castella *et al.* 2005, for an analysis of such difficulties in a neighbouring Vietnamese province). Because of its topography, the Ha Giang province could be divided into four zones: 1) a mountainous area in the northeast (Dong-Van and Meo-Vac districts); 2) a second mountainous area in the southwest (Xin-Man and Hoang Su Phi districts); 3) a medium hilly zone (Yen-Minh, Quang-Ba, north Vi-Xuyen, north Bac-Me districts) and; 4) a south-central plain area with the remaining districts.

In Vietnam, 50 ethnic communities are identified. The Kinh ethnic community is the main one, representing 80 % of the population (General statistics office of Vietnam: http://www.gso.gov.vn/default_en.aspx?tabid=494&itemid=1598&idmid=). Twenty-four ethnic communities inhabit the Ha Giang province where the most important are the H'mong and Dao ethnic communities that belong to the Miao-Yao ethno-linguistic family and, the Tay, Nung and Giay ethnic communities from the Thai-Kadaï ethno-linguistic family. It is known through historic records (Savina, 1924) that the Tay, Nung and Giay communities were the first inhabitants of the Ha Giang province and live in the plain while the H'mong and the Dao ethnic communities arrived latter, around 200 years ago, and live on the mountain tops.

Survey methodology

Since this study was a component of a multi-approach conservation project, interviews were conducted simultaneously with biological sample collection. Therefore, the main constraints for survey were due to the organisation of animal sampling (grazing, biological sampling agreement, unrelated animals between farms, etc.). In such a context, a stratified sampling protocol was applied, with 3 strata: district / commune / village. All 11

63

districts of the province were surveyed, with 2 to 5 communes randomly chosen within each district, 2 to 8 villages per communes and 6 to 8 farmers per village. After 14 months of survey, we interviewed 1,570 farmers, from 246 villages distributed in 37 communes (Figure 1).

Data collection

Two sets of data were collected. The first one originated from the Province Agriculture Department and consisted of official census results on crops, livestock and the population in each commune for year 2004. The second set concerned the information collected through single-subject interviews conducted by two trained investigators using a semi-structured questionnaire that was previously tested during a one-week survey. Interviews were carried out in Vietnamese or the ethnic language, according to the farmer's speaking language.

Information was obtained through direct questions for crops and herd structures. Most of the farmers did not know the exact area cultivated or the yield. Therefore crop information was collected in terms of kg of seeds sown or kg rhizomes planted per farm and per year, on the one hand, and kg of seeds or rhizomes harvested per farm and per year, on the other hand. Finally, the crop yield was estimated as the ratio of kg harvested over kg sown or planted.

Farmers group animals from many farms to pasture freely but, in some cases, they practice stalling. So both practices were considered. Also, the fodder was from two origins: leaves cut in the forest and crop residuals. For pigs, seven kinds of feedstuff were recorded: leaves from the forest; "cam" (mushy of vegetables and crop residues); "Bong rio" (residues of rice or maize from alcohol production); rhizome residues (canna or cassava); rice residues; maize residues and vegetable residues.

A way to apprehend management practices and to assess the existence of gene flow was to question farmers about their knowledge of the pedigree and/or the origin of each animal. Eight categories of animal origin (O) were then considered: Farm (f), i.e. the animal was born on the farm where the dam and grand-dame were raised; Village (v), i.e. another farm within the same village; Commune (c); District (d); District Market (dm), i.e. the animal was bought

in the district market, with no information about its farm origin; Different District (*dd*); Different Province (*pd*); Project (*p*), because some poverty alleviation projects have given buffaloes or cattle to farmers, with no information about the origin of the animals. For the Becoming (*B*) of animals, the same categories were used plus one category: Buyers (*buy*), meaning that in some cases buyers go directly to the village or the commune in order to buy animals that will be brought to slaughter houses or to a big market for sale at a higher price. For the origin of the mating male, two specific categories were added: unknown males (*Pu*) and "recognisable related" such as parent-offspring or full-sib (*Pr*).

With the origins of animals established as previously explained, it was possible to established a "genealogy" and therefore possible to estimate age at the 1st parturition, interval between two parturitions, and the average number of offspring per parturition for each reproducing female from the farm.

Data Analysis

Means and standard errors were computed for the quantity of seeds sown and harvested, yield, and reproductive performances. Pairwise differences of herd sizes estimates from official census between districts were tested by the Student test. Significant differences in frequencies of feeding types between ethnic communities were tested by Chi^2 test for each species. Differences in performance values were tested by a linear mixed regression model.

Multivariate clustering analyses have been widely used to characterise typologies of farming systems (Solano *et al.* 2001; Köbrich *et al.* 2003; Zwald *et al.* 2003, Usai *et al.* 2006). The multivariate approach allowed exploiting the large amount of recorded variables in a most efficient way. It was applied to the Farm characteristics table and the breeding-management tables. For each species, a table of frequency counts was done where lines were farmers and columns were origin categories. Statistical analyses were done with the ADE4 package (Chessel *et al.,* 2004; Dray & Dufour, 2007) and Cluster package (Struyf *et al.* 1996) from R software (R development core team, 2007). For each species, these analyses were carried out in two steps: first, with a Factorial Discriminant Analysis (FDA) and, secondly, with a Hierarchical Agglomerative Cluster Analysis (HACA). The FDA will be used to determine which variables separate the communities (i.e. ethnic communities) the

best and which variables are correlated or linked to each other. In order to do that, within- and between-group variability will be considered. It will summarise variable information by creating new synthetic factors. The coordinates obtained from the first three axes of the FDA analysis will be used to perform the HACA analysis in order to build a dissimilarity matrix. The distance between two clusters will be the average of the dissimilarities between the points in one cluster and the points in the other cluster. The HACA will build an agglomerative tree for which the vertical coordinate of the junction of two branches is the dissimilarity between the corresponding clusters. The agglomerative coefficient (*Ac*) (Rousseeuw, 1986) is a quality index of the agglomerative tree obtained. To summarise, the FDA will performthe best synthetic variables for description of farming systems that will be used by the HACA analysis in order to see which communities were similar.

In order to consider the multi-species information and to see if there is differentiation between ethnic communities for all species, we performed a Multivariate Coinertia Analysis (MCOA). The MCOA will combine the information from the tables of species-management and farm characteristics to draw a synthetic position. The MCOA supply three indexes for qualification: 1) the variance within the species-management table, 2) the Cos^2 which express the similarity of the table structure with the synthetic structure and, 3) the typological value (*TV*) which is the discriminative power (i.e. pseudo-eigenvalues).

Results

Official statistics

Population

In 2004, there was on average 36 families per km^2 in the Ha Giang province, ranging from 9 (Bac-Me district) to 194 (Ha-Giang prefecture) (Table 1). The H'mong ethnic community was found to be mainly represented in the northeastern districts of the province (Table 1 and Figure 1). The Dao community was found to be mainly represented in the central districts. In the southwestern mountainous districts, the main represented ethnic community was the Nung community. The Tay community was found to mainly inhabit the plain area.

Crops

From the 78,840,000 ha of the Ha Giang province, approximately 18.7 % were used for family-consumption crops. Main cultures were maize (25.8 %) and paddy rice (20.3 %). The Maize crop was encountered in the northern-eastern districts, with a cultivated area 9 (Dong-Van) and 11 times (Meo-Vac) higher than for paddy rice. However, in these districts, the maize yield was lower than the province average (31 ton/ha), varying from 10 (Dong-Van) to 12 ton/ha (Meo-Vac). Fifty percent of paddy rice was cultivated in the Bac-Quang and Vi-Xuyen districts. The average production in the province was about 56 ton/ha. Other crops were also cultivated in specific areas but with less importance. Cassava was mainly found in the Meo-Vac district (36 % of the cultivated area), canna in the Dong-Van district (37 %), and soya bean in the Xin-Man (19 %) and Hoang Su Phi (17 %) districts.

Livestock

For draught power, the main animal was the swamp buffalo (130,000 heads in the whole province) compared to cattle (65,000). For meat production, the pig stock was found to be much larger (315,000 heads) than the goat stock (99,000). In the mountainous area, the average cattle herd size ranged from 0.7 to 1.4 according to the district (Table 2), whereas cattle were almost inexistent in the plain districts (from 0.0 to 0.1 head per farm). On the contrary, swamp buffalo was almost inexistent in the Dong-Van district (0.1 animal per farm) while the average buffalo herd size was significantly higher ($P<0.05$) than elsewhere in Bac-Me (2.0), Quang-Binh (1.9) and Vi-Xuyen (1.8). Farms with on average more than one goat were found in the three southern districts of Xin-Man, Hoang Su Phi, Vi-Xuyen and the northwestern districts of Bac-Me and Dong-Van. In Bac-Me, the mean number of goats per farm (1.9) was significantly higher than in the other districts. Five districts were found to have an average pig herd size significantly higher ($P<0.05$) than the other ones: Hoang Su Phi (3.5), Vi –Xuyen (3.3), Bac-Me (3.3), Quang-Binh (3.1), Quang-Ba (3.0).

Table 1. Farm and inhabitant density and ethnic community repartition in the Ha Giang province, in 2004 from official statistics, and in our sample.

District	Families/ km^2	Inhab./ km^2	Proportion (%) of inhabitants in the different ethnic communities					
			H'mong	Dao	Giay	Nung	Tay	Others
Dong-Van	30	136	90.8	1.1	0.9	0.2	0.7	6.3
Meo-Vac	16	51	75.8	9.5	7.2	0.4	2.1	5.0
Yen-Minh	19	92	55.6	12.3	9.6	8.4	12.6	1.5
Quang-Ba	18	86	69.1	10.5	2.0	4.9	7.9	5.6
Bac-Me	9	50	26.9	36.4	0.1	2.0	30.0	4.6
Ha Giang prefecture	194	686	2.6	12.3	9.2	0.9	24.5	50.5
Vi-Xuyen	18	82	16.9	34.0	0.9	6.1	27.7	14.4
Hoang Su Phi	24	109	17.3	23.6	0.0	36.1	11.5	11.5
Xin-Man	21	111	28.8	5.4	0.0	52.0	4.5	9.3
Quang-Binh	15	78	6.3	25.4	0.0	4.0	38.5	25.8
Bac-Quang	27	117	8.5	21.1	0.4	5.3	34.1	30.6
Proportion (%) in the whole province			36.1	17.7	2.4	12.4	18.3	14.0
Proportion (%) in our sample			40.4	17.9	2.5	13.0	19.1	7.1

Table 2. Average herd size (± se) per district in 2004 in the Ha Giang province according to the district, from official statistics.

District	Cattle	Buffalo	Pig	Goat
Dong-Van	1.1 ± 0.1	0.1 ± 0	1.5 ± 0.1	1 ± 0.2
Meo-Vac	1.4 ± 0.2	1 ± 0.3	1.8 ± 0.2	0.8 ± 0.1
Yen-Minh	1.2 ± 0.2	1.1 ± 0.3	2.2 ± 0.4	0.9 ± 0.2
Quan-Ba	0.9 ± 0.1	0.8 ± 0.1	3 ± 0.4	0.5 ± 0.1
Bac-Me	1.2 ± 0.1	2 ± 0.2	3.3 ± 0.3	1.9 ± 0.3
Ha Giang prefecture	0.1 ± 0	0.9 ± 0.4	1.6 ± 0.4	0.4 ± 0.2
Vi-Xuyen	0.1 ± 0	1.8 ± 0.1	3.3 ± 0.3	1.2 ± 0.2
Hoang Su Phi	0.8 ± 0.4	1.5 ± 0.7	3.5 ± 1	1.4 ± 0.6
Xin-Man	0.7 ± 0.1	1.1 ± 0.2	2.4 ± 0.2	1.3 ± 0.1
Quang-Binh	0 ± 0	1.9 ± 0.2	3.1 ± 0.3	0.6 ± 0.2
Bac-Quang	0 ± 0	1.2 ± 0.1	2.8 ± 0.2	0.4 ± 0.1
Average	0.7 ± 0.5	1.2 ± 0.6	2.6 ± 0.7	0.9 ± 0.5

Farming crop and livestock interview analyses

Population

Among the 1,570 farmers interviewed, the most represented ethnic community was the H'mong community (34.9 %) and the lowest one was the Giay community (2.3 %) (Table 1). The Chi^2 test was not significantly different when comparing the proportion of the ethnic communities in our sample to the official census. As a consequence, our sampling seemed to be representative of the ethnic composition of the Ha Giang province.

Crops

In the Ha Giang province, there was only one harvest per year for maize and rice crops. Zone 4, as a lowland area, had the best topography for growing paddy rice. Thus, the Tay community showed the highest value of cultivated quantity of paddy (42 kg of seeds sown per family) and the Giay community the highest yield (332.3). The H'mong community, from the upland, showed the highest value for the maize crop (15.9 kg of seeds sown per family) compared to other ethnic communities (Table 3b). The harsh slopes are inappropriate for paddy rice and so H'mong farmers grew only 7.1 kg of seed rice per family. For both crops, the H'mong community had the lowest yield, which may result from the lower fertility of the tops of the mountains.

Livestock

Cattle rearing was practiced by 72 % of the H'mong farmers and only by 15 % of Tay and 35 % of Giay farmers (Table 3d). The average cattle herd size in H'mong farms was about two heads per farm. On the contrary, for the other ethnic communities, more than 85 % of farmers reared swamp buffalo. The highest herd size was observed in Dao farms with an average of 3.6 buffalo per farm (Table 3e). Around 31 % of Giay farmers interviewed reared simultaneously cattle and buffalo, whereas it was less than 16 % of farmers for other ethnic communities. The main ethnic communities that reared goats were the Dao (57 %), the H'mong (46 %) and the Nung (43 %). The goat herd size ranged from 1.5 to 3.7 goats per farm. Rearing pigs seemed to be an essential activity for all ethnic communities, and so more than 90 % of the farmers owned pigs on their farm with a herd size ranging from 3.2 (H'mong) to 6.2 pigs (Dao) per farm. Pig husbandry could be divided into three groups: farmers only involved in reproducing sows and weaning piglets, farmers only involved in fattening pigs after weaning, and farmers involved in both activities. Half of H'mong farmers were only involved in fattening, whereas 58 % of Giay farmers were only involved in fattening. For the three other main ethnic communities, the majority of the farmers were involved in both weaning and fattening (Table 3c).

Table 3. Average (± se) of family size, main crops and percentage of pig husbandry types, according to the ethnic community.

	H'mong	Dao	Nung	Giay	Tay
No of Farmers	*548*	*278*	*262*	*36*	*446*
a) Farm characteristics					
Altitude (m)	955 ± 16	545 ± 14	745 ± 15	494 ± 68	192 ± 8
Family size	6.5 ± 0.1	6.3 ± 0.1	5.3 ± 0.1	5.6 ± 0.3	5.4 ± 0.1
No of Children	2.6 ± 0.1	2.2 ± 0.1	1.9 ± 0.1	1.8 ± 0.2	1.6 ± 0.1
b) Crops					
Maize sown*	15.9 ± 0.5	6.2 ± 0.5	5.6 ± 0.5	4.3 ± 0.6	4.4 ± 0.3
Maize Yield*	125.7 ± 6.1	124 ± 22.2	134.7 ± 8.7	170.9 ± 44.7	137.7 ± 9.3
Rice sown	7.1 ± 0.5	23.8 ± 1.0	27.6 ± 11.4	16 ± 3.0	41.9 ± 11.5
Rice Yield	87.2 ± 4.4	134.6 ± 6.7	178.2 ± 7.7	332.3 ± 30.9	110.4 ± 4.2
c) Pig Husbandry (%)					
Weaning only	23	27	24	58	28
Fattening only	50	15	23	28	31
Weaning and Fattening	27	58	52	14	42
d) Presence of (%)					
Cattle	72	18	22	35	15
Buffalo	44	93	85	87	95
Cattle and Buffalo	16	15	14	31	11
Goat	46	57	43	25	26
Pigs	94	94	93	89	95
e) Herd size					
Cattle	2.0 ± 0.1	0.5 ± 0.1	0.7 ± 0.1	0.9 ± 0.3	0.5 ± 0.1
Buffalo	0.9 ± 0.1	3.6 ± 0.1	2.2 ± 0.1	2.8 ± 0.3	2.6 ± 0.1
Pig	3.2 ± 0.1	6.2 ± 0.3	5.7 ± 0.3	5.9 ± 0.7	5.6 ± 0.2
Goat	2.6 ± 0.2	3.7 ± 0.3	1.7 ± 0.1	2.1 ± 0.7	1.5 ± 0.2

* Maize and Rice planted are in kg of seeds sown per family per year; Yield: kg seed

Reproduction traits

The only significant difference on reproductive traits of livestock following ethnic communities was found for the buffaloes reared by the Giay community, which had their 1st parturition (5.6 years old) significantly older ($P<0.05$) than buffaloes from H'mong (4.6) and Dao (4.5) communities. The age at 1st parturition for cattle ranged from 3.2 to 4.1 years, with an interval between two successive parturitions ranging from 16.2 to 22.6 months (Table 4). For buffalo, the interval between two parturitions ranged from 19.6 to 25.5 months. For goats, age at 1st parturition occurred between 1.4 to 2.5 years. In pigs, the 1st parturition occurred when the sows were about 1.3 years old, with an interval of 8.3 months between two successive parturitions (Table 4).

In goats, the litter size averaged 1.3 kids per goat. Farmers declared to generally obtain only one kid per goat at first parturition and to obtain rather frequently two kids per goat in the following parturitions. The litter size of sows ranged from 5.0 (Dao) to 6.4 (Giay) piglets per litter.

Table 4. Average values (± se) of some indicators of reproductive efficiency, according to the ethnic community.

	Age at 1st parturition (years)	Interval between two successive parturitions (months)	Litter size
Cattle			
H'mong	3.2 ± 0.2	18.3 ± 1.1	-
Dao	4.1 ± 0.3	22.6 ± 2.4	-
Nung	3.5 ± 0.3	19.9 ± 2.2	-
Tay	3.3 ± 0.3	16.2 ± 2.2	-
Buffalo			
H'mong	4.5 ± 0.1	21.5 ± 1.4	-
Dao	4.6 ± 0.1	19.6 ± 1	-
Nung	4.8 ± 0.1	22.3 ± 1.2	-
Giay	5.6 ± 0.3*	25.5 ± 2.4	-
Tay	4.7 ± 0.1	21.6 ± 1	-
Goat			
H'mong	1.5 ± 0.4	10.2 ± 0.6	1.39 ± 0.05
Dao	1.4 ± 0.4	10.5 ± 0.7	1.20 ± 0.05
Nung	2 ± 0.6	10.9 ± 1	1.28 ± 0.08
Tay	2.5 ± 0.5	8.7 ± 0.8	1.23 ± 0.06
Pig			
H'mong	1.21 ± 0.07	8.6 ± 0.4	5.8 ± 0.5
Dao	1.3 ± 0.06	9.2 ± 0.3	5.0 ± 0.4
Nung	1.29 ± 0.08	8.9 ± 0.4	5.3 ± 0.8
Giay	1.4 ± 0.2	7.2 ± 1.6	6.4 ± 0.8
Tay	1.29 ± 0.07	7.6 ± 0.4	5.8 ± 0.4

Typologies of farming systems

Farm characteristics

The FDA of variables of Table 3a-b-c showed high structuring. The 1st factor of the FDA explained 67.8 % of inertia and the 2nd factor explained 28.2 %. The 1st factor separates ethnic communities according to the inhabiting altitude. The 2nd factor was constructed by two groups of variables: (1) rice yield, number of buffaloes, number of pigs, pig breeding; (2) the number of cattle, kg of maize sown and pig fattening. The FDA separates ethnic communities into three clusters: the H'mong community on one side of the first axis, the Tay

community at the opposite side, and the remaining three ethnic communities overlapping at the centre (Figure 2). The H'mong community was mainly characterised by a high inhabiting altitude, with maize as the main crop and cattle as the main livestock as expected from Table 3. The Dao, Giay and Nung communities are defined by the highest rice yield associated with buffaloes and pigs as principal livestock resources, while the Tay community lived in lowlands (192.2 meters height on average).

Figure 2. Ethnic community differentiation on farm characteristics by FDA. Each point represents a farmer interviewed.
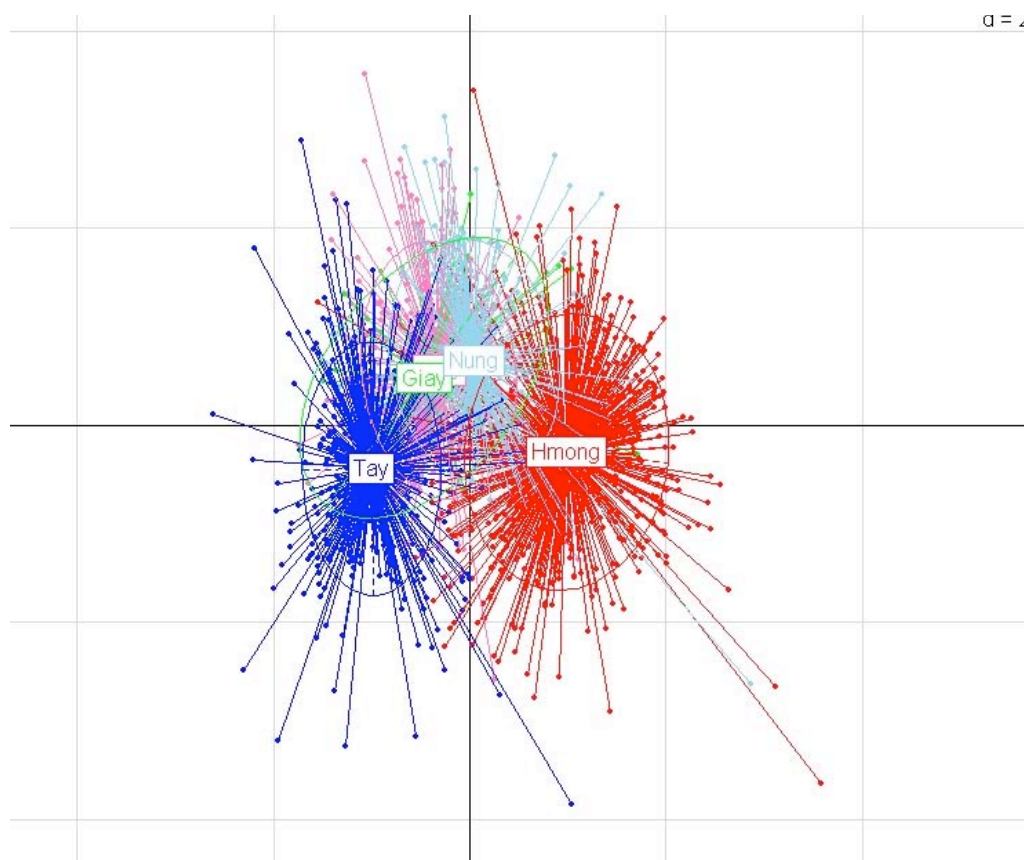
Feeding practices

Association of feeding and fodder records allowed the identification of five feeding typologies of cattle, buffaloes and goats which are the following: only pasture *(P);* pasture and forest leaves *(Pfor);* pasture and crop residuals *(Pcr);* stalling and forest leaves *(Sfor)*; and pasture with stalling and forest leaves *(PSfor).* The proportion of typologies encountered in ethnic communities are summarised in Table S1 and for pigs in Table S2. The Chi^2 test showed significant differences in feeding type frequencies between ethnic communities for

each species. Among all ethnic communities and species, the H'mong community was the one using the stalling practice for livestock the most. When animals were stalled, fodder originates from the forest. Farmers went in the surrounding forest every day to cut leaves for fodder. The Nung and the Tay farmers always supplemented grazing with forest fodder provision. The Dao and Giay communities almost only practiced free pasture without fodder provision. The crop residuals are mainly only used for pig husbandry. The H'mong ethnic community fed pigs with principally maize, forest fodder and canna or cassava residuals. The Tay, Dao and Nung communities fed pigs first with vegetable residuals then with rice and maize crop residuals. The Giay communities prepared "cam" that is a mushy of vegetable and crop residues.

Breeding practices

Average rates of variables used for description of breeding practices are summarised in Table S3 (a, b, c). The proportion of inertia explained by the two factorial components of FDA on breeding practices per species ranged from 80 % (pig husbandry) to 89 % (cattle husbandry).

Cattle management was characterised by the important rate of animals originating from a different province for the Tay community (15. 2 %); and their use for reproduction (2.2 %) on the contrary to the other ethnic communities that never used animals from other provinces for cattle breeding (Table S3). The Dao and Nung communities benefited the most from cattle offered by the Developing Project with 2.2 % and 1.6 % of their cattle livestock respectively. Also, since all mating occurred during free pasture, it was supposed that cows mate with a different bull each time that leads to the high *FL* ratio (0.95) compared to the Tay community (0.88) and H'mong community (0.85). The HACA showed that H'mong and Nung cattle breeding practices were the most similar whereas the Tay community was the most different (Figure 3a). The *Ac* value reached 0.37.

For buffalo breeding practices, the factors that contribute the most in the construction of the first axis were the reproductive unknown bull (*Pu)* and the sale at the district market (*Bmd*). The *Pu* values ranged from 0.712 (Tay) to 0.926 (H'mong), and *Bmd* values ranged from 0.001 (Tay) to 0.078 (Nung). The second axis separates ethnic communities according to the number of animals sold to buyers (*Bbuy*); the highest values were observed for the Tay (0.165) and the Dao community (0.117). Selection of bulls within the village (*Pv*) was also a

variable allowing differentiation between ethnic communities. This practice was applied at the highest rate (0.254) by the Tay community. A dendrogram clustered first H'mong and Nung people followed by the Tay community; the Giay community is the one who seemed the most different in breeding practices for buffaloes (Figure 3b). The *Ac* value was low (0.24).

For goat breeding, the Dao community raised goats over generation since almost a half (42%) of the goats originate from the family (*Of*). The Dao farmers sold their goats in the village (*Bv*) in 24.1 % of the cases. The Dao and H'mong communities rarely practiced selection by choosing the male and so *Pu* values averaged 0.705 and 0.732 respectively. The Nung community differed from other ethnic communities by selling their goats in the district market (*Bmd*) in 21.7 % of cases (Table S3). The H'mong and Tay communities clustered together with the Nung community and Dao lately clustered being the most different (Figure 3c) but the *Ac* value was low (0.23).

Figure 3. Agglomerative tree of ethnic communities according to their breeding- practices, for: a) cattle; b) buffaloes; c) goats; d) pigs.
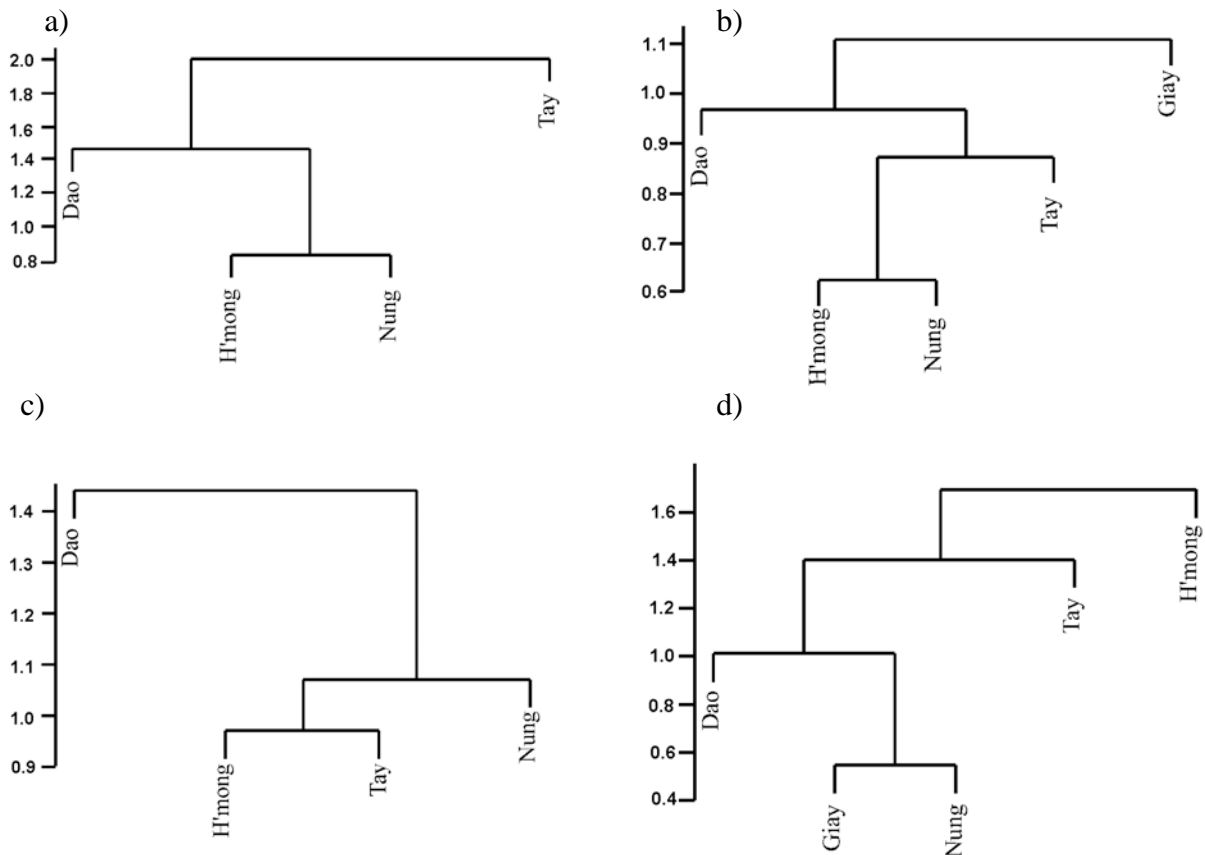
Figure 3.1 Distribution of livestock according to official data in the commune sampled. a) Proportion of cattle and buffaloes. b) Proportion of goats and pigs.
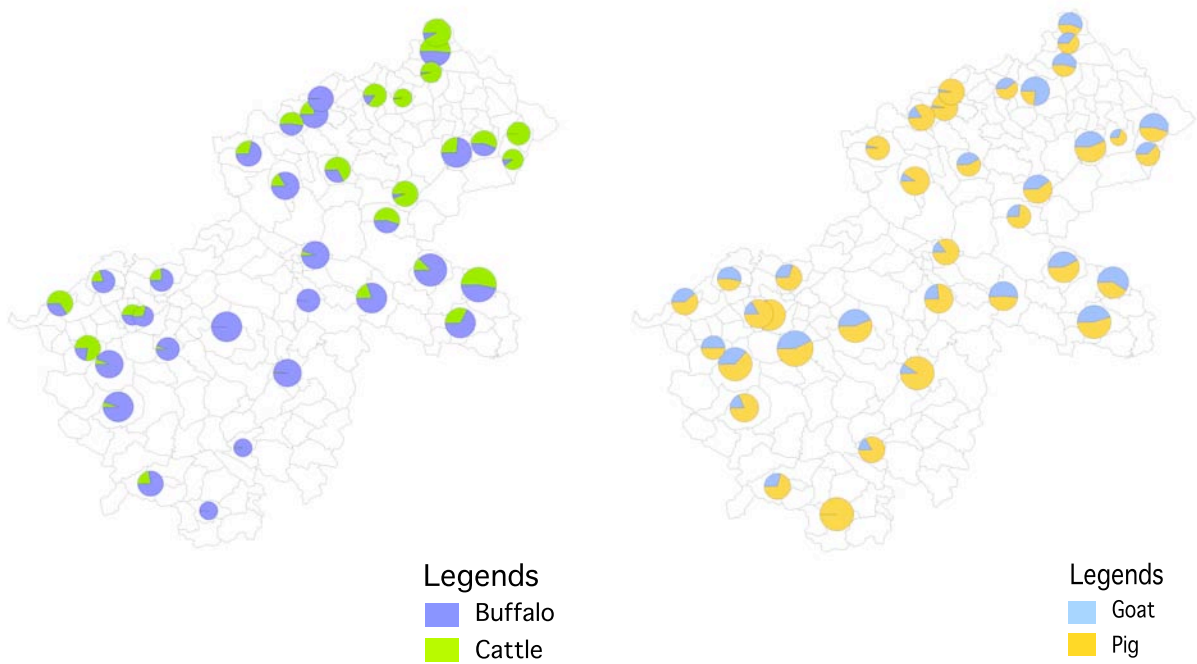
Figure 3.2 Proportion of pig husbandry types according to farmers interviews (in blue : fattening only; in light green: weaning only; in dark green: weaning and fattening).
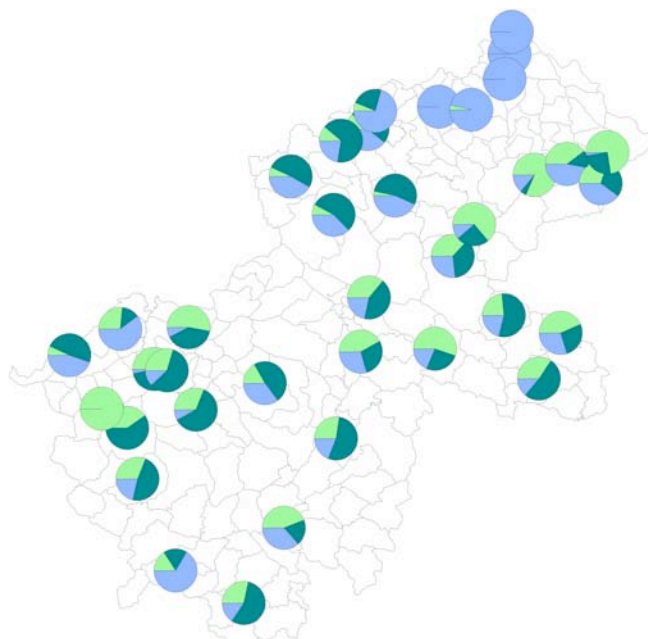
Four main factors seemed to differentiate pig breeding between ethnic communities. More than a half (52%) of pigs in Dao farms was born on the farm (*Of*). On the contrary, the H'mong community mainly exchanged in the district market that constitutes an important component of the pig husbandry: 12.7 % of pigs were bought in the market and 23.9 % were sold. The Tay and Giay communities bought and sold animals in the village at the same frequency that they keep animals born on the farm. The *Ac* value for pigs was the highest (0.42), and the dendrogram shows firstly the Giay and Nung communities clustering followed by Tay and Dao communities; here the most different community for pig husbandry was the H'mong community (Figure 3d).

The Giay community was removed from the MCOA analysis because they do not rear cattle. Then for the four other ethnic communities we combined the farm characteristics-pig husbandry types, on the one hand, and cattle, buffalo, goat and pig breeding practices, on the other hand. There was a good congruence between the MCOA synthetic position obtained for ethnic communities and the ones obtained within each of the five management tables with Cos^2 values ranging from 0.777 to 0.999. For the first axis, the highest typological values *TV* were observed for the Cattle management practices and the Farm characteristic tables (Figure 4a). For the second axis, discrimination of farming systems between ethnic communities was also due to the cattle management practices. For the third axis, discrimination of ethnic communities resulted from farm characteristics and goat breeding-management practices (Figure 4b). The buffalo breeding-management practice almost did not contribute to the differentiation of farming systems. The synthetic position of the ethnic farming systems, showed that considering the first and second axes, the H'mong and Nung overlapped and Tay and Dao community are well separated and on the opposite side (Figure 5a). Whereas, for axes 1 and 3, H'mong and Dao were grouped and Nung and Tay were far away in both directions (Figure 5b).

Discussion

General characteristic

One of the most interesting and of highest importance result for future development projects was that the surrounding forest supplies fodder for all ethnic communities, except the Dao community.

Figure 4. Percentage of variance, $\cos^2$ and *TV* for the first axis (a) and third axis (b) of the MCOA analysis (the picture for the second axis is very similar to the first one).
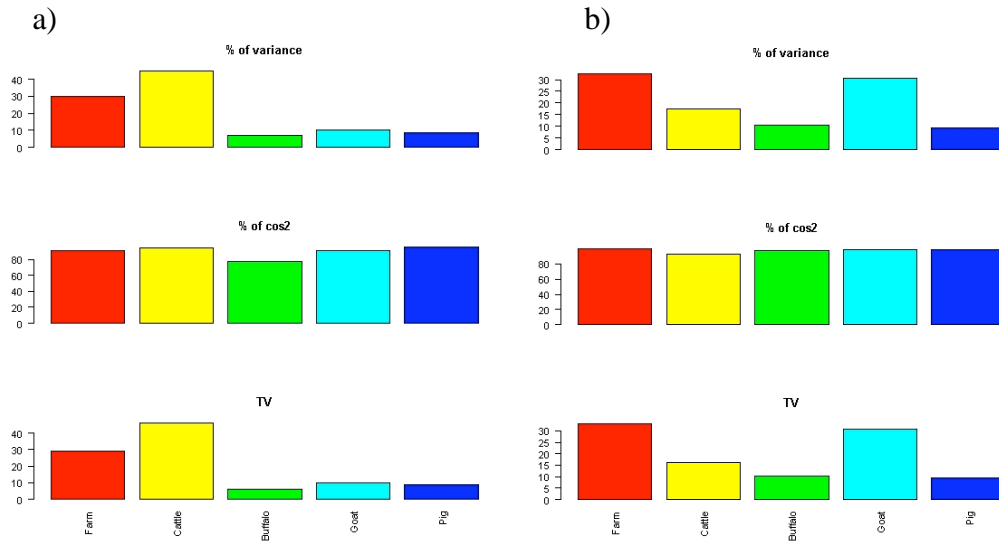
Figure 5. Synthetic MCOA representation of ethnic communities a) for the first and second axis, b) for the first and third axis. Each point represents the position of an ethnic community considering one of the five tables (Farms characteristics, Cattle, Buffalo, Pig and Goat management).

77
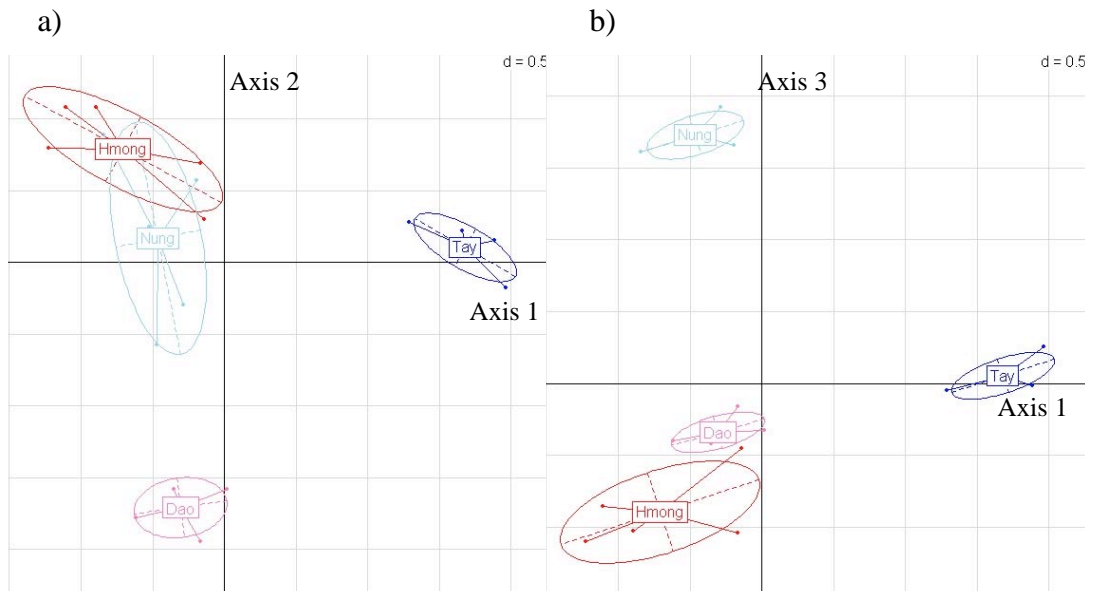
This might be associated with the fact that crop residues were mainly used for pig husbandry only. In fact, farmers in the Ha Giang province are not able to perform efficient storage of fodders and so have difficulties in supplying fodder for livestock. In order to offset such a constraint, farmers established a "produce-sharing" system. Farmers, which have enough to buy more than one animal but not enough to feed it, lend it to a poorer family without an animal and no savings to buy but with fodders to give. Then, they will share the offspring between the two families. This "produce-sharing" system established between farmers alleviates poverty. However, the efficiency of such a system could be increased by improving the storage of fodder, which also will allow limiting deforestation.

Breeding management through sire selection was rarely practiced. No breeding selection (*Pu*) always represented more than 70 % of matings in cattle, buffaloes and goats. In cases where sires were selected, the main criterion was the proximity and, as a consequence, sires were chosen within the village herd. Hence, for goats and pigs, reproduction between closely related animals (i.e. parent-offspring; full-sib) was also practiced. For pig breeding, 40 % of boars were unknown and almost all of the others (30-70 %) came from the same village. Such observations can influence in many ways conservation project policies. Widely, improvement projects proposed to establish a "sire centre" where farmers bring females for mating with proven sires. Such a project may not be very successful in the Ha Giang province as suggested by our results. Constraints in transport of females in remote areas and so by foot, or difficulties to observe oestrus may discourage farmers. Therefore, bringing semen (i.e. artificial insemination) directly to farmers or implementing farmer cooperatives within villages or communes with systems of sire exchanges might be more appropriate since it will not compel-farmers to travel too far.

However, even if our analysis showed common traits of practices between ethnic communities, it also showed important diversity that should be taken into account. Accessibility constraints might be an important consideration for development projects since both data sets (i.e. official and interviews) showed a stratigraphic distribution of ethnic communities. Michaud *et al.* (2002), studied the relationship between ethnicity and topography in the neighbouring Vietnamese province, namely Lao Cai, with a similar topology as the Ha Giang province. In other words, the crops and the livestock used will be

adapted to the ecological environment where the ethnic community lives. Projects involving cattle livestock should be addressed to H'mong or Nung communities, whereas buffalo projects should be implemented with other ethnic communities. Similarly, our data highlighted the cattle stalling practice in the H'mong community. Discussion with farmers revealed that many farmers keep their biggest bull stalled, since they believed that mating would negatively influence their docility and fieldwork. So, if a cattle cooperative is established in a village or commune, the bulls proposed for reproduction should not belong to any farmer who would use it for field working. Also, since the district market has a key role in commercial exchanges and in the social life of H'mong people, one possibility could be to establish a reproductive centre or other institutions close to district markets. The Bac-Me district, followed by Hoang Su Phi and Vi-Xuyen are the districts where the herd size for almost all species considered were the biggest. In these three districts, the Dao community is an important community, while the Tay community is the second one in the Bac-Me and Vi-Xuyen districts, and the Nung community is the most numerous in the Hoang Su Phi. Therefore, those ethnic communities from these districts are already acting as breeders and so, their needs for improving their husbandry may be different than in other places and with other ethnic communities.

In conclusion, we show that the H'mong community was the main community for cattle husbandry and that they are mainly pig fatteners. They may be the most dependent on this livestock resource. In addition, they used the district market for commercial exchanges, whereas other communities use buffaloes and practice weaning and fattening of pigs. In general, it seems that storage of fodder and animal performances need to be improved. This could be done through a cooperative at village or commune levels because the social and commercial life is already well established at those scales.

Acknowledgments

References

Castella, J-C., Manh, H.P., Kam, P.S., Villano, L., Tronche N.R., 2005, Analysis of village accessibility and its impact on land use dynamics in a mountainous province of northern Vietnam. Appl. Geogra., 25, 308-326

Granevitze, Z., Hillel, J., Chen, G.H., Cuc, N.T.K., Feldam, M., Eding, H., Weigend, S., 2007, Genetic diversity within chicken populations from different continents and management histories. Anim. Genet., doi:10111/j1365-2052.2007.01650

Köbrich, C., Rehman, T., Khan, M., 2003, Typification of farming systems for constructing representative farm models: two illustrations of the application of multi-variate analyses in Chile and Pakistan. Agr. Syst., 76, 141-157

Michaud, J., Turner, S, Roche, Y., 2002, Mapping ethnic diversity in highland Northern Vietnam. Geojournal, 57, 305-323

Mwacharo, J.M., Drucker, A.G., 2005, Production objectives and management strategies of livestock keepers in South-East Kenya: Implications for a breeding programme. Trop. Anim. Health Prod., 37, 635-652

Rousseeuw, P.J., 1986, A visual display for hierarchical classification, In: E Diday, Y., Escoufier, L., Lebart, J., Pagés, Y., Schektman, J., and R Tomassone (Eds), Data Analysis and Informatics 4 (North-Holland, Amsterdam), 743-748

Savina, F.M.,1924, Histoire des Miao. Hong Kong. Societé des Missions Etrangères Scherf, B.D., 2000, World Watch List for Domestic Animal Diversity, 3rd edn. Food and Agriculture Organization of the United Unions, Rome

Sokal, R.R., Rohlf, F.J., 1969, Biometry: the principles and practice of statistics in biological research. Ed: W.H. Freeman and Company, New York

Solano, C., León, H., Pérez, E., Herrero, M. 2001, Characterising objectives profiles of Costa Rican dairy farmers. Agr. Syst., 67, 153-179

Struyf, A., Hubert, M., Rousseeuw, P.J. 1996, Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, http://www.stat.ucla.edu/journals/jss/

Usai, M.G., Casu, S., Molle, G., Decandia, M., Ligios, S., Carta, A., 2006, Using cluster analysis to characterize the goat farming system in Sardinia. Livest. Sci., 104, 63-76

Taberlet, P., Velentini, A., Rezaei, H.R., Nader, S., Pompanon, F., Negrini, R., Ajmone-Marsan, P., 2007, Are cattle, sheep, and goats endangered species? Mol. Ecol., 17, 275-284

Zwald, N.R., Weigel, K.A., Fikse, W.F., Rekaya, R. 2003, Identification of factors that cause genotypes by environment interaction between herds of Holstein cattle in seventeen countries. J. Dairy Sci., 86, 1009-1018

Supplementary Material

Table S1. Feeding typologies rates per ethnic groups for cattle, buffaloes and goats

	P	Pfor	Pcr	Sfor	PSfor
Cattle					
HMong	0.50	0.07	0.00	0.39	0.03
Dao	0.93	0.04	0.00	0.02	0.02
Nung	0.75	0.25	0.00	0.00	0.00
Giay	1.00	0.00	0.00	0.00	0.00
Tay	0.76	0.24	0.00	0.00	0.00
Buffalo					
HMong	0.83	0.14	0.00	0.03	0.01
Dao	0.94	0.05	0.00	0.00	0.00
Nung	0.79	0.20	0.01	0.00	0.00
Giay	0.94	0.06	0.00	0.00	0.00
Tay	0.82	0.15	0.02	0.00	0.01
Goat					
HMong	0.74	0.06	0.01	0.19	0.00
Dao	0.93	0.05	0.01	0.00	0.00
Nung	0.71	0.28	0.00	0.00	0.00
Giay	1.00	0.00	0.00	0.00	0.00
Tay	0.89	0.10	0.00	0.00	0.00

P: pasture; *Pfor*: pasture and forest leaves; *Pcr*: pasture and crop residuals; *Sfor*: stalling and forest leaves; *PSfor*: pasture, stalling and forest leaves

Table S2. Pig feeding per ethnic groups

	Forest Leave	Canna sp./ Cassava	Maize	Rice	Vegetable	Bong rio	Cam
Hmong	0.28	0.12	0.33	0.08	0.12	0.04	0.03
Dao	0.15	0.01	0.22	0.26	0.31	0.01	0.03
Nung	0.17	0.05	0.22	0.27	0.28	0.00	0.01
Giay	0.25	0.05	0.25	0.14	0.12	0.00	0.18
Tay	0.07	0.00	0.24	0.33	0.35	0.01	0.01

Bong rio: crop residues from making rice or maize alcohol; cam: mushy of vegetables and crop residues

Table S3 Rates of Breeding-management practices per species per ethnic group: a) Origin (*O*) of the animals from de farm; b) Becoming (*B*) i.e. destination; c) Origin of the reproductive male (*P* for paternity).

a)

Species / Ethnic groups	ORIGIN							
	Of	Ov	Oc	Od	Op	Omd	Odd	Opd
Cattle								
Hmong	0.161	0.267	0.154	0.206	0.002	0.129	0.014	0.025
Dao	0.067	0.233	0.267	0.122	0.022	0	0.133	0.022
Nung	0.156	0.25	0.234	0.188	0.016	0.031	0.031	0
Tay	0.059	0.353	0.088	0.191	0	0	0.039	0.152
Buffalo								
Hmong	0.24	0.305	0.146	0.145	0	0	0.045	0.002
Dao	0.35	0.327	0.145	0.115	0.004	0.01	0.025	0.011
Nung	0.314	0.258	0.181	0.067	0.003	0.003	0.09	0.005
Giay	0.51	0.177	0.063	0	0	0	0	0
Tay	0.332	0.263	0.182	0.07	0.01	0.004	0.05	0.028
Goat								
Hmong	0.242	0.417	0.205	0.03	0.045	0	0	0
Dao	0.417	0.335	0.12	0.045	0.031	0.014	0.002	0.035
Nung	0.24	0.242	0.143	0.059	0.068	0.138	0.045	0
Tay	0.144	0.252	0.291	0.089	0.057	0.008	0.061	0
Pig								
Hmong	0.26	0.347	0.163	0.048	0	0.127	0.012	0.003
Dao	0.515	0.289	0.112	0.015	0	0	0.013	0.035
Nung	0.355	0.286	0.144	0.027	0	0.073	0.025	0
Giay	0.318	0.364	0.136	0	0	0	0.091	0
Tay	0.364	0.389	0.138	0.045	0	0.015	0.021	0.021

b)

Species	BECOMING						
Ethnic groups	Bv	Bc	Bd	Bmd	Bdd	Bpd	Bbuy
Cattle							
Hmong	0.256	0.159	0.185	0.173	0.036	0.02	0
Dao	0.333	0.1	0	0	0.033	0	0
Nung	0.13	0.211	0.125	0.128	0.031	0	0
Tay	0.235	0	0.059	0.059	0.059	0	0
Buffalo							
Hmong	0.236	0.136	0.066	0.025	0.016	0.021	0
Dao	0.336	0.102	0.085	0.001	0.038	0.029	0.117
Nung	0.194	0.201	0.088	0.078	0.02	0.023	0.015
Giay	0.313	0.031	0.073	0.063	0.083	0	0
Tay	0.278	0.16	0.076	0.011	0.047	0	0.165
Goat							
Hmong	0.119	0.03	0.071	0.096	0	0	0
Dao	0.241	0.105	0.047	0	0.027	0	0.039
Nung	0.071	0.104	0.019	0.217	0.003	0	0.019
Tay	0.141	0.158	0	0.026	0.012	0	0.024
Pig							
Hmong	0.174	0.122	0.033	0.239	0.001	0	0
Dao	0.276	0.094	0.052	0.015	0.02	0	0.033
Nung	0.312	0.062	0	0.208	0	0	0.008
Giay	0.364	0.182	0	0.182	0	0	0
Tay	0.371	0.164	0.007	0.011	0.004	0	0.016

84

c)

Species	PATERNITY									
Ethnic groups	Pu	Pf	Pv	Pc	Pd	Pdd	Pr	Pp	Ppd	FL
Cattle										
Hmong	0.622	0.034	0.237	0.017	0.008	0	0.017	0	0	0.85
Dao	0.836	0.067	0.022	0.044	0	0	0.03	0	0	0.95
Nung	0.745	0.018	0.206	0.031	0	0	0	0	0	0.95
Tay	0.787	0.044	0.074	0	0	0.039	0.034	0	0.022	0.88
Buffalo										
Hmong	0.926	0.032	0.031	0.011	0	0	0	0	0	0.98
Dao	0.823	0.003	0.141	0.029	0	0	0.003	0	0	0.93
Nung	0.825	0.02	0.132	0.02	0.003	0	0	0	0	0.98
Giay	0.825	0	0.175	0	0	0	0	0	0	1
Tay	0.712	0.015	0.254	0.015	0.003	0	0	0	0	0.98
Goat										
Hmong	0.732	0.097	0.11	0.02	0.03	0	0.01	0	0	0.89
Dao	0.705	0.06	0.158	0.014	0.014	0	0.021	0.014	0	0.92
Nung	0.547	0.032	0.274	0.048	0.022	0	0.057	0.007	0	0.92
Tay	0.598	0.049	0.211	0.121	0.006	0	0.015	0	0	0.90
Pig										
Hmong	0.432	0.032	0.305	0.031	0.007	0	0.085	0.002	0	0.84
Dao	0.429	0.122	0.358	0.005	0	0	0.057	0	0	0.84
Nung	0.337	0.063	0.421	0.013	0.02	0	0.105	0	0	0.89
Giay	0.227	0	0.773	0	0	0	0	0	0	0.76
Tay	0.272	0.058	0.439	0.097	0.034	0	0.046	0	0.005	0.82

f: farm; v: village; c: commune; d: district; dd: different district; md : market district; p : project; pd : different province; buy: buyers; u: unknown; r : related (parent-offspring, full-sib); FL: nb males/parturition

3.3 Conclusion

According to the farming systems analysis, some results may affect the genetic pattern of livestock breeds in a significant way:

- the reduced localisation of cattle in the tops of the mountains
- the wide distribution area of swamp buffalo which is reared by almost all ethnic communities
- the low human selection v.s. the high natural selection pressure
- the importance of the social organisation in village or commune where almost all sale/purchase exchanges of living animals occurred
- the importance of district market for the H'mong community in the eastern districts of the Ha Giang province.

CHAPTER 4

THE H'MONG CATTLE

REVEALING SUBPOPULATIONS AND POSSIBLE ADMIXTURE IN THE VIETNAMESE H'MONG CATTLE BREED FOR CONSERVATION PURPOSES

C. Berthouly, L. Pham Doan, X. Rognon, T. Nhu Van, G. LeRoy, H. Hoang Thanh, B. Bed'Hom, D. Laloë, C. Vu Chi, V. Nguyen Dang, E. Verrier, J-C. Maillard

4 THE H'MONG CATTLE
4.1 Taxonomic status, distribution and domestication of cattle

Whereas taxonomic classification of Bovidae is not yet totally clear, domestic cattle belong to the *Bovini* tribe from the family of Bovidae. The *Bovini* tribe is composed of three genera including the *Bison* with American (*Bison bison*) and European bison (*Bison bonasus*), the genera of *Poëphagus* for which the Yak is the only representative species and the genera *Bos* subdivided into two sub-genera: *Bibos* with the gaur, the banteng and the kouprey, and the sub-genera *Bos* to which domestic cattle belong (Figure 4.1). Within the domestic cattle, there are two recognised forms of domesticated cattle, the humpless *taurine* cattle of Europe, West Africa and northern Asia (*Bos taurus*) and the humped *zebu* cattle of southern Asia and Africa (*Bos indicus*).

Figure 4.1 Phylogeny of Bovini tribe modified from Hassannin & Ropiquet (2004).
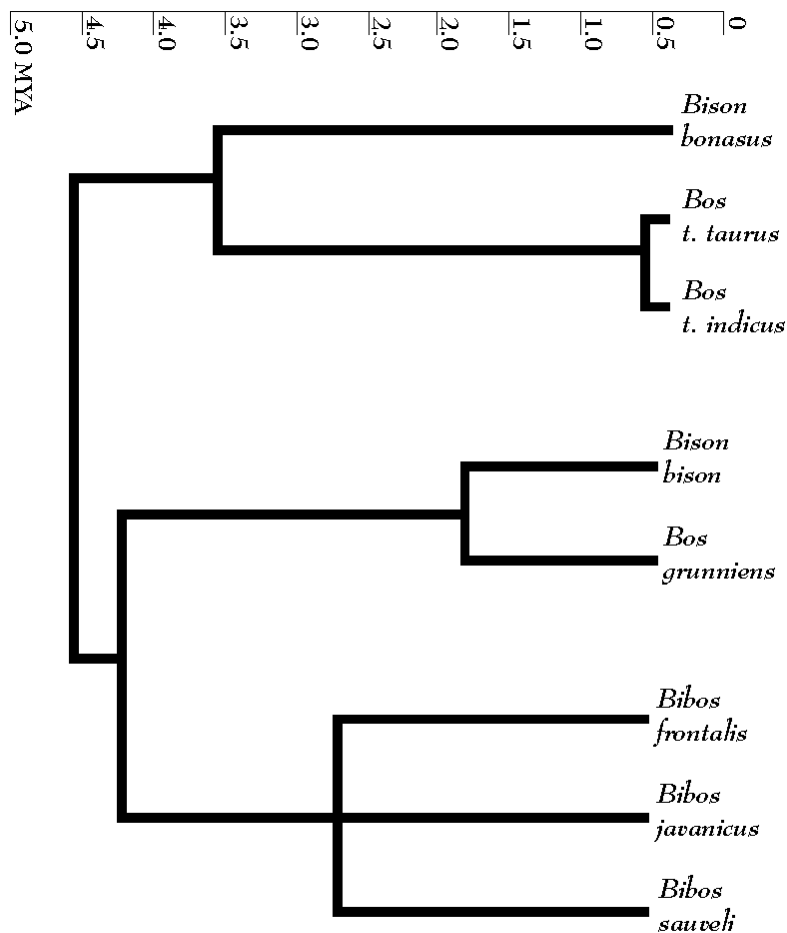
The Bovini are very similar in the number and gross morphology of their chromosomes. All species have 2n=60 with 58 acrocentric autosomes except the gaur which has 2n=58. The major morphological difference between chromosomes of cattle concerns the Y-chromosome. Taurine cattle, yak, gaur and banteng all have a submetacentric Y-chromosome (Baker & Manwell, 1991). However, humped zebu (*Bos indicus*) cattle and both species of bison have an acrocentric Y-chromosome (Gupta *et al.,* 1974).

The wild ancestor of all domesticated cattle was the extinct aurochs, *Bos primigenius* (Grigson 1980; Epstein & Mason 1984). The aurochs is thought to have evolved in Asia from a Pliocene ancestor *Bos acutifrons*. These aurochses spread from Asia, westward into Europe and southward into North Africa. The auroch had an extremely large geographical range from Britain to China. The large distribution area gave rise to at least three distinct geographical subspecies: the Eurasian subspecies (*Bos primigenius primigenius*); a North African subspecies (*Bos primigenius opisthonomus*) and a southern Asian subspecies (*Bos primigenius namadicus*) (Figure 4.2).

Figure 4.2 Distribution area of the European auroch (*Bos p. primigenius*); the Asian auroch (*Bos p. namadicus)* and the African auroch (*Bos p. opisthonomus).*
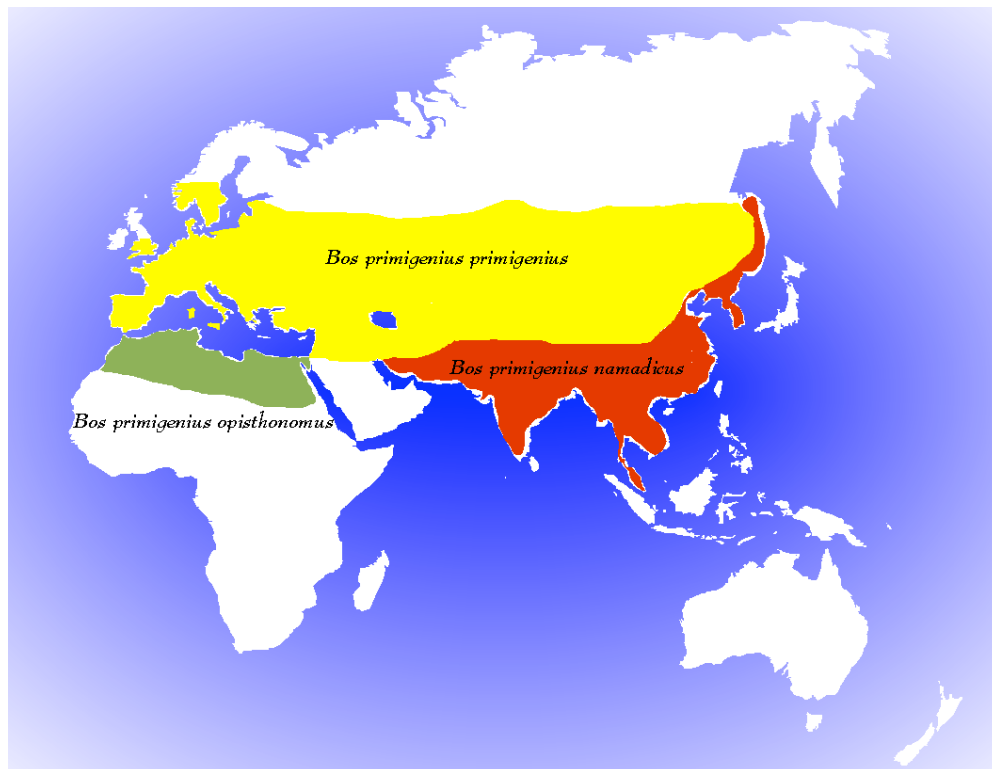
modified from Payne (1970)

Two major theories about cattle domestication have been proposed. The first one argues that both types, *B. taurus* and *B. indicus,* originated from a single common domestication event. The second hypothesis is that *B. indicus* were domesticated independently from *B. taurus* somewhere on the Indian subcontinent.

According to Epstein & Mason (1984), *B. taurus* cattle are descendent from the Asian aurochs (*B. primigenius namadicus)* resulting from domestication by Neolithic farmers in the Middle East approximately 8 000 years ago. These animals were originally longhorned and spread to many parts of the world following the migration of human groups. Subsequently, the morphometry of this cattle evolved into the shorthorn and humped type as an answer to ecological changes and economic pressures. To this first theory is opposed a second theory, suggested by Grigson (1980) based on considerable differences in craniometrist and other osteological measurements which can only be explained by separate domestications from two quite distinct groups of aurochses. She suggested that *Bos taurus* was originally derived from the northern Eurasian aurochs (*Bos primigenius primigenius*) and that the southern Asian aurochs (Bos primigenius namadicus) provided the ancestor for *Bos indicus* cattle.

A molecular clock approach based on mtDNA highlighted that separation between taurine/zebu preceded the domestication period of the Neolithic (Loftus *et al.,*1994a ,1994b ; Bradley *et al.,* 1996). The divergence time between the two main mtDNA clades was then estimated from 200 000 to 1 000 000 BC. Troy *et al.* (2001) had estimated sequence variations of the mtDNA control region of 392 animals sampled in Europe, Africa and the Near East. They identified four haplotypes. The African cattle were mainly represented by the haplotype T1, whereas the haplotypes T, T2 and T3 were observed in European and Near-East cattle populations. Moreover, the mtDNA of African cattle, barely present in the Near East, was phylogenetically different from the other *Bos taurus*. Those results lead authors to hypothesise the existence of a different domestication centre of *Bos taurus* in Africa (Bailey *et al*., 1996; Troy *et al.*, 2001).

Through the genetic diversity study of fifty African cattle populations, Hanotte *et al.* (2002) proposed the hypothesis of a possible domestication in Central Africa followed by two introgressions: a first one by *Bos taurus* from the Near East; then by zebu (*Bos indicus*) from the Horn of Africa through an Arabic terrestrial route or by commercial sea routes with India.

All African cattle shared taurine mtDNA (Bradley *et al.,* 1996). Introgression of zebu genes would have occurred by male's introgression and then spread from east to west (Hanotte *et al.* 2000). Freeman *et al.* (2004), through different admixture measures, assessed the hypothesis of a diffusion starting from the Horn of Africa to the East, but above all, the measurements provided further evidence of an African genetic taurine pool which has been gradually replaced by a European/Near-East taurine pool and an Indian zebu pool. This African genetic pool, which still remains in West Africa in the N'Dama breed for example, has persisted in the distribution area of the Tsetse mosquito solely because of his trypanotolerant aptitudes.

Earlier admixture ages were estimated for South-African breeds than for North-African breeds. This scenario implies that zebu (*Bos indicus*) introgression occurred before pastoralism spread southward, and these first hybrids were the progenitor of the South-African breeds (Freeman *et al.,* 2006).

While *Bos taurus* was domesticated between 8 000-10 000 years BC in the Near East and in Africa, archaeological studies showed proof of domestication in the Indus valley. This domestication seemed to be anterior of the Harappa civilisation. Evidence of domestication beginnings was found in the Baluchistan region providing an estimated date of around 9 000 BC (Meadow, 1989). Osteological analysis of the *Bos* material from Mehrgarh in Baluchistan indicated that these cattle were *Bos indicus*, whereas the known distribution of *Bos namadicus* during Pleistocene overlapped the Baluchistan region and this species would be the most likely candidate for the wild ancestor of the domesticated cattle found in this area.

Therefore, an earlier civilisation could have started the domestication process in the Baluchistan region around 10 000 – 70 000 BC, and then spread, migrating and probably interacting with the Harappan civilisation which appeared around 4 000 BC.

Although the origins of European, African and Indian cattle have been widely studied, it is only recently that several studies about cattle from East Asia were published. The older domestic cattle from North-East Asia were dated about 5000 years old, and incidentally several thousand years after the domestication of aurochs began (Payne & Hodges, 1997). Around the second century AD, cattle migrated from North China via the Korean peninsula to Japan, corresponding with rice cultivation introgression (Mannen *et al.,* 1998).

The majority of studies focused on the origin of Chinese cattle. This genetic pool is commonly subdivided into three categories:

- the Mongolian type in North China
- the Huanghuai type distributed in the central and south basin of the Yellow river (Huanh He) and the Huai He river basin
- the Changzhu type around the Yangtze and Zhujiang river from South China

Yu *et al.* (1999) studied cattle mtDNA from the Yunnan, Hainan, Guangdong and Guizhou provinces from South China. In these provinces, cattle belong to the Changzhu type. The study showed that 5 haplotypes were present in those breeds. The haplotype I and II belonging to the zebu and taurine type respectively were simultaneously observed in all breeds except in the Dehong cattle and Hainan cattle. Within these breeds, two new haplotypes IV and V were found. The haplotypes differed by four and one restriction sites with haplotype I respectively. Therefore, they are assumed to be zebu type. The haplotypes I and IV may have diverged 268 000-538 000 years ago, therefore prior to the beginnings of domestication which started approximately 8 000- 10 000 years ago. At the same time, Yu *et al.* (1999) revealed introgression of Yak mtDNA in female hybrids in the Dinqing cattle breed. The Dinqing cattle are reared in the mountainous (2800-3000 m) north-western region of Yunnan province, which is also the distribution area of yaks. Therefore Tibetan farmers have long reared yaks and cattle together. More recently, Mannen *et al.* (2004), studied the variability and frequencies of mtDNA from *Bos taurus* and *Bos indicus* in cattle breeds from Mongolia, Korea and Japan. They observed four haplotypes (T, T2-T4) from *Bos taurus* with a definite geographical distribution. Based on the absence of haplotype T4 in cattle populations from the Near East, Europe and Africa, Mannen *et al.* (2004) suggested an independent domestication in East Asia thereby joining with previous conclusions made by Yu *et al.* (1999). Taking into account these preliminary studies, Lei *et al.* (2006) studied 14 breeds from West and South-West China. The three haplotypes T2, T3 and T4 were present in the Chinese breeds and consequently confirmed the assumption of a domestication centre in central Asia. The mtDNA of *Bos indicus* was subdivided into two clades I1 and I2. The haplotype I1 was distributed from North to South China whereas the haplotype I2 was only observed in the breeds from the south and south-west of China. Because both haplotypes were also observed in cattle from Nepal, India and the Philippines, Lei *et al.* (2006) concluded that *Bos indicus* might have been gradually introduced into China. The south and the south-west of China, geographically nearest to India, received a more important contribution of genetic pool from *Bos indicus* than the rest of China. In order to better understand the phylogeny

of Chinese cattle, 22 breeds were simultaneously studied by Cai *et al.* (2007) for D-loop on one hand and by Lei *et al.* (2006) for Y-chromosome on the other hand. All breeds from the north of the Yellow river were found to be taurine type except for the Kazakh breed from the North East, which had been submitted to a zebu male introgression. Among the eight breeds from the region south of the Yangtze river, four breeds were of the zebu type for both mtDNA and Y chromosome while the remaining four had undergone an introgression of zebu female only. Thus, the *Bos taurus* cattle might have been introduced into China through Mongolia towards the huge and extensive area to the north of the Great Wall between 3799 and 3122 years before the present time, whereas *Bos indicus* migrated from India to the North, in the central plain area, and stopped by the Qinling mountain. Thus, *Bos indicus* interbred with *Bos taurus* in the central and south Chinese regions leading to the present day hybrid Chinese breeds.

Figure 4.3 Hypothetic migratory roads for taurine and zebu cattle.
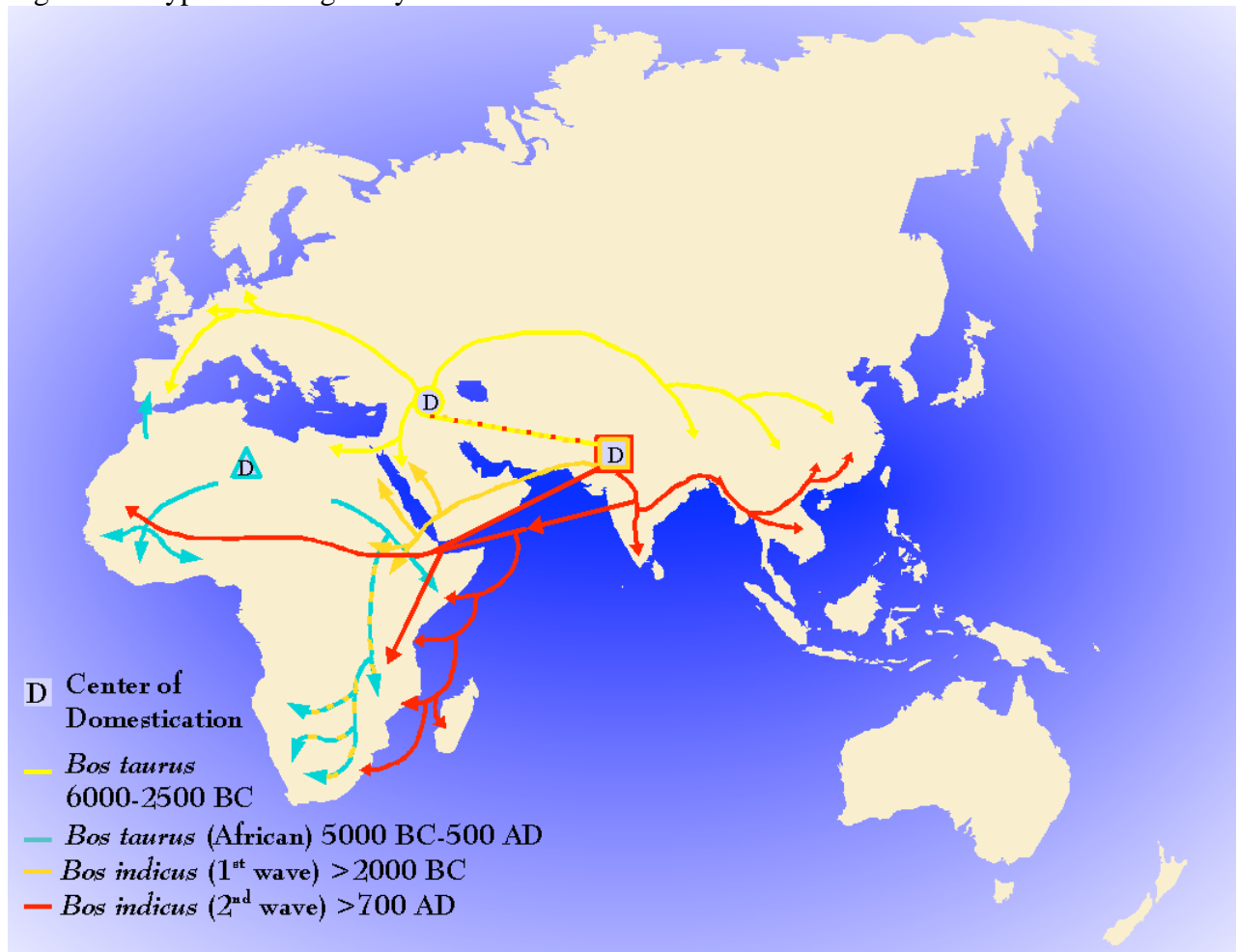
modified from Graphics unit, ILRI (2006)

94

4.2 Role of cattle

4.2.1 Role In the socio-economy of Asia

Both *B. taurus* and *B. indicus* types are found in Asia; nevertheless most local cattle breeds are of *Bos indicus* type. Asian cattle breeds are less productive than Western breeds, but they are adapted to harsh conditions. Moreover, they are for multipurposes: meat, milk, draught, manure… These last years, the world cattle population has increased. However this trend was not uniform in all cattle-rearing countries in Asia (Table 4.1). Although an important increase was observed in South-East Asia, countries in South Asia exhibited a decline in cattle population (Figure 4.4).

Table 4.1 Changes in cattle population in different regions of Asia between 1990-2002.

Region	Year		% Change
	1990	2002	
South Asia	260 999 632	255 265 040	-2.2
South-East Asia	33 895 418	39 686 868	17.1
Asia	519 940 054	442 560 390	-14.9
World	1 297 868 474	1 334 815 521	2.8

The most important decline was observed in Indonesia with an annual decrease of 0.48 % per year since 1990. The highest increase of cattle population was observed in Vietnam where the total herd was doubled. Indeed, cattle milk production in Vietnam has increased by nearly 500 % during the last 16 years and cattle meat production about 100 %. Similar increases were observed in countries such as China or Thailand. In the Philippines and Sri Lanka a decrease of about 20 % was observed in cattle milk production (Table 4.2).

4.2.2 Role in the socio-economy of Vietnam

With a land surface area of 301 000 km^2, 31 % of the land is used for agriculture (Figure 4.5). Vietnam mainly produces rice (paddy) on the plains (36 thousand million tons in 2005), maize (3.7 thousand million) and coffee (752 million) in mountainous areas. In 2006 Vietnam became the 2nd exporting country of rice and Robusta coffee in the world.

Figure 4.4 Evolution of cattle herd in some Asian countries between 1990-2006.
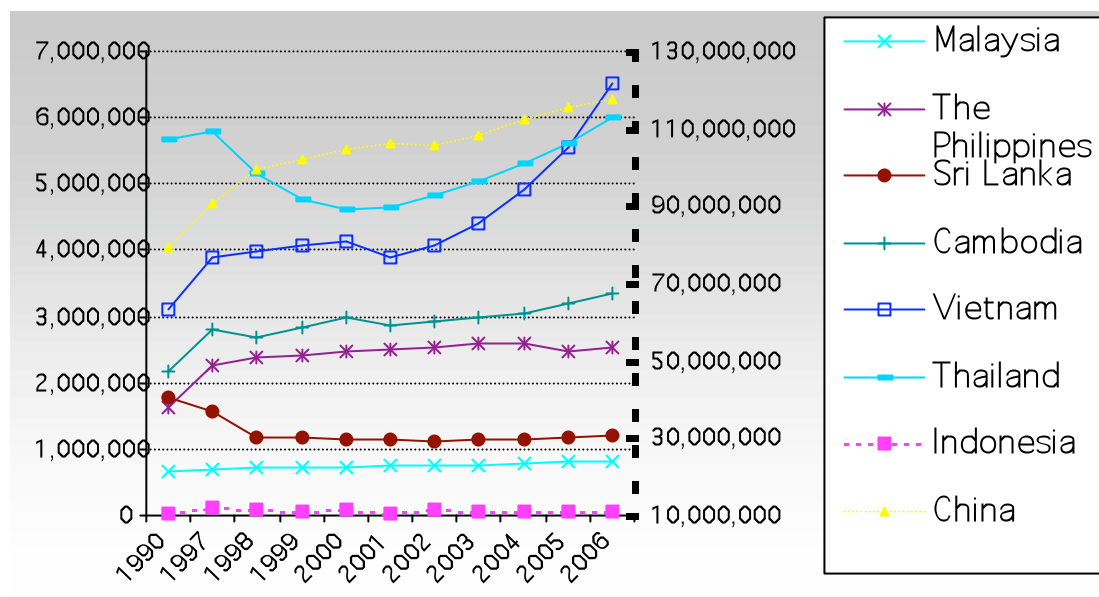
Source: FAOSTAT

Table 4.2 Cattle milk and meat production in some Asian countries between 1990-2006.

Country	Years				% Changes
	1990	1995	2000	**2006**	
Cattle milk production (tons)					
Cambodia	17 000	19 125	20 400	22 700	33.5
China	4 362 638	6 082 213	8 632 308	32 249 100	639
Malaysia	29 200	37 100	30 460	37 650	30
The Philippines	15 000	12 110	10 210	12 00	-20
Sri Lanka	182 600	214 059	127 740	139 340	-24
Thailand	130 278	307 229	520 115	826 465	534
Vietnam	36 000	40 800	54 456	215 00	497
Cattle meat production (tons)					
Cambodia	28 200	39 600	56 700	60 240	113
China	1 143 788	3 296 527	4 991 163	7 172 800	527
Malaysia	11 113	16 261	14 515	21 254	91
The Philippines	82 000	97 360	190 159	166 816	103
Sri Lanka	20 350	27 490	28 720	26 250	29
Thailand	180 129	254 281	171 490	175 989	-2.3
Vietnam	75 000	83 000	92 268	181 000	141

Figure 4.5 Land use in Vietnam in 2005.
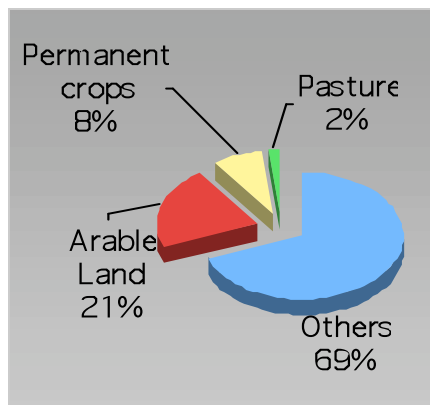
Vietnamese agriculture was under great pressure due to human density. This pressure was heavier on husbandry as only 2 % of lands are used for pasture. Excepted for the fish meat, if we consider the three principal meat sources, which are cattle, pigs and poultry, an average Vietnamese person passed from 65 g/day of meat consumption in 2001 to 86 g/day in 2005. While this average meat consumption was much lower compared to the French (163 g/day), the change in consumption habits put pressure on husbandry activities. This increase of meat consumption was equally spread among meat types, and so between 2001 and 2005, the proportion of cattle meat consumption remained equal to 10 % of the total Vietnamese meat consumption (Figure 4.6).

Figure 4.6 Evolution of proportion of meat type consumption between 2001-2005.
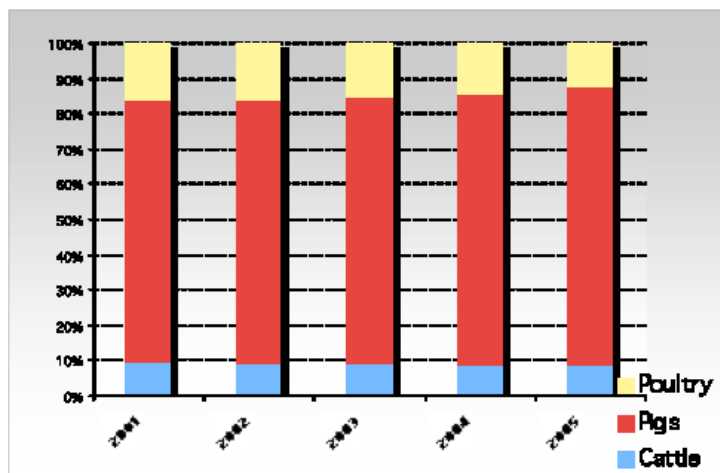
Cattle livestock was multiplied by 1.4 between 2001-2005, increasing from 3.9 million heads to 5.5 million. The highest density of cattle was localised in the southern provinces of Hanoi and around the Mekong river delta (Figure 4.7). The Thanh Hoa province reared 6 % of the Vietnamese cattle population, the two central provinces of Gia Lai and Binh Dinh reared 5 % and 5.2 % respectively. In the Northern mountainous provinces, cattle population was lower : Cao Bang (2.24 %), Ha Giang (1.31 %) and Lao Cai (0.35 %). Except for Da Nang province where cattle population decreased (Figure 4.7), on average the cattle population increased by 1.72 %. This increase was mainly effective in the Mekong river delta, whereas in northern provinces previously cited, cattle population remained stable (1.25; 1.20 and 1.09 respectively).

Figure 4.7 Density and herd size of cattle population in Vietnam. a) Density of cattle per km2 in Vietnam for 2005; b) Differences in herd size between 2001-2005 in Vietnam.
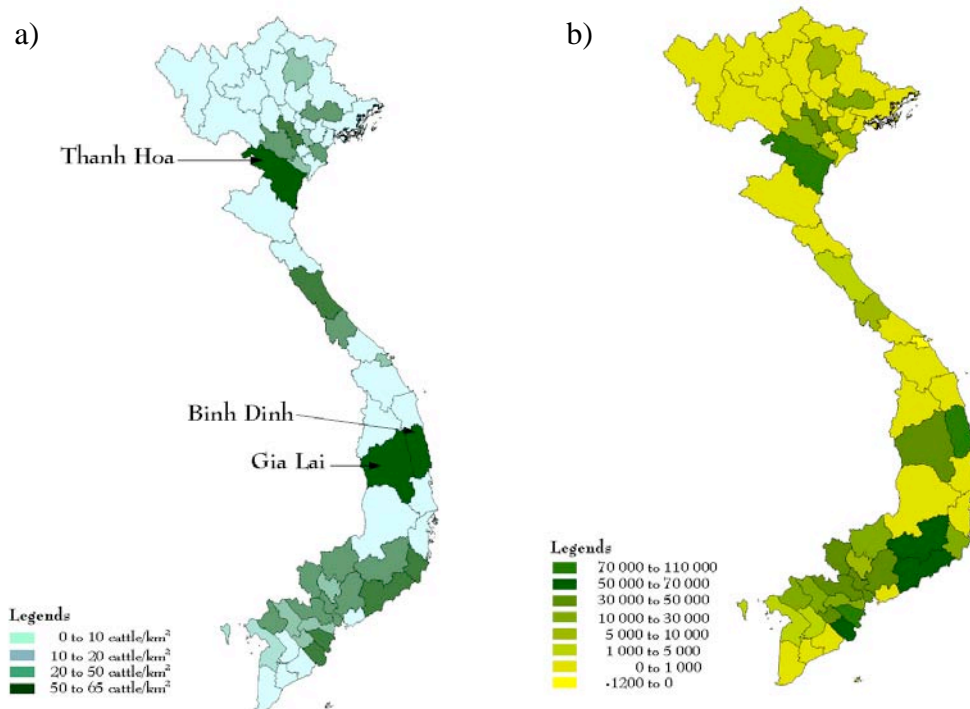
Source: Statistical Year Book of Vietnam

In 2005 Vietnam produced 142 160 000 tons of cattle meat. Meat was produced in the provinces where cattle is reared and so, mainly in the vicinity of Hanoi and Ho Chi Minh cities. The Binh Dinh province showed the highest value of meat production with more than 9 million tons (Figure 4.8.A). Since 2001, meat cattle production has increased however. This increase seemed to be insufficient to supply population needs as, simultaneously, the importation of cattle meat increased from 2002 and exportation strongly decreased in 2004 (Figure 4.8.B).

4.3 Diversity of cattle

Over the world, a total of 990 cattle breeds have been reported, 897 are classified as local or indigenous breeds and 93 as transboundary breeds, meaning distributed over the world. The 2/3 of local breeds are distributed in Africa, Europe and Asia, which possesses more than 25 % of the local breeds (Table 4.3). Among the 258 breeds reported in Asia, 11 % were classified at risk, 51 % were not in danger whereas the status of the remaining breeds (38 %) was unknown.

Figure 4.8 Vietnamese Cattle meat production. a) Cattle meat production (x 1000 tons) in 2005; b) Evolution of meat cattle production; importation and exportation between 2001-2005.

a)

b)
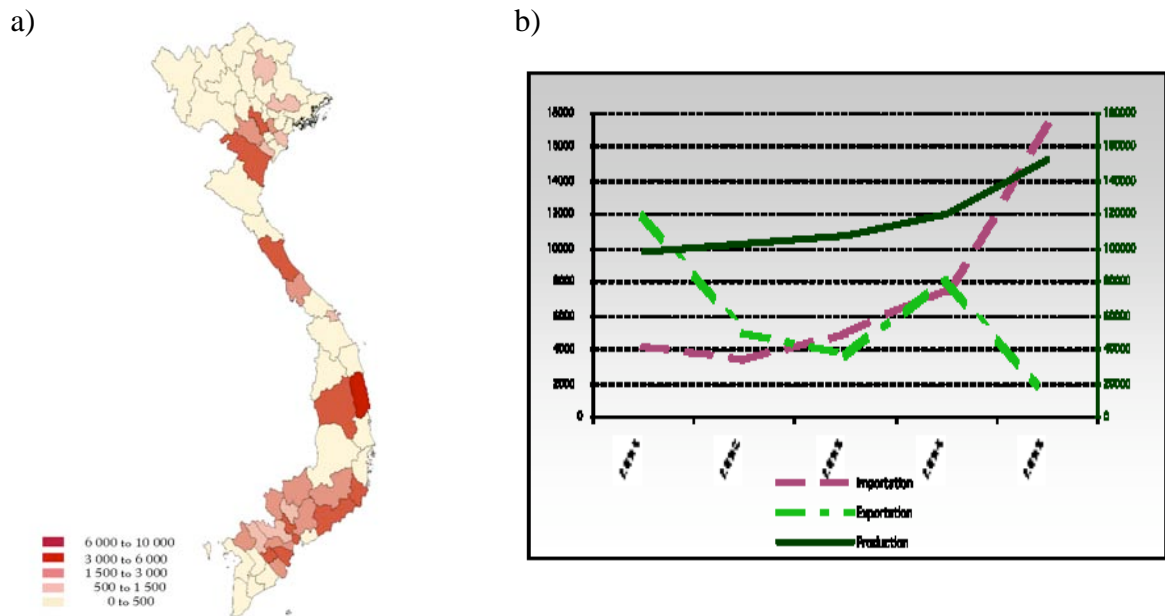
Table 4.3 Number and proportion per region of cattle breeds represented according to World herd.

Breeds		Africa	Asia	Europe	Near-East	Latino-America	North-America
Local Breeds	Number	154	239	277	43	129	29
	%	17.2	26.6	30.9	4.7	14.4	3.2
Transboundary Breeds	Number	35	19	28	0	8	3
	%	37.6	20.4	30.1	0	8.6	3.2

4.3.1 Asian breeds

The cattle of Asia can be subdivided into humped and humpless cattle. There are also sizable numbers of Bali cattle (domesticated Banteng, *Bos* [Bibos] *javanicus*) and Mithun cattle (domesticated gaur, *Bos [bibos] frontalis*) in South-East Asia. The humped zebu breeds are more prevalent in the southern regions of Asia, particularly in India and Pakistan. Humpless taurine cattle are found across most of the Asian continent in the northerly regions.

On the Indian subcontinent, cattle are venerated in Hindu society and occupy a position of special significance. Cattle in India are seldom slaughtered for meat and they are mainly used for

milk, draught and dung which is used for fuel and as a building material. Outside their home area, South-Asian breeds have been most successful in tropical Latin America and Africa. The Sahiwal, from Pakistan and India, has been introduced to 12 African countries. Also, South-Asian animals have made a major contribution to composite breeds used elsewhere in the tropics. These include the Santa Gertrudis (Shorthorn x Brahman crosses, found in 34 countries around the world), Brangus (Angus x Brahman), etc. However, South-East Asian cattle breeds, like Chinese breeds, have not been taken out of their home region.

In China, swamp buffalo is preferred for draught power in rice cultivation. However, in the mountainous area where other crops are cultivated, cattle is also used. With agricultural modernisation and haphazard crossing in the past 20 years, eight cattle breeds have become extinct and a few others are currently in danger. Chinese cattle are subdivided into three groups: the turano-mongolian type above the Yellow river which is of taurine type, the Changzhu type below the Yangtze river which is zebu type, and in the intermediary central area the Huanghuai or Yellow cattle type is reared, which is a hybrid of taurine and zebu. Coat coloration of Chinese cattle varies from light yellow to black, and horns are of small or medium size (Table 4.5). The zebu hump is smaller than in the Indian breeds, but dewlap is as well developed. Except for the Diqing breed, crossed with the Yak and measuring nearly 100 cm , and the Luxi breed at 140-150 cm; most of the Chinese breeds averaged between 110 - 130 cm height for males and between 105-120 for females. South-East Asia is inhabited by wild cattle (Bos (bibos) *spp.).* Indeed, the domestication of the banteng seems to have occurred in Indonesia and then spread across South-East Asia. Studies on the geographical distribution of the Hbb alleles of hemoglobin showed that the majority of South-East and East Asian breeds have resulted from three-way hybridisation, thereby possessing genes from *B. taurus*, *B. indicus* and *Bos (bibos) spp.* ancestors (Table 4.4).

Figure 4.9 Map of China (a) and Vietnam (a) showing the distribution area (1 colour per breed)
of the main cattle breeds.

a)   b)
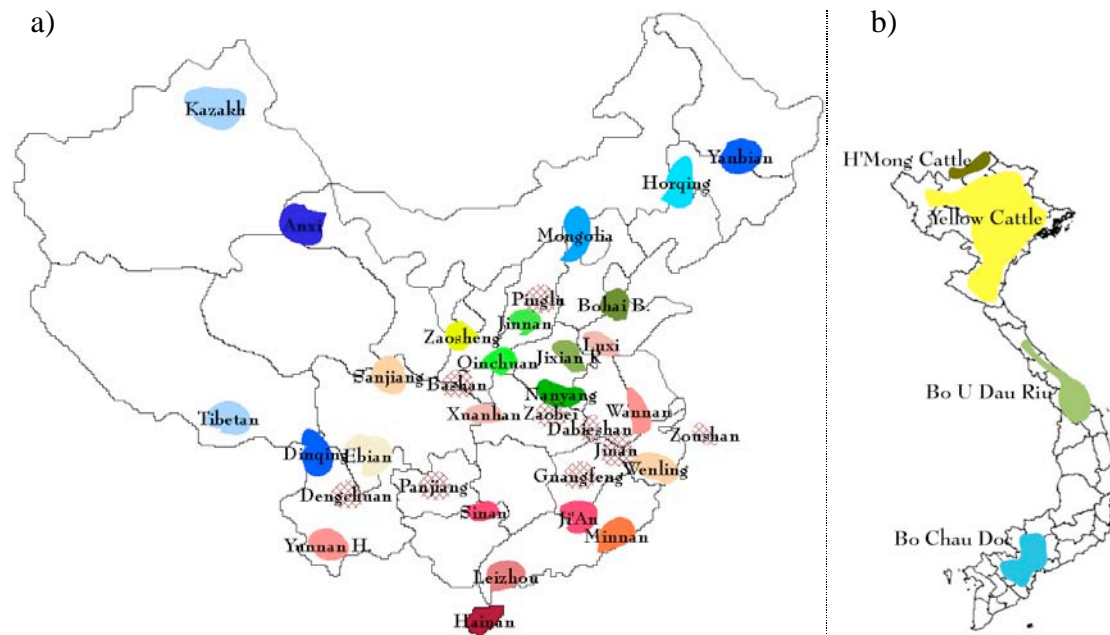
Chinese cattle breeds: Breeds in blue tones had Y Taurine type only; Breeds in red tones had Y

Zebu type only; Breeds in green tones are hybrids Y Taurine-Zebu type; and hatched breeds have

not been studied (Cai *et al.* 2006). B) Vietnamese cattle breeds

Table 4.4 Estimated proportion of genes derived from *Bos taurus*, *Bos indicus* and *Bos (bibos)*
type cattle (Namikawa, 1981)

Country	Cattle breed	Proportion of genes originating from		
		Bos taurus type	*Bos indicus* type	*Bos (bibos)* type
Indonesia				
Bali	Bali	0.083	0.292	0.791
Java	Javanese	0.374	0.395	0.231
Madura	Madurese	0.456	0.328	0.217
Sumatra	Sumatran	0.236	0.569	0.194
Japan	Japanese black	0.974	0.026	0.000
Korea	Korean native	0.799	0.185	0.015
Malaysia	Kedah-Kelantan	0.604	0.285	0.111
Philippines				
Luzon	Philippine Native	0.107	0.723	0.170
Palawan	Philippine Native	0.109	0.641	0.250
Taiwan	Yellow	0.030	0.864	0.106
Thailand				
North	Thai	0.123	0.823	0.054
Northeast	Thai	0.422	0.490	0.088
South	Thai	0.539	0.346	0.115

According to the FAO network, only a few breeds per country were identified. In fact, it is

often recognised that only one breed was reared in the country but bears different names

according to the geographical zone where it is reared. For example, the Philippines or Malaysia had only one indigenous breed. In Thailand, there was a variety of Thai breeds used as fighting cattle. In Cambodia, four breeds were reported where two, namely the lowland and highland Khmer, are crossbred with banteng.

4.3.2 Vietnamese breeds

Lucien Choquart, veterinary inspector since 1921 in Tonkin reported that:

" *Bien avant l'occupation européenne, le bétail de la haute région était amené dans le delta par troupeaux qui suivaient, non pas des routes, mais des sentiers.» Avec une moyenne de 15 a 20 Km par jour, les arrêts étaient longs et fréquents. […]De temps en temps, un repos complet d'1 ou 2 jours s'imposait. Il en résulte que des villages termes d'une étape journalière sont devenus petits à petits de véritables gîtes d'étapes, et que dans certains villages où un repos plus long était possible, des ventes d'animaux avaient lieu. »*

These "*gîtes d'étapes*" (stop-over-points to rest travelling cattle) gradually became big cattle markets. Choquart detailed the different routes taken by the herds and the multiple destinations (Fig. 4.10). We could then see that Viet-Tri constituted at that time an important market where animals originating from the Ha Giang, Lao Cai and Son La provinces gathered.

French military companies wrote many reports about the cattle from Indochina, and more specifically about the Tonkin cattle (Douarche, 1902). These documents reported that in the Tonkin, some taurine cattle could be found but the majority of the cattle was of zebu or hybrid types. Moreover, the taurine type seemed to originate from Cambodia or Laos.

The French inspectors classified the Tonkin cattle into four types:

- Cattle from Thanh Hoa province

- Cattle from Vinh province

- Cattle from Cao Bang

- Cattle from the Red river delta

However, it was considered that only three mountainous breeds really existed, the breed from the Red river delta being the result of multiple crossbreeding between the three others. Moreover, among the three mountainous breeds, two of them, the Thanh Hoa and the Vinh breed, may have come from bordering countries. These assumptions were corroborated by information that Choquart (1928) collected on the routes followed by herds (Figure 4.10).

However, it was surprising that the authors considered the Thanh Hoa and Vinh cattle as two distinct breeds when descriptions written by Choquart (1928) showed that Thanh Hoa cattle came from the Vinh province, and previously from Laos. Also, according to Douarche (1902), the third mountainous breed could be subdivided into subgroups identified under different names depending on the herd origin: the Chinese cattle, the Lang Son cattle, the Cao Bang or the Ha Giang cattle. The Ha Giang cattle were often used for provision of meat supplies for the military posted in the highlands; many herds were sent to Tuyen Quang province (Lunet & Lajonquiere, 1904).

Figure 4.10 The different routes taken by cattle during the French colonial period.

103
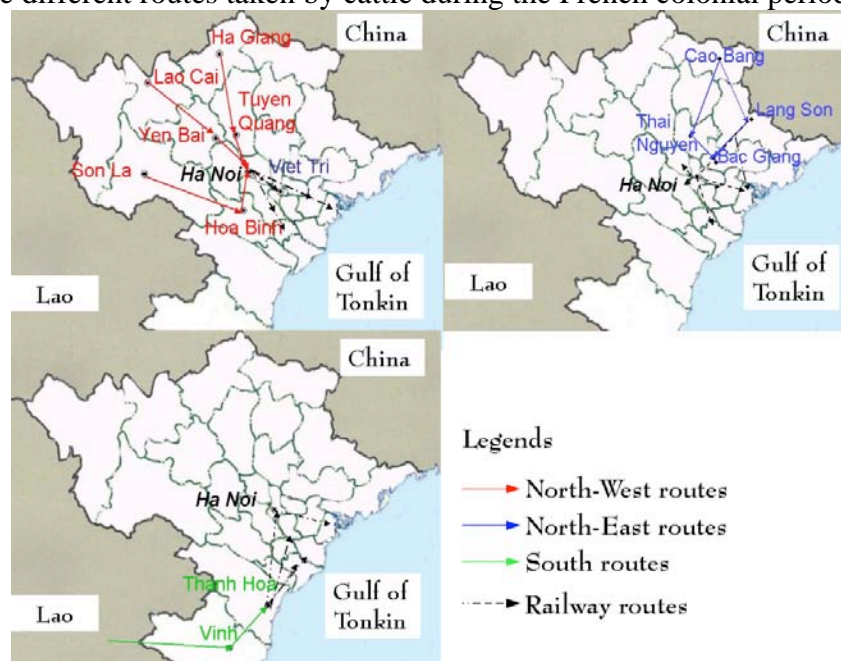
Table 4.5 List of main Asian local South-East Asian cattle breeds.

Country	Y-chromosome type	Colour	Horn	Hump	Male HW (cm)	Male Weight (kg)	Female HW (cm)	Female Weight (kg)
China								
Anxi	100 T			none				
Mongolian	100 T	reddish-brown	medium size	none	120	396	113	306
Horqing	100 T			none				
Yanbian	100 T	yellow		none	131	465.5	115	365.2
Tibetan	100 T	red -black		none	104	215	100	197
Zaosheng	93 T/ 7 I							
Kazakh	83 T/ 17 I	blackish-brown	medium size		124	500	112	330
Qinchuan	70 T/ 30 I	purplish red- red	shorthorn		135-145	480-710	120-130	310-455
Jinnan	62 T/ 38 I	light red-dark red	downward		135	610	117	340
Pinglu		yellowish-brown	shorthorn		110-120	325	105-115	270
Jinan		light fawn	shorthorn		127-137	375	110-120	230
Luxi	100 T	yellowish-brown	shorthorn	small hump	140-150	535-750	118-128	200-325
Bohai Black	58 T/ 42 I	black	medium size	small hump	125-135	260-335	110-120	260-335
Jiaxian Red	25 T/ 75 I	red	medium size		120-135	360-500	115-125	325-410
Nanyang	36 T/ 64 I	light yellow	medium size	small hump	135-150	650	120-135	325-495
Sanjiang	100 I	red	shorthorn	humped	115-125	315-435	100-112	225-300
Ebian	100 I	white spotted	medium size	humped	110-120	260-375	100-110	220-285
Dengchuan		dark brown	medium size	humped	130	270	105	250
Bashan		yellow-light red	medium size	humped	115-130	390-490	105-120	230-370
Sinan	24 T/ 76 I	yellow	medium size	humped	115	310	105	230
Zaobei		yellow, brown, red	medium size	small hump	125	445-460	115	260-350
Dabieshan		yellow-black	shorthorn	small hump	115	320	110	230-310
Wannan	100 I	dark brown	medium size	humped	115-125	290-375	105-115	215-280
Zhoushan		black	downward	humped	135	380-500	120	285-385
Wenling humped	100 I	yellowish-brown	downward	humped	125-130	360-485	105-125	245-330
Guanfeng		brown	shorthorn	humped	115	220-325	105	200-260
Minnan	100 I	yellowish-brown with black	longer horn	humped	110-125	285-370	105-115	230
Hainan	100 I	brown with blackish part	shorthorn	humped	115-125	260-300	110-115	230
Yunnan Humped	100 I	light tan to brown	medium size	humped	110-120	295-305	100-110	195-230
Panjiang (H'mong)		brown	brown	humped	115-120	185-350	108	230-260
Ji'an	100 I	yellowish-brown			107	223	104	233.7
Xhuanhan	8 T/ 92 I							
Diqing	18 T/ 82 I	black-yellow-brown			102	212.9	98	185.5
Leizhou	100 I							
Vietnam								
Bo U dau Ria				**humped**				
Bo Vang (Yellow)		**yellowish-brown**	**y**	**humped**	**104-110**	**275**	**102-104**	**175**
Chau Doc		**yellow**		**humped**	**128**	**230**	**114**	**210**
H'mong				**humped**		**385**		**260**
Cambodia								
Cambodian								
Moi								
Lowland Khmer	x Banteng	brown-red	medium size	humped		350		250
Highland Khmer	x Banteng	brown-red	shorthorn	humped		200		150

Country	Y-chromosome type	Colour	Horn	Hump	Male		Female	
					HW (cm)	Weight (kg)	HW (cm)	Weight (kg)
Lao								
Ngoua		black, yellow, red, brown		humped	115	220	110	170
Yellow		red, yellow-brown		humped	115		94	
Myanmar								
Katonta		white		humped	125	300	120	250
Shan Nwar		grey	longhorn	humped	110	244	104	226
Shwe Ni		fawn		humped	119	350	110	300
Thailand								
Thai		red, brown		humped	115	325	113	225
Thai fighting		red, brown, black,yellow		humped				
Khao Lumpoon		white		humped				
Malaysia								
Kedah-Kelantan		red-black	shorthorn	humped	102	250	96	200
The Philippines								
Philippine native		red-yellow		small hump				280-300

Y chromosome: % of T(taurine type)/I(indicine type) according to Cai et al. (2006); Source: DAD-IS; Felius (1995)

Picture 4.1 Pictures of H'mong cattle.

According to Choquart (1928), the cattle breed from Cao Bang-Lang Son provinces may be the smallest, weighing around 200-300 kg, whereas the breeds from Vinh and Thanh Hoa weighed between 400-500 kg and 360-400 kg respectively with an average meat yield of 45 % of the live weight.

Currently, 22 cattle breeds are present in Vietnam including 4 breeds identified as local breeds (Vo Van, 2004) (Figure 4.9.B):

- The Bo U Dau Riu breed: in the provinces of Bac Bo golf

- The Bo Vang (Yellow) breed which merges the types from Thanh Hoa, Thuy Hoa, Ba Ria and Cao Bang and is distributed across the country

- The Chau Doc breed: in the Mekong Delta

- The H'mong breed: in the northern provinces bordering China and more specifically in the Ha Giang province

Almost all Vietnamese breeds have a uniform coat colour varying from light yellow to red froment, except for the H'mong breed where some individuals have a black coat. Vietnamese cattle are shorthorn, and the dewlap is well developed; animals and mainly males have a hump and so are of zebu type. The shape of the hump has been used in Vietnam as a criterion for breed differentiation. According to the Atlas of Vietnamese breeds (Vo Van, 2004), the Yellow cattle breed seems to have a "hump like Indian type in opposition to the Chinese type", and it seems that the hump of this breed is in fact smaller. The Bo U dau Riu breed has a yellow hump that becomes black in the front, the hump is firm, as straight as on an axis. The H'mong breed has a less developed hump, not as straight, sloping down in the front.

4.4 Genetic studies on cattle
4.4.1 An overview

Over the years, domesticated cattle have been the subject of numerous surveys of genetic variation within and among populations. The first reports of molecular DNA-based surveys of genetic variation in cattle appeared in the 1990's (Loftus *et al.* 1992; Bradley *et al.* 1994; MacHugh *et al.* 1994, Ciampolini *et al.* 1995, Arranz 1996a, Moazami-Goudarzi *et al.* 1997). Studies have been undertaken on a broad scale to encompass populations from every continent

and also at a more localised level among closely related populations within particular regions. Arranz *et al.* (1996b) showed that the use of only five microsatellites was much more efficient than protein markers in order to estimate heterozygosity and genetic distances between the Brown Swiss breed and three other Spanish breeds. Russell *et al.* (2000) analysed with four microsatellites five Mexican criollo cattle populations, which were geographically isolated from each other. While the number of microsatellites was low and so the results obtained should be used with caution, the authors highlighted that their cattle population should actually be considered as a single population except for one which in fact was not the most isolated of all.

The Near East being the domestication centre of *Bos taurus*, neighbouring regions such as the Balkans, the Caucasus and Ukraine may represent major migratory corridors for the expansion of cattle out of the Near East. Also, Meng-Hua *et al.* (2007) studied twenty-one breeds from Finland to North-East Russia and three Turano-Mongolian breeds. They were able to show the wide spread of the international commercial dairy nuclear cattle into indigenous cattle because of current AI (artificial insemination) use. The Murano-Mongolian breeds, localised at the north border of China, showed the highest values of $N_e m$ (number of migrants per generation). Thus, migratory roads and gene flow occurred widely and at a high spatial level, not only between close countries but also across continents. Similarly, Kumar *et al.* (2003) showed, through microsatellite admixture analysis, that introgression of *Bos taurus* in the indigenous *Bos indicus* occurred at a higher rate and earlier than expected. Zhang *et al.* (2007) studied 27 indigenous Chinese cattle breeds. They confirmed introgression of *Bos taurus* from the north and *Bos indicus* in the south. But also, they highlighted that grouping breeds according to their geographical distribution was not always consistent because of human migratory movement or selection dictated by human needs or purposes. In China, the biggest cattle type regrouping most of all Chinese cattle breeds is known as Yellow cattle. However, according to the results of this study, this type of cattle may be derived from two independent centres of domestication: one in the South-East and the second one in the South-West. Figure 4.11 summarises gene diversity measures of some previous studies.

More than underlining genetic and historical relationships between breeds/populations, microsatellite analysis is useful to quantify the weight of each population towards global species genetic diversity. Such values are very interesting and useful in order to understand what constraints are suffered by the population and then to establish conservation measures

accordingly. As an example, Kantanen *et al.* (2000) studied twenty breeds from Northern Europe, and they pointed out a subdivision of these breeds into two genetically distinct groups. They also concluded that not only certain breeds were in danger of extinction but also groups of breeds. Therefore, many studies have tried to prioritise breeds using different approaches. For instance, recently 69 breeds were characterised with 30 microsatellites by the European Cattle Genetic Consortium (2006). All these studies have in common the assessment of genetic diversity as a prerequisite for the management and conservation of these animal genetic resources.

Figure 4.11 Concise map of expected Heterozygosity from published cattle breed studies.
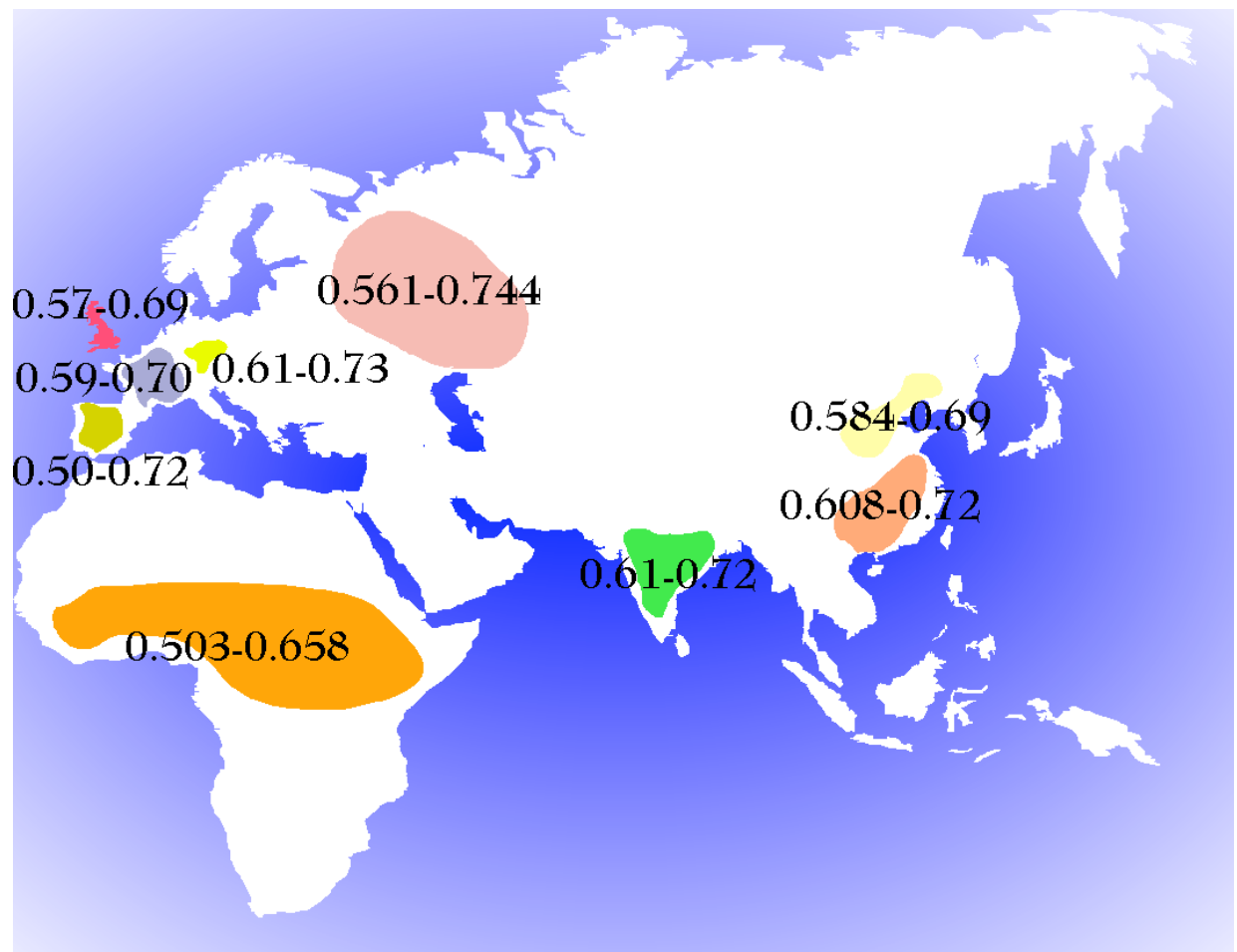
Source: MacHugh et al. (1997); Kumar et al. (2003); European Cattle Diversity Consortium (2006); Meng-Hua et al. (2007); Zhang et al. (2007)

4.4.2 Characterisation of the H'mong cattle in the Ha Giang Province

REVEALING SUBPOPULATIONS AND POSSIBLE ADMIXTURE IN THE VIETNAMESE H'MONG CATTLE BREED FOR CONSERVATION PURPOSES

Submitted to Journal of Heredity

Abstract

During the last decades, it seems there has been an acceleration of the loss of domestic animal biodiversity. For conservation purposes, the genetic diversity of the H'mong cattle, an indigenous local breed, has been studied. Genetic diversity was assessed through genetic data with twenty-five FAO microsatellites; and morphometric data with five body measurements of 408 animals sampled in eight districts from the Ha Giang province. The observed and expected heterozygosity ranged from 0.616 to 0.673 and from 0.681 to 0.729 respectively. Low genetic differentiation between districts was observed (F_{ST}= 0.0076). Multivariate analysis on morphometric and genetic data showed a separation of districts into two groups. These two groups separate from a south-west/north-east cline. Admixture analysis confirmed the two clusters and showed possible introgression from the Vietnamese Yellow cattle breed assessed through farmers interviews. Therefore, conservation policies could be established to conserve and limit crossbreeding of the H'mong cattle.

Introduction

During the last decades, it seems there has been an accaleration of the loss of domestic animal biodiversity, a phenomenon which may be illustrated, among other indicators, by the evolution of the number of breeds raised. Among the estimated 6 400 breeds of farm animal species worldwide, about 30 % are endangered and 1 % becomes extinct every year. Over the past 15 years, about 300 breeds of farm animals identified by the FAO became extinct and 1 350 face extinction in the near future (Scherf, 2000). The assessment of genetic diversity is a prerequisite for the management and conservation of these animal genetic resources. Rege (1999) found that 32 % of the sub-Saharan cattle breeds inventoried were endangered and 13 % already extinct. The major threats to livestock genetic diversity result from systemic, regional and global economic forces and from changing agricultural practices. Economic globalisation with the process of the extending market (Tisdell, 2003), agricultural policies (Anderson, 2003; Mendelsonh, 2003) and poverty have often driven farmers to abandon their traditional breeds and raise more productive industrial breeds. And yet, contrary to highly selected commercial breeds, indigenous breeds are raised on a semi-wild way. Most of them are managed through traditional farming systems and are mainly submitted to natural selection, become locally adapted to a wide range of environments and present valuable traits such as resistance to local infectious and parasitic diseases or adaptation to rough forage.

Vietnam, as a part of the Indochina Peninsula, has been considered as one of the world's biodiversity hot spot. Nevertheless, among the 220 farm animal breeds identified, less than 15 % had population data records. In recent years, the Vietnamese government has become aware of the importance of conserving their local breed diversity and has decided to invest in their conservation and improvement (Xuan *et al.,* 2006). One of the most emblematic local breeds is the H'mong cattle breed because it is associated to the H'mong ethnic group, and so has a strong cultural value. Moreover, this breed is considered to be endemic of the Northern part of the Ha Giang province (a Vietnamese province bordering China). Historic notes from French veterinarians from the beginning of the 20th century report that in the North of the Ha Giang province, there might be a strong and tall cattle breed but that it was not possible to confirm this fact because of the inaccessibility of the area (Choquart, 1928). The H'mong ethnic group used to live above 1 000 m altitude which made them semi-isolated and gave them limited interactions with other provinces until a few dozens of years ago. As agropastoralists, H'mong farmers raise one to three cattle per family for draught power used in maize crop production.

The cattle are raised in a semi-wild way without any selection or breeding management. However, the development of new road infrastructures has made the Ha Giang province more accessible and has allowed the development of a big cattle market reached by trucks. As a consequence, the flow of animals to/from other provinces has become possible over recent years. Therefore it is essential to characterise this breed either phenotypically and genetically to be able to keep its specificity before replacement or admixture occurs, and to design rational breeding strategies for its improvement and conservation.

Under such a context, microsatellites in combination with a fine-scale survey, multivariate procedure and a Bayesian-clustering approach were used to analyse the genetic structure and admixture of the H'mong cattle breed. In addition, morphological characterisation was done because it is important in breed identification and classification in ways to which farming communities can relate. Since there was no a priori substructure, we first checked structuring according to districts. Afterwards, genetic and morphometric analysis were done according to the clusters identified with the Bayesian approach. Management practices were also studied to ascertain the relevance of the biological results.

Materials and Methods

Sampling procedure

The remote Ha Giang province (22º08' - 23º19' N ; 104º33'- 105º33' E) is a mountainous area with elevations of 100 to 2400 m above sea level and a temperate climate. Twenty-four ethnic groups are geographically distributed, with the H'mong ethnic group being the main one followed by the Tay and the Dao ethnic groups. Sampling sites were distributed in the high cattle density area as shown in Figure 1. Blood and tissue samples were collected from 407 individuals from 133 villages within 25 communes belonging to 8 districts (Table 1).

Farmer interviews

Farmers were questioned about their knowledge of the pedigree of each animal and their origin. Eight categories of origin were considered: Farm (i.e. the animal is born on the farm where the dam and granddam were raised, bulls are rarely known); Village (another farm within the village); Commune; District; District Market (the animal was bought in the district market, with no information about its farm origin); Other district; Outside the Province; Project

(many poverty alleviation projects have given buffaloes or cattle to farmers, therefore a class was done for these "gifts" because their origin was not possible to determine).

Figure 1. Map of the Ha Giang province : a) density plot of the cattle population, b) commune sampled in light grey.
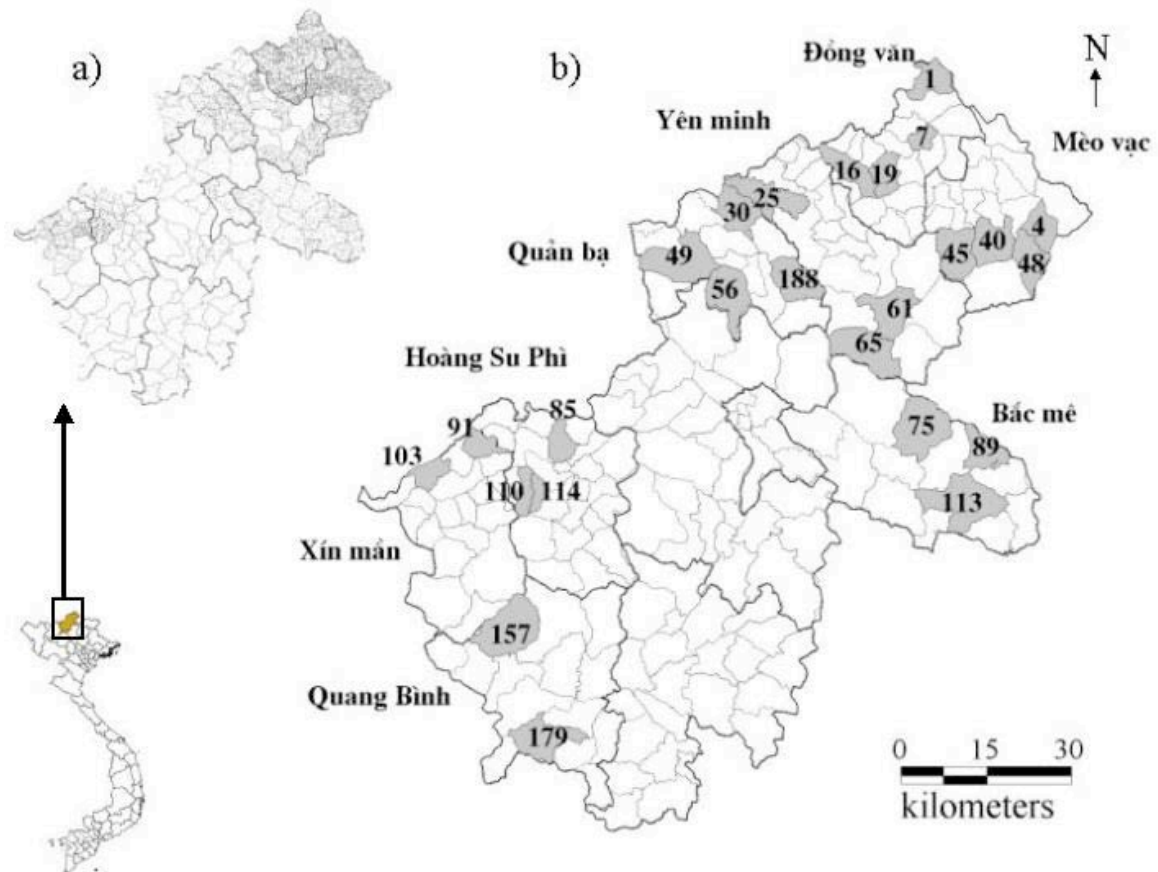
Microsatellites markers

Genomic DNA was extracted from blood and/or tissue samples using the QIAamp Kit from QIAGEN. Twenty-five microsatellites of the Secondary FAO Guidelines organised in 6 multiplexes were used (Table S1).

Morphological data

To avoid an age effect, only adult animals were sampled. For each sample, five measurements were taken with a cloth tape : the thorax depth; the height at wither (HW); the body length (BL); the heart girth (HG); and the ear length (EL). These measures were combined into 5 indices: the index of slenderness IGs = volume of substernal cavity/thorax depth (Lauvergne &

112

Souvenir Zafindrajaona 1992); the body length index I_{BL} = BL/HW; the heart girth index I_{HG} = HG/HW and the Ear index I_{EL} = EL/HW.

Table 1. Summary of sampling sites characteristics

Districts (code)	Commune (code [a])	Nv s	Ni	Main Ethnic group	Altitudes range (m)	Mean herd size[b]
Quang-Binh (QBn)	Tan-Nam (157)	4	5	Tay	150-350	2.8
	Xuan-Giang (179)	2	4	Tay	116-210	2.7
Hoang Su Phi (HSP)	Po-Lo (85)	3	9	H'mong	642-1057	1.9
	Po Ly Ngai (110)	4	13	Nung	825-914	2.0
	San Sa Ho (114)	4	12	Nung/Tay	646-731	1.6
Xin-Man (XM)	Nan-Xin (91)	3	6	La Chi	888-1395	3.8
	Chi-Ca (103)	8	45	H'mong	950-1410	2.5
Quang-Ba (QB)	Bat Dai Son (30)	6	19	H'mong	606-1105	2.4
	Tung-Vai (49)	6	11	H'mong	890-1010	2.6
	Quet-Tien (56)	5	12	H'mong	689-936	1.9
	Lung-Tam (188)	5	12	H'mong	429-975	2.7
Yen-Minh (YM)	Na-Khe (25)	6	14	Dao	457-806	1.7
	Lung-Ho (61)	7	23	H'mong	486-756	3.1
	Du-Gia (65)	4	15	H'mong	395-916	3.1
Dong-Van (DV)	Lun-Pu (1)	7	25	H'mong	1288-1427	3.5
	Tai Phin Tung (7)	6	24	H'mong	1095-1363	2.2
	Pho-Cao (16)	8	18	H'mong	1148-1530	2.4
	San-Tung (19)	6	21	H'mong	1284-1511	1.8
Meo-Vac (MV)	Lung-Cu (4)	7	28	H'mong	968-1210	2.6
	Tat-Nga (40)	5	14	H'mong	375-799	2.7
	Nam-Ban (45)	3	8	Giay	312-946	3.0
	Khau-Vai (48)	8	32	H'mong	461-1245	2.6
Bac-Me (BM)	Giap-Trung (75)	3	4	Dao	127-678	2.9
	Yen-Phong (89)	6	19	Tay	171-445	4.3
	Yen-Cuong (113)	7	9	H'mong/Dao	132-392	3.3

[a] Number code of communes have been randomly given by MAPINFO software for cartographic purposes; [b] mean farmer herd size; Nv: number of villages sampled; Ni: sample size;

Statistical analysis

Molecular diversity, within and between populations

The presence of null alleles was tested using FreeNA (Chapuis & Estoup, 2007). We considered only loci with r ≤ 0.2 to be potentially problematic for calculations. Allele frequencies, number of alleles, expected and observed heterozygosity were calculated using GENETIX 4.4 (Belkhir *et al.*, 2000). The GENEPOP software (Rousset & Raymond, 1995) was used to compute F_{IS}, F_{ST} (Weir & Cockerman, 1984), the pairwise genetic differentiation between populations and departure from Hardy-Weinberg equilibrium using exact tests. Test significance was corrected with sequential Bonferroni correction on loci. For further analysis,

we removed samples from the QBn district because of the low number of samples (9) compared to other districts. The effects of migration and gene flow on the genetic structure of the hypothesised district populations were estimated between each pair of district populations according to an island model under neutrality and negligible mutation. The $N_e m$ value indicates the average number of effective migrants exchanged per generation to produce the observed F_{ST} under the island model and were obtained according to the following formula (Slatkin, 1993):

$$N_e m = \frac{1}{4}\left(\frac{1}{F_{ST}} - 1\right)$$

Pairwise genetic distances (D_R; Reynolds, 1983) between district populations were estimated performing 1000 bootstraps with the POPULATION v.1.2.28 software package (Olivier Langella; available at http://bioinformatics.org/project/?group_id=84). Dendrograms were constructed from a distance matrix using the neighbour-joining algorithm.

In domestic species, the local management and use of the breed may lead to fragmentation and isolation reducing the effective population size. Caballero & Toro (2002) showed that minimising coancestry in a subdivided and undivided population is equivalent to maximising the effective population size. We estimated the molecular coancestry f_m with Molkin v2.0 (Gutiérrez *et al.,* 2006). Moreover, because cryoconservation was one of the Vietnamese government targets, we estimated the contribution of each strain to the eventual germplasm in order to maximise its genetic diversity.

We investigated the genetic structure and individual assignments using a Bayesian clustering procedure implemented in STRUCTURE (Pritchard *et al*., 2000), with the admixture method and correlated allele frequency version of the programme (Falush *et al.,* 2003). The implemented Bayesian approach has the advantage to infer multiple ancestral populations, and has shown efficiency in detecting subtle traces of historical admixture (Kumar *et al.,* 2003; Freeman *et al.,* 2004). Assuming that the data set could be represented by K separate genetic clusters (*K*=1-10), we then did 15 runs for each different values of K with 10^5 iterations following a burn-in period of 30 000, without any prior information. The values for the number of clusters (K) were assessed following Evanno *et al.* (2005), by comparing the estimated posterior probability of data for different K values and the standard deviation between runs for the same *K*. The programme estimates the posterior distribution (*q*) of each individual's

admixture coefficient. The contour maps of the admixture coefficient *q* were constructed using MAPINFO software.

Morphometric Analysis

Morphological data were analysed using single trait linear hierarchical mixed models with the SAS software. For a given trait, the model was:

$$y_{ijkl} = \mu + b \times alt_{ijkl} + s_i + d_j + c_{ijk} + e_{ijkl}$$

where *y* is the observation, μ the overall mean, *alt* the altitude where the animal is raised and *b* the coefficient of regression of the performance over altitude, s_i the fixed effect of sex *i* (*i* = 1 or 2), d_j the fixed effect of district *j*, c_{ijk} the random effect of the commune nested into district and e_{ijkl} a random error. For analysis following genetic structure using molecular data and the Bayesian approach clustering procedure, district populations were replaced by cluster populations and the random effect of the district nested into the cluster was added to the model. For some measures, transformations were needed in order to obtain normality and homoskedasticity of model residuals (Table S3). The P-values of pairwise mean adjusted comparisons were corrected with Bonferroni correction.

A linear discriminant analysis was used on the whole set of markers and measurements. Since this set contains a mix of quantitative and qualitative variables, we used an approach similar to that of Hill & Smith (1976). It is a combination of an internal correspondence analysis on markers (Cazes *et al,.* 1998; Laloë *et al.,* 2002) and of principal component analysis on quantitative measures. This approach is based on functions available in the ade4 package (Chessel *et al.,* 2004; Dray & Dufour, 2007) of the R software (R development core team, 2007).

Results

Marker analysis

No loci had r values below 0.2, therefore it was assumed that null alleles were absent from our dataset. A total of 195 alleles were detected. The number of alleles ranged from five (INRA063) to thirteen (ILSTS006 and HEL9), with a mean number of 7.8 alleles per locus. The observed heterozygosity per locus averaged 0.643 and ranged from 0.293 for locus HEL1 to 0.821 for locus HAUT27 (Table S1). To check HWE at 25 loci across eight districts, a total

Figure 2. Locus F_{ST} values for each district populations.
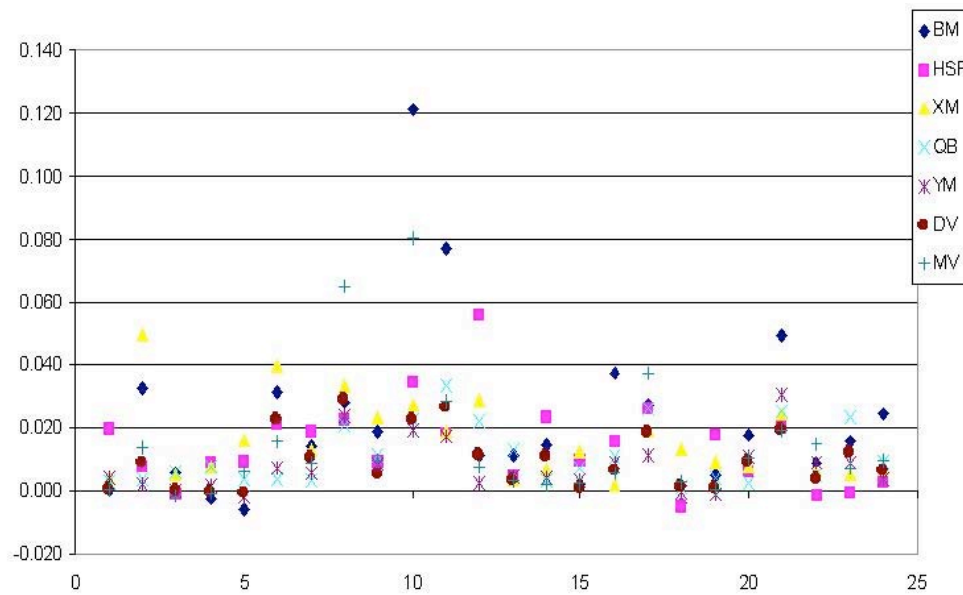
Table 2. Summary of District populations and their polymorphism measures

District	H_{Exp}	H_{Obs}	A	A_e	A_{HWE}	F_{IS}	F_{ST}	f_m
QBn	0.695	0.633	*4.7*	*2.9*	*0*	0.085	0.016	/
HSP	0.716	0.667	6.8	3.4	4	*0.070*	0.014	0.247
XM	**0.729**	**0.673**	7.4	**3.6**	5	0.075	0.016	0.248
QB	0.724	0.63	7.4	3.5	**6**	**0.128**	0.011	*0.214*
YM	0.721	0.657	7.2	3.5	2	0.087	*0.007*	0.223
DV	0.713	0.638	**7.7**	3.0	4	0.103	0.009	0.242
MV	0.703	0.634	7.5	3.3	4	0.095	0.013	0.244
BM	*0.681*	*0.616*	6.7	3.0	3	0.093	**0.020**	**0.285**

in bold: maximum values, *in italic*: minimum values

Table 3. Pairwise district populations N_em estimates above the diagonal *in italic* and average N_em overall districts **in bold** on the diagonal; D_R genetic distance within districts below the diagonal

District	HSP	XM	QB	YM	DV	MV	BM
HSP	**24**	*51*	*23*	*30*	*20*	*11*	*8*
XM	0.006	**21**	*28*	*20*	*14*	*9*	*7*
QB	0.013	0.011	**33**	*86*	*32*	*19*	*11*
YM	0.010	0.014	0.005	**54**	*119*	*43*	*23*
DV	0.014	0.019	0.01	0.003	**59**	*125*	*44*
MV	0.025	0.027	0.015	0.007	0.003	**50**	*96*
BM	0.033	0.038	0.026	0.013	0.007	0.004	**31**

of 200 tests were conducted. Twenty-eight district-locus combinations were not at equilibrium after Bonferroni correction. Out of the 25 loci studied, 9 showed a lack of equilibrium in one or more district populations. The HEL5 locus was in disequilibrium except in the QBn district, and ETH185 locus in all the districts except in QBn and BM district.

According to the multilocus $F_{ST,}$ only 0.76 % of the total genetic variability was explained by subdivision of populations among districts. When plotting F_{ST} loci values per district, the highest differentiation was observed for loci HEL1 and HEL13 which discriminated the BM and secondly the DV district (Figure 2).

Molecular diversity within and between district populations

The QBn district with only 9 individuals sampled showed 4.7 average number of alleles per loci, whereas for other districts this average ranged from 6.7 to 7.7 (Table 2). There were, on average, 3.3 effective alleles per district. The observed and expected heterozygosity averaged over all loci ranged from 0.616 to 0.673 and from 0.681 to 0.729 respectively. The average F_{IS} in the Ha Giang province reached 0.092 varying from 0.07 (HSP) to 0.128 (QB). Except for the QBn district, all district populations had more than two loci deviating from HWE, and the maximum value of 6 loci deviated from HWE was reached for the QB district. Pairwise genetic differentiation between districts ranged from 0.002 (DV-MV) to 0.0356 (XM-BM). All the pairwise were statistically significant, with the exception of those involving QBn and other districts. The mean within district genetic diversity within individual (GD_{WI}) was found to be equal to 0.347 and the between individual diversity (GD_{BI}) was 0.411. The average molecular coancestry f_m ranged from 0.214 in the QB district to 0.285 for the BM district; then within a district, individuals are nearly full-sib (Table 2). Following Caballero & Toro (2002), if a synthetic population of cattle from Ha Giang that maximises genetic diversity needs to be created, it appeared that the synthetic population that would maximise the genetic diversity would be composed of 70 % of cattle from the QB district , followed by 27 % from the YM district and the 3 % remaining coming from the HSP district.

The number of individuals $N_e m$ that migrate between each pair of sampling localities per generation are summarised in Table 3. The MV and DV district had the highest value of $N_e m$ between them (125), whereas the lowest values were obtained between BM and XM (7) and BM and HSP (8). On average, the DV district was the one with the highest $N_e m$ value (59).

Indeed, more than one hundred individuals per generation are exchanged with the two bordering districts YM and MV. On the contrary, XM showed the lowest mean value (21) and reached a maximum of 51 individuals with the bordering district of HSP.

The neighbour-joining tree based on D_R genetic distance without the QBn district is presented in Figure 3. The dendrogram reveals two groups with bootstrap values higher than 97 %: firstly, the two geographically close districts of HSP and XM and subsequently the QB district, and secondly BM and MV on which is linked the bordering DV district. The YM district, which in fact had a central geographic position between the first groups, did not cluster with one another in a preferential way. Further results using the Bayesian approach showed that two northern communes and the two southern communes did not belong to the same genetic subpopulation. Therefore if we draw a new dendrogram where YM district samples are divided into two groups, we see that the southern commune (YM2) clusters with high bootstrap values to the northern cluster (Figure 4). We can then draw a south-west/north-east line passing through the YM district between the two groups.

Morphometry and Multivariate analysis among district populations

Mean body traits between districts are summarised in Table S2 and variables effects in Table S3. There were significant differences between males and females for all measures except for the length of ears (*EL*) and the index body length I_{BL} (BL/HW). Because crop varieties, yield and qualities depend on the altitude, this variable will translate in some way the quality and the quantity of the forage. Our results show that altitude had an effect on the heart girth (*HG*) and the body length index I_{BL}. Pairwise comparison between districts showed that only the height at withers (*HW*) was significantly different after Bonferroni correction. In fact, cattle from the MV district were significantly smaller than the cattle from the three western districts: HSP, XM and QB. The smallest heart girths were observed in the MV district for both sexes (134.2 cm; 142.4 cm) whereas the biggest females were observed in the XM district (144) and males in the YM district (155 cm). Males and females from the BM district were the shortest, the longest females were found in the YM district and the longest males in the QB district.

The two first axes of the discriminant analysis on the whole set of variates (markers + measures), explained 22.7 % and 19.6 % of the observed inertia (Figure 4). On the 25 markers plus the 5 morphometric traits, the two markers ETH3 and BM2113 are the ones that

Figure 3. Dendrogram of district populations with D_R genetic distances and bootstrap values
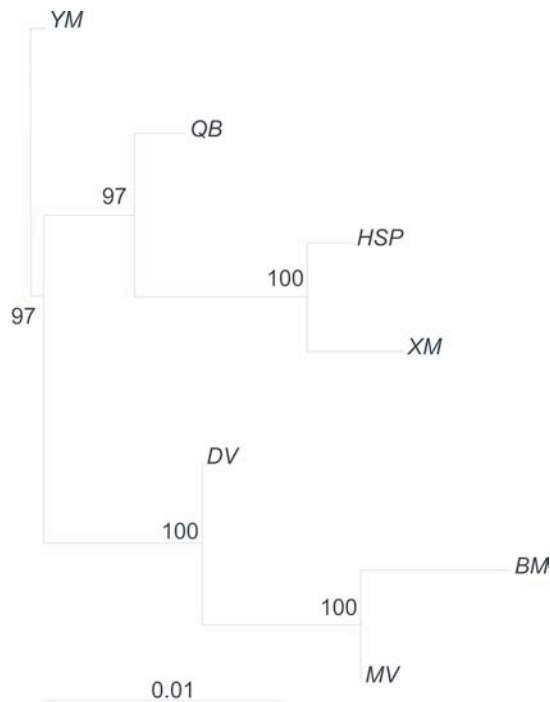
Figure 4. Dendrogram of district populations with D_R genetic distances and bootstrap values where the Yen-Minh district was divided into two sub-groups
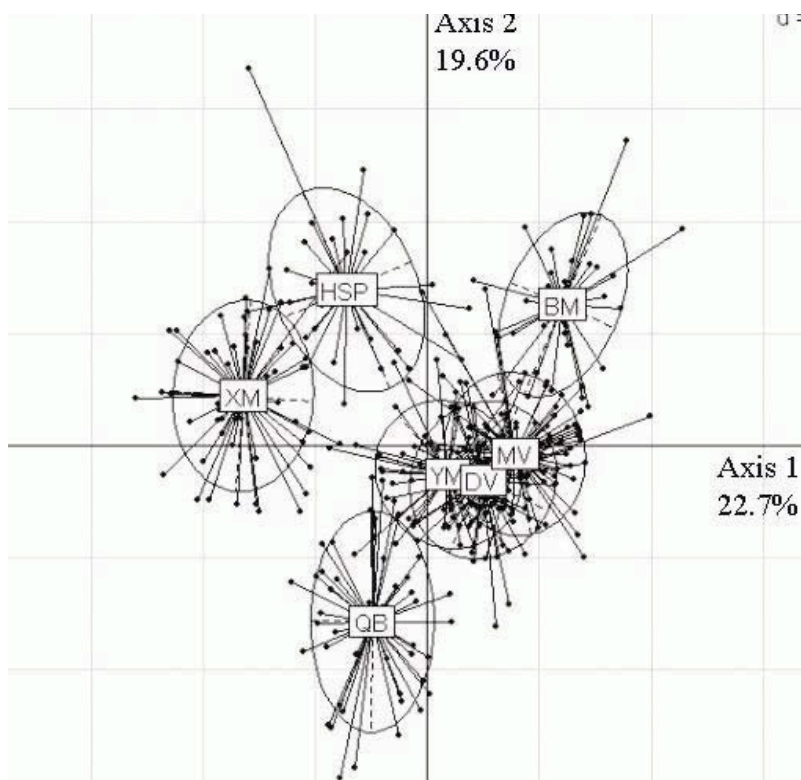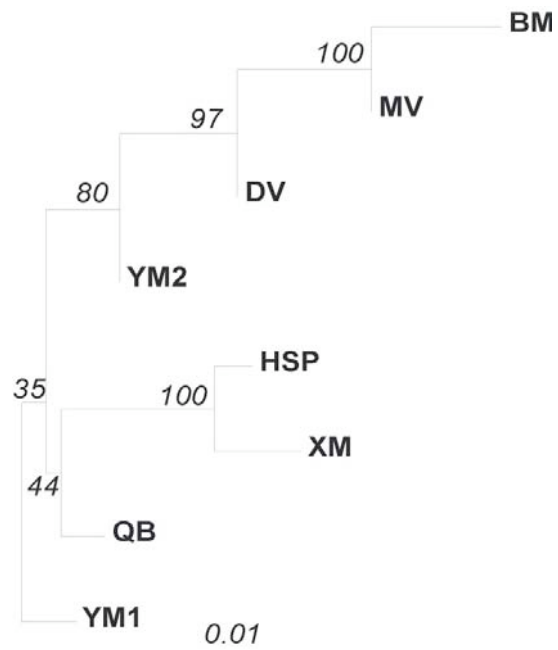
Figure 5. Representation of district populations using discriminant anlysis with the Hill&Smith procedure

contributed the most to the construction of the first axis for 7.6 % and 6.7 % respectively. The latter were followed by the height at withers HW, which contributed for 4.8 %. For the second axis, SPS115 (6 %) and ILSTS006 (5.9 %) are the most important followed by the index of body length I_{BL} with 4.34%. The first axis separated district populations into two groups (Figure 4). The bordering districts HSP and XM are positioned on the left and further on the QB district. On the right, the three geographically close districts of YM, DV and MV are overlapped, whereas the BM district was separated to the top. Therefore, the first axis coincides with the east/west cline.

Genetic admixture and exchanges of animals

The Bayesian approach implemented in STRUCTURE indicated that two sub-populations (K=2) are the most likely figures for the cattle population of the Ha Giang province: the value of the criterion used (D.ΔK) was about 49 for K=2, about 4 for K=3 and less for higher values of K. Communes were attributed to a preferential cluster with probabilities ranging from 0.51 to 0.88. A grid map of individual probabilities of admixture to cluster 1 (q_1) is represented in Figure 6. It is therefore possible to see that the admixture rate with cluster 1 is decreasing from the West to the East. The HSP and XM district had a mean admixture with cluster 1 of 0.77 and 0.83 (Table 4). A zoom at the commune level, shows that in those districts, only one commune over five had less than 50 % of its cattle with less than 80 % of their genome admixed with cluster 1. For the QBn district which is composed of two communes, commune 157 strongly admixed with cluster1 and commune 179 to cluster 2 with probabilities equal to 0.86 and 0.73 respectively. Of the four communes from the QB district only commune 56 had half a proportion of its genome admixing with cluster 1 and cluster 2, whereas all others had a higher proportion with cluster 1. Therefore, cluster 1 could be identified as the south-west province. Starting from DV, then MV and BM, we observed an increasing gradient of admixture with cluster 2, the maximum being observed in commune 113 with a mean admixture value of 0.88. Cluster 2 could be considered as the north-east province.

Then, we compared morphometric traits assuming the two genetic clusters obtained with the Bayesian procedure. A significant effect of altitude was observed for the heart girth and its index. It appeared that the two clusters differentiated significantly for the height at withers. Cattle from the south-west Cluster (HSP, XM and QB) are more than 5 cm taller than cattle from the north-east. Also, we could assume the two clusters as two distinct units that need to be conserved. Commune 103 from the XM district and commune 49 from the QB district

contributed the most in the potential germplasm for cluster 1 with 40 % and 36 % respectively. For cluster 2, three communes 65, 4 from the MV district and 16 from the DV district contributed each one for more than 20 %.

A total of 814 pedigree-origin interviews has been done and percentages of answers per category are summarised in Table S4. "Family", which includes the animal being raised on the farm for three generations, constituted in most of the communes (22/25) a relatively important origin. This means that farmers keep and reproduce animals on their own farm, and without

Figure 6. Grid map of individual probabilities (*q values*) to belong to cluster 1 ranging from *q* <0.03 in white to *q*>0.96 in dark
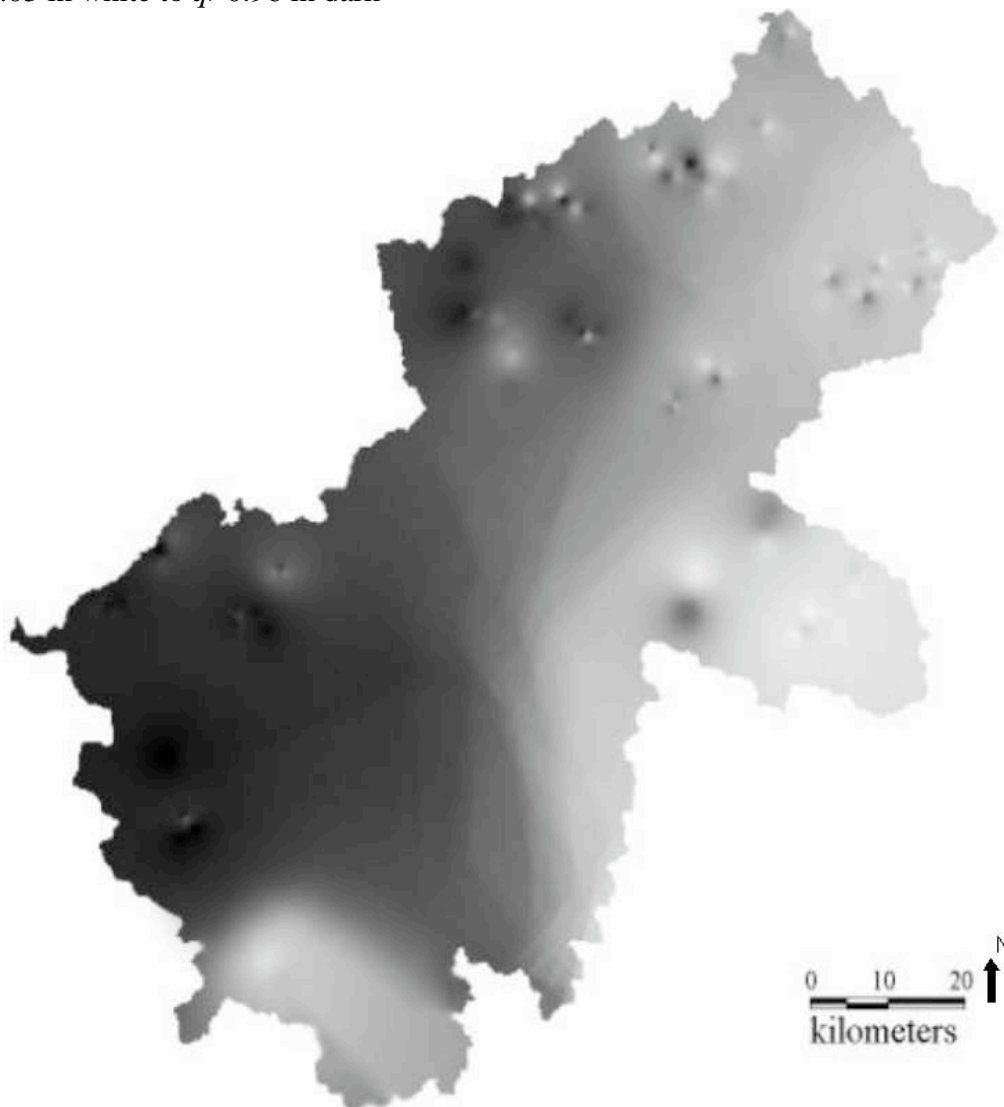
reproduction management, a bull could easily reproduce with related animals. In general, exchanges occurred mostly within the commune, in the commune market or within a village,

but also with bordering communes. Nevertheless, in MV and DV districts where two important cattle markets, are held, percentages of animals bought at district markets ranged from 3.1 % to 64 %. In commune 179 from the QBn district, nearly 67 % of cattle were given by a development project (Table S4). This also happened but at a lower degree in commune 113 and 89 from the BM district. The BM district is also the one where there is the most important rate of cattle coming from outside the province (6.5 % - 23. 2 %). In DV and XM, animals included in the category "Other province" came in fact from China.

Discussion

Genetic and Morphometric Variation

We used 25 FAO microsatellites to study population genetics and the structure of a local Vietnamese zebu breed. The expected heterozygosities per locus were found to be on the range reported in the Roslin database (http://www.databases.roslin.ac.uk/cadbase/). Three microsatellites, HEL5, ETH185 and HEL13 were in heterozygotes deficit nearly for all district populations. Such observations question the neutrality of these markers. However, Ibeagha-Awenu & Erhardt (2004) observed that on the 9 zebu breeds from their study, 7 were not in HWE equilibrium for HEL5, and 3 for ETH185. Because we did not find null alleles, we supposed that such disequilibrium could be a characteristic of zebu breeds. For Ha Giang cattle, the mean f_m value (0.243) was lower than the values observed in previous studies as for the conserved Xalda sheep breed (0.324; Alvarez *et al.,* 2007) or the Iberian pig breeds (0.394; Fabuel *et al.,* 2004). Individual herds in our study area were quite small ranging from 1.7 to 4.3 cattle per household. Pastures are used communally by 1 to 3 farmers and breeding animals stay in the herd for prolonged periods, implying that the relationship of the animals within a herd and even within a village is narrow. As a consequence of such management, one could expect a higher molecular coancestry or kinship value than half-sib. However, if new unrelated animals are regularly introduced on the farm or the village via the cattle market for example, it may strongly reduce kinship within a village, commune or district.

Table 4. Estimated proportion of admixture (q_i) of the 25 commune samples in each of the two inferred clusters

District	Commune	N	Mean $q(1)$	Mean $q(2)$	% $q(1)$ >0.5	% $q(2)$ >0.5	% $q(1)$ >0.8	% $q(2)$ >0.8
BM		34	0.18	**0.82**	5.9	**94.1**	0.0	**73.5**
	75	4	0.41	**0.59**	25.0	**75.0**	0.0	**25.0**
	89	19	0.14	**0.86**	0.0	**100.0**	0.0	**78.9**
	113	9	0.12	**0.88**	0.0	**100.0**	0.0	**88.9**
MV		82	0.28	**0.72**	15.9	84.1	2.4	42.7
	4	28	0.28	**0.72**	17.9	**82.1**	0.0	**46.4**
	40	14	0.31	**0.69**	21.4	**78.6**	7.1	**35.7**
	45	8	0.26	**0.74**	0.0	**100.0**	0.0	**25.0**
	48	32	0.28	**0.72**	15.6	**84.4**	3.1	**46.9**
DV		88	0.40	**0.60**	36.4	**63**.6	4.5	**30.7**
	1	25	0.42	**0.58**	40.0	**60.0**	4.0	**28.0**
	7	24	0.35	**0.65**	29.2	**70.8**	0.0	**41.7**
	16	18	0.48	**0.52**	55.6	44.4	16.7	**22.2**
	19	21	0.34	**0.66**	23.8	**76.2**	0.0	**28.6**
YM		52	0.46	**0.54**	44.2	**55.8**	9.6	**21.2**
	25	14	**0.58**	0.42	**64.3**	35.7	21.4	14.3
	61	23	0.38	**0.62**	34.8	**65.2**	0.0	**34.8**
	65	15	0.47	**0.53**	40.0	**60.0**	13.3	6.7
QB		54	**0.68**	0.32	**79.6**	20.4	**35.2**	3.7
	30	19	**0.74**	0.26	**89.5**	10.5	**36.8**	0.0
	49	11	**0.80**	0.20	**100.0**	0.0	**63.6**	0.0
	56	12	**0.51**	0.49	50.0	50.0	8.3	8.3
	188	12	**0.66**	0.34	**75.0**	25.0	**33.3**	8.3
QBn		9	**0.60**	0.40	**66.7**	33.3	44.4	33.3
	157	5	**0.86**	0.14	**100.0**	0.0	**80.0**	0.0
	179	4	0.27	**0.73**	25.0	**75.0**	0.0	**75.0**
HSP		34	**0.77**	0.23	**91.2**	8.8	**61.8**	2.9
	85	9	**0.63**	0.37	**77.8**	22.2	**33.3**	0.0
	110	13	**0.78**	0.22	**92.3**	7.7	**69.2**	7.7
	114	12	**0.86**	0.14	**100.0**	0.0	**75.0**	0.0
XM		54	**0.83**	0.17	**94.4**	5.6	**68.5**	0.0
	91	6	**0.72**	0.28	**83.3**	16.7	**50.0**	0.0
	103	45	**0.84**	0.16	**95.6**	4.4	**68.9**	0.0

in bold: highest mean q_i values; N: number of individuals

Conservation program on farm animals mainly consist in the improvement of animal performances or valorisation through the developpement of high value products and market chains. Therefore, there is a need to identify specific phenotypic characteristics, which is difficult when animals are raised in a semi-wild way without information records and in a wide range of environments. Skeletal measurement such as ulna length, body height, length and chest depth are less affected by nutrition and thus indicate inherent size better than measures related to girth measurements (Kamalzadeh *et al.,* 1998; Hall, 1991). Then, our body measurements were useful to compare the H'mong cattle breed to the six other local breeds currently well identified in Vietnam. The H'mong cattle seems to be taller but shorter than the Yellow cattle breed, which has an average value for height of about 104-110.5 cm in males and 102-104 cm in females and has a body length higher than 113 cm (Le Vie, 1999). The H'mong cattle is smaller than the Chau Doc breed (128 cm for males and 104 cm for females) but of the same size as the Ba Ria and Lai Sin breeds which have been crossbred with the Red Sindhi breed from India.

The height at withers (*HW*) was the only significantly different body trait between districts and it differentiates the Meo-Vac district from the others. Additionally, we enlighten an effect of altitude on the heat girth (*HG*) and the body length index (I_{BL}). *HW* and I_{BL} contributed in a significant way to the district population's differentiation using multivariate analysis. A stratigraphic crop repartition was observed in the Ha Giang province, maize is planted at a high altitude with low clay soil whereas paddy rice field with higher yield is cultivated in the plains. Since this might influence the quality of forage, it also expresses the use of animals: where maize is cropped, cattle are preferred for draught power; on the contrary where there is a paddy, swamp buffalo is used and cattle are raised for additional income. It seems that farmers from the Ha Giang province prefer a more compact (tall and short) cattle for draught power in rocky mountains than farmers from less mountainous provinces like Cao Bang where the Yellow cattle is smaller and longer. This kind of consideration about the required morphology of draft cattle was reported in another tropical country namely Kenya (Mwacharo & Drucker, 2005).

By using both morphometric data and genetic data in a multivariate procedure we observed a south-west/north-east cline. Within the south-western cluster, the Quan-Ba district seems to have specificities which make it more separate from Hoang Su Phi and Xin-Man districts. This relation was confirmed by the D_R dendrogram. In the north-eastern group, three districts: Yen-

Minh, Dong-Van and Meo-Vac largely overlap. However, they did not cluster exactly in the same way using the D_R genetic distance. One possibility is that body traits are more similar within these districts than genetic inducing an overlapping when mixing both data. However the three methods genetic matrix, multivariate analysis and Bayesian approach lead to the same structure in two clusters: the south-west cluster and the north-east cluster.

Genetic admixture and commercial exchanges

Using multivariate analysis, Bac-Me district separates well from other districts. This district was also the one for which the F_{ST} value was the highest and particularly for the locus HEL1. In fact, the Bac-Me population is nearly fixed for allele 181 with a frequency of 0.9. Analysing admixture levels, it appeared that 73.5 % of this population had a probability of more than 0.8 to belong to cluster 2. In the Bac-Me district, the three sampled communes were mainly inhabited by the Tay ethnic group. The Tay ethnic group used to crop paddy rice and raised cattle for additional income. Therefore, they had the highest mean herd size in the province. According to pedigree interviews, cattle for the Yen-Phong, Giap-Trung and Yen Cuong communes respectively, were bought in another province and more specifically in the Cao Bang province where there is an important cattle market. The cattle population from the Cao Bang province is considered to belong to the Yellow Cattle breed, which is small and used for meat production. This was in accordance with the observed body measures showing that Bac-Me cattle is one of the smallest in the Ha Giang province. Therefore, the Bac-Me district could be an important area of introgression of the Yellow cattle breed into the Ha Giang province.

In the Meo-Vac district, a weekly cattle market occurs in the center and a second one in the Lung-Phin commune in the Dong-Van district. Meo-Vac farmers used to buy their herd in the district market. The Meo-Vac cattle market is considered to be the most important market in the province and it can be reached by trucks coming from other provinces. In fact, the main road comes from the Cao Bang province. Therefore, despite a lack of accurate information on the importance of the cattle from Cao Bang in the Meo-Vac market, our results support the hypothesis of migration within the H'mong cattle of Yellow cattle, this breed being the one raised the most in the Cao Bang province. As Yellow cattle are the Vietnamese breed with the smallest body size, this is also consistant with the smallest body size observed on Meo-Vac cattle. Thus, introgression of Cao Bang Yellow cattle could be more important than in Bac-Me,

however market practices have acted as a melting pot leading to a mean admixture coefficient lower (0.72) than in the Bac-Me district (0.82).

Similarly, the communes of Thai Phin Tung and Sang-Tung from the Dong-Van district and bordering with the Lung Phin market had the highest rate of market buying and admixture coefficients. The Lung Phin market depends on the lunar calendar, therefore it was not held on the same day or at a same regular interval as the Meo-Vac market. In the Sang-Tung commune, farmers used to go to both markets depending on the day and the price they could obtain, thus more than 8 % of cattle, according to pedigree interviews, were bought outside the Dong-Van district. Such practice was in accordance with the large number of migrant estimate N_em (125 per generation) between these two districts. However, migration also occurred between the Dong-Van and Yen-Minh districts with 119 migrants per generation. This was the case for example in the Pho-Cao commune where 9.4 % of cattle were bought in the Yen-Minh district. In summary, the most important exchanges occurred between Meo-Vac and Dong-Van and between Dong-Van and Yen-Minh, therefore this was well in accordance with the overlapping of these three districts observed with multivariate analysis.

In the south-western districts such as Quan-Ba, Quang-Binh, Hoan Su Phi and Xin-Man, no large cattle market is established. Farmers preferentially buy animals in the village or in the commune market. Thus admixture is lower and commune samples belong to cluster 1 with a mean probability higher than 0.6. However, there are two exceptions: the Xuan-Giang and Quyet-Tien communes. In the Xuan-Giang commune, 75 % of the cattle belong to cluster 2 with a probability higher than 0.8 and none to cluster 1. Interviews on the origins showed that 66.7 % of cattle were given by a Development Project. It was not possible to know exactly where the animals were bought by the project. However, according to farmers, they may have come from the north-eastern districts. Such assumption was in accordance with our observations of admixture rate.

Conservation policies

For monitoring future conservation programmes, it is useful to estimate the contribution of each population in order to create a pool of sire or germplam for cryoconservation. Thus, we estimated such a contribution according to Caballero & Toro (2002) and found that the population from the Quan-Ba district should contribute for 70 % and Yen-Minh cattle population for 27 %. Quan-Ba and Yen-Minh are the ones with nearly 50 % of genomes from

cluster 1 and cluster 2, therefore by conserving such a cattle population, one would be able to conserve the genome from both clusters. Northern provinces of Vietnam as Ha Giang and Cao Bang were initially inhabited by the Tay and Giay ethnic group. They migrated thousands of years before from the Guanxi province of China. The H'mong ethnic group had migrated from South Sichuan, Yunnan and mainly Guizhou Chinese provinces. Wiens (1954) wrote that the H'mong ethnic group initiated the migratory movements from 1698 to 1855. Bonifacy (1904) reported more precisely two migratory waves of H'mong ethnic groups arriving with their farm animals into the Ha Giang province: the first one occurred around 1800 and the second one in 1850 through the Dong-Van district and towards the South-West. Cattle from Quan Ba, Hoang Su Phi and Xin-Man could represent the original H'mong cattle breed, this was also assessed by the low N_em migrant estimates between these districts and Bac-Me. Therefore, considering that the conservation programme target will be to conserve H'mong cattle, and as enlightened by previous admixture results and history, the animals should come from the Xin-Man district and mainly from the commune of Chi-Ca and from the Tung-Vai commune from the Quan-Ba district. Moreover, it would be interesting to extend this kind of study to the Cao Bang province in order to ensure admixture in Ha Giang and to indicate whether it is a one or two-way H'mong-Yellow cattle admixture process between these two bordering provinces. Therefore, new conservation programmes could be established specifically for Yellow cattle.

As pointed out by Taberlet *et al.* (2007), domestic species could be considered as endangered from a genetic point of view. It is then important to take measures that promote management of the animal genetic resources. Many studies have evaluated conservation policies and prioritisation of breeds comparing within each other as for Eurasian cattle breeds (Li *et al.,* 2007), Riverine buffalo breeds (Kumar *et al.,* 2006) or Iberian Pigs (Fabuel *et al.,* 2004). One of the interests of our study was the very fine-scale spatial analysis across the area where the breed is raised. With fine-scale sampling we were able to establish a fine-admixture grid of the Ha Giang province which will allow to carefully select sires for the conservation project. Also, we believed that such approach will be useful on genetic studies of wild species. Indeed, there is a whole complex of interbreeding wild, domestic and feral cattle in North-East Asia (Grigson, 1988): two wild species of the genus are present, the gaur *Bos gaurus* and the banteng *Bos javanicus*, in addition there are also domestic cattle, zebu *Bos indicus*. These taxa interbreed and are sporadically distributed in south-east Asia, and many are conserved in protected areas. Moreover, as for banteng, in the Xe Pian (Steinmetz ,2004) protected area in Lao, or gaur in the Cat Tien National Park (Pedrono M, personal communication), the

population seems to be fragmented within a protected area. One of the most problematic threats to wild populations is the fragmentation reducing the effective population, which below 50 compromises the short-term survival of a population (Franklin, 1980). Therefore efficiency of corridors have been questioned in conservation programs. In domestic species, the local management and use of the breed may lead to fragmentation and isolation whereas commercial market exchanges may act as a corridor similarly than for wild animals. Previous studies have shown that bovine markers could be applied on wild bovids such as gaur (Nguyen *et al.*, 2007). Then if bovine markers and fine-scale surveys which have shown to be efficient to retrace subtle origins and exchanges (i.e. corridors) are used on wild species such as the gaur, it could be possible to estimate exchanges or fragmentation between populations and quantify the efficiency of corridors between protected areas. Such estimates are of primary interest when conservation policies need to be established.

Acknowledgments

References

Alvarez I, Royo LJ, Gutiérrez JP, Fernandez I, Arranz JJ, Goyache F, 2007. Relashionship between genealogical and microsatellite information characterizing losses of genetic variability: Empirical evidence from the rare Xalda sheep breed. *Livest. Sci.,* doi:10.1016/j.livsci.2007.06.009

Anderson S, 2003. Animal genetic resources and sustainable livelihoods. *Ecolog. Econ.*, **43:** 331-339

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F, 2000. GENETIX, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions: CNRSUPR. 9060. Université Montpellier II, Montpellier, France.

Bonifacy ALM, 1904. Les groupes ethniques de la Riviere Claire, *Revue Indo-Chinoise*, pp 823-838

Caballero A, Toro MA, 2002. Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Gen.,* **3:** 289-299

Cazes P, Chessel D, Doledec S, 1988. L'analyse des correspondences internes d'un tableau partitionné. Son usage en hydrobiology. *Rev. Statistique Appliquée*, **36 :** 39-54

Chapuis MP, Estoup A, 2007. Microsatellites null alleles and estimation of population differentiation. *Mol. Biol. Evol.*, **24:** 621-631

Chessel D, Dufour AB, Thioulouse J, 2004. The ade4 package-I :One –table methods, R-News 4: 5-10.

Choquart L, 1928. Les Marchés de bestiaux et le commerce du bétail au Tonkin : Thèse pour le doctorat vétérinaire. eds. "M. Lac", Paris

Dray S, Dufour AB, 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22(4) http://www.jstatsoft.org/v22/i04/

Evanno G, Regnaut S, Goudet J, 2005). Detecting the number of clusters of individuals using the software STRUCTURE, a simulation study. *Mol. Ecol.,***34:** 2611-2620

Fabuel E, Barragan C, Silio L, Rodroguez MC, Toro MA, 2004). Analysis of genetic diversity and conservation priorities in Iberian pigs based on microsatellites markers. *Heredity*, **93:** 104-113

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

Freeman AR, MegHen CM, MacHugh DE, Loftus T, Achukwi D, Bado A *et al.*, 2004. Admixture and diversity in West African cattle populations. *Mol. Ecol.,* **13:** 3477-3487

Franklin IR, 1980. Evolutionary changes in small populations. In: Conservation Biology:An Evolutionary-Ecological Perspective, eds Soule ME, Wilcox BA, Sinauer & Associates, Sunderland, Massachusetts, pp 135-149.

Grigson C, 1988. Complex cattle. Letter to *New Scientist*, 4 August 1988 : 93-94

Gutiérrez JP, Royo LJ, Alvarez I, Goyache F, 2005. Molkin V2.0 , A computer program for genetic analyses of population using molecular coancestry information. *J. Hered.*, **96:** 718-721

Hall SJG, 1991. Body dimensions of the Nierian Cattle, sheep and goats. *Anim. Prod.*, **53:** 61-69

Hill MO, Smith JE, 1976. Principal component analysis of taxonomic data with multi-state discrete characters. *Taxon*, **25:** 249-255

Ibeaha-Awemu EM, Jann OC, Weimann C, Erhardt G, 2004. Genetic diversity, introgression and relationships among West/Central African cattle breeds. *Genet. Sel. Evol.*, **36:** 673-690

Kamalzadeh A, Koops WJ, Van Bruchem J, 1998. Feed quality restriction and compensatory growth in growing sheep, modelling changes in body dimensions. *Livest. Prod. Sci*, **53:** 57-67

Kumar S, Gupta J, Kumar N, Dikshit K, Navani N, Jain P, Nagarajan M, 2006. Genetic variation and relationships among eight Indian riverine buffalo breeds*, Mol. Ecol.,* **15:** 593-600

Laloë D, Moazami-Goudarzi K, Chessel D, 2002. Contribution of individual markers to the analysis of the relationships among breeds by correspondence analysis. 7th World Congress on Genetics Applied to Livestock Production, August 19-23, 2002.Communication 26-06. Vol 33, p 493-496. Montpellier, France

Lauvergne J, Souvenir-Zafindrajaona P, 1992. Comparaison de deux populations de Zébu Malgache par les distances génétiques biométriques. *Rev. Elev. Med, Vet.,* **45:** 167-174

Lawrence TLJ, Pearce J, 1964. Some effetcs of wintering beef cattle on different planes of nutrition I, Live weight gain food consumptio and body measurement changes during the winter period and the subsequent grazing period. *J. Agr. Sci.*, **63:** 5-20

Le Vie L, 1999. Yellow Cattle, Animal Conservation in Vietnam, eds NIAH, Hanoi, p 92-118

Li MH, Tapio I, Vilkki J, Ivanova Z, Kiselyova T, Marzanov N, *et al.*, 2007. The genetic structure of cattle populations (*Bos Taurus*) in northern Eurasia and the neighbouring Near Eastern regions: implications for breeding strategies and conservation. *Mol. Ecol.,* **16:** 3839-3853

Mendelsohn R, 2003. The challenge of conserving indigenous domesticated animals. *Ecolog. Econ.*, **45:** 501-510

Mwacharo JM, Drucker AG, 2005. Production objectives and management strategies of livestock keepers in South-East Kenya: Implications for a breeding programme. *Trop. Anim. Health Prod.*, **37:** 635-652

Nuyen TT, Genini S, Bui LC, Voegeli P, Stranzinger G, Renard JP *et al.*, 2007 Genomic conservation of cattle microsatellite loci in wild gaur (*Bos gaurus*) and current genetic status of this species in Vietnam. *BMC Genetics*, **8**: doi:10.1186/1471-2156-8-77

Pinheiro JC, Bates DM, 2006. Nlme, Linear and nonlinear mixed effects models. R Foundation for Statistical Computing, Vienna, Austria.

Pritchard JK, Stephens M, Donnelly P, 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945–959

R Development Core Team, 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rousset F, Raymond M, 1995. GENEPOP (version 3-4): population genetics software for exact test and eucumenicism. *J. Hered.*, **86,** 248-59

Rege JEO, 1999. The state of African cattle genetic resources I. Classification framework and identification of threatened and extinct breeds. *Anim. Genet. Resour. Inf.*, **25:** 1-25

Reynolds J, Weir BS, Cockerham CC, 1983. Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics,* **105:** 767-779

Scherf BD, 2000. *World Watch List for Domestic Animal Diversity*, 3rd edn. Food and Agriculture Organization of the United Unions, Rome

Slatkin M, 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution,* **4:** 264–279.

Steinmetz R, 2004. Gaur (*Bos gaurus***)** and banteng (*B. javanicus*) in the lowland forest mosaic of Xe Pian Protected Area, Lao PDR: abundance, habitat use, and conservation. *Mammalia,* **68:** 141-157.

Taberlet P, Velentini A, Rezaei HR, Nader S, Pompanon F, Negrini R, Ajmone-Marsan P, 2007. Are cattle, sheep, and goats endangered species? *Mol. Ecol.*, doi:10.111/j.1365-294X.2007.03475.x

Tisdell C, 2003. Socioeconomic values of loss of animal genetic diversity: analysis and assessment. *Ecolog. Econ.,***45:** 365-376

Xuan DT, Szalay I, Su VV, Tieu HV, Vang ND, 2006. Animal Genetic resources and traditional farming in Vietnam. *AGRI* (ed. FAO), **38:** p103-110

Weir BS, Cockerman CC, 1984. Estimating F-statistics for the analysis of population structure. *Evolution,* **38:** 1358-70

Wiens H, 1954. China's march into the tropics. Emden (Conn.): Shoe String Press

Supplementary Material

Table S1. Summary of loci and their polymorphism measures: H_{Exp} Roslin: range record in Roslin database; A: number of alleles, H_{Exp}: unbiased expected heterozygosity, H_{Obs}: observed heterozygosity, Dis_{HWE}: number of district populations deviated from HWE equilibrium after Bonferroni correction, F_{ST}

Loci	Range	H_{Exp} Roslin	A	H_{Exp}	H_{Obs}	Dis_{HWE}	F_{ST}
BM1818	256-270	0.270-0.826	7	0.760	0.708	0	0.004
BM1824	178-194	0.441-0.875	6	0.690	0.657	0	0.019
BM2113	122-140	0.205-0.901	8	0.700	0.699	0	0.000
ETH10	208-220	0.160-0.809	7	0.765	0.750	0	0.003
ETH152	190-202	0.210-0.836	7	0.550	0.537	0	0.007
ETH185	222-240	0.583-0.813	7	0.766	0.456	6	0.023
ETH225	136-164	0.362-0.842	9	0.848	0.795	1	0.009
ETH3	111-125	0.465-0.811	6	0.576	0.542	0	0.035
HAUT27	127-151	0.406-0.812	10	0.846	0.821	0	0.01
HEL1	99-109	0.402-0.800	5	0.443	0.293	3	0.032
HEL13	181-191	0.279-0.801	6	0.696	0.512	5	0.019
HEL5	149-165	0.459-0.873	8	0.793	0.396	5	0.018
HEL9	140-166	0.210-0.883	13	0.868	0.815	0	0.005
ILSTS005	176-192	0.142-0.745	7	0.508	0.431	1	0.007
ILSTS006	272-302	0.436-0.807	13	0.821	0.755	1	0.005
INRA005	133-145	0.442-0.747	6	0.786	0.749	0	0.008
INRA023	195-215	0.478-0.897	7	0.751	0.707	0	0.022
INRA035	101-123	0.111-0.790	6	0.768	0.693	0	0.005
INRA037	113-133	0.524-0.820	9	0.591	0.583	0	0.005
INRA063	170-184	0.157-0.806	5	0.696	0.698	0	0.008
MM12	106-122	0.534-0.862	8	0.770	0.690	0	0.014
SPSS115	243-259	0.357-0.854	9	0.835	0.751	2	0.018
TGLA122	137-167	0.265-0.869	12	0.821	0.799	0	0.010
TGLA126	117-127	0.418-0.804	6	0.707	0.632	0	0.010
TGLA227	72-96	0.100-0.905	8	0.648	0.620	0	0.004

Table S2. Summary of body traits for District populations: HW: height at withers, IGs: index slenderness, EL: ear length, I_{EL}: EL/HW, BL: body length, I_{BL}:BL/HW, HG: heart girth, I_{HG}: HG/HW.

Sexe	District	HW	IG_s	EL	I_{EL}	BL	I_{BL}	HG	I_{HG}
Female	HSP	112±6.8	0.85±0.09	18.6±1.6	*0.17±0.01*	103.6±8.1	*0.093±0.07*	138.6±9.3	*1.25±0.07*
	XM	**115.5±4.9**	0.82±0.09	**19.4±1.8**	0.17±0.02	107.7±7.3	0.94±0.07	**144.1±7.5**	1.25±0.05
	QB	111.6±5.3	*0.80±0.08*	19.2±2.2	0.17±0.02	105.8±6.7	0.95±0.06	142.4±6.8	1.28±0.07
	YM	107.9±5.4	0.84±0.11	19.2±2.2	**0.18±0.02**	**107.9±7.8**	**1.01±0.06**	139.7±8.2	**1.30±0.06**
	DV	109±6.6	0.86±0.11	19.2±2.1	0.18±0.02	106.7±8.8	0.99±0.07	141.0±9.4	1.29±0.06
	MV	*106±5.8*	**0.88±0.09**	18.9±2.5	**0.18±0.02**	102.3±8.3	0.97±0.08	*134.2±6.7*	1.27±0.06
	BM	106±5.8	0.79±0.10	*18.4±1.8*	0.17±0.02	*99.9±7.6*	0.94±0.06	136.8±8.4	1.29±0.06
Male	HSP	118.6±4.5	0.81±0.07	18.8±1.4	0.16±0.01	105.8±9.0	*0.89±0.06*	147.0±7.9	*1.24±0.05*
	XM	**121.8±8.9**	*0.77±0.11*	**19.6±1.6**	0.16±0.01	111.6±12.3	0.92±0.07	150.8±15.6	1.25±0.05
	QB	119±7.7	0.77±0.08	19.2±2.7	0.16±0.02	**111.9±9.5**	0.95±0.06	**155.0±13.8**	1.30±0.08
	YM	116±8.2	0.80±0.13	19.3±2.4	0.17±0.02	108±10.4	**0.99±0.08**	146.1±16.7	1.28±0.10
	DV	114.1±7.3	**0.84±0.11**	18.8±1.9	0.16±0.01	109.0±9.5	0.96±0.07	145.6±14.9	1.28±0.07
	MV	*111.6±8.6*	0.84±0.11	19.0±2.1	0.17±0.02	104.1±10.3	0.95±0.09	*142.4±15.0*	1.27±0.07
	BM	114.1±9.4	0.74±0.12	*17.3±1.4*	*0.15±0.01*	*105.2±9.7*	0.92±0.06	147.5±16.5	**1.29±0.08**

in bold: maximum values, *in italic*: minimum values

Table S3. Summary of *p*-values for variables and co-variables on body traits and their average for cluster 1 (South-West) and the cluster (North-East) obtained with STRUCTURE software.

Body trait	Transformation	Sexe	Altitude	Cluster	Male		Female	
					South-West	North-East	South-West	North-East
HW	none	*<0.0001*	0,1594	*0,0352*	119±8.1	113.6 ± 8.5	112.7±6	107.6±6.2
IGs	Racine	*<0.0002*	0,1356	0,5269	0.78±0.09	0.82±0.12	0.82±0.09	0.85±0.11
BL	Ln	*0,0031*	0,5372	0,8514	109.9±11.5	108.9±12.5	106.2±8.3	104.9±9.5
I_{BL}	Racine	*<0.0002*	0,8255	0,1718	0.92±0.07	0.95±0.07	0.94±0.07	0.98±0.07
HG	none	*<0.0001*	*0,0109*	0,5219	152.4±12.9	144.6±14.1	141.4±8.3	138.4±9
I_{HG}	racine	*<0.0001*	*0,0437*	0,0923	1.27±0.07	1.28±0.08	1.26±0.07	1.29±0.06
EL	none	0,2956	0,3236	0,3867	19±2.1	18.8±2	19.2±1.9	18.9±2.2
I_{EL}	none	*<0.0001*	0,9722	0,9722	0.16±0.02	0.17±0.02	0.017±0.02	0.18±0.02

in italic: significant p-values

Table S4. % of cattle origins per commune

District	Commune	Family	Village	Commune	District	District Market	Province	Project	Other Province
BM	75	10.0	30.0	10.0	30.0	0.0	10.0	0.0	10.0[a]
	89	12.5	37.5	10.7	10.7	0.0	3.6	1.8	23.2[a]
	113	12.9	41.9	6.5	22.6	0.0	0.0	9.7	6.5[a]
MV	4	16.0	6.0	2.0	12.0	64.0	0.0	0.0	0.0
	40	22.0	19.5	7.3	4.9	46.3	0.0	0.0	0.0
	45	6.7	33.3	20.0	20.0	20.0	0.0	0.0	0.0
	48	19.2	7.7	32.7	23.1	17.3	0.0	0.0	0.0
DV	1	6.7	37.8	31.1	6.7	13.3	0.0	0.0	4.4[b]
	7	0.0	22.4	10.3	31.0	36.2	0.0	0.0	0.0
	16	3.1	12.5	40.6	31.3	3.1	9.4	0.0	0.0
	19	2.2	19.6	6.5	30.2	30.4	8.7	0.0	2.2[b]
YM	25	26.3	36.8	23.7	7.9	0.0	5.3	0.0	0.0
	61	33.3	19.3	24.6	22.8	0.0	0.0	0.0	0.0
	65	46.7	13.3	20.0	20.0	0.0	0.0	0.0	0.0
QB	30	0.0	20.8	33.3	25.0	0.0	16.7	4.2	0.0
	49	33.3	20.0	33.3	6.7	0.0	6.7	0.0	0.0
	56	4.3	30.4	26.1	34.8	0.0	4.3	0.0	0.0
	188	25.0	28.1	21.9	21.9	0.0	3.1	0.0	0.0
QBn	157	0.0	0.0	20.0	20.0	0.0	60.0	0.0	0.0
	179	4.8	0.0	0.0	9.5	0.0	19.0	66.7	0.0
HSP	85	5.0	10.0	30.0	50.0	0.0	5.0	0.0	0.0
	110	6.5	41.9	16.1	9.7	6.5	6.5	12.9	0.0
	114	8.3	41.7	8.3	33.3	0.0	8.3	0.0	0.0
XM	91	0.0	25.0	50.0	25.0	0.0	0.0	0.0	0.0
	103	4.5	39.4	27.3	22.7	0.0	1.5	0.0	4.5[b]

[a]: from Cao Bang; [b]: from China

4.4.3 Complementary results

1   The MCOA analysis on Ha Giang cattle population

The MCOA analysis of the Ha Giang cattle population showed high structuring values: the first two axes explained 87 % of the total inertia. The Cos^2 values, which expressed the similarity of structuring observed within each locus table with the synthetic representation, ranged from 0.615 to 0.976 for the first axis; and from 0.237 to 0.921 for the second axis (Figure 4.12). However, only three loci had high *TV* values for the first axis: HEL13 (0.033), SPS115 (0.026) and HEL1 (0.025). For the second axis, the loci with highest *TV* values were HEL5 (0.014), MM12 (0.009) and HAUT27 (0.009). Variance and *TV* of loci were significantly correlated to F_{IS}, F_{ST} and F_{IT} values when performing Pearson's test (Table 4.6).

Figure 4.12 Distribution across loci of values for the components to 1st axis (A) and 2nd axis (B), in percentages
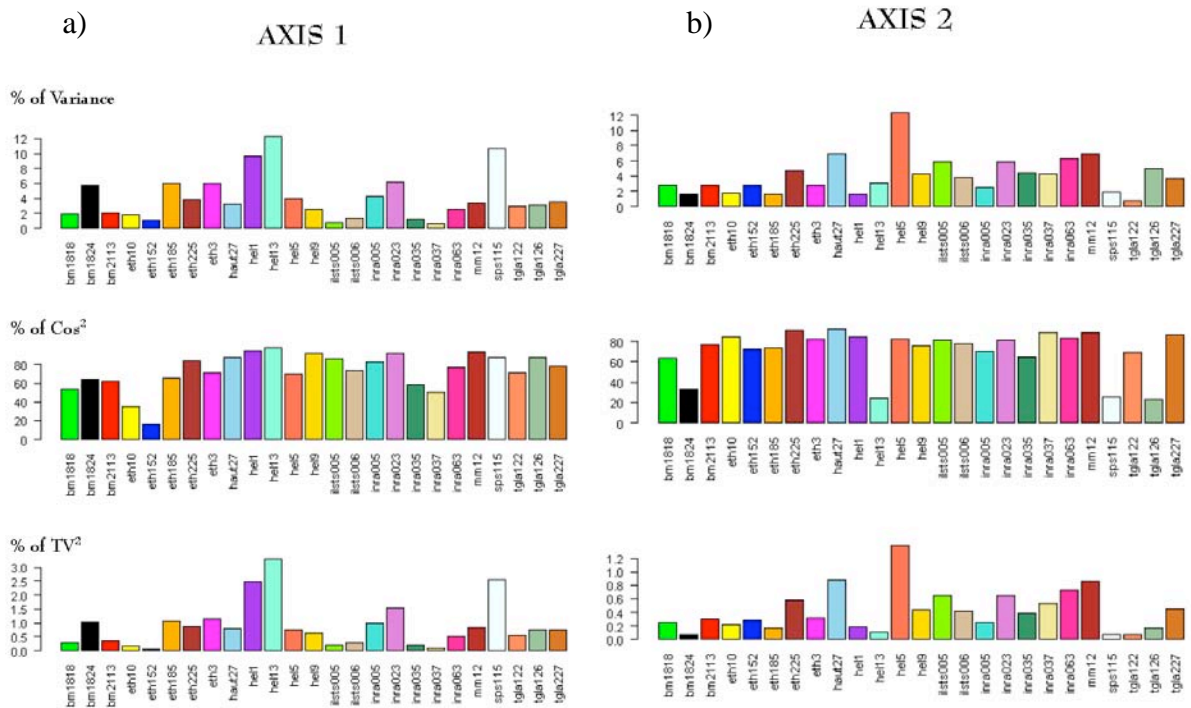
Table 4.6 Correlation values of Variance, Cos^2 and *TV* for the 1st axis with F-statistics values.

Variables	F_{IS}	F_{ST}	F_{ST}
Variance	0.412*	0.724**	0.445*
Cos^2	ns	ns	ns
TV	0.396*	0.67**	0.627*

ns: no significant,* p<0.05, **p<0.0001

The 1st axis allowed the differentiation of the cattle population from QBn district, while the 2nd axis differentiated districts according to an east-west cline (Figure 4.13).

Figure 4.13 Synthetic MCOA representation of district cattle populations.
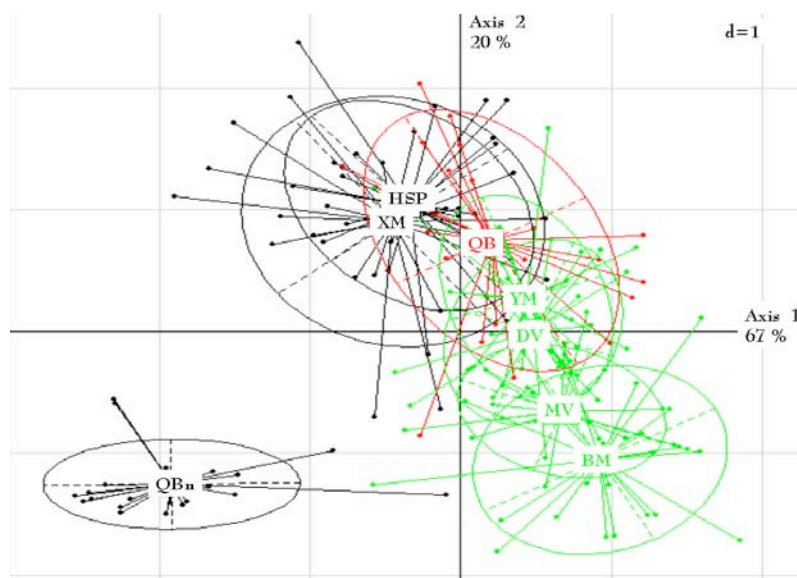
2   Live weights and phenotypes

Live weight estimated from linear regression (see p48) for bulls ranged from 231 kg (HSP and DV) to 283 kg (QBn); and for cows from 204 kg (MV) to 230 kg (QBn) (Table 4.7).

Table 4.7 Averages of cattle live weights (average ± sd)

District	Male	Female
BM	247 ± 90	208 ± 37
MV	218 ± 83	204 ± 32
DV	231 ± 83	227 ± 42
YM	249 ± 98	223 ± 27
QB	283 ± 70	230 ± 32
HSP	231 ± 48	218 ±4 0
XM	253 ± 85	236 ± 30

The Ha Giang cattle population had a uniform coat colour in more than 90 % of cases except in the Xin Man district where 17.5 % of animals had two colours, mainly black and red (Table 4.8). The main colour of Ha Giang cattle is yellow (50 % -76 %). The test of Chis^2 showed that there was significantly more yellow-coated cattle in Hoang Su Phi than in other districts. Black-coated animals are mainly found in XM (21,1 %), HSP (19.6 %) and MV (19.8 %)..

Table 4.8 Proportion of cattle's coat colours per district

District	Yellow	Black	Brown-Red	No uniforme
BM	50.0	14.2	25.8	4.8
MV	66.4	19.8	13.8	5.2
DV	60.4	13.9	25.7	5.0
YM	62.1	9.1	28.8	10.6
QB	62.3	9.8	27.9	8.2
QBn	60.0	5.0	35.5	10.0
HSP	76.1	19.6	4.3	8.7
XM	50.9	21.1	28.1	17.5

The Chi^2 test on horn shape distributions showed that in north-eastern districts there were significantly more frontward horns(45 % - 62.9 %) whereas in the south-western districts, mainly vertical horns were observed (41.8 % - 47.4 %) (Table 4.9).

Table 4.9 Proportion of cattle's horns position.

District	Frontward	Horizontal	Vertical
BM	50.0	22.2	27.8
MV	45.0	28.3	26.7
DV	54.2	18.8	27.1
YM	62.9	12.9	24.2
QB	35.8	22.4	41.8
QBn	20.3	33.3	46.4
HSP	39.0	16.0	45.0
XM	25.0	27.6	47.4

4.5 Conclusion

Laloë *et al.* (2007) applied MCOA analysis to African and European breeds. They found that loci HEL13 and HEL1 had the highest *TV* for axis 1, MM12 for axis 2 and HEL 5 for axis 3. In our study, we found the same discriminant power for most of those loci. For the 1st axis, HEL13, SSP115 and HEL1 had the highest *TV* values, and for axis 2: HEL 5, MM12 and HAUT27. Therefore, such results confirmed the power of those loci in discriminating the genetic structure of cattle populations. Moreover, high significant correlations between descriptive values of loci power and F-statistics were obtained.

Genetic distances maximising genetic differences between breeds, MCOA analysis taking into account within- and between-diversity of populations, and Bayesian individual admixture analysis, are all methods that revealed a genetic structure in two clusters of the Ha Giang population. This genetic differentiation was also observed in a few phenotypic traits, such as the height at withers and horn shape.

The combining of genetic data with farmer's management practices enabled us to further understand the genetic structure of the cattle population in the Ha Giang province. Therefore, conservation policies could be adapted for both aspects: the genetic status and the farmer's practices.

CHAPTER 5

THE SWAMP BUFFALO

GENETIC AND MORPHOMETRIC CHARACTERISATION OF A LOCAL VIETNAMESE SWAMP BUFFALO POPULATION

C. Berthouly, X. Rognon, T. Nhu Van, H. Thanh Hoang, B. Bed'Hom, D. Laloë, C. Vu Chi, E. Verrier, J-C Maillard

5 THE SWAMP BUFFALO
5.1 Taxonomic status, distribution and domestication of the Water Buffalo

Within the family Bovidae, the tribe of Bovini consists of the genera *Bos, Bison, Pseudoryx, Bubalus* and *Syncerus* (Figure 5.1). The latter two genera belong to the group of buffaloes, the Asian and the African buffaloes. Amano *et al.* (1994) analysed RFLPs in nuclear ribosomal DNA and mtDNA from cattle and water buffalo. They demonstrated that these two taxa diverged from approximately 4 to 6 MYA whereas around 1.85 to 4.93 MYA were estimated by Ritz *et al.* (2000) using microsatellites. Both results demonstrated a lineage separation at the end of the late Miocene. During the Pleistocene, the Bubalus genus was distributed from southern Asia to Europe. With an increasingly dry climate the area of distribution decreased to India, Indonesia, and parts of South-East Asia.

The genus Bubalus has been divided into two groups : wild and domestic buffaloes. Three species of wild buffalo exist: (i) the Asian water buffalo (Bubalus bubalis arni), (ii) the tamarao (Bubalus mindorensis), (iii) and the anoa (Bubalus depressicornis). Little information is reported on the wild buffalo which is considered to be critically endangered (IUCN list). Their area distribution of this species is small and it has been observed in western Thailand, eastern Nepal, Assam and southern Bhutan (Choudhury 1994; Heinen & Singh 2001). Observations of buffaloes in Thailand and Cambodia have been reported, however the wild or feral status of those animals was impossible to establish. The tamarao is encountered only on the Mindoro island of the Philippines while the anoa is found only on the island of Sulawesi in Indonesia. Previous studies on mtDNA showed that the anoa could be classified into two subspecies: the lowland: Bubalus depressicornis; and the mountain anoa: Bubalus quarlesi (Kikkawa et al., 1997). The water domestic buffalo of Asia (Bubalus Bubalis) has been classified on morphological and behavioural criteria into two types (Macgregor 1939): the river buffalo of the Indian subcontinent and west to the Balkan, and the swamp buffalo of South-East Asia, from Assam and Nepal east to the Yangtze valley of China (Figure 5.2).

Buffaloes were introduced to Italy from central Europe in the sixth century or by the bay of Tunis in the seventh century at the time of the Arab conquest (Salerno, 1974). Importation of water buffaloes to Africa, Australia or America took place only recently.

Figure 5.1 Systematic phylogeny of Buffalo groups modified from Groves (1981).
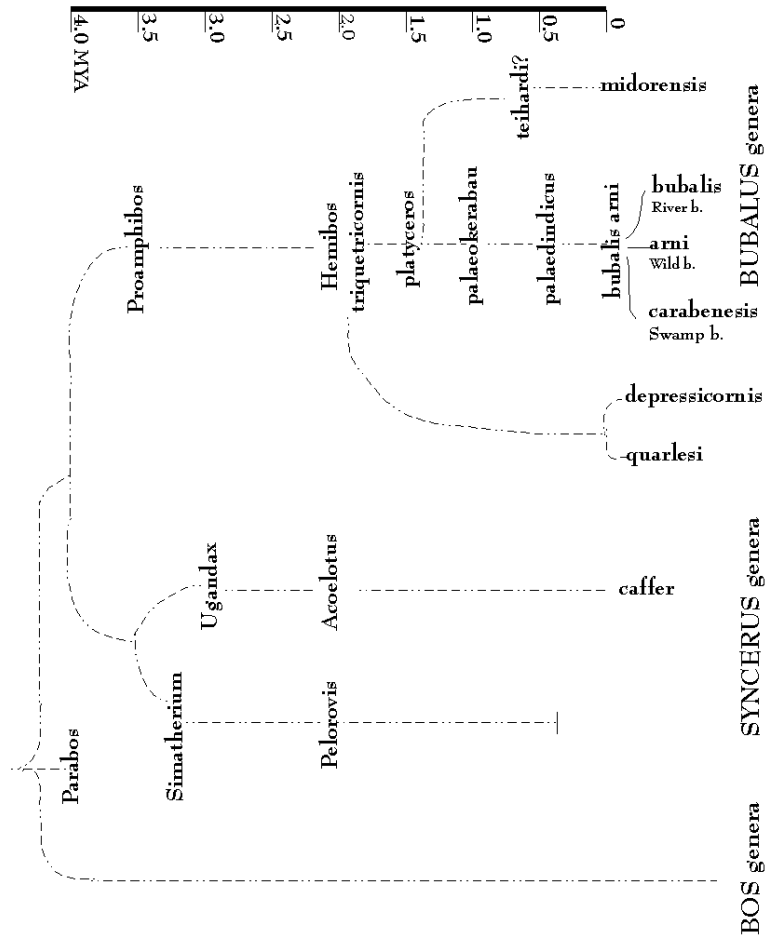
Figure 5.2 Distribution of *Bubalus sp.* in Asia and neighbouring countries.
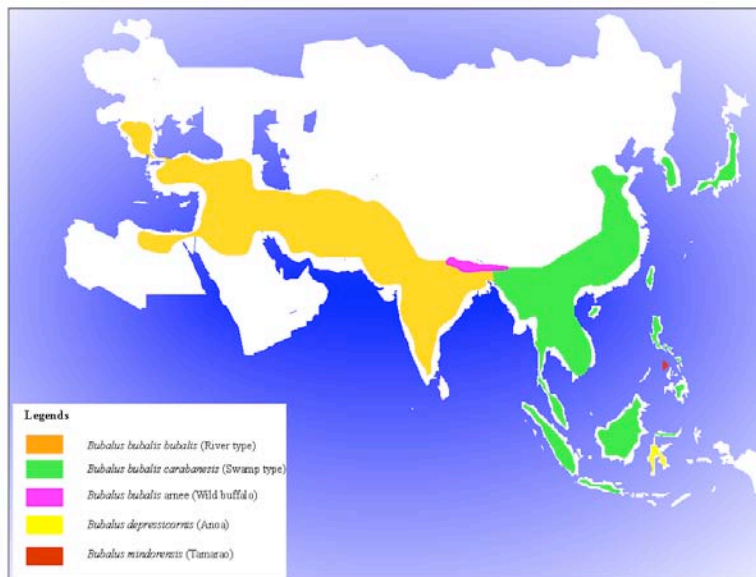
Therefore, river and swamp buffaloes are distributed West and East of Bangladesh. Previous cytogenic and mtDNA studies suggest that Bangladesh was a hybrid zone for the two domestic types of water buffalo (Amano et al. 1984; Kikkawa et al. 1997). In addition to different distribution areas, cytogenetic variations in chromosome number made it possible to distinguish Bubalus sp. The tamarao has 2n=46 (Custodio et al.,1996), while the status of anoa is unclear because animals with a number of chromosomes from 2n=38 to 2n=48 were observed (Burton et al., 2005). The wild buffalo, phenotypically similar to the swamp type, had a chromosome number of 2n=50 similarly to the river type but different to the swamp type which had 2n=48 (Ulbrich & Fischer 1967; Fischer & Ulbrich 1968). However, the wild and domestic water buffaloes are interfertile leading to a F1 and F2 offspring with 2n=49. Therefore, many studies have been conducted in order to clarify the taxonomic status and domestication history of the water buffalo.

Lau *et al.* (1998) had previously studied the cytochrome *b* of 12 swamp buffalo populations and two river populations from South-East Asia. They observed six haplotypes: four were found in swamp only and two in river buffalo only, one haplotype being ancestral of both types. The same conclusions were established from D-loop phylogeny. Among the 33 D-loop haplotypes observed, all the ones found only in river buffalo occurred as terminal tips (i.e. most recent) in the minimum spanning tree. Then, Lau *et al.* (1998) hypothesised that the species originated in mainland South-East Asia and then spread north to China and west to the Indian subcontinent where the river type might have evolved. Divergence time between the domestic river and swamp buffalo was estimated to be within the last one million years (Amano *et al.* 1994; Kikkawa *et al.* 1997) or within the last two million years according to Ritz *et al*. (2000). Based on the D-loop data, Lau *et al.* (1998) suggested that the two types of water buffalo diverged between 0.028 and 0.079 MYA only. However, Kumar *et al.* (2007a) estimated a divergence time between 128 000-271 000 years ago depending on the mutation rate assumed. According to Cockrill (1981), wild buffaloes were domesticated about 4500 years ago and the centre of domestication was either Mesopotamia, during the Akkadian dynasty, or the valley of the Indus, during the Harappa/Mohenjo Daro civilizations. It seems possible that the valley of the Indus was the earliest centre of domestication of river buffaloes. The swamp buffaloes are assumed to have been domesticated in China around 7000 years ago (Chen & Li, 1989). Then, and as proposed by Barker *et al.* (1997), domesticated swamp buffaloes dispersed following two routes, from China south through Thailand and Malaysia,

and from China through the Philippines to the eastern islands of Borneo (Sabah and Sarawak) and Sulawesi (Figure 5.3).

Figure 5.3 Dispersal routes of *Bubalus bubalis* according to Lau *et al.* (1998).
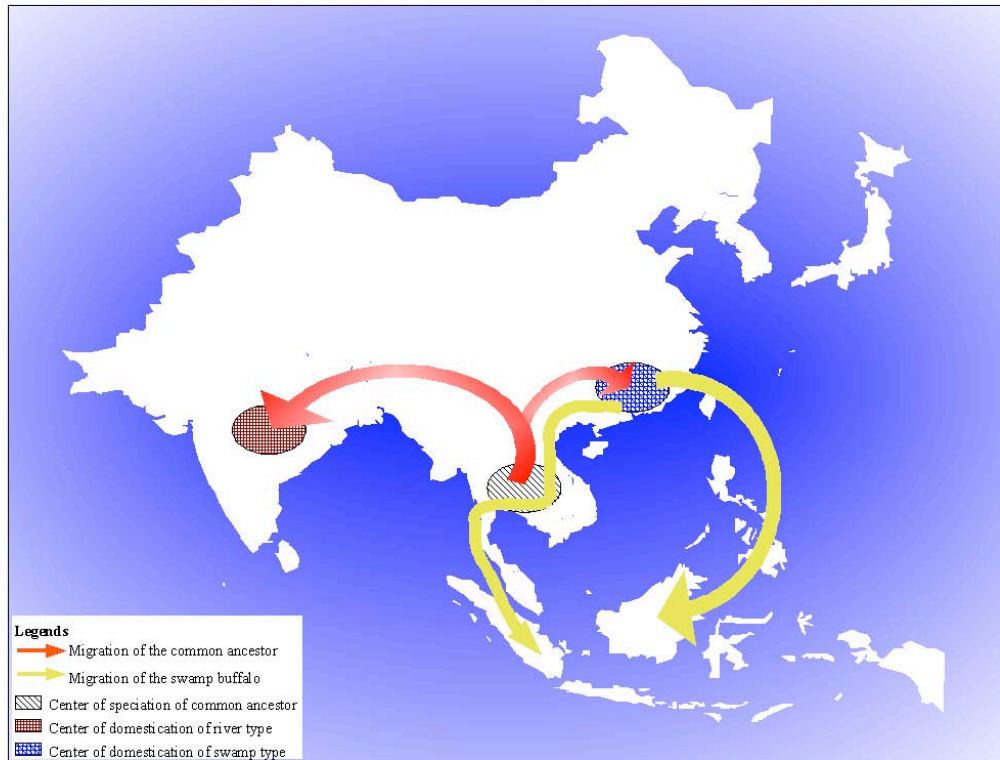
Such an hypothesis was questioned by Kierstein et al. (2004) when they analysed the entire D-loop region of the river and swamp buffaloes from Brazil and Italy. They concluded that the two types descended from a single domestication event some 5000 years ago. However, no samples from Indian or Chinese buffaloes were used in this study. Loftus et al. (1994) reported an average sequence divergence of 7.9 % between the mtDNA D-loop of B. taurus and B. indicus; while the corresponding value for the river and the swamp buffaloes was 8.6 % estimated by Kumar et al. (2007a) when combining Indian river haplotypes and haplotypes from Kierstein et al. (2004). Therefore these results support the classification of river and swamp buffalo as two subspecies, i.e. B. bubalis bubalis and B. bubalis carabenesis and the hypothesis of an independent domestication event of these two types. Moreover, Lei et al. (2007) obtained two clusters with a neighbour-joining tree on mtDNA on Chinese, Indian and Brazilian/Italian river and swamp buffaloes. This suggests a second domestication event for swamp buffalo in China that later diverged into two maternal lineages A and B. The mtDNA analysis on 8 Indian buffalo breeds (Kumar et al. 2007b) suggested that the river buffalo might have been domesticated in the Western region of India and the nearby Pakistan

146

area. At this point, river and swamp buffaloes should be considered as two subspecies descending from a common ancestor, possibly the wild buffalo B. bubalis arnee; the common ancestor might have been domesticated in two regions: in western India leading to the river buffalo; and in China for the swamp buffalo.

5.2 Role of buffaloes
5.2.1 Role in the socio-economy of Asia

Most Asian countries are agrarian with 60-80 % of the population related to farm operations in one way or another. In Asia, buffalo played a pivotal role in overall social development through contributions to milk, meat, hide production and draught power for agricultural activities. The river buffalo cows have high lactation yields and bulls are suited to ploughing and draught on dry plane land. The swamp buffalo is more suited to muddling and rice cultivation, hence their diffusion through South-East Asia with the spread of rice culture. The world buffalo population has increased these last years, but this growth was not uniform in all buffalo rearing countries in Asia (Table 5.1). Although important increases were observed in South Asia, countries in eastern Asia exhibited a decline in buffalo populations.

Table 5.1 Changes in buffalo population in different regions of Asia between 1990-2002

Region	Year		% Change
	1990	2002	
South Asia	102 689 700	123 326 100	+ 20.1
South-East Asia	18 170 490	14 804 550	-18.52
Asia	143 302 100	161 456 000	+12.67
World	148 183 900	166 419 000	+12.31

Source: FAOSTAT

The most important decline was observed in Thailand with an annual decrease of 5 % since 1995 (Figure 5.4). The Sri Lanka buffalo population experienced a similar situation to Thailand. Vietnam had a non or all pattern of buffalo population growth: on one hand, an important decline was observed around the Red and Mekong delta rivers and, in the other hand, an increase was observed in the coastal area.

There could be many reasons for the decline of the buffalo population in South-East Asia. The foremost reasons could be: (i) increased agricultural mechanisation; (ii) increased urbanisation, industrialisation decreasing paddy areas and also the need of buffalo for draught.

During the last decade, the importance of buffalo as a dairy animal in Asia has increased. More than 5 10^6 x tons (Table 5.2) are produced in India, which represent 65 % of the worldwide buffalo milk production, are produced in India whereas only 3.5 % are produced in China. From 1990 to 2006, India has increased its production by 80 % and Vietnam by 30 %.

Figure 5.4 Evolution of Buffalo herds in some Asian countries between 1990-2006.
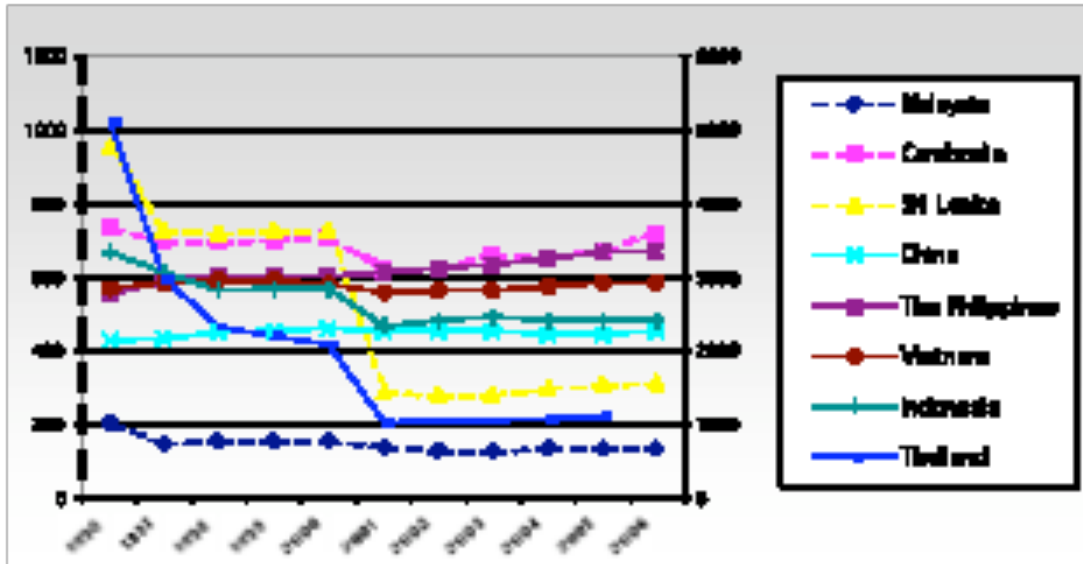
Source: FAOSTAT

However in Sri Lanka buffalo milk production decreased by 58 %. River buffalo produce more milk than swamp buffalo, which produces less than 2 kg/day during lactation periods and around 3.7 kg/day when it is crossed with Murrah river buffalo. Buffalo milk is richer in saturated fatty acids than cattle milk. Swamp buffalo milk has even higher fat, protein and lactose contents. Therefore, incomes from buffalo milk are higher than from cattle milk in countries such as India where the selling price of the milk depends on fat contents.

Buffalo meat in many cases comes from culled animals and extra males. Only 5.5 % of the total meat in South-East Asia comes from buffalo. China shows the highest increase in buffalo meat production during these last decades (Tab. 5.3) whereas buffalo meat production in Thailand was nearly halved. In 2006, India produced nearly 50 % of the buffalo meat worldwide production while China produced only 11 %. Buffalo meat contains 40 % less cholesterol, 55 % less calories, 11 % more protein and 10 % more minerals in comparison to cattle meat (UPSA, 1996). Nevertheless in some countries such as Vietnam a myth exists about the doubtful quality of buffalo meat and this has made it unpopular.

148

Table 5.2 Buffalo milk production in 1000 tons between 1990 and 2006.

Country	Years			
	1990	1995	2000	2006
China	1 900 000	2 200 000	2 650 000	2 850 000
India	29 057 000	35 870 000	43428000	52 100 000
Sri Lanka	63 891	67 448	25 520	26 730
Malaysia	9 545	7 590	6 785	7 475
Viet Nam	24 000	25 000	30 000	31 000

Source: FAOSTAT

Table 5.3 Rates of change (%) between 1990 and 2006 in buffalo meat production; and proportion (%) of the worldwide buffalo meat production in 2006 in different countries.

Country	% Change	% World
Cambodia	-4.3	0.3
China	121.7	11
India	37.9	46.7
Indonesia	-10.83	1.2
Laos	125.7	0.6
Malaysia	11.6	0.1
The Philippines	64.5	2.2
Sri Lanka	-50	0.1
Thailand	-54.2	2
Viet Nam	15.7	3.2

Source: FAOSTAT

For more than 5000 years, buffaloes have been used for draught power. They are particularly suited to work on wet fields. Because they love water, they can stay longer than mules while ploughing fields for rice plantation. Buffaloes are more docile than cattle and are, therefore, easy to handle, even by children, train and milk. Training for work starts at approximately 2 years old and until 11 years old on average. Buffaloes can work about 3-4 hours per day in summer and up to 6-8 hours in winter. Based on profits produced over and above the investment, a pair of buffaloes is 2.6 times more efficient than a tractor (Thu *et al,* 1995). Moreover, dung is a valuable product from buffalo farming, and is used as a fertiliser, as fuel for cooking, or for making biogas. Crop residues constitute 70 % of the roughages fed to buffalo. Buffalo grazing in empty rice and wheat fields pick up left over/fallen seeds, green grasses and weeds. Therefore buffaloes are an integral part of the socio-economic network; especially for small holders in several agrobased developing countries in Asia.

5.2.2 Role in the socio-economy of Vietnam

As in other Asian countries, buffalo play a key role in the socio-economy of Vietnam. The swamp buffalo is the main draught power for rice cultivation in the small Vietnamese farms.

As their importance of buffaloes has been recognised for a long time, many studies have been done in order to evaluate their performance and their economic efficiency (Thu, 1997). With 2.9 million heads, Vietnam represents 1.64 % of the worldwide buffalo population. Although buffalo population in the world is increasing by about 1.3 % per year, in Vietnam it is decreasing by 0.73 %. Since 1995, there has been a reduction of buffalo numbers due to increasing mechanisation in addition to a decrease in grazing land as a result of a decrease in the area under cultivation, but also because of increasing mechanisation. This situation has developed in the lowland, particularly in the Mekong and Red river delta provinces where buffaloes used to graze extensively. Such a decrease leads to the situation observed in 2006 as shown in Figure 5.5a. In almost all provinces, there were less than 10 buffaloes per km^2; the highest buffalo density being observed in three provinces from the North: Bac Giang, Hoa Binh and Thanh Hoa.

Figure 5.5 Density and herd size of buffalo population in Vietnam. a) Density of buffaloes per km^2 in Vietnam for 2005; b): Changes of herd size between 2001-2005 in Vietnam.
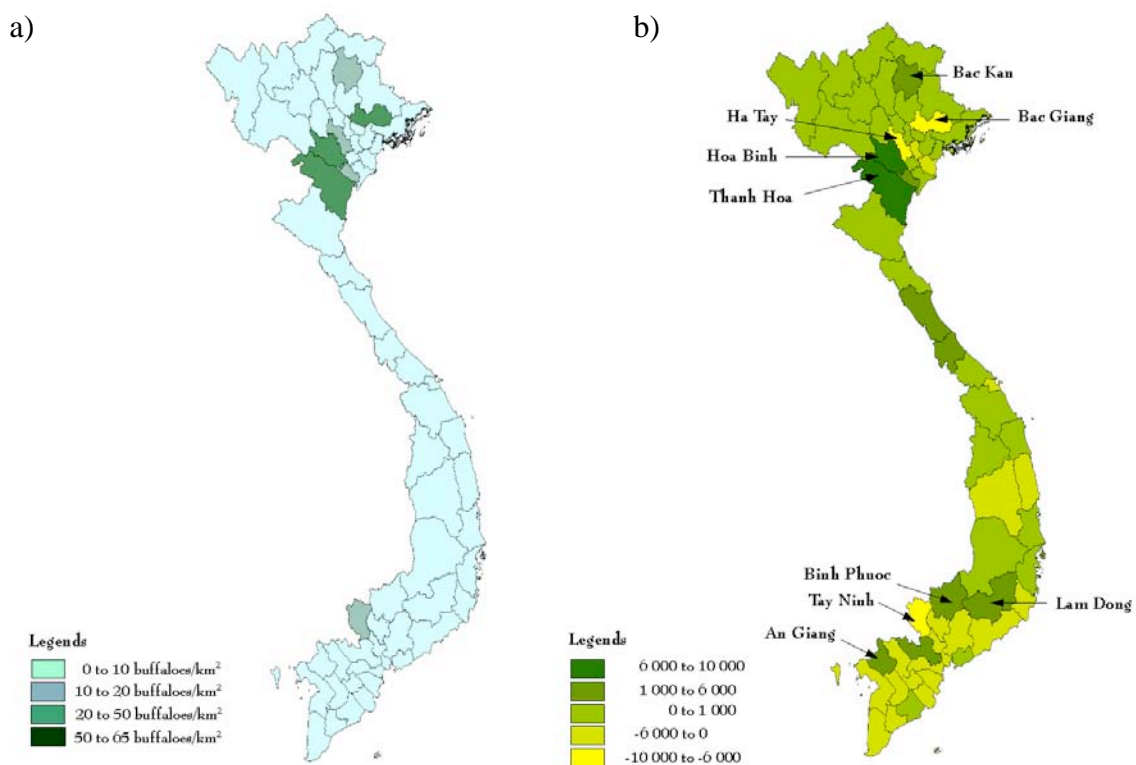
Source: Statistical Yearbook of Vietnam

From 2000 to 2005, the loss in number of heads was more important in the South around the Mekong River delta (Figure 5.5b), whereas in central and northern provinces of Vietnam, an increase in buffalo population was observed. However, two northern provinces, namely Bac Giang and Ha Tay showed an important decrease in their herds. The highest growth rate

150

was observed in An Giang province where buffalo population was multiplied by 1.76 between 2001 to 2005. An Giang province is nearly followed by Binh Phuoc (x 1.27) and Lam Dong (x 1.25) provinces.

Figure 5.6 Buffalo meat production (x 1000 tons) in 2005 in Vietnam.
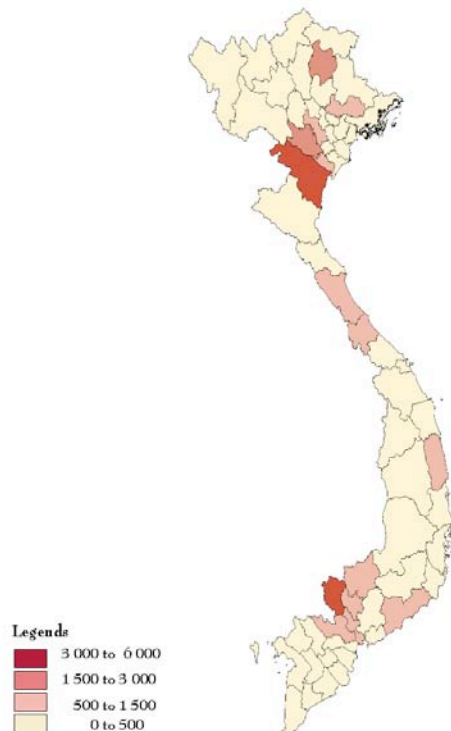
In 2005, Vietnam produced more than 256 000- 370 000 tons of meat (buffalo and cattle), where 40 % comes from buffalo production. Buffalo meat production is centralised in mainly two provinces: Than Hoa province for North Vietnam and Tay Ninh province in the South. Between 2001 and 2005, buffalo meat production increased by 6.6 % (Figure 5.6). However in Vietnam, farmers traditionally raised buffaloes for draught power and techniques are mostly primitive. Management practices are based on extensive systems and buffaloes graze freely on natural grass land, forests, roadsides, canal banks, rice fields after harvesting and dikes, etc … Farmers normally do not fatten buffaloes before selling for meat but middle men occasionally buy and fatten them. Until 1960, the slaughter of buffaloes below the age of 12 was prohibited. Even today, buffalo are mainly slaughtered when they are not able to work due to old age or after an injury. This induces a low yield for buffalo meat production and a bad reputation about buffalo meat quality.

5.3 Diversity of the Water buffalo

5.3.1 Asian breeds

While the river buffalo has been selected to form improved breeds with high milk yield, the swamp buffalo has not been divided into special types. It has retained the low milk yield

151

and the primitive horn shape and white marks of the wild water buffalo. The buffalo breeds in different areas have sometimes been given local names but they cannot be labelled a breed as defined by Western countries. Regarding Indochina, Le Louet (1931) emphasised that all buffaloes belong to the same breed. In Vietnam, there are a number of varieties bearing the name of the district in which they are bred but differing little from each other (Allarverdiev, 1962). In China, variations of conformation, horn shapes and colours are considerable but there is no clear subdivision into different breeds. Such remarks from the beginning or middle of the last century, about low differentiation between buffalo populations, seem to be still relevant now. A good clue of this phenomenon is expressed by the little information per breed available through the DAD-IS/FAO website, summarised in Table 5.5.

A total of 16 well-documented riverine breeds are reported in India, most of them being distributed in the western part of the country (Figure 5.7a). Most breeds are black-coloured; the live weight ranging from 380 kg (Toda) to 800 kg (Murrah) and the height at withers varying from 114 cm (Jeragi) to 150 cm (Toda). In Thailand, Swamp buffalo is considered to belong to one breed, whereas in Cambodia two breeds are differentiated according to the ecozone in which they evolved: mountain or plain area. In China, a total of 23 breeds are described in the FAO database, distributed in the south provinces (Figure 5.7b). Breed colour varies from brown, blue-grey to black-grey. The live weight ranged from 344 kg (Xinfeng Mountainous) to 500 kg (Dehong, Dongliu) and height at withers varies from 116 cm (Xinfeng Mountainous) to 133 cm (Xiajiang).

Figure 5.7 Map of India (a) and China (b) showing centres of origin of main buffalo breeds.

a)
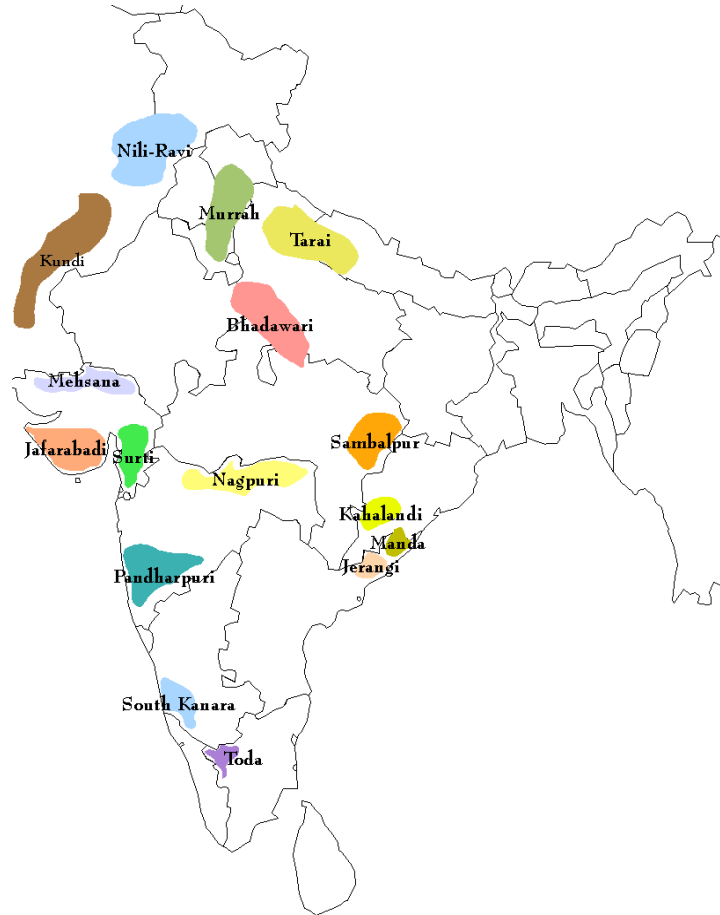
b)
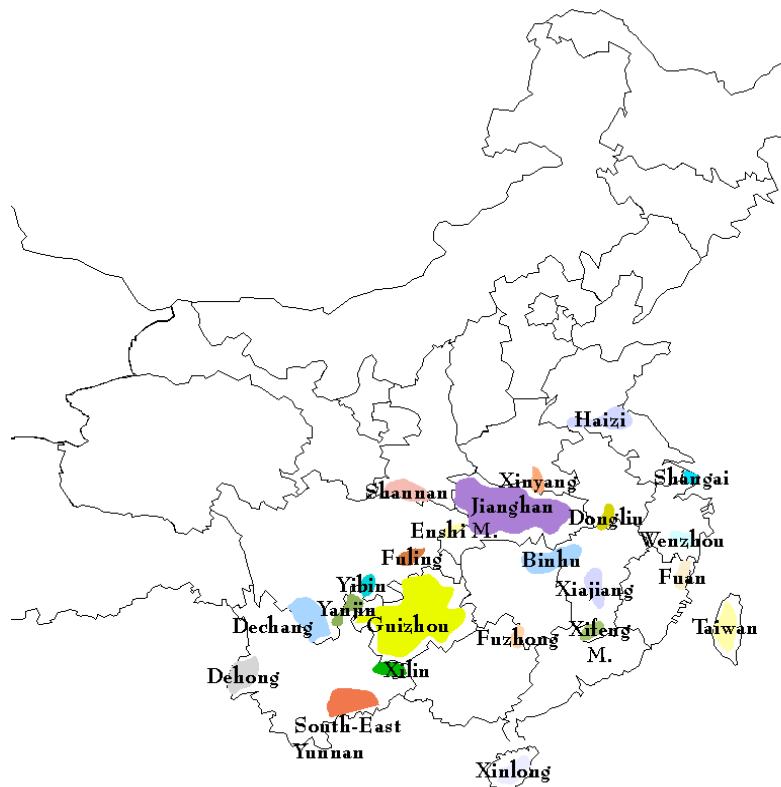
5.3.2 Vietnamese breeds

It is generally considered that two breeds are raised in Vietnam: the Ngo and the Noi breeds. However no reliable measurements are available of those breeds. The Vietnamese swamp buffalo is grey with white markings on the face, white socking and with none, one or two white chevrons. Horn shape is semi-circular and can vary in size. Albinoid buffalo are mainly found in northern provinces but their frequency can widely vary within a province. It seems that buffalo size increases from South to North as summarised in Table 5.4.

Table 5.4 Body measurements of Vietnamese swamp buffalo.

Region	Mekong delta	South Eastern Provinces	Central Area	Northern Provinces
Female				
Height at Withers	127	126	121	118
Length of body	141	132	128	127
Height girth	192	193	191	180
Live-weight	458	434	417	358
Male				
Height at Withers	131	129	129	121
Length of body	143	135	136	133
Height girth	198	197	198	190
Live-weight	494	464	494	428

Sources: Thu (1987); Hoang (1978), Thac (1979)

Table 5.5 Summary of Asian water buffalo breeds characteristics.

Country	Breed name (type)		Colour	% albinoids	white markings	Chevrons	Horns	Weight	HW
India	Murrah	River like	black		forehead, tail (but undesirable)		short and curled	450-800	142
	Nili-Ravi	River like	blue		forehead, tail		ort and curled	483	134
	Kundi	River like	black		forehead, tail			450	
	Surti	River like	black or brown		forehead, tail	yes	medium length and sickle-shaped	640-730	131
	Mehsana	Surti x Murrah	variable		forehead, tail		sickle-shaped		
	Jafarabadi	River like	black					590	142
	Bhadawari	x Murrah	black and brown		tail	yes	compact	425	124
	Tarai	x Murrah	black		tail		long and flat	325	120
	Nagpuri	River like	black		forehead, tail		long and flat	522	142
	Pandharpuri	River like	black		tail				
	Manda	River like	brown or grey				broad and semi-circular		132
	Jerangi	River like	black						114
	Kalahandi	River like	grey		tail				
	Sambalpur	River like	black or grey				short horn		
	Toda	similar to Swamp	grey					380	150
	South Kanara								
Nepal	Neapalese	x Murrah, Swamp ancestry	dark grey or black			yes		450	121
Thailand	Thai	Swamp like	grey, black, white	30 % in the north; 5 % in the centre and south	head. Tail and socks	yes		343	122
Sri-Lanka	Lankan	River but similar to Swamp b.	fawn to black					306	
	Mannar	Lankan variety	red						
	Tamankaduwa	Lankan variety							
The Philippines	Carabao	Swamp like	black	3 %	sickle-shaped	yes		420-500	127-137
Lao	Khouay	Swamp like						400	105
Indonesia	Spotted	Swamp like	black with white						
	Sumbawa	Swamp like	grey to black					350	100
	Tedong	Swamp like	black with white					400	
Cambodia	Krabey-beng (Mountain)	Swamp like	brown			yes	long horn	400	140
	Krabey-leu (plain)	Swamp like	brown			yes	short horn	350	129
Malaysia	Borneo		grey			yes	short horn		
	Kerbau-Sawah	Swamp like	grey	3 %		yes	handsomes horns	363	130

Country	Breed name (type)		Colour	% albinoids	white markings	Chevrons	Horns	Weight	HW
Vietnam	**Trau Ngo**	**Swamp like**	**grey**			**yes**			
	Trau Noi	**Swamp like**	**grey**			**yes**			
China	Binhu	Swamp like	black-grey, white				plate horn	473	125
	Dechang	Swamp like	grey				wide horn	490	128
	Dehong	Swamp like	black, grey and white					500	127
	Dongliu	Swamp like	black-grey, black-brown					500	126
	Enshi Mountainous	Swamp like	blue-grey				short horn	422.4	126
	Fuan	Swamp like	blue-grey, black-grey or light brown					456	125
	Fuling	Swamp like	brown, white				wide and flat	446.5	125
	Fuzhong	Swamp like	blue-grey			yes		415	125
	Guizhou	Swamp like	grey, white			yes		411	120
	Haizi	Swamp like	blue					496	134
	Jianghan	Swamp like	blue, blue-grey					519	127
	Shanghai	Swamp like	grey					616	141
	Shannan	Swamp like	black or grey			yes	flat	399	125
	South-east Yunnan	Swamp like	grey, white						116
	Taiwan	Swamp like	black			yes		389	127
	Wenzhou	Swamp like	blue-grey					496	124
	Xiajiang	Swamp like							133
	Xilin	Swamp like	grey-black					400	120
	Xinfeng Mountainous	Swamp like	black or brown			yes	long horn	344	116
	Xinlong	Swamp like	grey-black, black			yes		457	124
	Xinyang	Swamp like	yellow-grey or black-grey					490	124
	Yanjin	Swamp like	blue, grey, brown					393	121
	Yibin	Swamp like	grey, white			yes			

HW: height at withers; live weight (kg) and HW (cm) are only given for males. Sources: FAO/DAD-IS and reported in Cockrill (1974)

5.3.3 Diversity of Phenotypes

1   Colour types

Buffaloes phenotypes varies for coat colour, white markings and horn shape. A narrow range of coat colours is observed on water buffaloes compared to cattle. The main colours are black and grey. The swamp buffalo is usually dark grey and it has white or pale socks and one or two light bands (chevrons): one just below the jaw and the other below the neck on the front of the brisket. The pale chevrons are absent from the river buffalo except for the Surti and Bwadari breeds (Table 5.5). In the river buffalo, black is the most common colour but unimproved animals may be grey with white socks like the swamp type. Patches of white hair over the black skin may occur in the forehead, face and in the feather of the tail. Piebald, with extensive white and grey areas, described as tapir pattern type by Macgregor (1939), occurred in animals kept by the Toradjas community of Sulawesi Island. Red-pied animals have been observed by Fischer (1958) in Sumba, and are characterised by patches of skin lacking the normal black-brown pigment and appearing red because of photosensitisation on the unpigmented areas of skin.

Water buffalo can also sometimes be brown but this is uncommon. It occurs in some river breeds like the Surti or Manda Indian breeds. In Cambodia, the swamp buffalo is described as the brown colour type. Red colour has been reported for the Manna breed in Sri Lanka and tan or red-flowered animals in China. The swamp buffalo is sometimes all white but this is rarely reported for the river buffalo. This is not true albinism since pigmentation is still present in the iris, mucosa, horns and hooves. In Thailand, the frequency of white animals varied from 30 % in the north to less than 5 % in the central and southern parts of the country. White buffaloes have been observed in Laos, Malaysia (3 %) and the Philippines (3 %). In China, the frequency of white animals varied among breeds, however, most of the breeds where white animals have been reported are from the southern provinces of China. Also, in Vietnam, white animals are observed in the northern provinces bordering China. Rife (1962) suggested the following genetic hypothesis to explain the known facts of colour inheritance:

Phenotype	*Genotype*
White	bbW-
Black	B-wwR-
Grey	bbww
Greyish white	B-W-
Brown	rr

This assumes that white (W) and grey (w) are two alleles of one locus and black (B) is another locus. White is dominant compared to grey , however, it is partially dominant (or epistatic) to black. However, the white is recessive to black in the river type.

2   White Markings

White markings in Swamp buffalo are found around the eyes and along the side of the nose on the cheeks and mouth. The white markings tend to become more apparent in adults, may fade with advancing age, and are not necessarily visible on new calves.

The white socks and chevrons of the grey animals disappear when crossed with black and are therefore presumably recessive (or hypostatic) to self-colour (or black) (Rife, 1962). White socks and chevrons are visible on buffalo foetus at about the ninth month of gestation (Cockrill, 1974). A white tail is observed in the wild water buffalo (Heinen, 1992),whereas the black River buffalo has a white blaze on the face and a white tail. This pattern is retained in the crosses with the grey Swamp buffalo and is therefore presumably dominant (or epistatic) to self-colour (Rife, 1962).

3   Horns

Swamp buffaloes have horns of the same shape growing out horizontally from the forehead and sweeping backward in a wide semi-circle, but they vary in size. River buffaloes have mainly spirally curled horns. They are very tightly coiled in the Murrah group breed.

5.4 Genetic variation studies of Water Buffalo
5.4.1 An overview

Compared to other domestic species such as cattle, goat or pig;, there are only a few studies on the genetic diversity of buffalo breeds. This is maybe because buffaloes are only reared in developing countries which, until recent decades, did not have sufficient infrastructures to carry out such studies. Moore et al. (1995) tested 80 bovine microsatellites on swamp and river buffaloes. On these 80 markers, 21 were used by Barker et al. (1997) on water buffalo. Navani et al. (2002) tested 108 bovine microsatellites on river buffalo and Kumar et al. (2006) a total of 495 bovine microsatellites where 27 were highly polymorphic. Since 2004, thirty microsatellite markers are recommended by FAO and ISAG for domestic buffalo diversity studies (Hoffmann et al. 2004).

Barker et al. (1997) were the first to analyse genetic diversity of South Asian buffalo breeds using protein coding loci and twenty-one microsatellites (Figure 5.8).

Figure 5.8 Map and dendrogramme of south Asian breeds published by Barker *et al.* (1997).
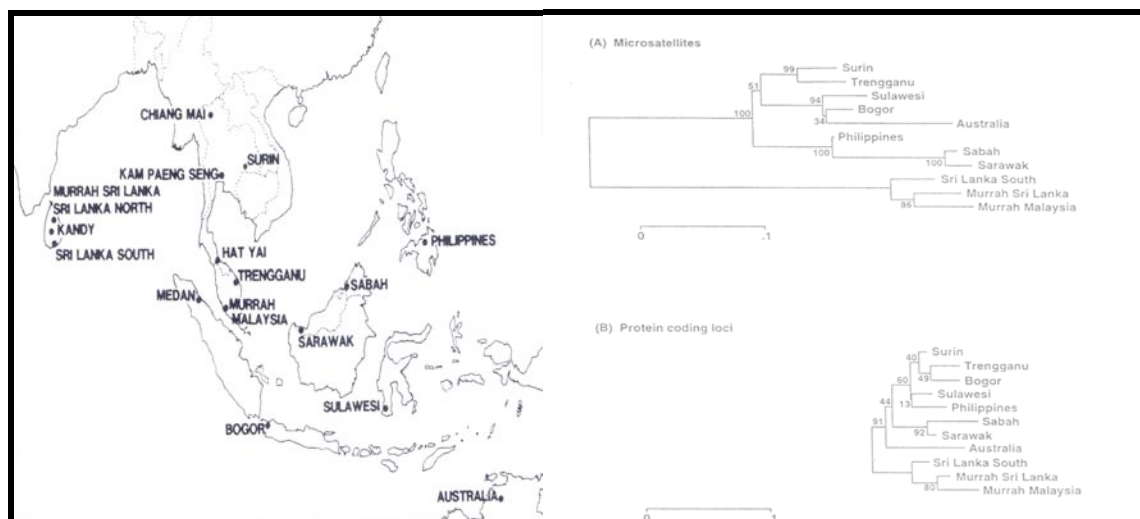
They observed a clear differentiation between the two types, river (Murrah breed) and swamp buffaloes that was confirmed by Zhang *et al.* (2007) when comparing Murrah and Nili-Ravi River breed with Chinese Swamp buffaloes. South Asian swamp buffalo breeds (Barker *et al.,* 1997) which are reared in a traditional way showed high values of expected heterozygosity averaging 0.506 without departure from HWE. Higher expected

159

heterozygosity (0.535) was observed in eighteen indigenous Chinese breeds with 30 microsatellites. Expected heterozygosity of swamp buffalo seems to be lower than for river buffalo from India (0.71-0.78; Kumar *et al.* 2006); Egypt (0,600); Greece (0.594) and Italy (0.577) (Moioli *et al.* 2001). Within Asian swamp buffalo or within Chinese swamp buffalo, effect of isolation by distance was not observed. However, in both cases, genetic distances were found to be correlated with geographical distances. Within South Asian swamp buffaloes, 16.8 % of genetic diversity occurred between breeds (Barker *et al.* 1997) which is a higher differentiation than for North European cattle breeds (10.6 %, Kantanen *et al.* 2000) or in South European cattle breeds (6.1 %, Jordana *et al.* 2003). However, on a single country scale, such as China, only 2.8 % of genetic differentiation was observed between indigenous breeds (Zhang *et al.* 2007). A similar result (3.4 %) was obtained by Kumar *et al.* (2006) on Indian river buffalo. Moioli *et al.* (2001) showed that 3.1 % of genetic differentiation occurred between the Italian and the Greek river buffalo, whereas it reached 7 % when compared to the Egyptian river buffalo (Figure 5.9).

Figure 5.9 Map summarizing expected heterozygosity and F_{ST} (%) of water buffalo.
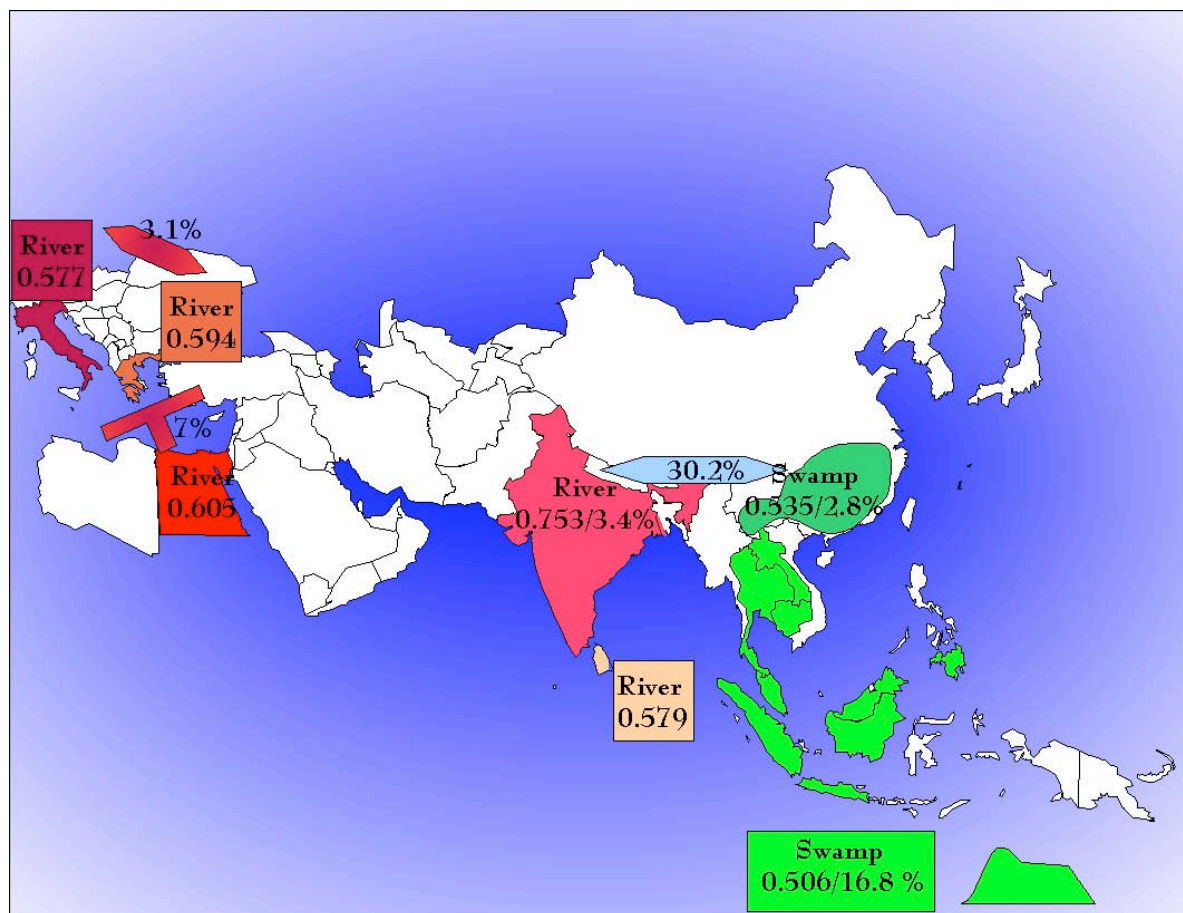
Source: Barker et al. (1997); Moioli et al. (2001); Kumar et al. (2006); Zhang et al. (2007)

Microsatellites are considered to be highly effective in recounting population history or breed relationships. Moreover they are less subjected to bottleneck effect because there is less effect on heterozygosity and on genetic distance estimates. As noted by Barker *et al.* (1997), because many domestic livestock breeds are possibly descended from a small founding population, microsatellite-based trees are better in assessing genetic relationships among breeds. A good example of such phenomenom is that microsatellites were powerful enough to confirm the origin of the Australian breed from Timor-Java despite the severe bottleneck suffered by this population and so proved to be superior to protein-coding distance trees. Moreover, microsatellites seem to be efficient in detecting hybridisation or crossbreeding. Flamand *et al.* (2003) were able to detect, using 10 microsatellites and STRUCTURE genetic software, hybrids from wild buffalo and river buffalo from Nepal. Furthermore, Zhang *et al.* (2006) noted that the Dehong breed, a western Chinese breed, was genetically more different compared to the other indigenous swamp buffalo breeds. In fact, they highlighted that this breed may have river buffalo introgression through Myanmar swamp buffaloes which are known to have been widely crossbred with the river type. According to geographical distribution, it was considered that a total of 24 Chinese buffalo breeds could be classified into four types (Wenping, 1998): seashore type with large body size, plain and lakeside type with middle body size, plateau flatland type with middle body size and hilly types with middle and small body size. However, when Zhang *et al.* (2007) genetically analysed 18 of the indigenous breeds, they first showed that the White Guizhou breed, which only differs from the Guizhou breed by its white coloration, was not genetically different (F_{ST}=0) and therefore should be considered as one unique breed. Secondly they showed that Chinese breeds were clustered into two groups separated by the Yangzte river which acts as a geographical barrier, and each group was subdivided into two subclusters. The total four subclusters obtained were not in accordance with the classification on geographical zones and morphology previously proposed by Wenping, (1998).

Proportions of genome from the Murrah river breed in other breeds revealed by microsatellites (Kumar *et al.* 2006), showed the importance of this breed in improving the other Indian river breeds. Similarly to Chinese breeds, Indian breeds used to be organised into four main groups proposed by Cockrill (1984). However, genetic analysis has illustrated that, within some groups, breeds significantly differed from each other such as the Jafarabadi and Meswana breeds, considered to belong to the same group, and that the Meshwana breed was genetically closer to the Ngapuri group.

5.4.2 The Ha Giang swamp buffalo

In the Ha Giang province, two main crops are grown: paddy rice and maize. While the cattle is used for maize crops, the swamp buffalo is used for paddy crops. As paddy rice is cultivated in the low mountainous slopes and plains, swamp buffalo are therefore reared by almost all ethnic groups as shown in Figure 5.10a.

Figure 5.10 Characterisation of white buffalo. a) Proportion of the Ethnic groups of farmers rearing swamp buffaloes per commune. b) Density map of swamp buffaloes.
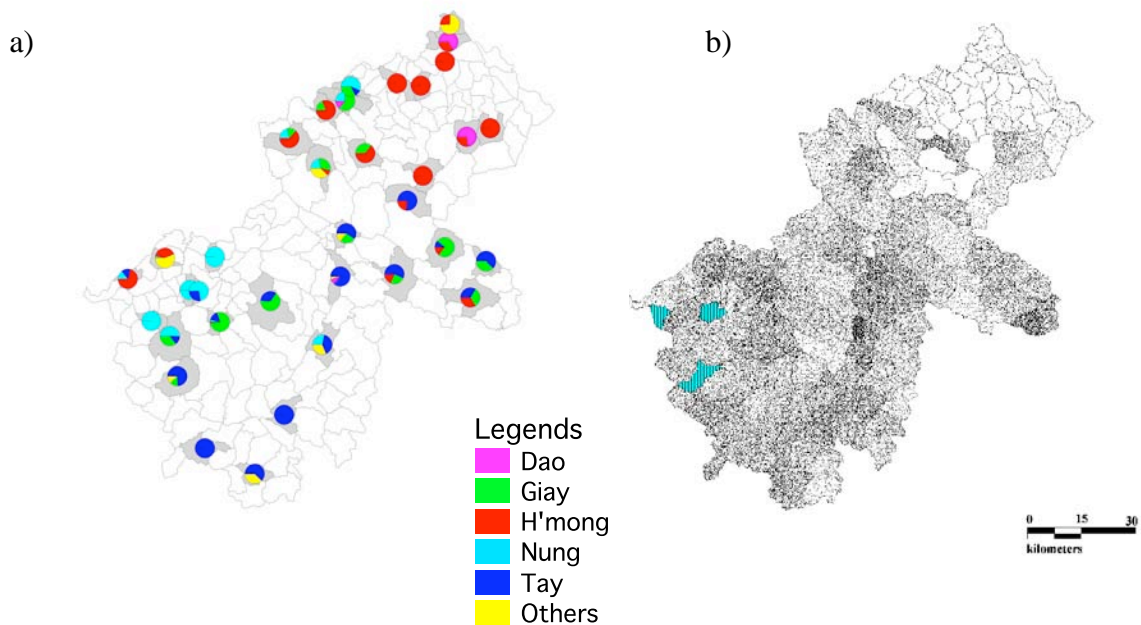
in blue hatched: no data available,b) 1 point represent 5 buffaloes

There is a North-South density gradient of swamp buffaloes in the Ha Giang province (Figure 5.10b). Farmers from mountainous districts (Dong-Van, Meo-Vac and Yen-Minh) in the northeast generally did not rear swamp buffalo, in contrast to their counterparts in the mountainous districts in the southwest (Xin-Man and Haong Su Phi). According to farmer interviews, in the northeast part, there is less than 1 buffalo per farm: 0.3 in average for Dong-Van district and 0.9 for Meo-Vac district. The biggest herd size was about 3.7 buffaloes per farm in Vi-Xuyen district.

1   Characterisation of the swamp buffalo in the Ha Giang province

GENETIC AND MORPHOMETRIC CHARACTERISATION OF A LOCAL VIETNAMESE SWAMP BUFFALO POPULATION

Submitted to Journal of Animal Breeding and Genetics

Abstract

The water buffalo had a key role in the socio-economy of South-East Asia since is the main draught power for paddy rice cultivation. While in Indian continent, the water buffalo is of the riverine type, in South-East Asia the majority of buffaloes are of the swamp type. In a poor remote northern Ha Giang province of Vietnam, improvement of the swamp buffalo breed may be one of the best ways to increase sustainability of farming systems. Therefore, analysis of the genetic structure of the province buffalo population is a prerequisite to any conservation or improvement project. First, 1 122 animals were described for 11 body and horn measurements. From this samples set, 744 animals were genotyped for 17 microsatellite markers. Also 17 animals from southern provinces of Vietnam were genotyped as a comparative sample. The obtained results showed that genetic diversity in the Ha Giang was high although inbreeding value were important. The F_{ST} values within the province and across Vietnam were low indicating that most of the population variation was explained by individual variability. Bayesian clustering approach did not highlight the presence of subdivided populations. The results are useful for the implementation of a conservation and improvement strategy of the swamp buffalo in order to fulfill farmer needs for sustainability of the farming systems.

Introduction

The water domestic buffalo (Bubalus Bubalis) has been classified according to morphological and behavioural criteria into two types (Macgregor 1939): the river buffalo in the Indian subcontinent and West to the Balkan, and the swamp buffalo of Southeast Asia, from Assam and Nepal East to the Yangtze valley of China. In Asia, buffalo have played a pivotal role in overall social development through contributions to milk, meat, hide production and draught power for agriculture. From 1990 to 2002, the Southeast Asian buffalo herd decreased by about 18.5 % (FAOSTAT). Therefore, it is of prime interest to conserve and well manage water buffalo breeds. Asian water swamp buffaloes have already been studied (Barker et al. 1997, Kumar et al. 2006, Zhang et al. 2007), but the Vietnamese swamp buffalo has not.

The Ha Giang province bordering China, is one of the poorest Vietnamese provinces. It is characterised by a temperate climate and mountainous relief. As a consequence, mechanisation or intensification of farming systems might be more difficult than in plain areas. Therefore, genetic management (characterisation, conservation and improvement) of the swamp buffalo might be a sustainable way to improve farming systems and alleviate poverty in such mountainous areas. Characterisation of the buffalo population from the Ha Giang province either genetically and phenotypically will be useful as a preliminary study for establishing rational breeding strategies which could later be applied to the other mountainous provinces of Vietnam.

In such a context, a fine-scale survey for genetic and phenotypic sampling was carried out. A multivariate procedure and a Bayesian-clustering approach allow the characterisation of the genetic structure of the swamp buffalo from the Ha Giang province.

Materials and Methods

Sampling procedure

The remote Ha Giang province (22º08' - 23º19' N ; 104º33'- 105º33' E) is a mountainous area with elevations of 10 to 2400 m above sea level. From an administrative point of view, this province is split into districts (11 in total), communes (193) and villages (1 715).

Twenty-four ethnic groups inhabit the Ha Giang province and the H'mong ethnic group is the main one (36.1 %), followed by the Tay (18.3 %) and the Dao ethnic groups (17.2 %). Sampling sites were distributed over the entire province as shown in Figure 1. Samples (1 122 animals) from 3 to 25 villages per commune (189 in total), and from 2 to 4 communes per district (30 in total) among the 11 districts were described for phenotypic characterisation. Caution was taken to not sample related animals, according to information provided by farmers.

Among these animals sampled in the Ha Giang province, 744 were genotyped for a set of microsatellite markers (see next). In addition, animals were sampled from two other provinces, namely the Nghe An (AN; 5 animals) and the Tay Ninh (TN; 12 animals) provinces (Figure 1). The geographical origin of genotyped animals is shown in Table 1. Each district sample was considered as an *a priori* basic unit for polymorphism analyses.

Figure 1. Geographic distribution of Ha Giang buffalo samples.
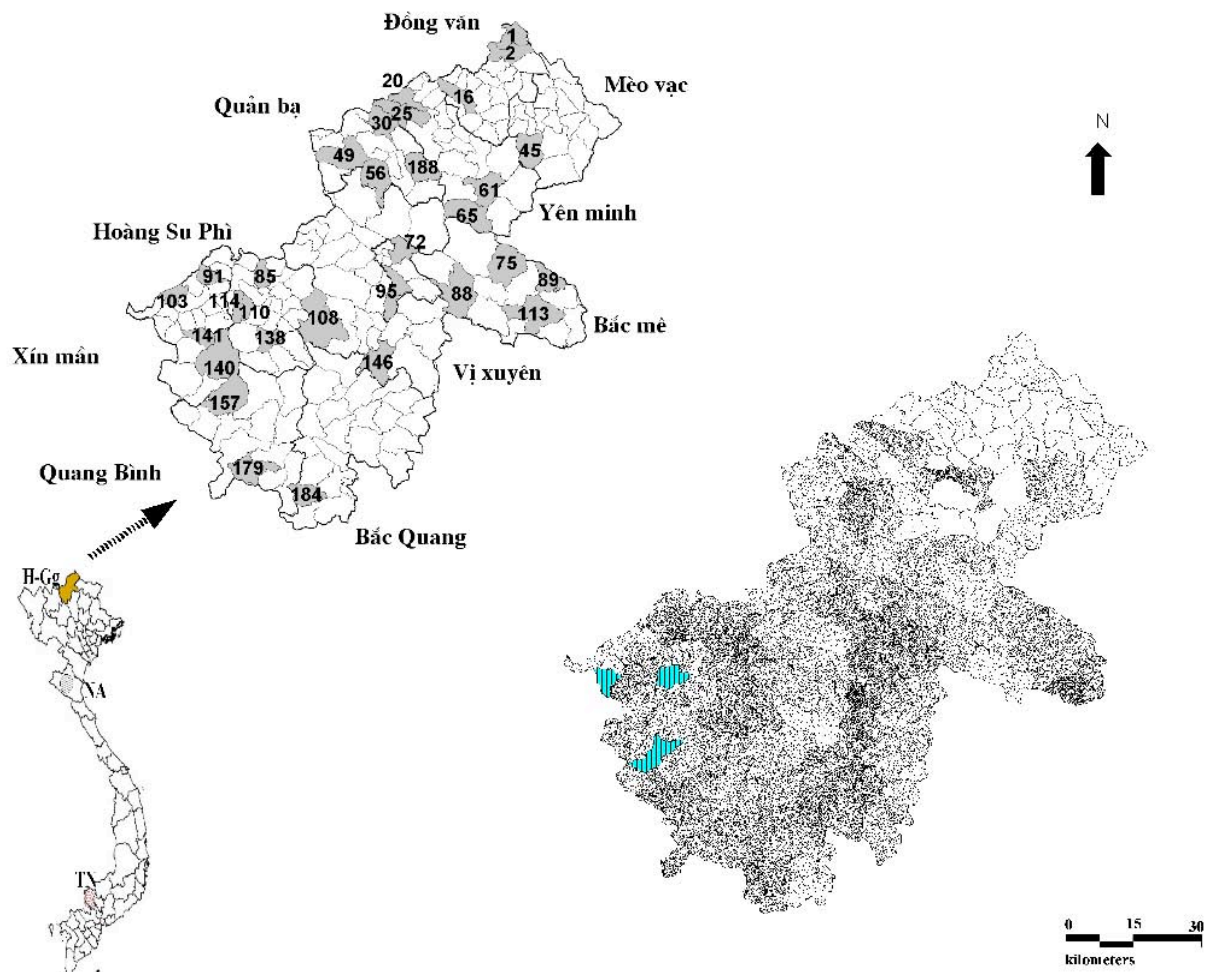
Genomic DNA was extracted from blood and/or tissue samples using the QIAamp Kit from QIAGEN. A set of twenty microsatellite markers was used (Table 1). The PCR products were labelled with fluorescent dyes and genotyped using a capillary sequencer (Beckman & Coulter CQ8000).

Statistical analysis

The presence of null alleles was tested using FreeNA (Chapuis & Estoup, 2007): loci with the estimated frequencies of null alleles $r \geq 0.2$ were considered to be potentially problematic for calculations. Allele frequencies, number of alleles, expected and observed heterozygosity were calculated using GENETIX 4.4 (Belkhir $et\ al.$ 2000). GENEPOP 3.2 (Rousset & Raymond 1995) was used to compute F-$statistics$ (Weir & Cockerman 1984). Test significance was corrected with sequential Bonferroni correction on loci.

Caballero & Toro (2002) proposed a molecular coancestry f_m applying the Malécot definition, but considering genes identical by state instead of by descendant, for subdivided populations. Using f_m, they estimated the total genetic diversity that is made up of three components: the genetic diversity within individuals (GD_{WI}); the genetic diversity between individuals (GD_{BI}) from the same subpopulation, and the genetic diversity between subpopulations (GD_{SB}). Calculations were computed with Molkin 2.0 (Gutiérrez $et\ al.$ 2006).

Pairwise genetic distances (D_R, Reynolds 1983) between district populations were estimated by performing 1000 bootstraps with the POPULATION v.1.2.28 (Olivier Langella; available at http://bioinformatics.org/project/?group_id=84). Dendrograms were constructed from the distance matrix using the neighbour-joining algorithm.

Genetic analysis of the 11 districts was also analysed through a Multivariate Coinertia Analysis (MCOA) in order to estimate genetic structure similarity across markers (Laloë $et\ al.$ 2007, Berthouly $et\ al.$ 2008a). This approach is based on functions available in the ADE4 package (Chessel $et\ al.$ 2004; Dray & Dufour 2007) of the R software (R development core team, 2007)

The genetic structure and individual assignments were investigated using a Bayesian clustering procedure implemented in STRUCTURE (Pritchard $et\ al.$, 2000), with the admixture method and correlated allele frequency version of the programme (Falush $et\ al.$,

2003). This approach has the advantage of inferring multiple ancestral populations. Assuming that the data set could be represented by *K* separate genetic clusters (*K*=1-15), we then did 15 runs for each different value of *K* with 10^5 iterations following a burn-in period of 30 000, without any prior information.

Morphological diversity

Morphological data

All the 1 122 animals were chosen to be older than 2 years old, in order to avoid an age effect. The age of each animal was estimated by the farmer For each animal, six body measurements were taken with a cloth tape: the thorax depth; the height at wither (*HW*); the body length (*BL*); the heart girth (*HG*); the neck girth (*NG*) and the ear length (*EL*). These measurements were combined into 4 indices: the length index I_{BL}= *BL/HW*; the heart girth index I_{HG} = HG/HW; the neck girth index I_{NG}= *NG/HW* and the Ear index I_{EL} = *EL/HW* plus the index of slenderness I_{GS} as described by Lauvergne & Souvenir Zafindrajaona (1992). Six horn measurements were done: length between the two tips (*TT*); external length of the horn (*ExL*); internal length of the horn (*IL*); chord distance between the tip and the base of the horn (*TB*); horn base girth (*BG*). These measurements were combined into three ratios: *TT/ExL*; *IL/ExL*; *CO/IL*. Thus, 11 measurements were analysed: *HW*, the five body indices, the 3 horn indices plus *TT* and *BG* values (Tables 3 and 4).

Statistical analysis

Morphological data were first analysed through MANOVA. The MANOVA will simultaneously take into account all the measurements to highlight morphometric differentiation between districts. Secondly, in order to identify precisely which trait allows morphometric differentiation, single trait analysis by linear hierarchical mixed models was performed with the SAS software. For a given trait, the model was the following:

$$y_{ijkl} = \mu + b \times alt_{ijkl} + s_i + d_j + c_{ijk} + e_{ijkl}$$

where *y* is the observation, μ the overall mean, *alt* the altitude where the animal is raised and *b* the coefficient of regression of the performance over altitude, s_i the fixed effect of sex *i* (*i* = 1 or 2), d_j the fixed effect of district *j*, c_{ijk} the random effect of the village nested into district and e_{ijkl} a random error. Normality and homoskedasticity of model residuals were tested. The *P*-values of pairwise mean adjusted comparisons were corrected with Bonferroni correction.

A linear discriminant analysis was used on the whole set of markers and measurements. Since this set contains a mix of quantitative and qualitative variables, we used an approach similar to that of Hill and Smith (1976). It is a combination of an internal correspondence analysis on markers (Cazes *et al.* 1998; Laloë *et al.* 2002) and of principal component analysis on quantitative measures. This approach is based on functions available in the ADE4 package (Chessel *et al.* 2004; Dray & Dufour 2007) of the R software (R development core team, 2007).

Results

Marker analysis

Three loci, namely CSSM46, ILSTS052 and ILSTS068, had *r* values below 0.2 so it was assumed that null alleles could be present for these loci. As a consequence, these loci were removed for the genetic analysis. Among the 17 loci remaining, a total of 97 alleles were detected. The number of alleles per locus ranged from 2 (ILSTS025) to 12 (CMSS019), with a mean number of 5.7 per locus. The observed heterozygosity H_{Obs} per locus averaged 0.560 and ranged from 0.282 for locus ILSTS005 to 0.826 for locus ILSTS056 (Table 1). Two loci that did not seem to have null alleles were found to deviate from HWE mainly for all districts: CMSS057 (10 districts) and CMSS022 (8 districts).

 Molecular diversity within and between district populations

Forty-eight karyotypes were done in order to ensure that buffalo from Ha Giang were only of swamp type (2n= 48) compared to crossbred (2n=49) with river type (2n=50) (data not published).

The NA province with only 5 animals sampled showed 3.4 alleles on average per loci and TN province (12 animals) had 4.3 alleles. Accross the 11 Ha Giang districts, the number of alleles per locus ranged from 4.9 to 7.7 (Table 2). The mean number of effective alleles within the whole Ha Giang province was 1.94. The observed and expected heterozygosity across all populations ranged from 0.531 to 0.607 and from 0.637 to 0.677 respectively. The highest F_{IS} value was obtained in the TN province (0.190). In the Ha Giang province, average F_{IS} was 0.141 varying from 0.100 (BQ) to 0.175 (QBn).

Table 1. Summary of genetic diversity measures.

Province	District	Nc	Nv	Ni	H_{Exp}	H_{Obs}	A	Ae	F_{IS}	D_{HWE}	F_{ST}	fm
Ha Giang												
	Bac-Quang (BQ)	1	4	17	0.673±0.15	0.608±0.186	5	2.1	0.1	0	0.024	0.277
	Quan-Binh (Qbn)	2	14	87	0.677±0.154	0.559±0.158	5.7	2.1	0.175	7	0.031	0.275
	Vi-Xuyen (VX)	2	15	104	0.655±0.166	0.57±0.157	5.7	1.9	0.132	4	0.019	0.284
	Bac-Me (BM)	4	26	114	0.637±0.167	0.556±0.154	5.7	1.8	0.13	6	0.034	0.292
	Hoang Su Phi (HSP)	4	22	105	0.664±0.161	0.552±0.165	5.7	2	0.164	8	0.034	0.278
	Xin-Man (XM)	4	23	92	0.672±0.153	0.563±0.163	5.7	2.1	0.161	6	0.038	0.276
	Quan-Ba (QB)	4	25	76	0.65±0.174	0.548±0.174	5.6	1.9	0.158	6	0.025	0.288
	Yen-Minh (YM)	3	19	73	0.659±0.157	0.569±0.173	5.5	1.9	0.137	5	0.029	0.281
	Dong-Van (DV)	3	5	17	0.661±0.203	0.583±0.228	4.8	2	0.127	5	0.039	0.273
	Meo-Vac (MV)	1	6	26	0.638±0.158	0.531±0.198	4.9	1.8	0.17	2	0.032	0.286
	Ha-Giang prefecture (HG)	2	8	33	0.672±0.149	0.601±0.143	5.4	2	0.102	1	0.009	0.280
Nghe An	NA			5	0.637±0.172	0.57±0.36	3.4	1.8	0.104	0	0.008	
Tay Ninh	TN			12	0.65±0.22	0.543±0.268	4.3	1.9	0.19	2	0.008	

Nc: number of communes; Nv: number of villages; Ni: number of animals sampled; A: number of alleles; Ae: number of effective alleles; D_{HWE}: number of loci in heterozygote deficiency

Table 2. Genetic polymorphism measures of microsatellite markers.

Loci	No A	Range	H_{Exp}	H_{Obs}	F_{IS}	D_{HWE}	F_{ST}
BMC1013	7	234-250	0.641±0.071	0.465±0.156	0.211	3	0.015
CSRM060	3	91-95	0.469±0.044	0.38±0.123	0.099	4	0.008
CSSM008	7	178-194	0.705±0.047	0.596±0.09	0.178	1	-0.002
CSSM019	12	129-160	0.785±0.043	0.723±0.089	0.282	8	0.004
CSSM022	3	206-210	0.474±0.096	0.365±0.202	0.161	4	0.001
CSSM029	6	176-194	0.567±0.062	0.525±0.106	0.017	0	0.007
CSSM033	5	154-174	0.722±0.035	0.628±0.079	0.095	2	0.008
CSSM041	4	129-141	0.595±0.058	0.568±0.109	0.116	1	0.027
CSSM043	5	239-251	0.721±0.04	0.452±0.154	0.338	7	0.004
CSSM047	5	125-148	0.763±0.018	0.648±0.106	0.175	3	0.002
CSSM057	7	111-129	0.799±0.064	0.480±0.141	0.356	10	0.007
ETH003	5	94-106	0.770±0.045	0.75±0.13	0.043	1	0.003
ILSTS005	5	178-186	0.285±0.138	0.275±0.142	0.086	1	0.010
ILSTS025	2	112-114	0.474±0.063	0.448±0.149	0.060	0	0.023
ILSTS028	5	158-166	0.729±0.049	0.622±0.145	0.144	2	0.007
ILSTS056	7	152-164	0.823±0.035	0.828±0.078	0.043	1	0.004
ILSTS061	9	131-157	0.86±0.031	0.774±0.094	0.099	2	0.006

No A: number of alleles, D_{HWE}: number of loci in heterozygote deficiency

The number of loci per population not at HWE ranged from 0 (HGp and NA) to 8 (HSP) with an average of 4 loci per district population. According to multiloci F_{ST}, only 0.57 % of the total genetic variability was explained by between district variability. The multiloci country F_{ST} only reached 0.74 %, also taking populations from the two southern provinces of Vietnam into account. The F_{ST} values of the BQ district were equal to 0 or below 0.001 except with the four eastern districts (BM, YM, DV and MV). The average molecular

coancestry f_m ranged from 0.273 in the DV district to 0.292 for the BM district (Table 2). The genetic diversity within individual (GD_{WI}) averaged 0.300 and the between individual diversity (GD_{BI}) was 0.370.

Low values of bootstraps where observed for the D_R distance tree. Only the DV and MV districts clustered with high bootstrap values (74 %). Afterwards, the districts were clustering following geographic distribution (Figure 2a). A rooted tree with the two southern provinces sampled (NA and TN) showed that these two provinces clustered together with high bootstrap values (82%) and all samples from the Ha Giang province formed one separate group where districts were added following the East-West line as for a previous tree (Figure 2b).

Figure 2. Neighbour-joining dendrograms of buffalo population using D_R genetic distance: a) within Ha Giang province; b)within Vietnam.
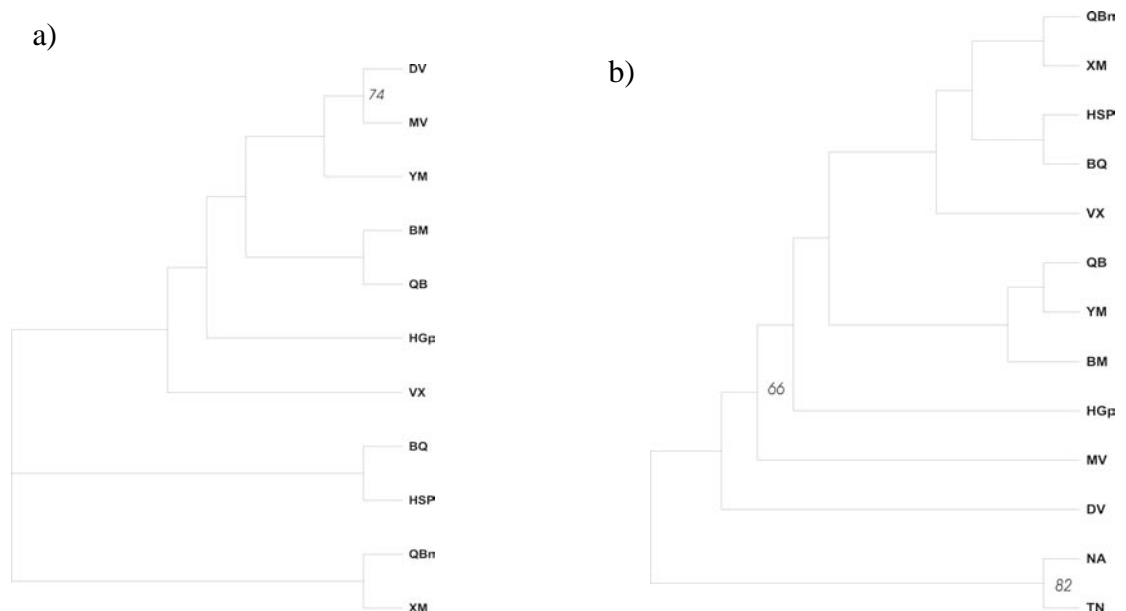
Common congruence values (Cos^2) ranged from 0.027 to 0.936 for the 1st axis. Only 3 loci had low Cos^2 values: CSRM60, CSSM022 and ILSTS025 (Figure 3). Thus, 14 loci had similar genetic structure for the Ha Giang buffalo population. The *TV* values were low and ranged from 0 to 0.014 for the 1st axis and from 0 to 0.004 for the 2nd axis. The synthetic position of the district populations showed that DV, MV and BQ districts separate well from other districts whereas other districts were overlapping (Figure 4a). In order to zoom in, we performed the same analysis without DV, MV and BQ district populations. This new analysis showed higher Cos^2 values (0.133-0.936 for 1st axis; 0-0.822 for 2nd axis) and *TV* values (0-0.016 for 1st axis, 0-005 for 2nd axis). The CMSS041 locus had the highest *TV* value and

contributed the most in the construction of the 1st axis. This was due to the presence of a rare allele from locus CMSS041 that separates district populations according to an East-West cline. The 2nd axis allowed the differentiation of the HGp district (Figure 4b). The same analysis without CMSS041 (data not show), was performed in order to see if the East-West cline was only due to this specific locus. In fact, a similar synthetic position was obtained when the CMSS041 locus was removed, and the East-West cline distribution remained.

A Bayesian approach of individual admixture probabilities using STRUCTURE did not show relevant genetic structure within the Ha Giang province (data not show)

Figure 3. Distribution across loci of values for the components to 1st axis, in percentage: a) Variance; b) Cos^2; c) *TV*.
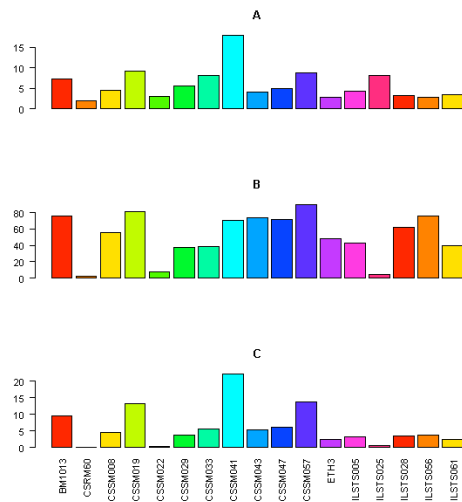
Figure 4. Synthetic position obtained by MCOA analysis for the 11 districts (a); and after DV and MV districts removed (b).
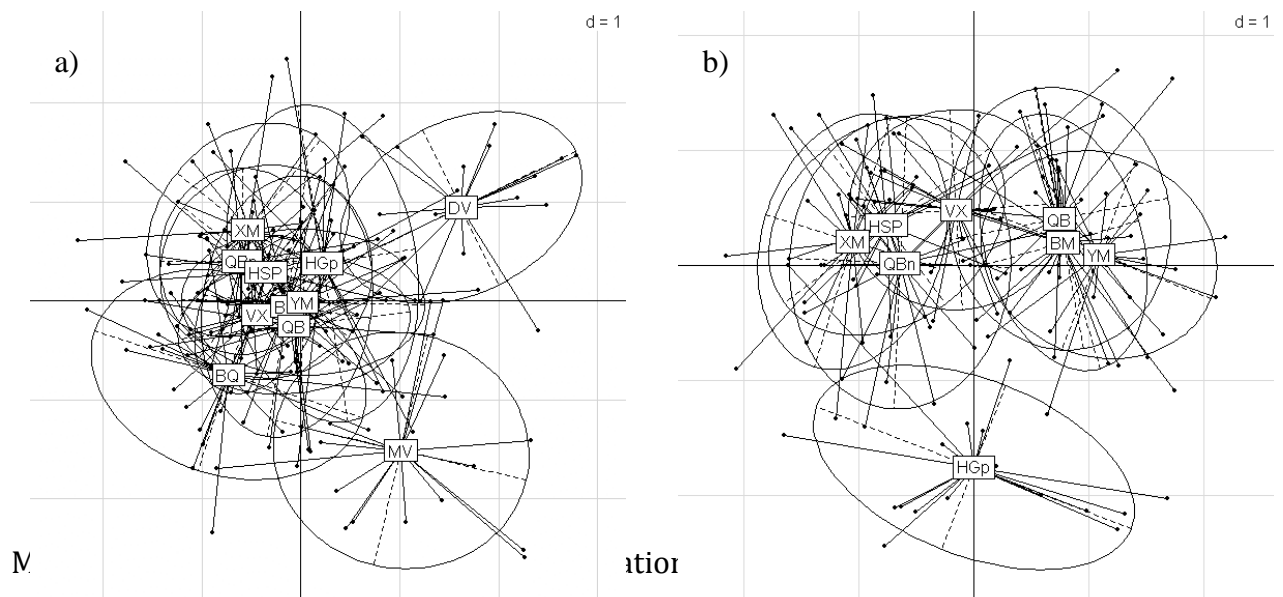
Mean body and horn measurements are summarised in Tables 3 and 4. Among the 55 MANOVA pairwise comparisons, only 4 were not significantly different: DV-MV, DV-YM, DV-QB, MV-YM. Therefore, buffaloes from DV and MV are not significantly different from buffaloes from YM and QB districts.

For body measurements, sexual dimorphism was only observed for *HW* with an average of 120.3 cm for bulls and 119.4 cm for cows, and for I_{BL} with an average of 0.92 for bulls and 0.94 for cows (Table 3). Significant differences between bulls and cows were observed for all horn measurements (Table 4). The buffaloes from the districts of YM and MV had significantly lower values of *HW* and I_{BL} than the western ones. The QB district had significant lower values of *TT* compared to VX and BQ districts and also for *SS/Exl* compared to BQ, VX, XM and HSP districts.

Due to their similarities (see above), the DV, MV and BQdistricts were removed from MCOA, in order to zoom in on other districts differentiation. Therefore, we removed these three districts from the Hill & Smith analysis that combined genotypes and measurements. The first two axes of Hill & Smith analysisexplained 25.7 % and 19.2 % of the observed inertia respectively. The two markers CSSM041 and CMSS019 are the ones that contributed the most to the construction of the 1st axis for 10.8 % and 10.3 % whereas I_{GS}, in the 6th position, was the measurement contributing the most and it only reached 6.3 % in the construction of the 1st axis. For the 2nd axis, CSSM019 (11.3 %) and CSSM33 (8.9 %) are the most important followed by the height at withers *HW* with (8.7 %). Figure 5 shows that district populations widely overlapped. However, it was possible to recognise geographic and ecotype groups: the two north-mountainous districts (QB-YM); the north-plain districts (HGp-BM); the south-mountainous districts (HSP-XM) and the central-plain districts (QBn-VX).

Table 3 Summary of body traits per district and sex (mean±sd)

Sex	Districts	Body traits					
		HW	*I$_{GS}$*	*I$_{HG}$*	*I$_{NG}$*	*I$_{BL}$*	*I$_{EL}$*
Male		**120.3±4.4**	**0.73±0.03**	**1.41±0.03**	**0.19±0.00**	**0.92±0.03**	**5.43±0.14**
	BQ	123.0±7.8	0.71±0.08	1.40±0.09	0.19±0.02	0.88±0.07	5.42±0.50
	HSP	119.5±5.6	0.73±0.09	1.38±0.07	0.19±0.02	0.89±0.05	5.38±0.48
	XM	123.0±9.4	0.68±0.11	1.40±0.07	0.18±0.01	0.93±0.06	5.44±0.43
	Qbn	123.1±6.4	0.71±0.08	1.40±0.06	0.18±0.01	0.91±0.06	5.54±0.36
	VX	123.4±6.1	0.71±0.08	1.41±0.13	0.18±0.01	0.95±0.07	5.56±0.45
	HG	118.4±7.4	0.76±0.08	1.40±0.06	0.19±0.02	0.92±0.06	5.29±0.49
	BM	121.1±6.1	0.68±0.07	1.39±0.12	0.18±0.02	0.92±0.08	5.57±0.57
	QB	118.1±5.4	0.75±0.10	1.44±0.08	0.18±0.02	0.94±0.07	5.58±0.64
	YM	116.1±8.4	0.76±0.11	1.38±0.13	0.19±0.02	0.96±0.1	5.23±0.69
	DV	117.6±8.5	0.79±0.08	1.46±0.09	0.19±0.02	0.94±0.06	5.26±0.43
	MV	108.5±3.1	0.76±0.07	1.43±0.14	0.18±0.02	1.00±0.12	5.60±0.49
Female		**119.4±3.1**	**0.72±0.03**	**1.41±0.01**	**0.19±0.00**	**0.93±0.02**	**5.43±0.10**
	BQ	121.4±6.1	0.73±0.08	1.45±0.08	0.18±0.01	0.92±0.07	5.49±0.38
	HSP	119.9±6.5	0.71±0.09	1.41±0.07	0.19±0.01	0.91±0.07	5.35±0.43
	XM	123.5±7.4	0.67±0.08	1.42±0.07	0.18±0.01	0.93±0.07	5.46±0.42
	Qbn	120.8±5.1	0.73±0.08	1.41±0.06	0.19±0.01	0.94±0.06	5.42±0.37
	VX	121.8±6.0	0.68±0.07	1.43±0.10	0.18±0.01	0.94±0.07	5.54±0.36
	HG	118.9±6.5	0.73±0.09	1.43±0.09	0.19±0.02	0.93±0.06	5.36±0.46
	BM	120.6±5.6	0.69±0.08	1.43±0.08	0.18±0.02	0.94±0.08	5.59±0.5
	QB	117.6±6.6	0.73±0.09	1.45±0.10	0.19±0.02	0.95±0.07	5.46±0.57
	YM	113.9±6.8	0.74±0.09	1.44±0.09	0.19±0.02	0.97±0.08	5.25±0.54
	DV	120.8±7.0	0.73±0.11	1.43±0.14	0.19±0.03	0.92±0.05	5.35±0.81
	MV	114.2±5.9	0.77±0.09	1.46±0.08	0.19±0.02	1.00±0.09	5.31±0.53

Discussion

Qualification of microsatellite marker set

On the twenty microsatellites used, three showed null alleles and one of these three markers, namely CMSS046, is proposed in the FAO guidelines (Hoffmann et al. 2004). Most of the microsatellite loci were highly polymorphic except for ILSTS025 that only had two alleles. Two microsatellites, CMSS057 and CMSS019 were in heterozygote deficit nearly for all district populations. Such observations could traduce the presence of null alleles, however, we had previously already tested this assumption. Another hypothesis could be a hitchhiking effect for these markers, which could be linked to genes under selection and so question the a priori neutrality of these markers. Multivariate analysis as MCOA allowed identifying characteristics and differences between loci. It may detect the efficiency of each locus for population consensual structuring (Laloë et al. 2007; Berthouly et al. 2008a). Our results exhibited high values of Cos^2 meaning a good congruence of each locus district

structuring and the synthetic (i.e. multiloci) one. When removing the CMSS041 locus because of its rare allele, all remaining loci had more homogenous TV values, indicating that

Table 4. Summary of horn measurements per district and sex (mean±sd)

Sex	Districts	Horn measurements				
		TT	TT/ExL	IL/ExL	CO/IL	BG
Male	**Mean districts**	**49.0± 5.0**	**1.2±0.1**	**0.78±0.02**	**0.89±0.03**	**29.2±1.5**
	BQ	53.2±10.3	1.46±0.32	0.82±0.06	0.93±0.08	27.4±4.8
	HSP	45.2±7.3	1.18±0.31	0.78±0.07	0.88±0.06	28.5±3.8
	XM	49.0±10.3	1.19±0.42	0.78±0.05	0.88±0.07	29.0±4.7
	Qbn	52.6±7.6	1.31±0.29	0.77±0.06	0.91±0.07	30.9±4.3
	VX	54.3±11.1	1.26±0.39	0.8±0.09	0.9±0.09	30.9±3.8
	HG	49.4±8.0	1.16±0.27	0.8±0.1	0.87±0.09	29.0±4.6
	BM	47.8±7.7	1.08±0.28	0.76±0.06	0.88±0.08	30.1±3.9
	QB	42.6±10.6	0.98±0.31	0.75±0.07	0.88±0.1	29.7±3.4
	YM	50.0±12.2	1.22±0.4	0.79±0.14	0.9±0.11	28.0±5.0
	DV	45.7±8.2	1.27±0.33	0.81±0.07	0.9±0.04	28.8±5.5
	MV	37.5±10.9	0.96±0.34	0.78±0.08	0.82±0.03	26.0±2.5
Female	**Mean districts**	**41.4±3.5**	**0.9±0.1**	**0.77±0.02**	**0.86±0.01**	**23.2±0.4**
	BQ	47.4±12.4	1.10±0.32	0.79±0.06	0.88±0.06	23.3±2.1
	HSP	40.1±11.2	0.95±0.31	0.77±0.05	0.86±0.08	23.0±2.1
	XM	41.1±9.9	0.92±0.25	0.75±0.05	0.85±0.06	23.5±2.5
	Qbn	42.1±10.7	0.97±0.30	0.8±0.09	0.86±0.08	23.0±1.9
	VX	45.6±13.0	1.03±0.34	0.78±0.09	0.87±0.10	24.0±3.0
	HG	43.1±11.7	0.99±0.30	0.78±0.07	0.86±0.07	23.1±2.9
	BM	42.0±11.3	0.94±0.31	0.76±0.07	0.84±0.09	23.7±4.3
	QB	38.0±12.7	0.84±0.31	0.75±0.05	0.85±0.07	23.0±2.5
	YM	40.6±14.2	0.92±0.37	0.77±0.06	0.87±0.11	22.7±3.0
	DV	34.3±13.7	0.74±0.29	0.72±0.07	0.84±0.11	22.8±1.7
	MV	42.3±17.3	0.90±0.34	0.78±0.08	0.88±0.12	23.2±2.4

Figure 5. District positions obtained with the Hill & Smith analysis.
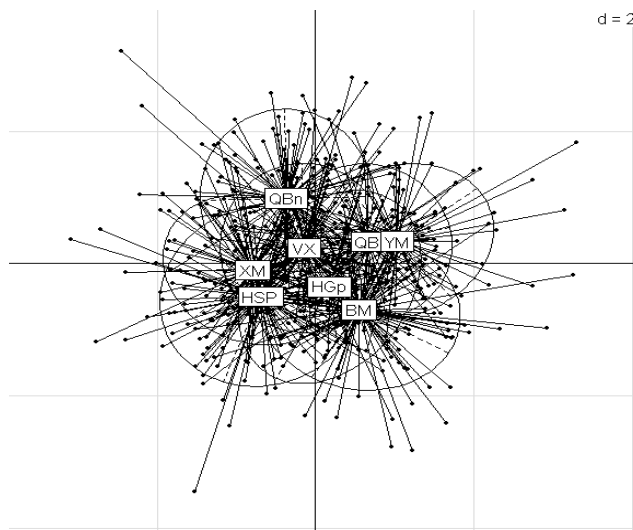
they have similar discriminant power. Therefore, our microsatellite set shows good congruence, and so might be powerful to highlight genetic structure of swamp buffalo populations.

Diversity and Management of swamp buffaloes

Barker et al. (1997) studied South Asian swamp buffalo breeds that were reared in an extensive traditional way. They estimated a gene diversity averaging 0.506. Higher gene diversity (0.535) was observed in 18 indigenous Chinese breeds studied by Zhang et al. (2007). These values are lower than the one observed in the Ha Giang swamp buffalo population (0.657). However, gene diversity of Vietnamese swamp buffalo was lower than for the river buffalo from India (0.71-0.78; Kumar et al. 2006). If we only considered the 11 loci in common between our study and the one from Barker et al. (1997) and excluding the river breeds, only the Surin population had higher expected heterozygosity and number of alleles than the Vietnamese buffalo. However, the observed heterozygosity of the Vietnamese population (0.518) is slightly lower compared to South Asian swamp buffalo populations ranging from 0.387 (Australian originating from Timor-Java) to 0.613 (Surin).

Swamp buffalo plays a key role in Asia and Southeast Asia. It provides draught power for paddy rice crops in plain areas. Swamp buffalo was so valuable that until 1960 slaughtering before 12 years old was forbidden in Vietnam. Every family has at least one buffalo and so, the swamp buffalos distribution is continuous across the plain area; they are reared mostly for draught power without strict and systematic selection. Recent studies showed that in traditional unselected populations, there is a strong relationship between geographic distances and genetic distances (Mwacharo et al. 2007, Zhang et al. 2007) and so only the largest geographical distances or geographical barriers might lead to population differentiation. Swamp buffalo populations distributed across China, only showed 2.8 % of genetic differentiation between populations (Zhang et al., 2007). The Ha Giang, province is a limited area with mountainous districts and plain districts. The Ha Giang cattle population is mainly reared by one ethnic group, namely the H'mong group, and in the mountains, and so showed a disrupted distribution. Thus, Berthouly et al. (2008b) were able to reveal two genetic subpopulations within the Ha Giang cattle population. On the contrary, buffaloes are reared by almost all ethnic groups and distribution in the province is continuous enabling

frequent genetic exchanges. As a consequence, genetic differentiation into subpopulations did not occur and districts clustered following the East-West cline distribution.

However, high inbreeding values were observed in the Ha Giang buffalo population. Farmers often had two to three generations overlapping on the farm, and do not pay attention to selecting unrelated animals for reproduction (Berthouly et al., 2008c). Therefore, conservation or improvement projects should focus on decreasing inbreeding rates through management of reproduction. Such objectives could be easily implemented with a "clock-reproduction scheme" at the commune or district level or even at the province level since there was no genetic subpopulations.

Acknowledgments

References

Barker J.S.F., Moore S.S., Hetzel D.J.S., Evans D., Tan S.G., Byrne K. (1997) Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and comparison with protein coding loci. *Anim. Genet.*, **28,** 103-115

Belkhir K., Borsa P., Goudet J., Chikhi L., Bonhomme F. (2000). GENETIX, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions: CNRS.UPR. 9060. Université Montpellier II, Montpellier, France.

Berthouly C., Bed'Hom B., Tixier-Boichard M., Chen C.F., Lee Y.P., Laloë D., Legros H., Verrier E., Rognon X. (2008a) Using molecular and multivariate methods to study the genetic diversity on local European and Asian chickens breeds. *Anim. Genet.*, doi:10.1111/j.1365-2052.2008.01703.X

Berthouly C., Doan Phan L., Rognon X., Nhu Van T., Leroy G., Hoang Than H., Bed'Hom B., Laloë D., Vu Chi C., Nguyen Dang V., Verrier E., Maillard J.C. (2008b) Revealing subpopulations and possible admixture in the Vietnamese H'mong cattle for conservation purposes. *J. Hered*, submitted

Berthouly C., Rognon X., Nhu Van T., Guillouzo A., Hoang Thanh H., Vu Chi C., Verrier E., Maillard J.C. (2008) Typological analysis of farming systems in regards to ethnic communities and their geographical location. *Agri. Syst.,* submitted

Caballero A., Toro M.A. (2002) Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Gen.,* **3,** 289-299

Cazes P., Chessel D., Doledec S. (1988) L'analyse des correspondences internes d'un tableau partitionné. Son usage en hydrobiology. *Rev. Statistique Appliquée*, **36,** 39-54

Chapuis M.P., Estoup A. (2007) Microsatellites null alleles and estimation of population differentiation. *Mol. Biol. Evol.*, **24,**621-631

Chessel D., Dufour A.B., Thioulouse J. (2004) The ade4 package-I :One –table methods, R-News 4 (2005) 5-10.

Dray S., Dufour A.B. (2007). The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22(4) http://www.jstatsoft.org/v22/i04/

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

FAOSTAT available at http://faostat.fao.org

Gutiérrez J.P., Royo L.J., Alvarez I., Goyache F. (2005) Molkin V2.0 , A computer program for genetic analyses of population using molecular coancestry information. *J. Hered*., **96,** 718-721

Hoffmann I., Marsan P.A., Barker J.S.F., Cothran E.G., Hanotte O., Lenstra J.A., Milan D., Weigend S., Simainer H. (2004) New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a joint ISAG/FAO working group. In*: Proceedings of the 29th International Conference on Animal Genetics*, Tokyo, Japan

Kumar S., Gupta J., Kumar N., Dikshit K., Navami N., Jain P., Nagarajan M. (2006) Genetic variation and relationships among eight riverine buffalo breeds. *Mol. Ecol.*, **15,** 593-600

Laloë D., Moazami-Goudarzi K., Chessel D. (2002) Contribution of individual markers to the analysis of the relationships among breeds by correspondence analysis. 7th World Congress on Genetics Applied to Livestock Production, August 19-23, 2002.Communication 26-06. Vol 33, p 493-496. Montpellier, France

Laloë D., Jombart T., Dufour A.B., Moazami-Goudarzi K. (2007) Consensus genetic structuring and typological value of markers using Multiple Co-inertia Analysis. *Genet. Sel. Evol.,* **39,** 5545-569

Lauvergne J., Souvenir-Zafindrajaona P. (1992). Comparaison de deux populations de Zébu Malgache par les distances génétiques biométriques. *Rev. Elev. Med, Vet.,* **45,** 167-174

Macgregor R. (1939). The domestic buffalo. Thesis presented to the Royal College of Veterinary Surgeons (quoted by I.L. Mason in Cockrill (1974), 1-47

Mwacharo J.M., Nomura K., Hanada H., Jianli H., Hanotte O., Amano T. (2007). Genetic relationships among Kenyan and other East African indigenous chickens. *Anim. Genet.*, **38,** 145-150

2   Complementary Results

a) Karyotypes

All karyotypes showed $2n=48$ chromosomes as for the Swamp buffalo (Figure 5.11), then no F1 or F2 hybrids ($2n=49$) were found in the Ha Giang province.

Figure 5.11 Karyotype of one swamp buffalo from the Ha Giang province.
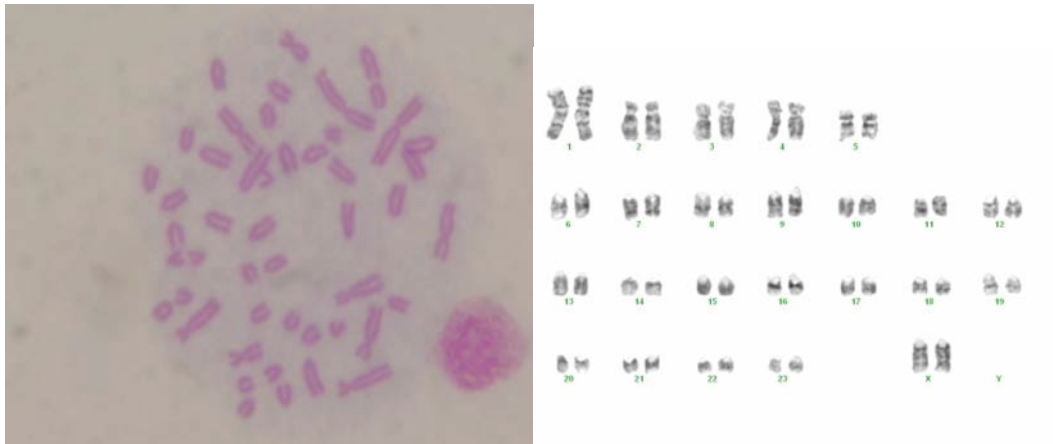
b) Coat colour and white markings

In the Ha Giang province, the proportion of white buffalo per commune ranged from 0 to 35 % according to the commune. White buffaloes are mainly observed in the districts of QB and YM. However, when performing a PCA analysis on body measurements (Figure 5.12b) and a FCA on genotypes (Figure not shown), the white buffaloes did not cluster separately from the others.

Figure 5.12. Characterisation of white buffalo. a) Proportion of white buffaloes (in pink) sampled per commune. b) Position of white buffaloes (in red) when performing discriminant analysis on body traits.

a)

b)
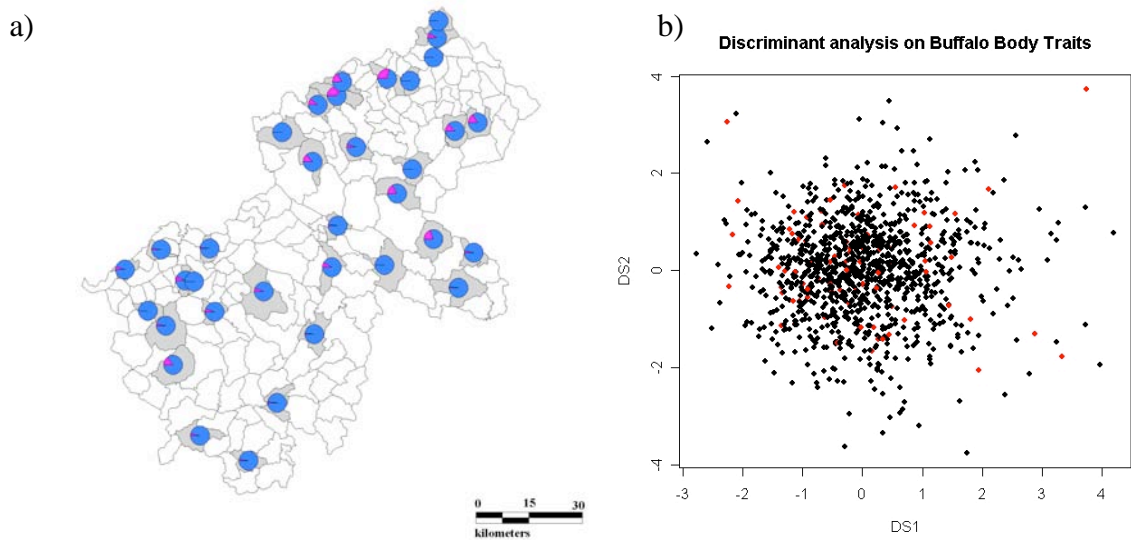
The grey coat colour is characterised by white markings on the face, the neck and white socks on the legs. The first two axes of the FDA on white markings explained 77 % of the inertia. The marking contributing the most in the construction of the 1st axis was the number of chevrons on the neck. For the 2nd axis, the most important marking was the presence of white markings on the cheeks (Table 5.6). No clear separations of districts based on the white markings was possible as shown in Figure 4.

Table 5.6 Scores of traits for the two discriminant axes.

Traits	CS1	CS2
Chevrons	-0.867	0.196
Eyes	-0.147	0.254
Cheeks	-0.2	-0.867
Socks	0.851	0.218

Figure 5.13 Districts representation obtained with FDA on white marks.
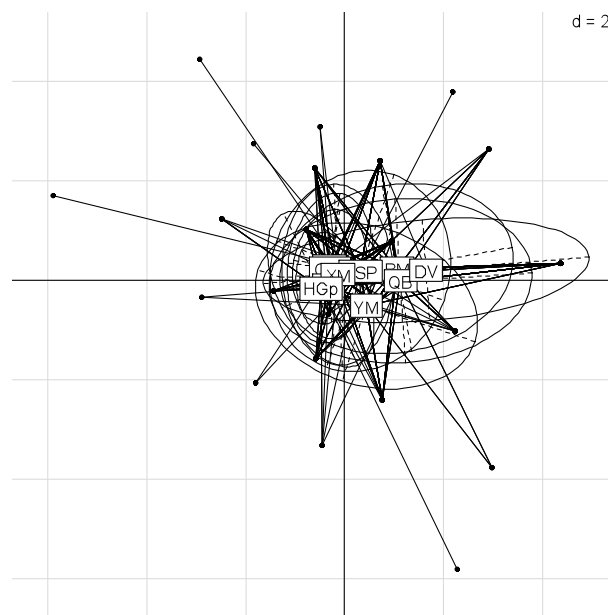
c) Live weights of swamp buffaloes

Live weight were estimated by linear regression (see p 48). In the Ha Giang province, buffalo bull weight averaged 363 kg and 342 kg for cows. In average, the biggest buffalo bulls were found in two western districts of VX and XM (411 kg, 404 kg). The biggest buffalo cows were also found in the western part of the province : QB (365 kg) and XM (404 kg) as showed in Table 5.7. Weight estimates of buffalo in the Ha Giang province were lower than in the other region of Vietnam (428-494 for bulls, 358-458 for cows) (Thu 1987; Hoang 1978; Thac 1979).

Table 5.7 Estimated live weight per district (average ± sd) of swamp buffaloes from the Ha Giang province.

Districts	Males	Females
BQ	385 ± 97	365 ± 53
HGp	341 ± 79	338 ± 46
QBn	382 ± 68	342 ± 69
VX	411± 70	357 ± 42
BM	371 ± 61	350± 71
HSP	338 ± 72	332± 58
XM	404 ± 85	366± 47
QB	371 ± 66	325 ± 49
YM	298 ± 113	311 ± 60
DV	351 ± 50	353 ± 62
MV	344± 46	316 ± 41

181

5.5 Conclusions

The swamp buffalo is the livestock species used for draught power on paddy rice crops. It evolved in the plain area, and so it was present in more than 2/3 of the Ha Giang province. Only the two mountainous districts of Dong-Van and Meo-Vac were shown to have low buffalo populations. Consequently, no discontinuity in its distribution area was observed. Although the H'mong ethnic groups rarely reared buffalo, this was not the case of all the other ethnic groups.

The buffaloes from Ha Giang province seemed to be of the swamp type as all karyotypes showed 2n=48 chromosomes. They also have high gene diversity compared to South–East Asian and Chinese buffaloes. However, genetic structure into subpopulation did not appear,: buffaloes from the Ha Giang province constitute an unsubdivided genetic population. All analyses showed that the clustering order of districts followed geographical proximity.

Whereas gene diversity was high, so was inbreeding. Breeding practice analysis showed that more than half of the buffalo came from family or were bought within the village. Also, in almost 30 % of cases farmers sold their animals within the village. Therefore the probability of mating with a related animal might be substantial which leads to high inbreeding values. However this is counterbalanced by the fact that buffaloes from various villages graze freely for several weeks during outside of the working season, and so allow continuous gene flow.

Similarly to Zhang et al. (2007) who did not find genetic differences between the White Guizhou buffalo population and the grey ones, we also did not find morphometric or genetic differentiation of white animals in the Ha Giang province. In the same way as for genetic data, low phenotypic differentiation on body and horn measurements and coat colour pattern was observed.

Therefore, our results suggested that there was a need to manage reproduction in the swamp buffalo population of the Ha Giang province in order to decrease the inbreeding. This would be easy to implement as only one genetic population is present and so reproductive bulls could be chosen all over the province from a genetic point of view.

CHAPTER 6

THE CHICKEN BREEDS

Paper 4: *published in Animal Genetics*

USING MOLECULAR MARKERS AND MULTIVARIATE METHODS TO STUDY THE GENETIC DIVERSITY ON LOCAL EUROPEAN AND ASIAN CHICKENS BREEDS

C. Berthouly, B. Bed'Hom, M.Tixier-Boichard, C.F. Chen, Y.P. Lee, D. Laloë, H. Legros, E. Verrier, X. Rognon.

Paper 5: *submitted to Animal Genetics*

GENETIC ANALYSIS OF LOCAL VIETNAMESE CHICKENS PROVIDES EVIDENCE OF GENE FLOW BETWEEN DOMESTIC AND WILD POPULATIONS

C. Berthouly, G. Leroy, T. Nhu Van, H. Hoang Thanh, B. Bed'Hom, B. Trong Nguyen, C. Vu Ci, F. Monicat, E. Verrier, J-C. Maillard, X. Rognon

6 THE CHICKEN BREEDS
6.1 Taxonomic status, distribution and domestication of chickens

The domesticated chicken (*Gallus g. domesticus*) belongs to the genus *Gallus*. The genus *Gallus* is actually composed of four species (Figure 6.1). The *Gallus varius* (Green Jungle Fowl) inhabit the islands of Java, Bali, Komodo and Sumbawa. The *Gallus lafayettei* (Ceylon Jungle fowl) evolves mainly in Sri Lanka. The *Gallus sonneratii* (Grey Jungle Fowl) can be found in the bamboo forests from south and west India. The last species, *Gallus gallus* (Red Jungle Fowl), is assumed to be the wild ancestor of the domesticated chicken. The *Gallus gallus* inhabited a very large area: the Asian mainland stretching from north-eastern India in the west to the western coast of China in the east.

Figure 6.1 Phylogeny of the *Gallus* genus.
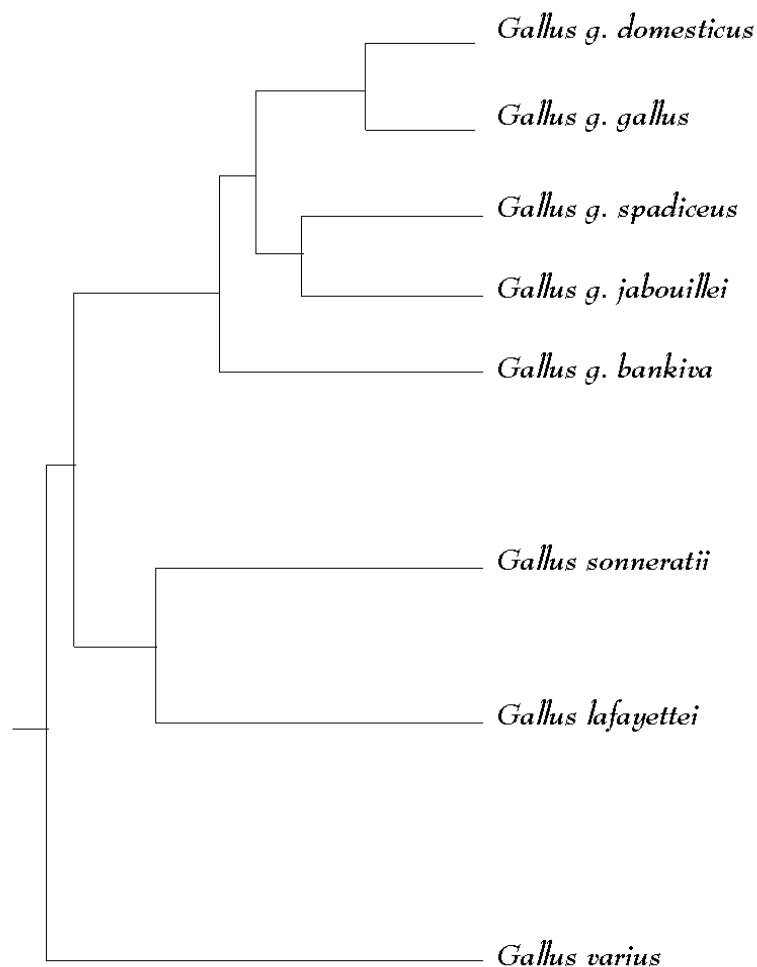
While taxonomy status of a few subspecies is still in question, it is assumed that the *Gallus gallus* is subdivided into five subspecies that are locally distributed (Figure 6.2). The *Gallus g. bankiva* comes from the island of Java and from North India where it shares the same environment with *Gallus sonneratii*. *The Gallus g. spadiceus* is found in Myanmar and the Chinese Yunnan province, whereas the *Gallus g. jabouillei* is found in Vietnam and southern China, and the *Gallus g. murghi* comes from India. The *Gallus g. gallus* is distributed in South-East Asia: mainly Thailand and some areas of Vietnam (Del Hoyo 1994, West & Zhou, 1988).

Figure 6.2 Distribution area of *Gallus sp.*.
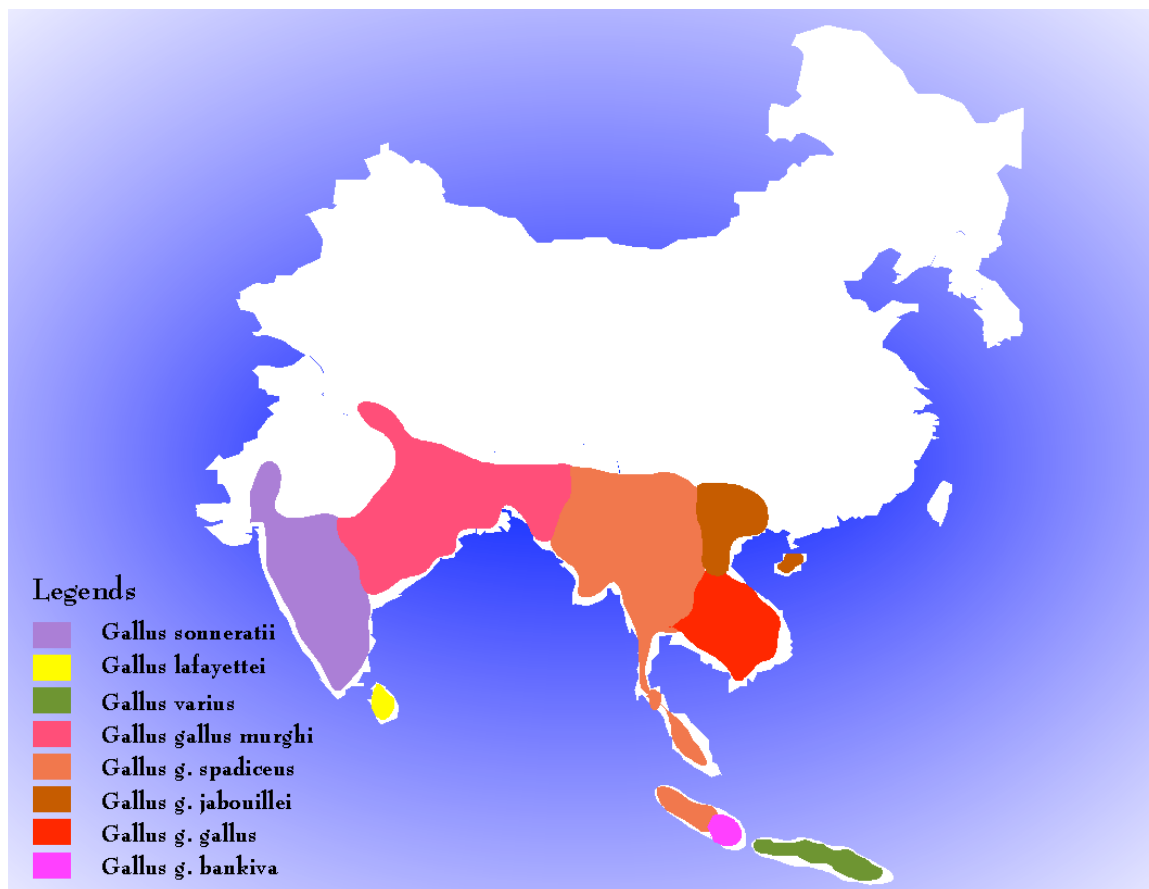
Source: according to West & Zhou (1988)

The first archaeological evidence of domestication to date is from China about 5400 BC, in geographically widespread sites such as Cishan (Heibei province, 5300 BC), Beixin (Shandong province, 5000 BC), and Xian (Shaanxi province, 4300 BC). Domesticated chickens appeared at Mohenjo-Daro in the Indus valley by about 2000 BC (West & Zhou, 1988).

Akishinonomiya *et al.* (1994,1996) were the first to study the evolutionary history of the genus *Gallus*. They used the D-loop regions of mtDNA from all species of *Gallus* and three subspecies of *G. gallus*, namely *G. g. gallus*, *G. g. spadiceus* and *G. g. bankiva.* Firstly they showed that a 60-bp-long unit containing the tetradecamer, AACTATGAATGGTT, was only duplicated in the genus *Gallus*. This tetradecamer was duplicated two to four times in *G. varius*, three times in the *G. sonneratii* and *G. lafayettei*, and two times in *G. gallus* and domesticated chickens. Therefore, they concluded that *G. varius* diverged first, and that afterwards *G. sonneratii* and *G. lafayettei* had diverged simultaneously from *G. gallus*. Also that domesticated chickens descend from *G. gallus*. Secondly, Akishinonomiya *et al.* (1994) used RFLP polymorphism within the 1200- and 1300-Base Control region. The substitution of nucleotides can be of two types: transversion and transition. Transversion mutations had lower occurrence than translations leading to higher divergence time estimates. Among the 23 substitutions observed within *G. g. bankiva* and the other *G. g. gallus,* 9 were transversions; then Akishinonomiya *et al.* (1994) hypothesised that *G. g. bankiva* was not involved in the origin of the domesticated chicken. Between *G. g. spadiceus* and *G. g. gallus*, only one transversion was found hence a total of 4 to 4.5 % of sequence divergence, while only 1.25 % between the two *G. gallus* subspecies. This proportion of divergence within *G. g. gallus* was wider than the range of sequence divergence observed between the wild *G. g. gallus* and the domesticated chickens (0.5 %-1.25%). Thus, the authors concluded that *G. g. gallus* suffices as the matriarchic ancestor of all domestic breeds. Moreover, because all domestic breeds clustered with the *G. g. gallus* originating from Thailand and none with the one originating from Sumatra, they proposed Thailand as the domestication centre of chickens. However, it was argued that because no samples from the two other subspecies, *G. g. murghi* and *G. g. jabouillei,* nor any domestic breeds from China were used in these studies, the results may not be totally relevant.

Similarly to Akishinonomiya *et al.* (1994), Niu *et al.* (2002) analysed mtDNA Control Region on five Chinese breeds and the White Leghorn breed. They showed that the two egg layers, the Chinese Xianju and the White Leghorn, clustered with the *G. g. gallus* and *G. g. spadiceus* from Thailand. The remaining four Chinese meat breeds first grouped together before clustering with the *G. gallus* group and so the authors assumed that these breeds might be closer to the Chinese red jungle fowl than to the Thai red jungle fowl. However, no sample of the Chinese *G. g. gallus* or the *G. g. jabouillei* were used. A similar comparison was done more recently by Liu *et al.* (2006). These authors analysed the mtDNA hypervariable segment I (HVS-I) for 834 domestic chickens across Eurasia as well as 66 wild red jungle fowls and found 169

haplotypes divided into 9 clades: 8 continental and 1 island clade only composed of *G. g. bankiva.* Breeds from the Chinese Yunnan province were found in 7 of the 8 continental clades. Moreover, the *G. g. jabouillei* and *G. g. spadiceus* were mainly distributed in the three clades where the Yunnan and surrounding area breeds fall into; while the *G. g. gallus* animals were mainly found in three other clusters with Indian, Indonesian and Vietnamese samples. Thus, Liu *et al.* (2006) showed that at least three subspecies of *G. gallus* were enrolled in the origin of the domestic chicken breeds, but also that there may be at least two domestication centres: one in South-Southeast Asia (Yunnan, Vietnam, Myanmar, Thailand) and one in the Indian subcontinent (Figure 6.3).

At this point, molecular studies seemed to prove that domestic chickens come from multiple domestication centres enrolling various subspecies of *G. gallus* but also other species such as *Gallus sonneratii.*

Figure 6.3 Possible domestication centres and dispersion routes.
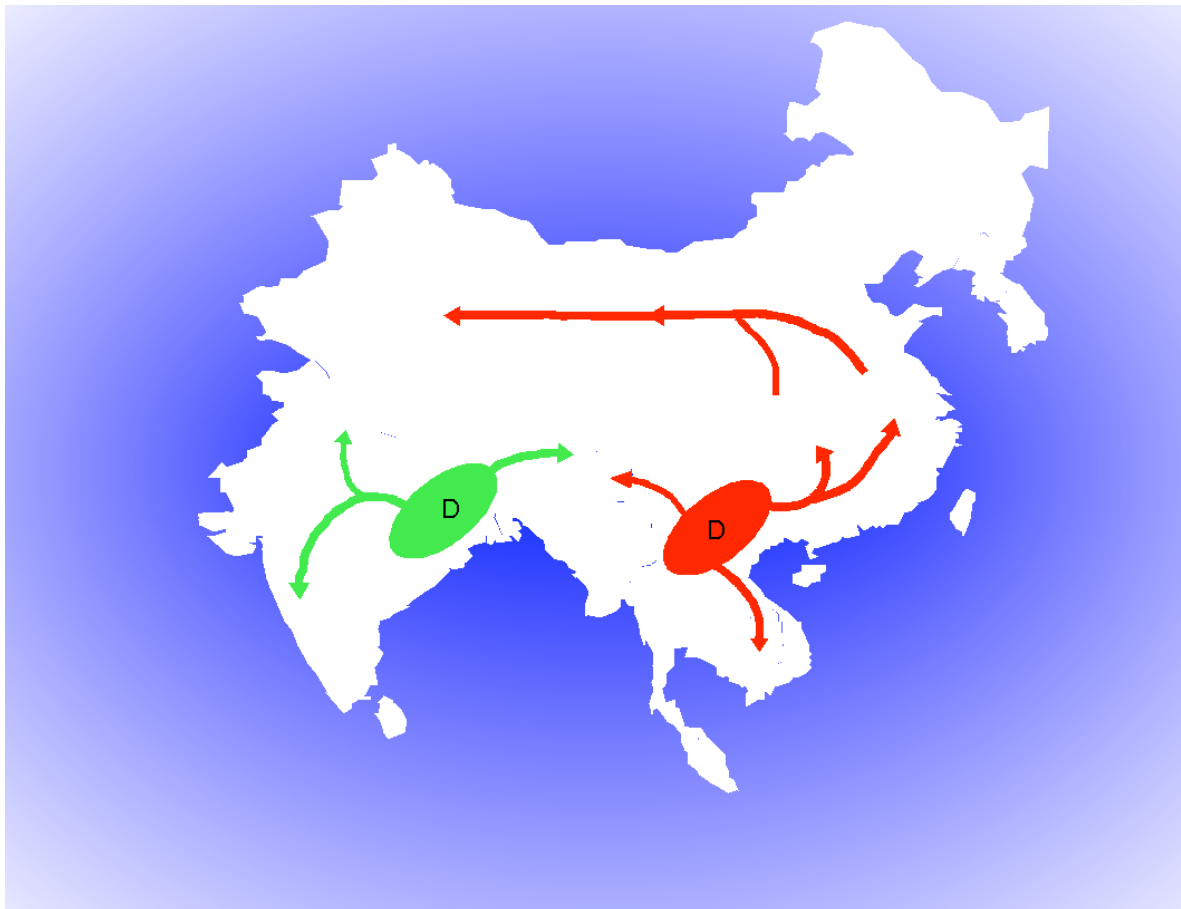
according to West & Zhou (1988) and Liu et al. (2006)

6.2 Role of the chicken

Family poultry (which comprises extensive and small-scale, intensive production) is still important for low income farmers from Africa, Asia, Latin America and the South Pacific. The contribution of family poultry to total poultry production in these countries, with some exceptions, is more than 70 %. In Bangladesh, even landless families which represent 22 % of total farmers keep 5-6 chickens (FAO, 2001). Poultry also plays a significant role in the cultural life of rural gifts, starting capital for young people and sacrifices.

6.2.1 Role in the socio-economy of Asia

Except for Sri Lanka, Asian and South-East Asian countries increased their poultry flock between 2001-2003 (Table 6.1). The highest increase was observed in Laos (38.5 %). This increase is also observed on egg and meat production. However, three countries showed a decrease in chicken meat production: Cambodia (-8.8 %), Sri Lanka (-0.2 %) and Vietnam (-0.2 %). Myanmar had a special pattern, as flock size increased only about 14.4 %, and egg and meat production increased about 28 % and 47 % respectively. This could be due to an improvement of poultry farming methods leading to a higher yield. In 2003, the high Avian Influenza H5N1 epidemic occurred and consequences were noticeable in the evolution of flock sizes. Four countries seemed to be much more affected than the rest: Cambodia (-5.8 %), Indonesia (-2.4 %), Thailand (-22.3 %) and Vietnam (-13.6 %). This seemed to affect meat production more seriously than egg production. In summary, between 2001-2005, two countries showed negative annual mean growth rate for poultry production: Cambodia and Thailand, while Myanmar didn't seem to be affected by the H5N1 epizooty and showed a substantial increase.

6.2.2 Role in the socio-economy of Vietnam

In Vietnam, poultry production is a traditional occupation, and chickens and ducks are common components of mixed farming systems in rural areas. It is estimated that 80 % of rural households participate in poultry production through backyard and garden raising. Figure 6.4 shows that poultry rearing is very important in rural area, the highest proportions of farmers rearing poultry was observed in the rural northern mountains.

Table 6.1 Rates of change: (%) in chicken population, meat and egg production and annual growth rate in South-East Asia from 2001 to 2005.

Countries	Flock size			Chicken Eggs			Chicken Meat		
	2001-2003	2003-2005	A	2001-2003	2003-2005	A	2001-2003	2003-2005	A
Cambodia	5	-5.8	-0.2	11.1	2.3	2.7	-8.8	-9.6	-3.5
China	5.6	8	2.8	11.3	8	4	5.9	7.5	2.8
India	13.2			4.3	14.3	3.8	28	18.8	10.4
Indonesia	25.4	-2.4	4.5	13.8	8.6	4.7	10.3	0.7	2.2
Laos	38.5	1.7	8.2	11.3	-2.3	1.7	29.7	11.1	8.8
Malaysia	10.3	12.1	4.7	5.9	4.9	2.2	12	12.4	5.2
Myanmar	14.4	29.4	9.6	27.9	38.7	15.5	47	47.7	23.4
Philippines	10.9			12.4			8.2	2.3	2.1
Sri Lanka	-8.3	19.1	1.8	-6.5	-2.3	-1.7	-0.2	10.4	2
Thailand	3.6	-22.3	-3.9	11.7	-15.4	-1.1	-0.2	-22.6	-4.6
Viet Nam	16.7	-13.6	0.2	21	-7.3	2.4	21	-13.6	0.9

A: annual growth rate between 2001-2005

Figure 6.4 Proportions (%) of householders keeping poultry by region (1998).
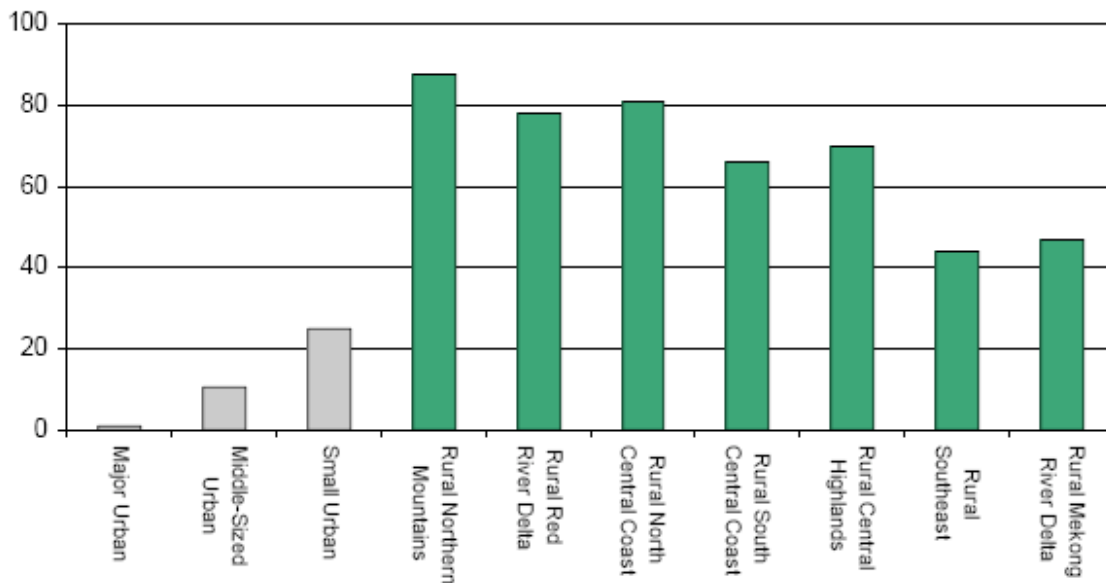
Source: FAOSTAT

Poultry livestock farms are present throughout Vietnam. However, variation should be noted in the density of livestock farming: 3.1 farms/km^2 in the Kom Tum province, and 236.6 farms/km^2 in the Thai Binh province (FAOSTAT, 2001). In the Ha Giang province, there were between 12-20 poultry farms/km^2. The FAO has suggested classifying poultry production systems in four sectors: (1) intensive, large scale, industrial poultry production, (2) commercial

poultry production, (3) semi-intensive, market oriented, (4) traditional extensive backyard poultry production. Traditional extensive household poultry (sector 4), with small production units (less than 50 birds) represents the largest portion of production systems with 94.8 % of farms, on average, throughout Vietnam and represent more than 80 % of the production of meat and eggs. Flocks are reared by scavenging or supplemented with limited amounts of home-produced grains such as paddy rice or maize, and kitchen waste. The amount of feed given to birds does not focus on production efficiency, it depends on the availability of grains that farmers have in storage for personal use. Chick replacements are hatched from own-stock eggs, but farmers can also buy replacements from local markets. Since it is considered as a side-line activity, attention to bird health and safety is limited and mortalities can be high, in bad weather conditions as high as 40 to 50 %.

The intensive production system farms (sector 1) have emerged over the last 10 years in Vietnam, and currently produce about 10 % of Vietnam's chickens. These farms are mainly implemented in the Red River delta (30.3 %) and the South-East region (24.6), while in the North-West region to which the Ha Giang province belongs, there is only 2.2 % of the Vietnamese industrial farms. As shown by Figure 6.5, the traditional farming system is the main supplier of chicken meat and eggs, and the contribution of industrial production systems is still limited.

Figure 6.5 Proportion of chicken by production system for meat (a) and eggs (b) in 2006.
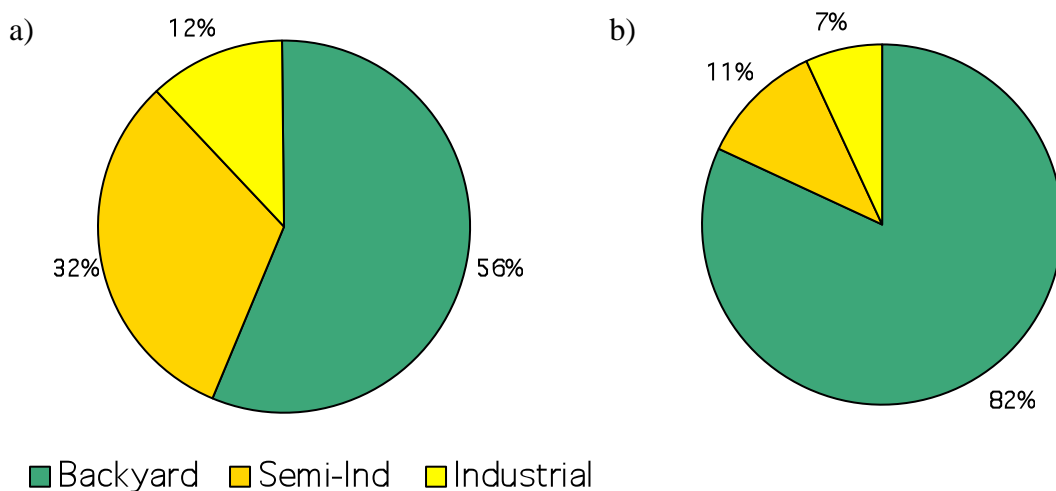
Sectors 2 and 3 are grouped into Semi-Industrial production system

Between 2001 and 2003, the number of poultry increased from 160 to 185 million birds (Table 6.2). Poultry meat and egg production increased about 8.4 % and 8.9 % respectively. In 2003, the epidemic of Avian Influenza started, and as a consequence the national flock decreased (by 6.7 %) as did chicken meat (-6.7 %) and eggs production (-9.3 %).

Table 6.2 Chicken population, meat and egg production in Vietnam (2001-2005).

	2001	2002	2003	2004	2005
Chicken (millions)	160.2	169.7	184.7	159.3	159.9
Chicken meat (tons)	307 671	338 402	372 721	316 409	321 890
Eggs (millions)	4 161	4 722	4 852	3 939	3 948

Source: FAOSTAT

6.3 Asian and Vietnamese chicken breeds

Over the world, a total of 734 chicken breeds are listed (Scherf, 2002) from which 26.6 % are not in danger and for 19.2 % status is unknown. Therefore more than 50 % of chicken breeds listed are categorised at risk. Europe is rearing more than 64.2 % of the world chicken breeds, while 17.7 % are encountered in Asia and the South Pacific. India had 22 breeds listed, and China 111, grouping indigenous and transboundary breeds.

For Asian chicken breeds, little information is available. In most South-Asian countries, one to two indigenous breeds are recorded, except for Vietnam (Table 6.3). For Vietnamese breeds there were some discrepancies between data sources (DAD-IS and National Breed Atlas).

Most of the Vietnamese breeds are reared for meat production. There is one fighting cock (Choi Breed) from the south-central area. There is also one fancy breed reared as pets for children, the Tre breed. This chicken, which looks like a wild chicken but is dwarf-sized. In Vietnam, black meat is assumed to have medicinal properties, and so black meat chicken are very valuable. Currently, three breeds are considered to have black meat: the Ac, the O ke and the H'mong breeds. The Ac breed is reared around Hanoi and Ho Chi Minh, the O Ke breed is reared in the mountainous province of Lao Cai with the H'mong breed, which is also reared in the neighbouring provinces of Son La and Ha Giang (Figure 6.6).

Table 6.3 Summary of the main South-East Asian indigenous chicken breeds.

Countries	Breed names	Production	Traits	Weight (kg)		Nb of eggs/year
				cock	hen	
Cambodia						
	Monn Khmer	meat		3	2	
Laos						
	Kai Lat			1.5	1.2	
Malaysia						
	Ayam Kampong	meat		2	1.5	
Thailand						
	Chon	fighting				
	Puen Muang	meat				
Myanmar						
	Tain Nyin	meat				
	Myogyi	meat		2.2	2	120
Vietnam						
	Ac	**meat**	**black meat**	**1**	**0.8**	**95**
	Choi	**fighting**		**3.5**	**2**	
	Dong Tao	**meat**		**4.5**	**3.5**	**60**
	Ho	**meat**		**4.4**	**2.7**	**30-60**
	Mia	**meat**		**3**	**2.3**	**58**
	Ri	**meat**		**2.7**	**1.2**	**40-75**
	Tre	**Fancy**	**dwarf**	**1.2**	**0.8**	**40-70**
	H'mong	**meat**	**black meat**	**2.3**	**1.8**	**40-75**
	O ke	**meat**	**black meat**	**2.5**	**1.2**	**16-30**
	Te	**meat**		**1.6**	**1.3**	**45-72**

Source: DAD-IS

6.4 Genetic comparison of breeds

6.4.1 Previous studies

As chicken may be the first animal protein source throughout the world (accounting for meat and eggs), there is particular interest in managing chicken genetic biodiversity. The first studies on microsatellite polymorphism were carried out on commercial breeds or highly inbred lines (Crooijmans *et al.* 1996, Zhou and Lamont 1999). Chicken breeds are managed with different levels of selection status. Highly selected commercial breeds have been created for meat or egg production such as the White Leghorn breed. Traditional selected breeds or fancy breeds were submitted to lower selection pressures, mainly for some morphological traits. The last type but the most extended is represented by the local unselected breeds. Many developing countries do not establish selection schemes for their poultry, which is raised as scavenging chickens in backyards.

Figure 6.6 Distribution of main Vietnamese indigenous chicken breeds.

Legends

⬛ (gray)	H'Mong breed
⬛ (light blue)	Tre breed
🟥 (red)	Choi breed
▤ (horizontal lines)	O Ke breed
🟩 (green)	Ac breed
🟪 (purple)	Te breed
▨ (diagonal lines)	Mia breed and Dong Tao breed
🟨 (yellow)	Ho breed
▧ (boxed diagonal)	Ri breed
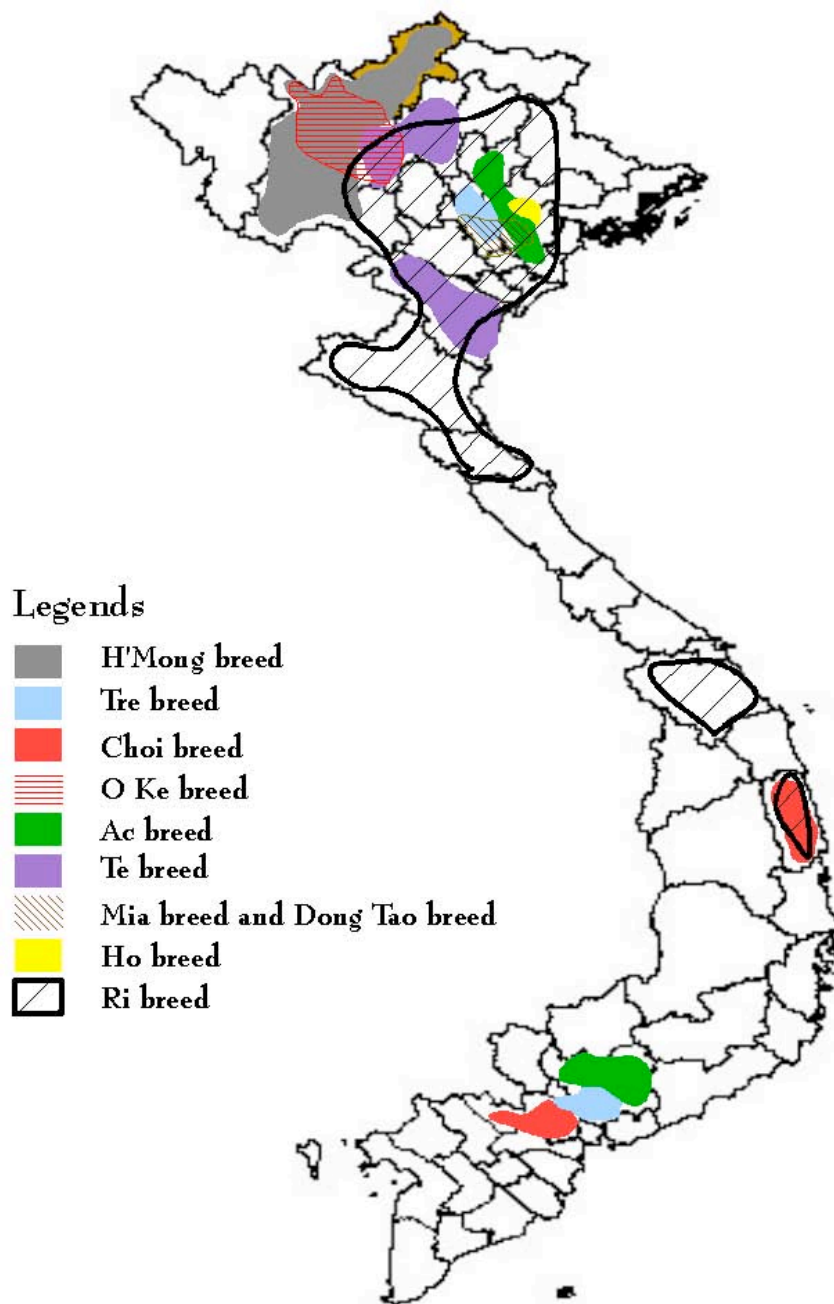
Therefore, the worldwide chicken flock will contain: chickens under high human selection pressure (may be more than in mammals species due to the short generations interval) on one side and chickens under high natural selection pressure on the other side. Consequently, it has been of prime interest to study and compare the effect of different selection types on genetic diversity. One of the first studies with such an objective was carried out by Hillel *et al.* (2003) who assessed the genetic diversity of 52 European chicken breeds using 25 microsatellite

194

markers (AVIANDIV project). Within this set of breeds, commercial, traditional selected and unselected breeds were compared to the wild ancestor *Gallus g. gallus*. These authors showed that (i) the traditional unselected breeds had the highest gene diversity (0.56), (ii) the traditional selected had slightly higher diversity on average (0.46) and, (iii) egg layers on commercial lines were found to have lowest gene diversity (0.45). Gene diversity values within this type was heterogeneous and could be explained by the complexity of histories and breeding practices. More recently, Granevitze *et al.* (2007) used some of those breeds with additional local Asian and African breeds. They confirmed the high genetic diversity of local unselected breeds but they showed that commercial selected breeds had higher genetic diversity than traditional selected (i.e. fancy) breeds. These authors also showed that there was a relationship between inbreeding values and the types of management.

In order to have a global insight into the genetic diversity of chicken breeds, FAO/ISAG proposed a set of microsatellite markers mainly including the ones previously used in the AVIANDIV project. This will allow the comparison of genetic polymorphism between breeds but it will not compare breeds for prioritisation and for setting conservation policies. For targets such as these, a set of breeds that represents all selection types and different breed histories is needed. Therefore we used the previously used AVIANDIV breed set in order to compare worldwide distributed chicken breeds and we used a multivariate approach to asses the usefulness of the microsatellite marker and breed sets used.

6.4.2 Comparison of European and Asian breeds

USING MOLECULAR MARKERS AND MULTIVARIATE METHODS TO STUDY THE GENETIC DIVERSITY OF LOCAL EUROPEAN AND ASIAN CHICKEN BREEDS

Published in Animal Genetics

Using molecular markers and multivariate methods to study the genetic diversity of local European and Asian chicken breeds

C. Berthouly*,†,‡, B. Bed'Hom†, M. Tixier-Boichard†, C. F. Chen§, Y. P. Lee§, D. Laloë¶, H. Legros**, E. Verrier*,† and X. Rognon*,†

*AgroParisTech, UMR1236 Génétique et Diversité Animales, 16 rue Claude Bernard, 75321 Paris Cedex 05, France. †INRA, UMR1236 Génétique et diversité animales, 78352 Jouy-en-Josas, France. ‡CIRAD, UR 22, Campus International de Baillarguet, 34398 Montpellier Cedex 05, France. §National Chung-Hsing University, Department of Animal Science, Taichung, Taiwan. ¶INRA, UR337 SGQA, 78352 Jouy-en-Josas, France. **LABOGENA, 78352 Jouy-en-Josas, France

Summary

French and Asian subsets of chicken breeds were first analysed using 22 microsatellites and then compared to the AVIANDIV European set using 14 loci. Positive correlations were observed between F_{IT} or F_{ST} and typological values or variance of markers using the multivariate analysis MCOA. The first axis of the multivariate representation separated Asian from European breeds, revealing breeds with Asian ancestor. Using all or 14 loci, correct assignation rate was always higher than 93%. The Weitzman index and the aggregate diversity D were calculated using 22 loci within French and Asian breeds. The French breed Coucou de Rennes and the Hua-Tung breed seemed to contribute the most to the global diversity of each subset. This approach on French-only breeds and then on French with AVIANDIV domestic breeds (14 loci) showed that the Marans breed contributed the most. The AVIANDIV framework could be useful to evaluate the genetic diversity of local breeds and to help in connecting national and regional conservation policies.

Keywords chicken, genetic diversity, microsatellite, multivariate.

Introduction

Livestock populations have been subjected to a variety of evolutionary forces during their histories. The cumulative effect of foundation events, genetic drift and natural or artificial selection has led to the formation of distinct breeds. In the process of evaluating genetic diversity to develop conservation programmes, it is of interest to assess genetic variation between domestic stocks by using powerful tools as genetic markers. The discriminatory power of microsatellites has been measured by performing an assignment test, and the assignment error rate was used as an estimate of their power (Blott *et al.* 1999; Rosenberg *et al.* 2001; Bjornstad & Roed 2002). Moreover, multivariate techniques represent an alternative and complementary approach to study genetic relationships between populations because they do not make any evolutionary hypotheses (Moazami-Goudarzi & Laloë 2002). Thus, they can be applied to populations

where admixture is known to have occurred, and they can be compared to classical methods such as phenetic analyses.

One of the primary objectives of this study was to characterize the diversity of a set of traditional chicken breeds from two distant regions (France and Taiwan) that were definitively different from many points of view. Because local breeds represent both a heritage and a reservoir of variability that deserves to be explored and properly managed, we analysed the contribution to genetic diversity of those breeds, as proposed by Ollivier & Foulley (2005), to highlight helpful breed characteristics. A second approach was to analyse these two sets of geographically distant populations in the framework of a set of European breeds previously studied in the AVIANDIV project (Rosenberg *et al.* 2001; Hillel *et al.* 2003). The data set provided by this EU-funded project could be considered a reference data set to study the diversity of additional populations.

Materials and methods

Breeds and genotypes

We used three data sets, including a set of 14 traditional French breeds (Table 1) selected for morphological traits,

Table 1 Summary of breeds and their polymorphism measures.

Breed name	Breed code	Breed category	H_E 22 loci	H_O 14 loci	H_O 22 loci	Average no. per locus (22 loci) Alleles	Effective alleles	Private alleles	F_{IS}	F_{ST}	HWE deficiency[1]
Bresse Blanche	B99	TSP, France	0.553	0.517	0.555	3.64	2.19	0	−0.004	0.221	0
Gauloise Grise	GLG	TSP, France	0.519	0.467	0.479	3.41	2.04	0	0.073	0.231	1
Gauloise Noire	GLN	TSP, France	0.514	0.464	0.473	3.77	2.02	0	0.089*	0.184	1
Géline de Touraine	GLT	TSP, France	0.599	0.567	0.581	3.91	2.43	0	0.052	0.205	0
Barbezieux	BAZ	TS, France	0.517	0.503	0.526	4.00	2.04	1	−0.035	0.195	0
Bourbonnaise	BNA	TS, France	0.472	0.500	0.487	3.23	1.87	0	−0.020	0.243	0
Coucou de Rennes	COU	TS, France	0.625	0.586	0.583	4.73	2.59	0	0.054	0.175	0
Creve-Coeur	CRC	TS, France	0.510	0.536	0.528	3.32	2.00	0	0.002	0.220	0
Gasconne	GAS	TS, France	0.540	0.491	0.538	3.64	2.13	0	0.010	0.201	0
Gauloise Dorée	GLD	TS, France	0.435	0.370	0.409	3.36	1.75	0	0.094	0.266	1
Gournay	GOU	TS, France	0.529	0.543	0.542	3.59	2.08	1	−0.038	0.256	0
Houdan	HOU	TS, France	0.429	0.424	0.426	2.68	1.72	0	−0.029	0.246	0
Marans	MR	TS, France	0.530	0.580	0.492	3.59	2.08	2	0.049	0.180	1
Noire de Challans	NC	TS, France	0.608	0.378	0.445	4.68	2.49	1	0.353*	0.217	5
Hsin-Yi	HY	TU, Taiwan	0.523	0.462	0.534	3.86	2.07	0	−0.018	0.216	0
Hua-Tung	HT	TU, Taiwan	0.570	0.559	0.584	3.36	2.29	0	−0.053	0.212	0
Ju-Chi	JC	TU, Taiwan	0.444	0.426	0.458	3.45	1.78	0	−0.053	0.278	0
Nagoya	NG	TS, Japan	0.429	0.415	0.438	3.23	1.74	0	−0.049	0.347	0
Quemoy	KM	TU, Quemoy Island	0.506	0.531	0.491	3.27	2.00	0	0.008	0.260	0
Shek-Ki	SK	TS, China	0.458	0.401	0.423	2.77	1.83	0	0.068*	0.286	0

TU, traditional unselected; TS, traditional standardized for morphological traits; TSP, traditional standardized breed selected for a production trait; CS, commercial selected; H_E, unbiased heterozygosity; H_O, observed heterozygosity.
*Significant value after 10 000 bootstraps.
[1]Number of loci deviating from Hardy–Weinberg equilibrium after Bonferroni correction.

four of which were moderately selected for a production trait. The second set was formed by six Asian breeds, kept under a conservation programme in Taiwan (Table 1). Four of them were traditional unselected breeds. The remaining two breeds from China and Japan were selected for quantitative traits. We compared these three sets to a reference set of 15 breeds taken from AVIANDIV project (Table S1), which included a Red Junglefowl population. The genotypes are available at http://charles.stanford.edu. For the French and Asian sets, a total of 687 individuals from 20 breeds (22–48 individuals per breeds) were genotyped for 22 microsatellite loci (Table S2) using a capillary sequencer (ABI PRISM 3100 Genetic Analyzer; Applied Biosystems). Genotypes are available upon request.

Statistical analysis

Allele frequencies, number of alleles, expected and observed heterozygosity values were calculated, and F_{IS} values were tested by performing 10 000 bootstraps using GENETIX 4.4 (Belkhir *et al.* 2000). The GENEPOP (Rousset & Raymond 1995) was used to compare F_{IS}, F_{ST} (Weir & Cockerman 1984), pairwise genetic differentiation between breeds and departure from Hardy–Weinberg equilibrium using exact tests and Bonferroni correction on loci. To establish whether any individual could be assigned or excluded from its own breed or the others, we used the Bayesian exclusion procedure of GENECLASS 2 (Piry *et al.* 2004).

The Multiple CO-inertia Analysis (MCOA) was carried out on breed-by-locus frequency data using the ADE-4 package (Chessel *et al.* 2005) implemented in R (R Core Development Team 2006). The MCOA simultaneously finds ordinations from the tables that are the most congruent. It does this by finding successive axes from each table of allelic frequencies, which maximizes a covariance function. This method permits the extraction of common information from separate analyses, in the setting up of a reference typology, and in the comparison of each separate typology. The efficiency of a marker is assessed by its typological value (Tv; the contribution of the marker to the construction of the reference typology), which is equal to the product of the variance (Var) multiplied by the congruence with the consensus Cos^2 (i.e. the correlation between the scores of individual locus tables and the synthetic variable of the same rank) (Laloë *et al.* 2007).

The matrices of D_R genetic distance (Reynolds *et al.* 1983) have been calculated to draw a NEIGHBORNET tree using SPLITSTREE 4.8 (Huson & Bryant 2006).

A measurement of breed contribution to genetic diversity was computed with WEITZPRO (Derban *et al.* 2002; http://

www-sgqa.jouy.inra.fr/article.php3?id_article=3) using the D_A genetic distance (Nei *et al.* 1983) calculated according to POPULATION (Olivier Langella; http://bioinformatics.org/project/?group_id=84). Weitzman values represented the contribution to between-breed diversity (CB). Within-breed (CW) contributions to diversity, as well as the aggregate diversity D index, were calculated as proposed by Ollivier & Foulley (2005). The D index was obtained after weighting CB by F_{ST} and CW by $(1 - F_{ST})$.

Results

Gene diversity within French and Asian breeds

A total of 161 alleles was detected, and the mean number of alleles per locus was 6.6 for French breeds and 5.6 for Asian breeds (Table 1). The observed heterozygosity per locus (H_O) averaged 0.500 and ranged from 0.299 for locus *MCW098* to 0.749 for locus *LEI234* (Table S2). All loci showed highly significant heterogeneity in allele frequencies (F_{ST}) between breeds. Given the multilocus F_{ST}, around 24.2% of the total genetic variation could be explained by breed differences.

French breeds showed the highest heterozygosity with a mean value of 0.505, whereas the heterozygosity averaged 0.477 for Asian breeds. The highest value of H_O in French breeds was observed for the COU breed (0.583), and the smallest value of H_O was found for the GLD breed (0.409). The NC breed presented the highest number of loci (5) deviating from Hardy–Weinberg equilibrium after Bonferroni correction. For Asian breeds, H_O ranged from 0.423 for the SK breed to 0.584 for the HT breed. Pairwise tests of genic and genotypic differentiation between breeds were highly significant.

The mean F_{ST} was lower within French breeds (0.193) than within Asian breeds (0.270). Considering the French subset, the BNA breed seemed to be the most different (0.233), whereas the COU breed had the smallest F_{ST} (0.155). Within Asian breeds, the NG breed showed the highest value of F_{ST} (0.351), and F_{ST} values averaged 0.174 within the four local Taiwanese breeds.

Analysis of French and Asian breeds within AVIANDIV framework

A set of nine reference samples was analysed with French and Asian samples to integrate these new genotypes with previous AVIANDIV data. Because all of these genotypes have been obtained in different places with different equipment, the reference samples were used to calibrate allele sizes between different studies. For a subset of 14 markers, the observed size difference could be easily corrected by a simple offset: the allele sizes were adjusted to be consistent with AVIANDIV data. In a few instances, the capillary technology identified unambiguously two close alleles with sizes differing only by one base (these were classified as a single allele in

previous studies); in these cases, new alleles were merged to match the previous one. But the match between new and previous alleles was impossible to establish for eight markers because of irreconcilable discrepancies between genotypes of reference samples with previous studies. As a consequence, the French and Asian breeds have been compared with the other European breeds using 14 markers only.

Assignment test

Bayesian assignment showed that 96.8% of individuals were assigned to their true breed when using 22 loci for French and Asian breeds. The mean probabilities of individual assignment to their true breed ranged from 0.494 for the NC breed to 0.858 for the SK breed (Table 2). Twenty individuals were assigned with higher probabilities to other breeds than to their true breed; 17 of these were assigned to the HT breed. Restricting to 14 loci led to a slight decrease in the percentage of French and Asian individuals correctly assigned (93.7% instead of 96.8%). Within the AVIANDIV framework, the success rate for the 36 breeds was still high and reached 92.1%. However, 73% and 55% of all individuals were assigned with a probability higher than 0.1 to the HT breed and to the Red Junglefowl wild ancestor respectively.

Table 2 Assignment analysis of 687 individuals from 20 breeds using 22 loci.

Breed code[1]	P-value[2]	% Correct[3]	Misassignment[4]
B99	0.647	100	
BAZ	0.6097	93.33	COU(1),HT(1)
BNA	0.6176	93.33	HT(2)
COU	0.5577	96.66	HT(1)
CRC	0.6338	100	
GAS	0.6935	96.66	HT(1)
GLD	0.5786	82.14	HT(3)
GLG	0.5281	100	
GLN	0.5446	96.66	HT(1)
GLT	0.5950	100	
GOU	0.6317	100	
HOU	0.8382	100	
MR	0.5062	100	
NC	0.4976	96.55	HT(1)
HY	0.5726	95.74	HT(2)
HT	0.8475	100	
JC	0.6740	91.67	HT(2), NG(1),HY(1)
NG	0.6049	93.75	HT(3)
KM	0.6868	100	
SK	0.8514	100	

[1]See Table 1 for breed codes.
[2]The average of individual assignment probabilities of individuals from one breed to their own breed.
[3]Percentage of individuals showing the highest probability of assignment to their true breed.
[4]The breed (and number of individuals) where the individual had the highest probability of assignment.
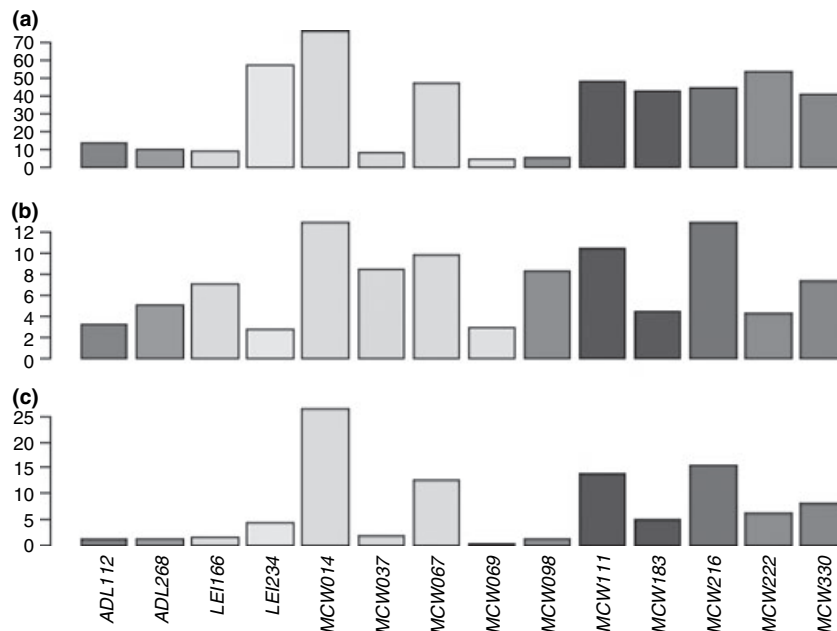
Figure 1 Distribution across loci of values for the components to first axis, in percentages: (a) Cos², (b) Var, (c) Tv.

Multivariate procedure

Of 14 loci, eight loci showed high common congruence values (Cos²); the highest ones were *MCW014, LEI234* and *MCW222* (Fig. 1a). Comparisons between Fig. 1a and b highlight that high variance at a locus implies a good differentiation between breeds but in a different scheme or typology than the consensus one. This situation is observed for locus *MCW098*. This locus is triallelic with a private allele. It highly differentiates breeds but the resulting typology is quite different from the consensus and so Tv of *MCW098* is low. However, combining high values of variance and Cos² will lead to a strong Tv as found for *MCW014* for the first axis (Fig. 1c) and *MCW330* for the second axis (data not shown). The *MCW014* locus, contributing about 25% to the construction of the first axis, was able to organize breeds in subgroups. High Tv values indicate good congruence between the breed structure observed for each locus and the synthetic position. When high Tv values are observed for all the loci then structuring is homogeneous among loci. However, Fig. 1c shows that Tv was quite different according to loci and made possible to select the best markers for future studies if necessary. For example, locus *MCW014* presented the highest values of F_{ST} and F_{IT}. It was the most discriminative and its breed differentiation was the most similar with the synthesis among all loci. In addition, for the first axis, positive correlations were found between Var or Tv and F_{ST} (0.54 and 0.60 for Var or Tv respectively) or F_{IT} (0.65 and 0.79 for Var or Tv respectively) and between Cos² and F_{IT} (0.63). The synthetic position of breeds (Fig. 2) shows that the first

axis separated all Asian breeds and a subset of European breeds from the remaining breeds, which suggests an east–west cline.
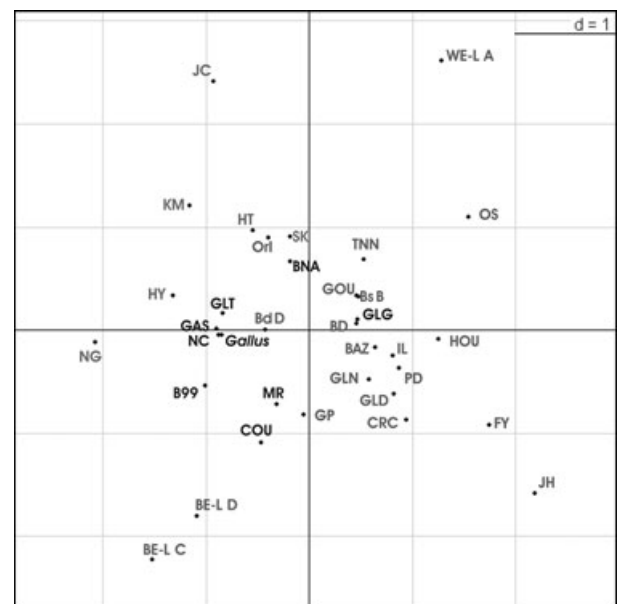
Figure 2 Synthetic representation of breed positions considering 14 loci analysed by MCOA. Breed codes are given in Table 1 for French and Asian breeds and in Table 3b for AVIANDIV breeds. The *Gallus gallus* breed is coded as *Gallus* in bold font. The colours in the online version of the figure denote the following: red, AVIANDIV breeds; blue, French breeds with Asian ancestor; violet, French breeds; green, Asian breeds.
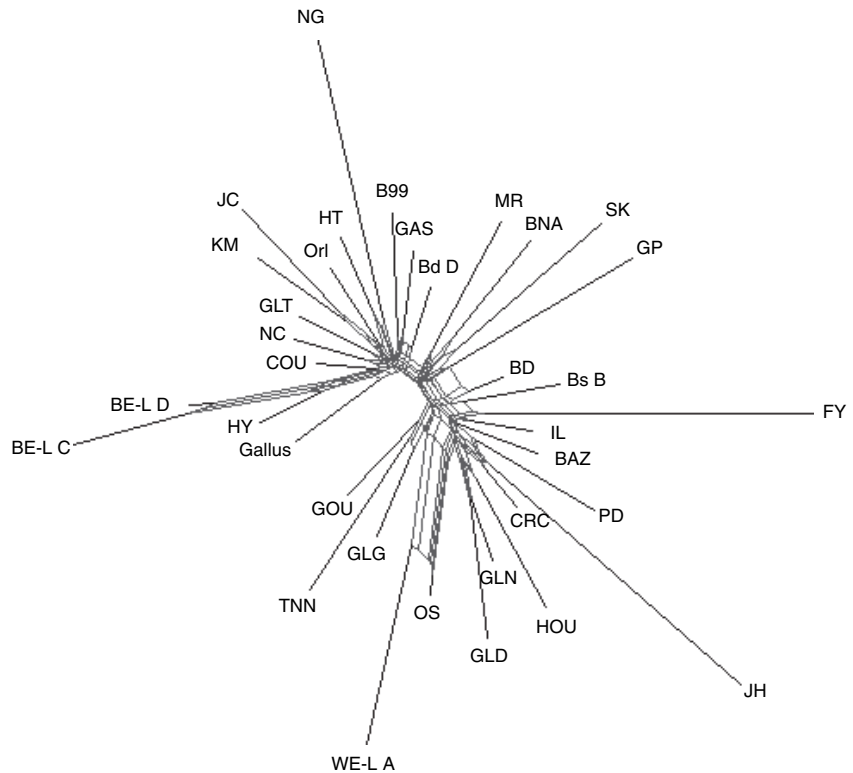
Figure 3 NEIGHBORNET tree of breeds on DR distances using 14 loci. Breed codes are given in Table 1 for French and Asian breeds and in Table 3b for AVIANDIV breeds.

Tree analysis

The NEIGHBORNET graph shows numerous splits indicating a reticulated structure between breeds (Fig. 3). The suggested east–west cline can be considered in light of the reticulated tree. Two blocks of highly related breeds were observed for White Leghorn layers (WE-L A and OS) on one hand and for commercial brown-egg layers (BE-L C and BE-L D) on the other hand.

Contribution of breeds to genetic diversity

The results listed in Table 3a show that the individual French breed contribution to within-breed diversity (CW) varied from −1.48% (HOU) to 1.44% (COU). Contribution to between-breed diversity (CB) is similar between French breeds, ranging from 5.53% to 10.58% for the HOU breed. Aggregating within- and between-breed diversity (D), the GLD breed seemed to participate less in the diversity of French breeds because of its low within- and between-breed diversity. The HOU breed had the second-least-important contribution because of its low CW, despite having one of the highest values of CB. The COU breed presented a high level of CW but had one of the smallest values of CB; it had the highest value of D. Considering the Asian subset, the HT breed showed the highest value of CW (3.32%), which resulted also in the highest value of D (7.55%). The JC breed, with a low value of CW (−1.82%) and the lowest

value of CB (9.8%), was the smallest contributor to the diversity of Asian breeds (1.32%).

If we compare the index measured depending on number of loci (14 vs. 22), the Wilcoxon rank test revealed significant differences for H_O and F-statistics but not for CW, CB or D. The MR breed had the highest D value using 14 loci, whereas the COU breed exhibited the highest value using 22 loci and for the Asian subset, the KM breed replaced the HT breed.

When comparing AVIANDIV breeds (without Red Junglefowl) with French breeds using 14 loci, the FY breed had the highest between-breed diversity contribution (Table 3b). In addition, the MR breed had the highest aggregate diversity over all breeds. The traditionally selected Orlov breed and the broiler line Bd D followed, each with an aggregate diversity of 1.6.

Discussion

Diversity analysis of two local breed subsets

Although many local breed populations were dramatically reduced after the Second World War, they have been restored or maintained with variable population sizes and number of breeders, which has led to heterogeneous rates of inbreeding. For the 20 local breeds studied here, heterozygosity values exceeded values estimated for highly inbred lines (Zhou & Lamont 1999). Nevertheless, mean values

Breed code[1]	For 22 loci			For 14 loci		
	CW	CB	D	CW	CB	D
(a)						
B99	0.39	9.30	1.92	−0.01	8.86	1.66
BAZ	−0.14	5.73	0.99	−0.47	5.29	0.77
BNA	−0.80	7.87	0.87	−0.38	9.14	1.67
COU	**1.44**	5.98	**2.31**	1.54	5.82	**2.46**
CRC	−0.27	6.70	1.07	0.13	5.02	1.18
GAS	0.18	7.46	1.58	−0.05	6.24	1.30
GLD	−1.35	5.79	*0.00*	−1.64	5.14	*−0.18*
GLG	−0.12	6.82	1.22	−0.24	7.60	1.45
GLN	−0.19	7.04	1.20	−0.15	*4.38*	0.83
GLT	1.06	6.44	2.09	1.05	6.43	2.21
GOU	0.04	6.48	1.28	0.13	9.07	2.06
HOU	*−1.48*	**10.58**	0.84	*−1.95*	9.36	0.49
MR	0.03	9.62	1.88	1.15	**13.53**	**3.82**
NC	1.20	*5.53*	2.03	0.88	5.29	1.83
HY	1.43	12.40	4.39	0.40	10.70	3.04
HT	**3.32**	18.99	**7.55**	3.03	18.89	**7.09**
JC	−1.82	9.80	*1.32*	−2.41	8.87	*0.48*
NG	−2.41	**33.56**	7.31	−2.84	33.77	6.54
KM	0.74	18.16	5.45	**3.20**	21.67	7.93
SK	−1.26	25.80	6.05	−1.37	27.93	6.14

Breed name	Breed code[1]	CW	CB	D
(b)				
	B99	0.2	3.1	0.9
	BAZ	0	2.6	0.6
	BNA	−0.1	3.7	0.8
	COU	**1**	2.6	1.4
	CRC	0.3	2.5	0.8
	GAS	0.3	2.8	0.9
	GLD	−0.6	2.5	*0.2*
	GLG	0.1	3.4	0.9
	GLN	0.1	*2.3*	0.6
	GLT	0.6	3.1	1.2
	GOU	0.3	3.3	1.0
	HOU	*−0.7*	3.5	0.3
	MR	0.7	**4.6**	**1.7**
	NC	0.7	2.7	1.2
Fayoumi	FY	−1.1	**6.7**	0.8
Bedouin	BD	0.5	2.8	1.1
Icelandic landrace	IL	0.6	2.4	1.0
Old Scandivanian reference population	OS	−0.2	2.4	0.4
Jaerhoens	JH	−1.5	5.5	0.2
Padovana	PD	−0.3	3.7	0.7
Transylvanian naked neck	TNN	−0.4	3.9	0.7
Green-legged partridge	GP	−0.4	4.8	0.9
Orlov	OL	0.6	4.5	**1.6**
White-egg layer line A	WE-L A	−1.3	4.9	0.2
Broiler sire line B	Bs B	0.4	3.2	1.1
Brown-egg layer line C	BE-L C	−1	2.3	−0.2
Brown-egg layer line D	BE-L D	0	*1.6*	0.4
Broiler dam line D	Bd D	**1.3**	2.6	**1.6**

Table 3 Contributions (%) of chicken breeds to genetic diversity: (a) for 22 and 14 loci within French breeds and within Asian breeds; (b) for 14 loci within domestic breeds from France and the AVIANDIV framework.

CW, contribution to within-breed diversity; CB, contribution to between-breed diversity using D_A distances; D, aggregate diversity.

For each of these categories, the lowest values are in italics and the highest values are in bold.

[1]Breed codes are given in Table 1 for French and Asian breeds.

were lower than those estimated for some local Asian, African and American breeds (Wimmers *et al.* 2000; Muchadeyi *et al.* 2007) or Chinese breeds (Qu *et al.* 2004). Considering the population management, deviation from Hardy–Weinberg equilibrium was rare; it may be due to genetic drift, inbreeding or selection, leading to population fragmentation. In the case of the Noire de Challans breed, the animals were sampled from different farms, and different patterns of genotypes were obtained depending on the farm. The farmers appeared to have few exchanges between them, and the resulting fragmentation of the breed has been detected by the use of molecular markers. The Gauloise Noire breed has been submitted to a genetic improvement programme after an initial sampling of 200 chicks from fancy breeders. Similarly, the Shek-Ki breed, originating from the Canton province in China, has been selected as a sire line to be distributed to farmers producing the 'Triple-Yellow' breed for meat consumption. Therefore, the observed heterozygote deficiency could be due to the selection performed on this breed.

In Africa, recent molecular studies on local chickens have shown that Zimbabwe populations, separated by 300–800 km, were not genetically differentiated, with an F_{ST} averaging 0.01 (Muchadeyi *et al.* 2007). Mwacharo *et al.* (2007) identified genetic subdivisions between the Kenyan/Ugandan chicken populations and the Ethiopian/Sudanese chicken populations, but, within a country, the F_{ST} values were always lower than 0.1. In our case, the Hua-Tung, Ju-Chi and Hsin-Yi breeds are scavenging chickens but they exhibited higher values of F_{ST} (0.174). This could be explained in part by some historic traits. The Hsin-Yi breed has been conserved by an ethnic tribe who believes that the pure breed has a medicinal effect (Lee 2006). In a similar way, the Ju-Chi breed is only found in mountainside orchards (Lee 2006) and exchanges with other populations are unlikely.

Heterozygosity and inbreeding will influence the diversity value of a breed. Values of aggregate diversity were more homogeneous for French breeds than for Asian breeds. The Coucou de Rennes breed had the highest value of aggregate diversity mainly because of its strong within-breed contribution. We could consider the Coucou de Rennes breed as having a representative set of genes for the French breeds. Some breeds, such as the Bourbonnaise and Houdan breeds, had low values of within-breed and aggregate diversity. Nevertheless, they had an important contribution to between-breed diversity. Those breeds could be a reservoir for some 'private' genes.

The Asian subset exhibited contrasting situations: the traditional unselected breed Ju-Chi exhibited low impact on global diversity, whereas all individuals studied could be reassigned to the Hua-Tung breed, a situation also found for Red Junglefowl. The Hua-Tung breed is a small fighting game chicken (Lee 2006), traditionally reared on the East coast and sold in Taipei (in the north) at a high price. Some

farmers may have decided to cross it with exotic game birds from Southeastern Asia to improve growth performance (Y.P. Lee, personal communication). This could explain in part, the high values for H_O and aggregate diversity, and therefore this breed could be genetically representative of all other breeds. Similarly, it was found within the AVIANDIV project that breeds which exhibited the highest heterozygosity level also exhibited the lowest mean genetic distance to the other breeds (Hillel *et al.* 2003).

Usefulness of multivariate analysis

All loci showed significant differentiation among breeds. Multivariate analysis (MCOA) allowed us to identify characteristics and differences between loci. It may detect the efficiency of each locus for breeds' consensual structuring because it takes into account the shared as well as the non-shared structure of data tables, i.e. loci, to build an axis (Laloë *et al.* 2007). Our results exhibited high differences between contribution values (Tv) in the construction of each axis showing that markers did not present the same discriminative power. Furthermore, heterogeneous values of similarity (Cos^2) with the representation of breeds according to a given axis showed that markers such as MCW098 had different discriminative behaviour and exhibited different typologies. Positive correlations found between Tv, Cos^2 and F-statistics are in agreement with results obtained in previous studies. Bowcock *et al.* (1994) showed that measures of marker variability, such as H_E and F_{ST}, might be indicators of the power of individual loci to discriminate among populations. Bjornstad & Roed (2002) have found a negative correlation between assignment error rate and F_{ST} or number of alleles. Discriminatory efficiency for loci with high heterozygosity has been reported by Rosenberg *et al.* (2001) using 27 AVIANDIV microsatellites; however, we did not found any correlation between Tv and H_E.

The positions of breeds in the synthetic space created by the MCOA analysis, as well as on their positions in the NEIGHBORNET graphs, may be used to unravel historic relationships between breeds. The Asian breeds (including the three breeds coming from Taiwan Island) were positioned in the left-upper part of the graphic, and were well separated from each other. It seems that the first axis separated Asia from Europe. The Géline de Touraine, Gasconne and Noire de Challans breeds clustered with the Red Junglefowl on the left side and the Marans breed was situated on the left side. It is well known (Coquerelle 2000) that the Géline de Touraine, the Noire de Challans and the Marans breeds had been crossed in their past with the Langshan breed, which originated from the North of China. For the Gasconne breed, crossbreeding with Langshan, Orpington or Barbezieux breeds has been hypothesized. Our results seem to show that crossbreeding with the Langshan breed is more probable. The Bourbonnaise breed is said to have the Brahma breed (originating from China and India) as an ancestor, and it

was grouped with the Chinese Shek-Ki breed. This was also the case for the Russian Orlov breed, which has Asiatic origins (Moiseeva *et al.* 1994). The use of the Langshan breed was also proposed to have taken place for the Coucou de Rennes breed (Coquerelle 2000), which seems to be supported by our results. At the beginning of the 20th century, breeders of the Bresse Blanche and the Gauloise Grise breeds decided to introduce some cocks from regions having imported Asian chickens (Voitellier 1925); indeed, the MCOA suggested that current Bresse Blanche breed may have relationships with Asian breeds, but this was not observed for the Gauloise Grise breed.

Combination of data sets

This study may be one of the first comparing traditional selected or unselected breeds of two distant regions, France and Eastern Asia, within the AVIANDIV framework. Such a study could be carried out using a common set of 14 markers exhibiting high polymorphism. AVIANDIV genotypes and DNA samples for calibration being available, other studies could use it without high supplementary cost. Because of high productivity policies and selection, many breeds are known to be endangered. There is a strong need to evaluate diversity for defining conservation policies. Classical methods such as the Weitzman approach are based on between-breed diversity only, but both the approach suggested by Ollivier & Foulley (2005) and the multivariate analysis propose to combine within- and between-breed diversity. Combining such approaches with the use of the AVIANDIV framework allowed us to show with 14 loci that the Marans breed had the highest contribution to diversity within French breeds, and had a similar importance at the European level represented by AVIANDIV breeds. The Coucou de Rennes breed also appeared remarkable at both scales, national and European. Thus, the AVIANDIV framework could be used to evaluate the genetic diversity of local breeds and to help connect national and regional conservation policies. Furthermore, it should be remembered that final conservation decisions consider additional information on economic traits and cultural values, including adaptative traits and current production systems associated with the breed, to be in line with a sustainable project.

Acknowledgements

References

Belkhir K., Borsa P., Goudet J., Chikhi L. & Bonhomme F. (2000) GENETIX, *Logiciel sous Windows TM pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions: CNRSUPR. 9060.* Université Montpellier II, Montpellier, France.

Bjornstad G. & Roed K.H. (2002) Evaluation of factors affecting individual assignment precision using microsatellite data from horse breeds and simulated breed crosses. *Animal Genetics* **33**, 264–70.

Blott S.C., Williams J.L. & Haley S.C. (1999) Discriminating among cattle breeds using genetic markers. *Heredity* **82**, 613–9.

Bowcock A.M., Ruiz-Linares A., Tomfohrde J., Minch E., Kidd J.R. & Cavalli-Sforza L.L. (1994) High-resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455–67.

Chessel D., Dufour A.B. & Thioulouse J. (2005) The ADE4 package-I: one-table methods. *R-News* **4**, 5–10.

Coquerelle G. (2000) *Les poules Diversité Génétique Visible.* INRA Editions, France.

Derban S., Foulley J.-L. & Ollivier L. (2002) WEITZPRO: *A Software for Analysing Genetic Diversity.* INRA, Paris.

Hillel J., Groenen M.A.M., Tixier-Boichard M. *et al.* (2003) Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genetics Selection Evolution* **35**, 533–57.

Huson D.H. & Bryant D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254–67.

Laloë D., Jombart T., Dufour A.B. & Moazami-Goudarzi K. (2007) Consensus genetic structuring and typological value of markers using Multiple Co-inertia Analysis. *Genetics Selection Evolution* **39**, 545–67.

Lee Y.P. (2006) Taiwan country chicken: a slow growth breed for eating quality. In: *Symposium, 2006 Scientific Cooperation in Agriculture between Council of Agriculture (Taiwan, R.O.C.) and Institut National de Recherche Agronomique (France)* (Ed. by C.W. Liao, B.L. Shih, M.L. Lee, A.L. Hsu & Y.S. Cheng), *Technical Bulletin of Livestock Research Institute* **103**, 121–32.

Moazami-Goudarzi K. & Laloë D. (2002) Is a multivariate consensus representation of genetic relationships among population always meaningful? *Genetics* **162**, 473–84.

Moiseeva I.G., Semenova S.K., Bannikova L.V. & Filippova N.D. (1994) Genetic structure and origin of an Old Russian Orloff chicken breed. *Genetike* **30**, 681–94.

Muchadeyi F.C., Eding H., Wollny C.B.A., Groeneveld E., Makusa S.M., Shamseldin R., Simianer H. & Weigend S. (2007) Absence of population structuring in Zimbabwe chicken ecotypes inferred using microsatellite analysis. *Animal Genetics* **38**, 332–9.

Mwacharo J.M., Nomura K., Hanada H., Jianli H., Hanotte O. & Amano T. (2007) Genetic relationships among Kenyan and other East African indigenous chickens. *Animal Genetics* **38**, 485–90.

Nei M., Tajima F. & Tateno Y. (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153–70.

Ollivier L. & Foulley J.L. (2005) Aggregate diversity: new approach combining within- and between-breed genetic diversity. *Livestock Production Science* **95**, 247–54.

Supplementary data

Table S1. Summary of AVIANDIV breeds and their polymorphism measures.

Breed name	Breed code	Breed category	H_{nb}	H_{obs}	Average number per locus (14 Loci)			F_{IS} (14 Loci)
			14 loci		Alleles	Effective Alleles	Specific Alleles	
Fayoumi	FY	TS, Egypt	0.347	0.339	2.50	1.52	0	0.024
Bedouin	BD	TU, Israel	0.560	0.541	4.50	2.23	1	0.034
Icelandic landrace	IC	TU, Iceland	0.582	0.528	4.14	2.27	3	0.072
Old Scandivanian reference population	OS	TS, Denmark[a]	0.568	0.449	3.36	1.82	0	0.022
Jaerhoens	JH	TS, Norway	0.459	0.301	2.43	1.40	0	-0.043
Padovana	PD	TS, Italy	0.289	0.363	2.79	1.78	2	0.190
Transylvanian naked neck	TNN	TS, Hungary	0.447	0.387	2.79	1.77	0	0.126
Green-legged partridge	GP	TS, Poland	0.442	0.362	3.21	1.76	2(1)	0.175
Orlov	Orl	TS, Russia	0.438	0.586	3.57	2.24	1	-0.041
White-egg layer line A	WE-L A	CS	0.563	0.318	2.29	1.47	0	0.018
Broiler sire line B	Bs B	CS	0.324	0.512	3.64	2.14	0	0.056
Brown-egg layer line C	BE-L C	CS	0.542	0.336	2.64	1.54	0	0.057
Brown-egg layer line D	BE-L D	CS	0.356	0.448	3.14	1.92	0	0.079
Broiler dam line D	Bd D	CS	0.486	0.650	4.86	2.81	2	0.009
Red Junglefowl	*Gallus*	Feral, Thailand	0.656	0.638	5.14	2.69	5	0.003

TU: Traditional Unselected, TS: Traditional standardised for morphological traits; TSP : traditional standardized breed selected for a production trait, CS: Commercial Selected. Unbiased (H_{nb}) and observed (H_{obs}) Heterozygosity; number of alleles per locus; F_{IS} & F_{ST}

a: Whitte Leghorn population use as control line for a selection experimetn in laying hens.

Table S2. Polymorphism measures: 14 first locus used on French ,Asian and AVIANDIV breeds and the remaining 8 loci on both French and Asian breeds.

Locus	No. of Alleles identified				Allele size	France and Taiwan			France, Taiwan and AVIANDIV		
	Asia	France	AVIANDIV	Total	Range	Mean of the H_{Obs} per breed (± standard error)	F_{IS}	F_{ST}	Mean of the H_{Obs} per breed (± standard error)	F_{IS}	F_{ST}
ADL112	3	6(1)	9(3)	10	120-134	0.447 ± 0.167	0.032	0.284	0.386 ± 0.180	0.042	0.268
ADL268	5	5	6(1)	6	104-116	0.577 ± 0.126	0.016	0.199	0.340 ± 0.150	0.005	0.254
LEI166	5	4	4(1)	6	356-370	0.561 ± 0.074	0.126	0.163	0.561 ± 0.156	0.124	0.297
LEI234	15(1)	18(4)	17(3)	23	216-354	**0.733** ± 0.069	-0.008	0.197	**0.681** ± 0.122	0.049	0.263
MCW014	5(1)	6(1)	8(2)	11	164-185	0.445 ± 0.138	0.063	0.334	0.238 ± 0.182	0.069	0.393
MCW037	3	3	4(1)	4	154-160	0.509 ± 0.188	0.015	0.224	0.498 ± 0.184	0.068	0.330
MCW067	4	5	4	5	178-186	0.465 ± 0.168	0.062	0.256	0.399 ± 0.188	0.047	0.327
MCW069	6	9(1)	8	9	158-176	0.505 ± 0.129	-0.031	0.212	0.588 ± 0.185	-0.019	0.203
MCW098	2	2	3(1)	3	261-265	*0.307* ± 0.178	0.020	0.229	*0.043* ± 0.170	0.029	0.263
MCW111	5	5(1)	7(2)	9	96-112	0.500 ± 0.173	0.010	0.235	0.499 ± 0.200	0.022	0.284
MCW183	9(1)	7(2)	7(1)	12	296-326	0.524 ± 0.145	-0.010	0.211	0.425 ± 0.193	-0.036	0.223
MCW216	6(1)	5	5	6	136-149	0.478 ± 0.163	0.033	0.29	0.642 ± 0.168	0.019	0.271
MCW222	4	4	4	4	220-226	0.408 ± 0.178	0.054	0.249	0.530 ± 0.184	0.053	0.331
MCW330	4	7(1)	11(4)	12	256-300	0.514 ± 0.182	0.024	0.304	0.120 ± 0.224	0.054	0.365
AVERAGE	**5.4**	**6.1**	**6.9**						**0.424**	**0.037**	**0.291**
ADL278	4(1)	5(2)	-	6	114-124	0.547 ± 0.135	0.007	0.275			
LEI094	10(2)	12(4)	-	14	247-287	0.645 ± 0.111	0.016	0.189			
MCW034	9(1)	13(5)	-	14	214-248	0.604 ± 0.145	-0.012	0.291			
MCW078	4	4(1)	-	5	135-143	0.487 ± 0.209	0.055	0.267			
MCW081	4	7(3)	-	7	114-135	0.584 ± 0.073	0.018	0.185			
MCW206	6(1)	7(2)	-	8	223-247	0.452 ± 0.178	0.045	0.261			
MCW248	4	5(1)	-	5	215-223	0.315 ± 0.215	0.043	0.282			
MCW295	6(1)	7(2)	-	8	88-106	0.557 ± 0.191	0.057	0.212			
AVERAGE	**5.6**	**6.6**				**0.507**	**0.027**	**0.242**			

Total number of breeds (number of private alleles with frequencies > 0,1); Observed (H_{obs}) Heterozygosity

Piry S., Alapetite A., Cornuet J.M., Paetkau D., Baudouin L. & Estoup A. (2004) GENECLASS 2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536–9.

Qu L., Yang N., Li X., Xu G., Chen K., Yang H., Zhang L., Wu G., Hou Z. & Xu G. (2004) Genetic diversity of 78 Chinese chicken breeds based on microsatellite analysis. Genetics session, paper G7. WPC, Istanbul.

R Core Development Team (2006) R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: http://www.R-project.org.

Reynolds J., Weir B. & Cockerham C. (1983) Estimation for the conancestry coefficient basis for a short-term genetic distance. *Genetics* **105**, 767–79.

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

Rousset F. & Raymond M. (1995) GENEPOP (version 3–4): populations genetics software for exact test and eucumenicism. *Journal of Heredity* **86**, 248–59.

Voitellier C. (1925) Les Races de Volailles. In: *Encyclopédie Agricole*, pp. 45–56. Librairie J.-B. Baillière et Fils, Paris.

Weir B.S. & Cockerman C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–70.

Wimmers K., Ponsuksili S., Hardge T., Valle-Zarate A., Mathur P.K. & Horst P. (2000) Genetic distinctness of African, Asian and South American local chickens. *Animal Genetics* **31**, 159–65.

Zhou H. & Lamont S.J. (1999) Genetic characterization of bio-diversity in highly inbred lines by microsatellite markers. *Animal Genetics* **30**, 256–64.

Supplementary material

The following supplementary material is available for this article online from http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2052.2008.01703.x.

Table S1 Summary of the AVIANDIV breeds and their polymorphism measures.

Table S2 Polymorphism measures on 14 loci in French, Asian and AVIANDIV breeds and on the remaining eight loci in both French and Asian breeds.

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors.

6.5 Genetic diversity of local chicken breeds

6.5.1 An Overview

The maintenance of high polymorphism among and within breeds ensures the capacity to supply future commercial demands and resistance against potential epidemics. Therefore local breeds may represent an important reservoir of genes, and this is why recent studies have tried to evaluate the genetic diversity of local breeds. Wimmers *et al.* (2000) were the first to analyse and compare genetic diversity within local chicken populations. The gene diversity values weere high for local chickens from Bolivia chicken (0.53-0.59),Tanzania chicken (0.54-0.67), Nigeria (0.55-0.64) and Cameroon (0.53). Such values were within the same range as those obtained by Qu *et al.* (2004) on Chinese local breeds (0.505 – 0.678).

Wimmers *et al.* (2000) also observed that most of the local chicken populations were not at Hardy-Weinberg equilibrium. They had sampled chickens from village backyards and regrouped animals according to ecotypes or within regions. Therefore, one assumption could be that local chickens were organised in subdivided populations.

Msoffe *et al.* (2005) studied the Tanzanian chicken population following nine ecotypes. Their findings indicate the existence of a main genotype within ecotypes. However the sharing of genetic materials between ecotypes was also evident. Gene flow between ecotypes could be non-negligible in some cases that it would not allow differentiation of subpopulations as it has been showed by Muchadeyi *et al.* (2007) for chicken populations from Zimbabwe. Ji *et al.* (2005) observed a genetic differentiation among 12 Chinese local chicken breeds of about 16.7 %, and so they were able to highlight two main groups: a light type group and a heavy one. However, some of the local breeds showed large genetic exchanges within each other such as the Xiaoshan and the Luyuan chickens, or appeared as a mixed population such as the Tibetan chickens. In East Africa, within single countries genetic differentiation of local chickens was lower (Mwacharo *et al.,* 2007) : Kenya (0.3-3.7 %), Uganda (2.7 %) and Ethiopia (4.7 %). Mwacharo *et al.* (2007) estimated that a distance higher than 600 km appeared to be a significant barrier to gene flow and so allowed genetic differentiation between chicken populations. This distance might be shorter in cases where topology reduces communication between valleys and therefore also animal exchanges, as is the case in the Ha Giang province. Therefore, the genetic structure of the Ha Giang chicken population was studied focusing on the indigenous H'mong chicken compared to wild, local Asian and commercial breeds.

6.5.2 Genetic analysis of local Vietnamese chickens

GENETIC ANALYSIS OF LOCAL VIETNAMESE CHICKEN PROVIDES EVIDENCE OF GENE FLOW BETWEEN DOMESTIC AND WILD POPULATIONS

Submitted to Animal Genetics

Abtracts

Chicken is one of the most important component of livestock in the world whereas 52 % of avian breeds are already extinct or endangered. In order to characterise the H'mong chicken breed i.e Black Chicken, genetic and phenotypic analyses of the chicken population from the Vietnamese Ha Giang province was realised. A total of 2 487 animals were described for 19 phenotypic characters. From this sample set, 1 082 animals were genotyped for 18 ISAG microsatellite markers. The Vietnamese population was jointly analysed with 14 other breeds including wild population (*G. g. gallus*), commercial and Asian breeds. The Ha Giang chicken population showed high gene diversity value (0.619) and high number of private alleles (33). Analysis of phenotypes revealed a geographic distribution of Black Chickens while the bayesian admixture analysis within the province did not show any genetic structure in the population. Analysis of the 14 breeds with the Vietnamese population showed that four communes from Ha Giang province had same admixture patterns as the wild population *G. g. gallus*.

Introduction

The domestic chicken is distributed over the world, it provides cheap animal proteins and it has many cultural or religious purposes. With such an importance, many studies had tried to elucidated the evolutionary history of the chicken. With studies on mitochondrial control regions, Akishinonomiya *et al.* (1994, 1996), concluded on a monophyletic and single domestication event of *Gallus g. gallus* that might have taken place in Thailand and has led to todays chicken breeds. Recently, this monophyletic and single domestication has been questioned by new molecular studies. Liu *et al.* (2006) analysed a mitochondrial DNA hypervariable segment and compared 106 chicken populations including four sub-species of *Gallus g.*. Two matriarchal lineages were observed: one in India in which the most represented wild ancestor was the *Gallus g. gallus*, and one in the Yunnan province, China, where the main ancestors were the *Gallus g. spadiceus* and *Gallus g. jabouillei.* Thus, most of the wild ancestors seemed to be enrolled in the genesis of modern domestic chickens and in multiple domestication origins in South and South-East Asia (Yunnan and surrounding areas). Moreover, recent studies showed a possible introgression of others *Gallus sp.* into domesticated chickens (Nishibori *et al.* 2003, Eriksson *et al.* 2008).

Granevitze *et al.* (2007) found a surprisingly high genetic diversity in the H'mong chicken breed from the northern Vietnamese province of Son La. This breed, also called the Black Chicken has black bones, black skin and black shank. However, the H'mong chicken breed is assumed to be mainly raised in the neighbouring province of Ha Giang bordered by China (Yunnan and Guanxi provinces). The Ha Giang province's terrain is strongly partitioned, resulting in reduced communication between villages (see, Castella *et al.* 2005, for an analysis of such reduction in a neighbouring Vietnamese province). In the Ha Giang province, the forest is an important component of the sustainability of the farming systems since it provides fodder for ruminants and where scavenging chickens can influenced. In order to characterise the H'mong chicken breed, we carried out a large survey in the Ha Giang province. In such a context, two approaches were used: i) the animals were phenotypically described and genotyped using ISAG microsatellites for genetic structure analysis; ii) we compared the Ha Giang chicken population with wild, Asian and commercial breeds.

Material and Methods

Sampling procedures

The remote Ha Giang province (22º08' - 23º19' N ; 104º33'- 105º33' E) is a mountainous area with elevations up to 2 400 m above sea level and a temperate climate. From an administrative point of view, this province is split into districts (11 in total), communes (193) and villages (1715). A total of 24 ethnic groups inhabit the Ha Giang province and the H'mong ethnic group is the main one, followed by the Tay and the Dao ethnic groups. H'mong people generally live in areas of higher altitude than others, up to 1 500 m. Sampling sites were distributed over all the province as shown in Figure 1. Blood samples were collected from 1 082 animals, among a total of 2 487 animals which were sampled for phenotypic characterisation and measurements. These animals were sampled from all districts of the province and from 30 communes. Within each district, 2 to 4 communes were chosen on the basis of official statistics, in order to represent all ethnic groups and to catch a wide range of environmental conditions and farming systems. The geographical origins of genotyped animals are shown in Table 1. An average of 5.6 animals per village and 1.4 animals per household were genotyped. Each commune sample was considered as an *a priori* basic unit for polymorphism analysis.

Figure 1. Geographic distribution of Ha Giang samples

.

Fourteen other populations were used for comparison: 3 wild populations of *Gallus g. gallus*, 7 local Asian breeds, previously studied by Berthouly *et al.* (2008), and 5 commercial breeds (Table 1).

Microsatellite markers and genotyping procedures

For all Vietnamese chickens, genomic DNA was extracted from blood samples using the QIAamp Kit from QIAGEN and 20 microsatellites from ISAG were used for genotyping in the NIAH laboratory, in Ha Noi. The animals from the 14 populations used for comparison were previously genotyped by the LABOGENA laboratory (France). In order to calibrate allele sizes between both laboratories a set of 17 references animals within the 14 external populations was analysed jointly with the animals from the Ha Giang province. Applying the same method as Berthouly *et al.* (2008), the difference in allele size observed between laboratories was easily interpretable for 18 markers. For the present study, only these 18 markers, for which allele sizes were adjusted to be consistent from one laboratory to the other, were used (Table S1).

Molecular polymorphism analysis

The presence of null alleles was tested using FreeNA (Chapuis & Estoup, 2007). We considered only loci with r ≤ 0.2 to be potentially problematic for calculations. Allele frequencies, number of alleles, expected and observed heterozygosity were calculated using GENETIX (Belkhir *et al.,* 2000). GENEPOP (Rousset & Raymond, 1995) was used to compute *F-statistics* (Weir & Cockerman, 1984) and departure from Hardy-Weinberg equilibrium using exact tests. Test significance was corrected with sequential Bonferroni correction on loci.

We performed a Multivariate Coinertia Analysis (MCOA) for the 30 communes from the Ha Giang province. This methodology permits the estimation of genetic structure similarity across markers (Laloë *et al*. 2007, Berthouly *et al.* 2008).

At first, we investigated the genetic structure of the Ha Giang population using a Bayesian clustering procedure implemented in STRUCTURE (Pritchard *et al*., 2000), with the admixture method and correlated allele frequency version of the programme (Falush *et al.,* 2003). Assuming that the data set could be represented by *K* separate genetic clusters (*K*=1 to 10), we did 15 runs for each different values of *K* with 10^5 iterations following a burn-in period of 30 000. The values for the number of clusters (*K*) were first assessed according to Pritchard *et al.* (2000) to ensure that *K* > 1 and then, according to Evanno *et al.* (2005). Then, we analysed the clustering of the Ha Giang population with the others fourteen breeds previously studied following the procedure from Rosenberg *et al.* (2001). We performed 100 runs from *K*=2 to

K=10 with a 60 000 iterations and a burn-in period of 40 000 using the admixture-correlated frequency model. We used CLUMPP (Jakobsson & Rosenberg, 2007) to compute similarities for the different *K*, using the similarity function G'. The mean Q-matrix over all runs was computed for each *K*.

Phenotypic description

A set of phenotypic traits was chosen on the basis of easy identification and breed standards (Table 2). For traits known to be determined by a single biallelic gene with one recessive allele, the frequency of this allele was estimated as the square root of the frequency of animals with the "recessive" phenotype. For more complex traits, only phenotype frequencies were used. Allelic frequencies and phenotype frequencies were analysed by Principal Component analysis (PCA) as implemented in ADE 4 (Chessel *et al.,* 2004; Dray & Dufour, 2007) of the R software (R development core team, 2007).

Table 2. Phenotypic traits described

Single Locus traits	Name	Alleles	Recessif Genotype	Recessif Phenotype
c	Recessive white	C+; c	[cc]	no coloration, white chicken
I	Dominant white	I; i+	[i+i+]	presence of black
Pg	Pattern feather	Pg; pg+	[pg+pg+]	no pattern
H	Silkiness	H+; h	[hh]	silkiness
F	Frizzled plumage	F; f+	[f+f+]	frizzled plumage
Cr	Crest	Cr; cr+	[cr+cr+]	without crest
vh	Vulture hocks	Vh+; vh	[vhvh]	without vulture hocks
Mb	Muffs and beards	Mb; mb+	[mb+mb+]	no muffs and beards
Double Loci traits	**Trait**	**Loci/Alleles**	**Genotypes**	**Phenotypes**
R & P	Comb type	R; r+	[r+r+,p+p+]	Simple Comb
		P; p+	[R-,p+p+]	Rose Comb
			[r+r+,P-]	Pea Comb
			[R-,P-]	Walnut Comb
W & Id	Shank colour	W+; w	[id+id+,ww]	green
		Id; id+	[id+id+,W+-]	blue grey
			[Id-,W+-]	white
			[Id-,ww]	yellow
Complex traits	**Code**		*6.5.3 Phenotypes*	
Black shank	BlackT		Black	
Eyes colour	BlackE, BrownE, OranE		Black / Brown / Orange	
Skin colour	BlackS, YellowS, WhiteS		Black / Yellow / White	
Comb colour	BlackCmb, CmbRB, CmbRv, CmbRc		Black / Red-Black/ Dark Red/ Light Red	
Ear Colour	BlackO, RedO, WhiteO, WRO		Black / Red / White / White and Red	
Shank without feathers	TnF		No feathers	
Polydactily	TnP		no polydactily	

Table 1. Summary of genetic diversity measures.

Code	Breed name	Origin/District	Nv	Ni	H_{Exp}	H_{Obs}	A	Ae	F_{IS}	D_{HWE}
BS_LD	Broiler sire Line D	Selected, commercial		30	0.46±0.2	0.45±0.21	3.2	1.8	0.01	0
BS_LC	Broiler Sire Line C	Selected, commercial		25	0.47±0.18	0.49±0.19	2.9	1.9	-0.03	0
BD_LB	Broiler dam Line B	Selected, commercial		25	0.47±0.21	0.47±0.22	3.0	1.9	0	0
BE_LC	Brown-egg Layer C	Selected, commercial		25	0.41±0.24	0.36±0.23	2.8	1.7	0.11	1
WE_LA	White-egg Layer A	Selected, commercial		25	0.27±0.22	0.27±0.22	2.0	1.4	-0.02	0
Gg1	*G. g. gallus*	Wild, Thailand		16	0.60±0.18	0.58±0.25	4.1	2.5	0.05	1
Gg2	*G. g. gallus*	Wild, Thailand		15	0.62±0.19	0.66±0.26	4.3	2.6	-0.06	0
Gg3	*G. g. gallus*	Wild, Vietnam		6	0.57±0.23	0.70±0.33	2.7	2.3	-0.27	0
HT	Hua-Tung	TUC, Taiwan		45	0.55±0.17	0.57±0.20	3.1	2.3	-0.02	0
JC	Ju-Chi	TUC, Taiwan		48	0.40±0.15	0.42±0.17	3.1	1.7	-0.05	0
KM	Quemoy	TUC, Taiwan		47	0.49±0.19	0.47±0.17	3.1	2.0	0.04	0
HY	Hsin-Yi	TUC, Taiwan		47	0.50±0.13	0.51±0.13	3.3	2.0	-0.02	0
SK	Shek-Ki	TUS, China		46	0.44±0.18	0.40±0.20	2.7	1.8	0.10	0
NG	Nagoya	TUS, Japan		48	0.42±0.18	0.43±0.19	3.1	1.7	-0.02	0
Ha Giang population - no managed population										
HG1	Lung-Pu	Dong-Van	6	36	0.63±0.19	0.49±0.19	5.8	2.7	0.22	4
HG2	Ma-Le	Dong-Van	2	9	0.61±0.18	0.54±0.23	3.8	2.6	0.13	0
HG4	Lung-Cu	Meo-Vac	7	44	0.62±0.17	0.52±0.19	5.7	2.6	0.17	4
HG7	Thai Phin Tung	Dong-Van	2	6	0.70±0.19	0.62±0.26	4.4	3.3	0.12	0
HG16	Pho-Cao	Dong-Van	8	39	0.57±0.19	0.52±0.19	5.1	2.3	0.08	1
HG20	Bach-Dich	Yen-Minh	8	23	0.64±0.18	0.59±0.16	5.5	2.8	0.08	1
HG25	Na-Khe	Yen-Minh	4	22	0.57±0.19	0.49±0.19	4.5	2.3	0.15	1
HG40	Tat Nga	Meo-Vac	7	26	0.69±0.13	0.57±0.17	6.0	3.2	0.18	1
HG48	Khau-Vai	Meo-Vac	8	47	0.61±0.15	0.53±0.16	5.7	2.5	0.13	1
HG49	Tung-Vai	Quang-Ba	8	52	0.61±0.16	0.55±0.17	5.5	2.5	0.09	1
HG56	Quyet-Tien	Quang-Ba	8	24	0.58±0.18	0.52±0.21	4.9	2.4	0.11	0
HG61	Lung-Ho	Yen-Minh	7	37	0.58±0.17	0.54±0.19	5.2	2.4	0.08	0
HG65	Du-Gia	Yen-Minh	5	9	0.68±0.14	0.59±0.27	4.2	3.1	0.14	0
HG72	Ngoc-Duong	HG prefecture	5	16	0.64±0.14	0.59±0.18	4.4	2.7	0.07	0
HG75	Giap-Trung	Bac-Me	8	52	0.61±0.17	0.53±0.20	5.7	2.6	0.14	4
HG85	Po-Lo	Hoang Su Phi	7	90	0.63±0.15	0.55±0.17	6.0	2.7	0.13	3
HG88	Minh-Ngoc	Bac-Me	5	32	0.64±0.15	0.57±0.20	5.8	2.8	0.11	3
HG91	Nan-Xin	Xin-Man	6	64	0.61±0.16	0.56±0.18	5.7	2.5	0.08	1
HG95	Phu-Linh	HG prefecture	3	32	0.61±0.14	0.51±0.19	5.1	2.6	0.17	1
HG103	Chi-Ca	Xin-Man	8	88	0.65±0.14	0.58±0.14	6.3	2.8	0.10	4
HG108	Thuong-Son	Vi-Xuyen	7	16	0.62±0.17	0.53±0.21	4.8	2.6	0.14	1
HG110	Po Ly Ngai	Hoang Su Phi	4	49	0.61±0.16	0.56±0.17	5.6	2.5	0.09	1
HG113	Yen-Cuong	Bac-Me	8	52	0.62±0.17	0.58±0.19	5.8	2.6	0.06	4
HG114	San Sa Ho	Hoang Su Phi	4	44	0.62±0.14	0.58±0.15	5.7	2.6	0.06	3
HG140	Quang-Nguyen	Xin-Man	7	25	0.62±0.16	0.55±0.21	4.8	2.6	0.10	3
HG146	Trung Thanh	Vi-Xuyen	7	14	0.61±0.16	0.54±0.21	4.8	2.6	0.13	0
HG157	Tan Nam	Quang-Binh	7	14	0.65±0.14	0.57±0.19	4.6	2.9	0.12	0
HG169	Quang-Minh	Bac-Quang	6	15	0.62±0.16	0.54±0.23	4.5	2.7	0.15	1
HG179	Xuan-Giang	Quang-Binh	9	88	0.64±0.13	0.55±0.15	6.2	2.8	0.14	3
HG184	Vinh-Phuc	Bqc-Quang	6	17	0.60±0.17	0.51±0.16	4.7	2.5	0.15	0

Results

Molecular polymorphism analysis

Genetic diversity among populations

No loci with r≤ 0.2 were found, and so it was assume that in the our set of marker, loci with null alleles were absent. Among the 44 chicken populations (30 Ha Giang communes plus 14 external populations), H_{Exp} ranged from 0.267 (WE-LA) to 0.697 (HG7) while H_{Obs} ranged from 0.272 (WE-LA) to 0.700 (Gg3) (Table 1). If we considered all the Ha Giang chickens as one population (HG), then on average, the global HG population showed the highest H_{Exp} (0.619) and the highest mean H_{Obs} was observed for wild populations (0.644). In the overall populations, 4.5 alleles on average were observed per population while the highest effective number of alleles was observed in the HG population (2.6). A total of 36 private alleles were found. Thirty-three belonged to the HG population. Moreover 17 alleles were only shared by the wild and the Vietnamese HG populations: 3 were shared by all wild and HG populations; 4 were shared by Gg3 and HG; and 10 were shared by Gg1, Gg2 and HG.

Within the Ha Giang province, F_{IS} averaged 0.121 with 0 to 4 loci in heterozygote deficiency. Among the 30 communes, only two were not significantly deviated from HWE (Table 1). The remaining Asian breeds and commercial breeds reached mean F_{IS} values of 0.004 and 0.15 respectively (Table 3).

Table 3. Heterozygosities, F_{IS} and F_{ST} and measures per breed categories

	Breed category	H_{Exp}	H_{Obs}	F_{IS}	F_{ST}				
					Total	Selected	Conserved	Wild	No-Managed
Total		0.576	0.528	0.078	*0.136*				
Selected	Commercial	0.415	0.408	0.015	0.290	*0.369*			
Conserved	Taiwan, China, Japan	0.468	0.466	0.004	0.222	0.317	*0.243*		
Wild	3 *G. g. gallus*	0.597	0.644	-0.091	0.177	0.326	0.247	*0.152*	
No-Managed	Vietnam	0.622	0.548	0.121	0.094	0.268	0.199	0.141	*0.037*

In the overall loci and breeds, genetic diversity occurred at about 12.9 % between breeds. For the 7 other Asian breeds, genetic differentiation between breeds averaged 24.3 % whereas for commercial lines it reached 36.9 % (Table 3). Within the Ha Giang province, only 3.7 % of the genetic diversity was due to differentiation between communes. The HG population had the lowest mean F_{ST} value with the wild populations and ranged from 0.117 (HG169) to 0.172

(HG16) whereas the smallest F_{ST} value within the wild population was observed between Gg1 and Gg2 (0.095). Four communes from HG populations (HG72, HG88, HG108, HG169) had F_{ST} values lower than 0.1 with Gg1 and only one commune (HG169) with Gg2 and none with Gg3.

MCOA analysis

The Cos^2 values, i.e. pseudo-eigenvalues, traduced similarity of population structure of the locus table with the multi-loci population structure. The mean Cos^2 values for the first two axes reached 0.387 and 0.305 respectively, indicating that a few loci exhibited discrepancies with the synthetic multi-loci positions of commune populations. The mean *TV* (Typological Value) values for the first two axes were 0.009 and 0.003 respectively. The locus MCW216 had the highest *TV* value (0.049) for the 1st axis and the locus MCW222 (0.009) for the 2nd axis. The synthetic position of commune populations did not show a clear regrouping of communes according to districts. However, the second axis separates them to an east-west geographic distribution (Figure 2).

Figure 2. Synthetic position of commune populations obtained by MCOA analysis
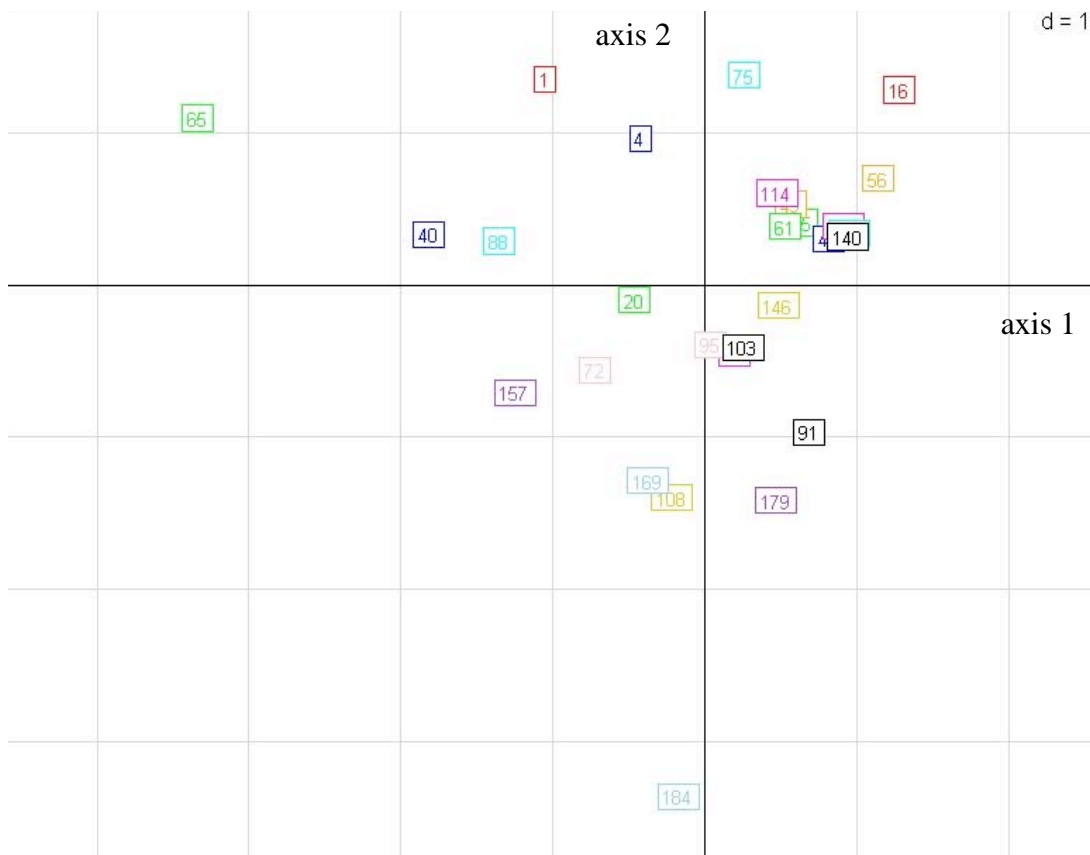
Cluster analysis

Considering only the HG population, genetic structure analysis without *a priori* with *STRUCTURE* found a K value equal to 4. However maximum mean q values per commune ranged from 0.352 (HG146) to 0.948 (HG65). Except for the commune HG65 for which all animals belong to one population, we found that within a given commune animals belong preferentially to 2 to 4 populations (Table S2).

Considering the 30 communes from the HG population plus the 14 breeds, results obtained by *STRUCTURE* clustering are displayed in Figure 3. Similarity coefficients between runs was highest when $K=2$ (0.99) as remarked by Leroy *et al.* (2008), for higher K values, it ranged from 0.70 ($K=8$) to 0.74 ($K=6$). The likelihood increased until K equaled 7-8 indicating that the most significant subdivisions were obtained for such values (Cañon *et al.*, 2006). At a $K=2$, two clusters appeared: the conserved-selected breeds (Cluster 1) and the no-managed Ha Giang population (Cluster 2), while the three wild populations admixed with both clusters. Starting from $K=3$, the two previous clusters remained and a new population (in Yellow) appeared representing a high ratio in the wild populations. This new third population admixed with some breeds from cluster 1 but mainly with the cluster 2 (i.e. the Ha Giang province). Since $K=4$, the Japanese NG breed separated from the other breeds. At $K=7$, two Taiwanese breeds showed high admixture rates: the HT breed admixed with the NG while the HY breed admixed with the commercial breeds. The remaining three Taiwanese breeds always clustered together. From $K=4$ to $K=7$, mainly undistinguished clusters appeared for the Ha Giang population but four communes (HG88, HG65, HG7, HG 40) always shared the same admixture pattern with the three wild species.

Phenotype analysis

The two principal components (PC) of the analysis on the 29 morphometric traits explained 53.6 % of the inertia. The values of normed scores are summarised in Table S3. The variables are clustered into four groups: three groups of colour and one group of feathers structure. On the top left of Figure 4, all variables describing black colour are grouped together except for eye colour (BlackE). The variable BlackE is grouped on the bottom left of the figure with the white colour variables and the wild alleles for shank colour (i.e. blue-grey). On the right, the yellow colour traits (yellow shank, yellow skin, orange eyes,..) are correlated with other wild alleles for feather colour (C^+, i^+, pg^+).

Figure 3. Clustering diagrams of chicken populations obtained from *K*=2 to *K*=8.
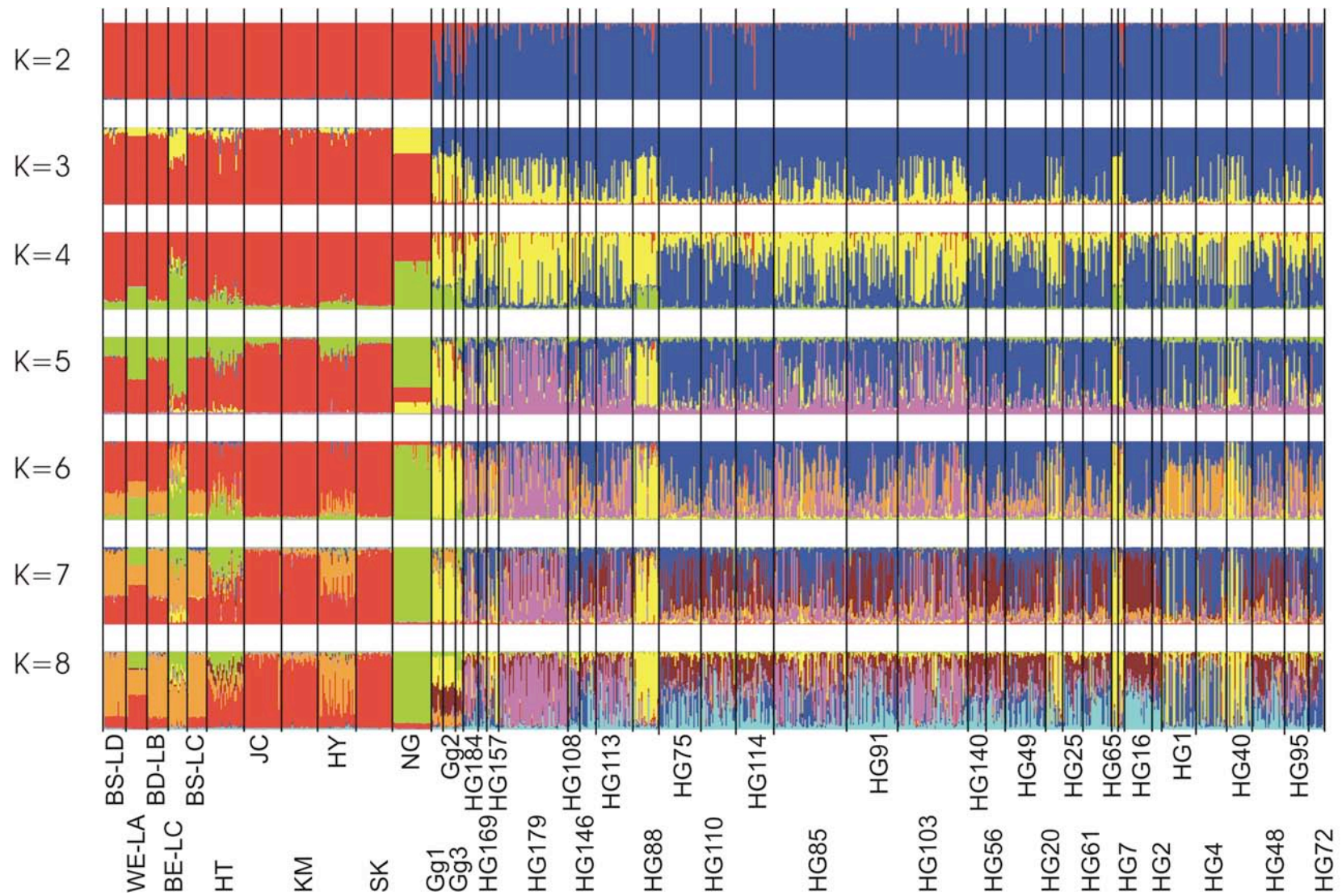
This could be considered as the description of the wild phenotype *G. g. gallus* except for the red ear trait. The red ears are described for *G. g. spadiceus* and *G. g. jabouillei* whereas *G. g. gallus* has white ears. The fourth group of variables situated at the top right described feathers structures. The trait of rose and dark-red comb are the ones contributing the most in the construction of the second PC.

Figure 4. Morphometric trait positions for the PCA analysis.
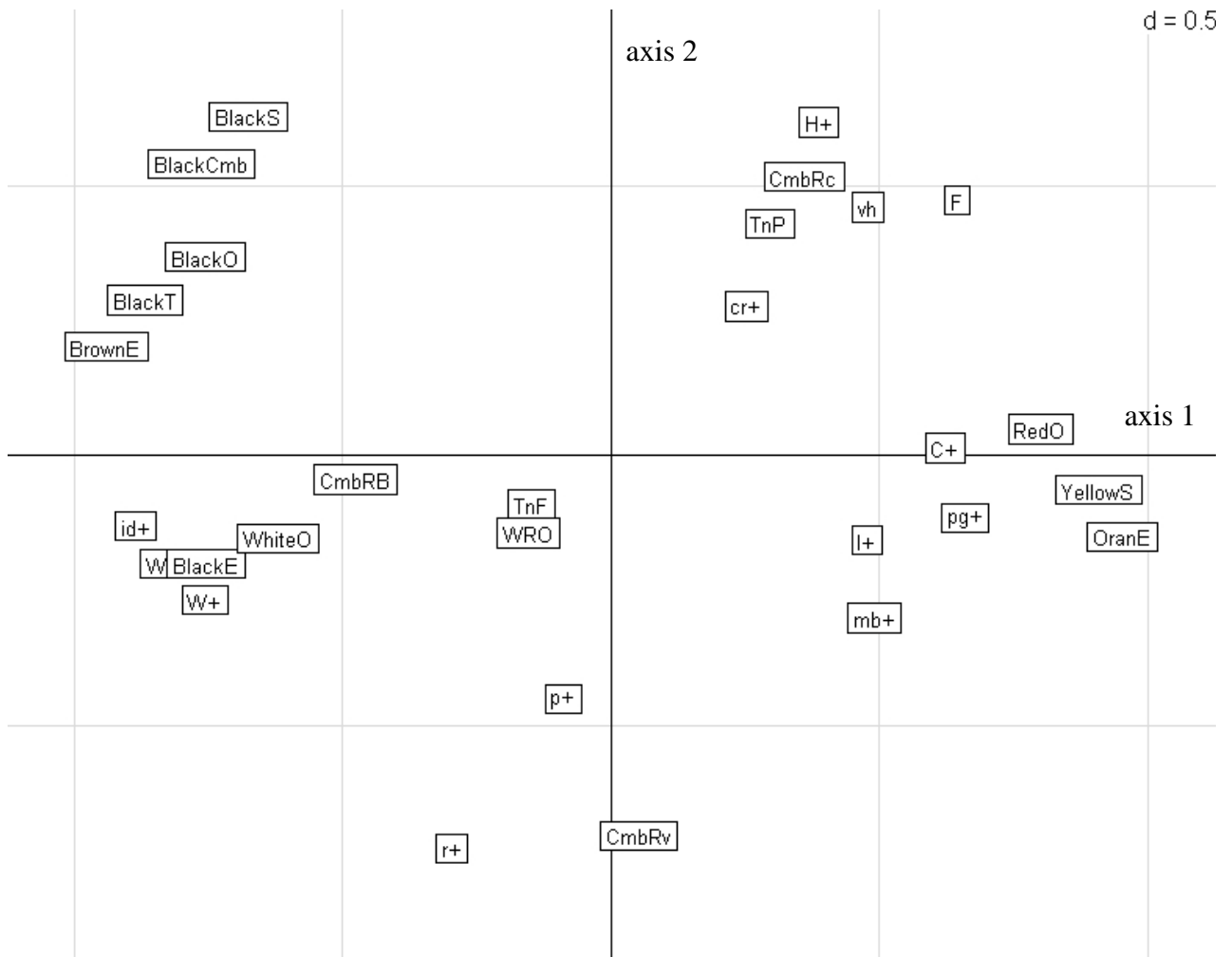
Regarding the positions of commune populations in the first factorial plan (Figure 5), the first axis acted as an East-West cline: communes from the eastern districts of Dong-Van and Meo-Vac are on the left, central communes are in the middle and communes from the western districts of the Ha Giang prefecture, Quang-Binh and Bac-Quang are on the right. Only the commune HG103 originally from the north-west part of the province clustered with the north-eastern communes (HG1-HG2) of the Dong-Van district. This is because the second axis separates communes according to phenotypes. Therefore, the Black chickens are mainly found in the HG1,

HG2 and HG103 communes. Chickens with white skin and blue-grey shanks and rare alleles such as silkiness or frizzled plumage are found in the south-eastern communes. Chicken with yellow skin and shank were in the south-west and well developed part of the province.

Figure 5. Position of commune populations obtained by PCA analysis. The colours in the online version of figure differentiate the 11 districts.
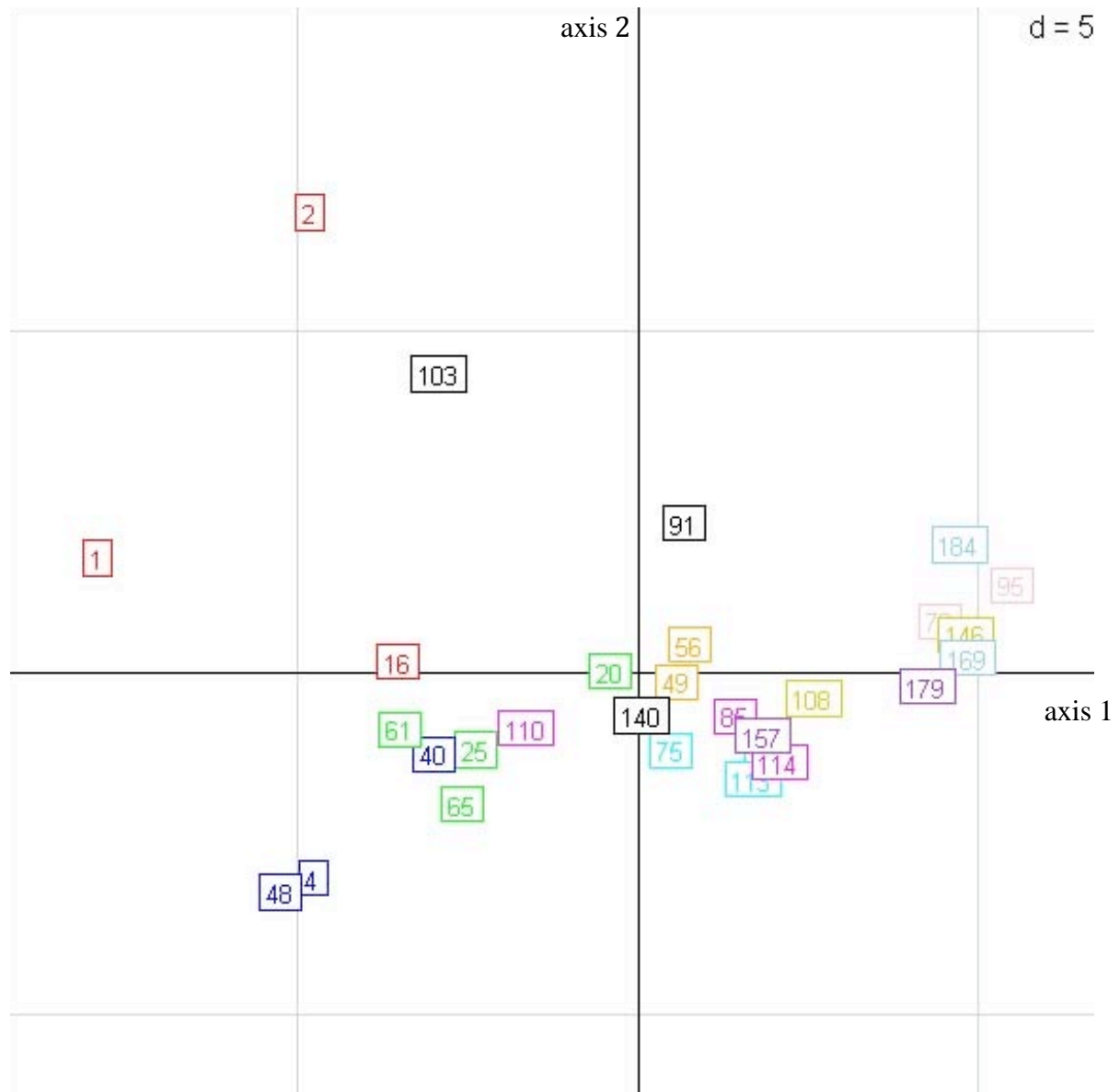
Discussion

Heterozygosity values of the local Ha Giang chicken population were in the same range as for other local breeds (Wimmers *et al.* 2000, Granevitze *et al.* 2007, Muchadeyi *et al.* 2007). Genetic differentiation of populations from a limited geographic area was low as in previous studies (Muchadeyi *et al.* 2007, Mwacharo *et al.* 2007),.

The H'mong breed, namely the Black chicken, is supposed to be raised by the H'mong ethnic group and secondly, it has phenotypic characteristics, namely black bones, skin and shank. Since the H'mong ethnic group migrated from China through the Ha Giang province (Savina, 1924) and moved southward to the Lao Cai and the Son La province, the potential area distribution of the H'mong chicken may be represented by these three provinces. Nevertheless, the Ha Giang province might be the main province for raising the H'mong chicken breed, because it is the first migratory province and the most inhabited by the H'mong community. Our results clearly enlighten two main points. Firstly, phenotypic traits allowed the separation of the Ha Giang chicken population into three geographically distributed groups: the black-skin chickens at the North-East and North-West, the white-skin, the silky and the frizzled chickens at the South-East and the yellow-skin chicken at the West. As a consequence, considering the black phenotype characteristic of the cultural description of the H'mong chicken, then three communes: Lung-Cu (HG1), Ma-Le (HG2) and Chi-Ca (HG103) showed the highest rate of black phenotypes i.e. H'mong chickens in their backyards.

Secondary, MCOA analysis did not cluster communes neither by districts nor by phenotypes. Only a East-West line was observed for the second axis. Moreover, an animal's clustering analysis showed that within a village or commune, individuals could belong to different hypothetic ancestral populations. In the hilly environment of the Ha Giang province, villages are distant and separated from each other by forest or wide land crops area. In fact, village poultry stock may behave as a small genetic unit, where related animals crossed over generations leading to high inbreeding. Since natural selection pressure is harsh, epidemic events occurred every one to three years after which householders started their new poultry stock with 2 to 8 individuals on average (householders personal communications). As a consequence, epidemic events will act as a severe bottleneck effect and as a founder effect on the village poultry stock. Whenever the number of survivors was too limiting, householders would buy a few individuals in the weekly

commune or district markets, introducing new animals that had suffered similar founding effects, into their village poultry stock. Because of the village's geographic isolation, no natural continual gene flow happened, but it is more a step-wise gene flow submitted to the rhythm of epidemic events. Therefore, the fine scale genetic analysis is a snapshot with sporadic bottleneck effects across the province in the background and highlighting recent locally introduced animals not yet admixed.

The black phenotype is due to one biallelic gene (Fm, for fibromelanosis) and the black trait is dominant, householders easily conserved this phenotype by not selling or not eating those animals. Therefore there is no need for a cross-selection scheme which could eventually lead to genetic differentiation of the black chicken. In summary, the phenotype was easy to conserve even when genome was admixing. As a consequence, and contrarily to the H'mong cattle for which genetic differentiation was found within the province (Berthouly *et al.,* 2008), no significant genetic differentiation of the H'mong chicken i.e. Black chicken was observed in the Ha Giang province. However, according to the phenotypic characterisation, if a reproductive stock for conservation needs to be done, chickens from the three previously cited communes should be chosen.

The Ha Giang chicken populations presented the highest mean number of alleles per locus. Moreover it had 33 private alleles and shared 17 alleles with the wild populations. Then, two remarks could be made. Firstly, as an older source, the Ha Giang population will typically display greater genetic diversity than a population derived more recently from it (Jorde *et al.,* 1998). A negative correlation between genetic diversity and distance from the domestication center supports the assumption of a domestication origin in the Yunnan province proposed by Liu *et al.* (2006) and the Ha Giang province as a bordering area. Secondly, individual admixture analysis revealed that four Vietnamese communes always clustered and had the same admixture pattern as the wild populations. Then our results, suggest that gene flow between the scavenging domestic chickens and wild ancestor populations occurres in the Ha Giang province. This is in accordance with the distribution area of the wild ancestors (West & Zhou 1988, Del Hoyo *et al.* 1994, Liu *et al.* 2006). However, further analyses on mtDNA should be realised in order to clarify the wild ancestor matriarchal lineage (i.e. *G. g. gallus, G. g. spadiceus, G.g. jabouillei*).

In conclusion, this fine-scale study shows direct effects of no-management practices on the local genetic populations, and shows evidence of current gene flow between wild and domesticated chickens. We believe that similar results could be obtained across South-East Asia which is the distribution area of four wild ancestors and that no-management of breeds is the main practice in these remote areas.

Acknowledgments

References

Belkhir K., Borsa P., Goudet J., Chikhi L. & Bonhomme F. (2000) GENETIX, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions: CNRSUPR. 9060. Université Montpellier II, Montpellier, France

Berthouly C., Bed'Hom B., Tixier-Boichard M., Chen C.F., Lee Y.P., Laloë D., Legros H., Verrier E. & Rognon X. (2008) Using molecular and multivariate methods to study the genetic diversity on local European and Asian chickens breeds. *Animal Genetics*, doi:10.1111/j.1365-2052.2008.01703.x

Berthouly C., Doan Phan L., Rognon X., Nhu Van T., Leroy G., Hoang Than H., Bed'Hom B., Laloë D., Vu Chi C., Nguyen Dang V., Verrier E. & Maillard J-C (2008) Revealing subpopulations and possible admixture in the Vietnamese H'mong cattle for conservation purposes. *Journal of Heredity*, submitted.

Cañon J., García D., García-Atlance M.A., Obexer-Ruff G., Lenstra J.A., Ajmone-Marsan P. & Dunner S. (2006) Geographical partitioning of goat diversity in Europe and the Middle East. *Animal Genetics*, **37:** 324-7

Castella J-C. Manh H.P., Kam P.S., Villano L. & Tronche N.R. (2005) Analysis of village accessibility and its impact on land use dynamics in a mountainous province of northern Vietnam. *Applied Geography*, **25:** 308-26

Chapuis M.P. & Estoup A. (2007) Microsatellites null alleles and estimation of population differentiation. *Molecular Biology and Evol*ution, **24:** 621-31

Chessel D., Dufour A.B. & Thioulouse J. (2004) The ade4 package-I :One –table methods, R-News 4: 5-10.

Del Hoyo J., Elliot A. & Sargatal J. (1994) *Handbook of the Birds of the World.* Vol 2. New World Vultures to Guineafowl. Lynx Edicions, Barcelona

Dray S. & Dufour A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22(4) http://www.jstatsoft.org/v22/i04/

Eriksson J., Larson G., Gunnarsson U., Bed'hom B., Tixier-Boichard M., Strömstedt L., Wright D., Jungerius A., Vereijken A., Randi E., Jensen P. & Andersson L. (2008) Identification of the Yellow Skin Gene Reveals a Hybrid Origin of the Domestic Chicken. *PLoS Genetics*,**4***: e1000010.doi:10.1371/journal.pgen.1000010.

Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE, a simulation study. *Molecular Ecology,* **34:** 2611-20

Falush D., Stephens M. & Pritchard J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics,* **164:** 1567-87

Akishinonomiya A., Miyake T., Sumi S., Takada M., Ohno S. & Kondo N. (1994) One subspecies of the red jungle fowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proceedings in National Academy of Science USA*, **91:** 12505-9

Akishinonomiya A., Miyake T., Takada M., Shingu R., Endo T., Gojobori T, Kondo N. & Onho S. (1996) Monophyletic origin and unique dispersal patterns of domestic fowls. *Proceedings in National Academy of Science USA*, **93:** 6792-5

Granevitze Z., Hillel J., Chen G.H., Cuc N.T.K., Feldam M., Eding H. & Weigend S (2007) Genetic diversity within chicken populations from different continents and management histories. *Animal Genet*ics, doi:10111/j1365-2052.2007.01650

Jakobsson M. & Rosenberg N. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics,* **23:** 1801-6.

Jorde L.B., Bamshad M. & Rogers A.R. (1998) Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *Bioessays*, **20:** 126-36

Laloë D., Moazami-Goudarzi K. & Chessel D. (2002) Contribution of individual markers to the analysis of the relationships among breeds by correspondence analysis. 7th World Congress on Genetics Applied to Livestock Production, August 19-23, 2002.Communication 26-06. Vol 33, p 493-496. Montpellier, France

Laloë D., Jombart T., Dufour A.B. & Moazami-Goudarzi K. (2007). Consensus genetic structuring and typological value of markers using Multiple Co-inertia Analysis. *Genetic, Selection & Evolution* **39:** 545-67.

Leroy G., Verrier E., Callèle L., Mériaux J.C., Ricard A. & Rognon X. (2008) Genetic of a large set of horse breeds raised in France assessed by microsatellite polymorphism. *Genetics Selection Evolution*, submitted

Liu Y.P., Wu G.S., Yao Y.G., Miao Y.W., Luikard G., Baig M., Beja-Pereira A., Ding Z.L., Palanichamy M. & Zhang Y.P. (2006) Multiple maternal origins of chickens: Out of the Asian Jungles. *Molecular Phylogenetics and Evolution,* **38:** 12-19

Muchadeyi F.C., Eding H., Wollny C.B.A., Groeneveld E., Makusa S.M., Shamseldin R., Simianer H. & Weigend S. (2007) Absence of population substructuring in Zimbabwe chicken ecotypes inferred using microsatellites analysis. *Animal Genetics*, **38:** 332-39

Mwacharo J.M., Nomura K., Hanada H., Jianli H., Hanotte O. & Amano T. (2007) Genetic relationships among Kenyan and other east African indigenous chickens. *Animal Genetics,* **38:** 485-90

Nishibori M., Shimogiri T., Hayashi T. & Yassue H. (2005) Molecular evidence for hybridization of species in the genus *Gallus* except for *Gallus varius. Animal Genetics,* **36:** 367-75

Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945–59

R Development Core Team, (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org./

Rosenberg N.A., Burke T., Elo K., Feldman M.W., Friedlin P.J., Groenen M.A.M, Hillel J., Mali-Tanila A., Tixier-Boichard M., Vignal A., Wimmers K. & Weigend S. (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics*, **159:** 699-713

Rousset F. & Raymond M. (1995) GENEPOP (version 3-4): population genetics software for exact test and eucumenicism. *Journal of Heredity*, **86:** 248-59

Savina F.M. (1924) Histoire des Miao. Hong Kong. Societé des Missions Etrangères.

Weir B.S. & Cockerman C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution,* **38**: 1358-70

West B. & Zhou B.X. (1988). Did chickens go north? New evidence for domestication. *Journal of Archaeological science*, **15:** 515-533

Wimmers K., Ponsuksili S., Hardge T., Valle-Zarate A., Mathur P.K. & Horst P. (2000) Genetic distinctness of African, Asian and South American local chickens. *Animal Genetics,* **31:** 159-65

Supplementary data

Table S1. Summary of polymorphic measures

Locus	Range	A	Ap	H_{Exp}	H_{Obs}	D_{HWE}	E_{HWE}	F_{IS}	F_{ST}	F_{IT}
ADL112	120-130	6	0	0.528±0.148	0.500±0.151	4	0	0.079	0.120	0.189
ADL268	100-114	6	0	0.629±0.100	0.628±0.102	0	0	0.014	0.190	0.201
ADL278	111-127	8	4 (2xHG,Gban,Gg1)	0.560±0.123	0.507±0.105	2	0	0.103	0.196	0.279
MCW034	214-248	16	4 (HG)	0.706±0.153	0.670±0.150	4	0	0.062	0.164	0.216
MCW037	151-157	4	0	0.622±0.130	0.541±0.130	3	0	0.139	0.106	0.230
MCW067	169-179	6	1(Gban)	0.638±0.147	0.601±0.150	1	0	0.067	0.100	0.160
MCW069	155-173	10	1(HG)	0.719±0.171	0.679±0.173	0	1	0.040	0.091	0.128
MCW078	134-144	6	2(HG)	0.433±0.183	0.360±0.178	10	0	0.229	0.147	0.342
MCW081	110-134	8	2(HG)	0.455±0.165	0.440±0.159	2	0	0.068	0.217	0.271
MCW098	252-258	4	2(HG)	0.275±0.125	0.250±0.120	0	0	0.037	0.087	0.121
MCW111	94-114	11	3(HG)	0.681±0.143	0.665±0.136	3	0	0.053	0.091	0.140
MCW183	293-333	18	8(HG)	0.550±0.165	0.526±0.164	0	0	0.070	0.107	0.170
MCW216	136-148	7	2(HG)	0.482±0.180	0.363±0.167	4	0	0.240	0.157	0.360
MCW222	216-224	5	1(HG)	0.468±0.153	0.375±0.140	4	0	0.187	0.135	0.297
MCW248	213-223	6	2(HG)	0.640±0.166	0.594±0.166	0	0	0.074	0.098	0.165
MCW295	84-106	12	2(HG)	0.712±0.129	0.631±0.122	2	0	0.102	0.100	0.192
MCW330	256-288	9	2(HG)	0.689±0.195	0.638±0.198	2	0	0.085	0.118	0.192
LEI166	344-354	6	3(HG)	0.562±0.134	0.539±0.134	1	0	0.042	0.094	0.131
Global		148	36			42	1	0.090	0.129	0.208

A: Number of alleles; Ap: number of private alleles (Breed); D_{HWE}: number of populations in Heterozygote deficiency; EHWE: number of populations in Heterozygotes excess; HG: Ha Giang chicken population

Table S2. Summary of individual attributions to the assumed ancestral population by *STRUCTURE.*

	Commune	Pop 1	Pop 2	Pop 3	Pop 4
HG1	Ni	10	2	2	22
	qNi	*0.94*	*0.56*	*0.54*	*0.82*
HG2	Ni	0	5	1	3
	qNi		*0.65*	*0.72*	*0.61*
HG4	Ni	9	6	3	26
	qNi	*0.93*	*0.69*	*0.55*	*0.81*
HG7	Ni	5	0	1	0
	qNi	*0.83*		*0.90*	*0.00*
HG16	Ni	0	36	2	1
	qNi		*0.80*	*0.37*	*0.61*
HG20	Ni	12	4	3	4
	qNi	*0.88*	*0.83*	*0.71*	*0.64*
HG25	Ni	0	14	2	6
	qNi		*0.71*	*0.56*	*0.70*
HG40	Ni	15	3	0	8
	qNi	*0.93*	*0.64*	*0.00*	*0.76*
HG48	Ni	1	27	7	12
	qNi	*0.93*	*0.70*	*0.55*	*0.76*
HG49	Ni	0	28	6	18
	qNi		*0.64*	*0.59*	*0.77*
HG56	Ni	1	17	3	3
	qNi	*0.52*	*0.68*	*0.82*	*0.48*
HG61	Ni	2	22	3	10
	qNi	*0.53*	*0.70*	*0.48*	*0.76*
HG65	Ni	9	0	0	0
	qNi	*0.95*			
HG72	Ni	0	4	1	11
	qNi		*0.61*	*0.39*	*0.83*
HG75	Ni	1	28	1	22
	qNi	*0.45*	*0.69*	*0.70*	*0.72*
HG85	Ni	10	21	27	32
	qNi	*0.63*	*0.60*	*0.80*	*0.68*
HG88	Ni	29	3	0	0
	qNi	*0.83*	*0.68*		
HG91	Ni	1	36	13	14
	qNi	*0.40*	*0.72*	*0.69*	*0.68*
HG95	Ni	4	6	8	14
	qNi	*0.48*	*0.64*	*0.67*	*0.75*
HG103	Ni	7	21	40	20
	qNi	*0.57*	*0.68*	*0.79*	*0.66*
HG108	Ni	2	3	4	7
	qNi	*0.74*	*0.57*	*0.59*	*0.64*
HG110	Ni	2	28	10	9
	qNi	*0.68*	*0.75*	*0.57*	*0.67*

	Commune	Pop 1	Pop 2	Pop 3	Pop 4
HG113	Ni	2	19	23	8
	qNi	*0.58*	*0.65*	*0.76*	*0.71*
HG114	Ni	0	21	7	16
	qNi		*0.67*	*0.51*	*0.63*
HG140	Ni	0	12	8	5
	qNi		*0.65*	*0.61*	*0.57*
HG146	Ni	3	5	1	5
	qNi	*0.50*	*0.60*	*0.60*	*0.66*
HG157	Ni	8	0	2	4
	qNi	*0.84*		*0.47*	*0.58*
HG169	Ni	1	1	3	10
	qNi	*0.78*	*0.58*	*0.71*	*0.76*
HG179	Ni	2	7	72	7
	qNi	*0.46*	*0.55*	*0.87*	*0.60*
HG184	Ni	1	0	11	5
	qNi	*0.84*		*0.78*	*0.73*
No	1082	137	379	264	302

Table S3. Normed scores of traits for principal component axes

Traits	PC1	PC2
TnP	-0.090	0.221
TnF	0.006	0.015
BlackT	0.241	0.171
CmbRv	0	-0.354
CmbRc	-0.110	0.245
CmbRB	0.125	0.006
BlackCmb	0.206	0.294
WhiteS	0.233	-0.115
YellowS	-0.261	-0.018
BlackS	0.189	0.331
OranE	-0.270	-0.097
BrownE	0.266	0.124
BlackE	0.221	-0.090
RedO	-0.219	0.042
WhiteO	0.165	-0.110
WRO	0.027	-0.006
BlackO	0.212	0.205
C^+	-0.188	0.034
i^+	-0.151	0.002
F	-0.195	0.242
H^+	-0.118	0.286
pg^+	-0.198	-0.033
cr^+	-0.084	0.160
mb^+	-0.143	-0.157
vh	-0.143	0.213
r^+	0.082	-0.357
p^+	0.026	-0.214
id^+	0.255	-0.054
W^+	0.221	-0.120

6.6 Conclusion

Firstly, we demonstrated the usefulness of combining different data sets in order to reveal breed histories and for establishing conservation policies. From a global point of view, when European and Asian breeds were jointly analysed, the Coucou de Rennes and the Hua Tung breeds were the ones contributing the most in the global diversity of each subset respectively. However, when only French and AVIANDIV breeds were jointly analysed, instead of the Coucou de Rennes, it was the Marans breed that contributed the most. As it was already noted that prioritisation of breeds should take into account both within- and between breeds diversity, attention should be paid to set of breeds compared. This strongly suggests that not only a set of markers should be proposed but also a set of breeds in order to be able to compare breeds. Therefore, since recent decades, the globalisation of conservation through organisms such as the FAO has become more important, and tools such as a microsatellite markers and breeds set for preliminary studies on genetic diversity will be helpful for establishing priorities in conservation projects. However, because a breed will only be conserved if it is actually reared, the cultural and economic value should also be taken into account. The H'mong breed is a good illustration of this problem. This breed has a strong cultural value as it is considered to: 1) be an indigenous breed, 2) be associated with an ethnic group and, 3) have medicinal properties. Our results showed that the H'mong breed is defined only by the phenotype as no genetic differentiation into subdivided population was observed in the Ha Giang province which is considered to be the province of origin of this breed. Therefore, conserving all the genetic diversity of the Ha Giang province through the conservation of the H'mong chicken phenotype could be a good option as the H'mong breed benefits already from strong cultural and economic value.

Secondly, because we used the similar set of markers and DNA reference samples, we were able to combine two sets of data from different laboratories: the Vietnamese set and the previous one which contained local Asian breeds but also wild populations. As the Ha Giang province is a part of the distribution area of the wild ancestor, our combined data set allowed us to give evidence of gene flow between the domestic chicken of Ha Giang province and the wild ancestor. Such results could not be obtained if the DNA reference samples were not used. However, it should be noted that only 18 markers among the 20 markers from the set were successfully calibrated between both sets. Therefore there is a need to increase this

success rate; possibly through standardisation of protocols (laboratory products,….), and there is a need to give priority to these markers if not all the FAO sets can be carried out due to financial reasons.

DISCUSSION
& CONCLUSIONS

7 DISCUSSION & CONCLUSIONS

Over recent decades, there has been a broad consensus on the loss rate of agricultural biodiversity. Estimates from the World Watch list of Domestic Animal Diversity (Scherf, 2000) indicate that 35 % of mammalian breeds and 63 % of avian breeds are at risk of extinction and that, at the World scale, one breed is lost every week. The population size of domestic species may be far greater than that of wild species, and it is henceforth recognised that the most important characteristic for the viability of a species may be the genetic diversity within the species. On this aspect, domestic species may be somehow more at risk as they have been submitted to intensive human selection in order to fulfil human needs, and so may be subjected to a higher loss rate of genetic diversity from one generation to the next. Therefore, it is now accepted that domestic breeds may be considered as endangered species from a genetic point of view (Taberlet *et al.,* 2008).

Asia is the cradle of many livestock species, and it is also one of the areas, with Africa, where agriculture is mainly performed in an extensive way. Both characteristics make Asia a privileged area for high biodiversity. Some Asian countries, because of their geographical location or political history, have experienced relatively few exchanges with western countries. As a consequence, their animal genetic resources may have remained in their original state. Indeed, that was the primary assumption when establishing a part of the BIODIVA project in the remote province of Ha Giang.

7.1 Diversity of ecotypes, diversity of ethnic communities, diversity of livestock

Genetic variation may be linked to farmer's decisions in two major categories: morphological characteristics, farm management practices. These decisions in turn are based on environmental and socio-economic influences. Mountainous areas of Vietnam are characterised by steep slopes, alternating valleys and mountains, thereby alternating good agricultural conditions with water and fertile soil with harsh slopes suffering from lack of water during winter. In the sino-indochinese mountains communities are organised in strata (Bruneau, 2002). This situation has remained stable across centuries as historical maps have shown (Michaud *et al.*, 2002). This stratification has also led to a stratification of political societies with "feudal" communities in the low- and middle-land and more "egalitarian" communities in higher altitudes (Bruneau, 2002). The "feudal" communities were found in the

ethnic groups from the Thai-Kadaï ethno-linguistic family. Their political systems were characterised by the setting up of states where principalities are the base unit: "muang". This unit is created by a conquering lord who governs various neighbouring valleys and middle-lands. A "muang" is therefore composed of various valley areas, and according to its importance it could also influence smaller surrounding "muangs".

First of all, as expected and confirmed by our data, environmental conditions in the lowlands allow farmers to grow paddy rice. This results in the use of buffalo as the main livestock for draught power. Therefore, ethnic communities from the low- and middle-lands are associated with buffalo husbandry.

Secondly, the geographical situation influences in part the political organisation, as it facilitates a "muang" type set-up where the base unit is a group of valleys. Thus, communication and commercial exchanges within a "muang" are expected to be important. With this kind of structure, a social communitarian organisation of farmers would generally appear. During our surveys, organised systems within a commune for reserving land for pasture was observed only in the lowland area. A seasonal migration of the commune buffalo livestock lasting several weeks was arranged and 1 or 2 persons were responsible for looking after the herd. Thus the environmental conditions and the social organisation would affect the genetic structure of the livestock allowing important gene flow between valleys and hence no division into subpopulation occurs. Continuous gene flow from one valley to another will translate into a clustering of populations according to geographical distribution as it was observed for the buffalo population in the Ha Giang province.

Environmental conditions and social organisation would also influence the population size. Paddy crops produce higher economic benefits, leading to better living conditions, first by increasing purchasing power and the option to buy more animals, and secondly, by enabling communities to keep lands for pasture and distribute fairly among farmers. Feeding constraints would be lower in arrangements such as these and as a consequence the herd size per family could be higher. Indeed buffalo herds were bigger than cattle herds, and generations overlapped in the farm.

The "egalitarian" communities from the highlands (Bruneau, 2002) function as autonomous village communities constituted by a few family lineages. There is no actual institutional authority, only the respect naturally due to the elders. In the highlands, poor soil fertility and lack of water compared to the lowlands, make maize the most adequate crop, as confirmed by our observations. Maize growth influences animal husbandry in two ways. First, the choice of animal for draught power, which is cattle, as they are good at working in rocky fields, and less sensitive to cold weather. Secondly, it affects herd size. Maize is less valuable than rice (selling prices could be 6 times lower than for rice), and as it is grown in poor soil, the yield is low. Consequently, farmers do not have enough income to invest in many animals, and are obliged to use every square metre of land for crop production, and so there is no land available for pasture. The basic unit of the social organisation being the village in the "egalitarian" communities, there is no organisation for communal pasture. To offset this lack of pasturelands, farmers use the forest to supply fodder. However, this requires an investment of time (3-4 hours per day for bringing fodder), and labour (mainly children or women) and so limits the quantity of fodder available and thus the size of the herd.

A broad social life in the highland communities is fostered through the district market. The district market plays a similar role to the communal pasture in terms of gene flows. However, the extended area of gene flows would be limited to the mountain plateau. In the Ha Giang province there are two mountainous areas separated by a plain area where no cattle is reared. As a consequence, continuous cattle gene flows throughout the province is not possible as it was for buffalo leading to a subdivision into two cattle subpopulations.

Our analyses highlighted a lack of interest in the selection of mating animals in order to improve the characteristics of animals. Highly related animals within a farm, or less related animals during free grazing easily mate, leading to the high inbreeding values observed for cattle and buffalo. This pattern may be intensified by the fact that there is a low number of bulls for both species, but for different reasons. For buffalo, in the same way that there are fighting cocks in Asia, there are also buffalo fighting contests. Bulls are known to be aggressive and fighting within a herd often occurs. Therefore farmers prefer to avoid problems by not rearing buffalo bulls. The opposite is observed for cattle; farmers prefer to rear bulls as they are more powerful for draught. However, farmers also believe that mating will negatively affect bull docility and ability to work in the field and so they keep the best bulls in cowsheds to limit mating opportunities.

The environment controls the choice of the crop and in one way the livestock species, but the social organisation and beliefs will influence and mould the genetic pattern of the livestock breeds. Such characteristics need to be included in the framework of conservation policies.

7.2 Diversity of genetic structures

Non-management or management in a traditional extensive style leads to similar genetic characteristics for the three species considered in this study. On one hand, livestock populations showed high gene diversity. Compared to the gene diversity recorded in the Roslin database, the cattle population from Ha Giang province had one of the highest gene diversity results. The chicken population showed extremely high gene diversity with an important number of private alleles and alleles shared with wild ancestors. Buffalo, which is mainly an Asian livestock, is in most cases reared according to tradition in an extensive way; therefore buffalo had higher genetic diversity than other livestock species. These observations confirmed the key role that local breeds could play in the future as important reservoirs of genes.

Wide surveys in a small distribution area of specific breeds such as the H'mong cattle and the H'mong chicken make it possible to obtain an accurate definition of the genetic pattern of those breeds. This accurate definition would not be possible with classical sampling surveys. It is clear that a fine scale survey in a reduced area is more expensive than classical surveys (30-50 animals per breed). However, such wide-scale surveys are more suitable when the final objective is to implement conservation projects on a specific value with high cultural value. With classical surveys we have an overview of the genetic diversity of breeds but no information about the genetic substructure of those breeds. On the other hand, the fine scale multi-approach protocol (microsatellites, measurements, phenotypes and interviews) offers the possibility of proposing more precise guidelines for the setting up of conservation projects.

7.3 The definition of a breed

The case of the H'mong chicken population illustrates the difficulty behind establishing a definition of a breed. For Vietnamese people (Kinh ethnic group) the H'mong breed is well defined. It is considered to be a breed raised by the H'mong ethnic group, and which exhibits

phenotypic characteristics: black meat, black skin and black shank. According to our results no genetic differentiation of a H'mong Black chicken was observed in the Ha Giang province where it is supposed to have originated.

Scherf (2000) judged that a breed should be considered as follows:

"*Either a subspecific group of domestic livestock with definable and identifiable external characteristics that enable it to be separated by visual appraisal from other similarly defined groups within the same species, or a group for which geographical and/or cultural separation from phenotypically similar groups has led to acceptance of its separate identity*"

But acceptance by whom? According to Köhler-Rollefson (1997) it should be "*the community that keeps it*". However, no H'mong farmer interviewed ever defined a specific H'mong chicken breed. Also another black meat chicken, the O ke breed, is considered to be raised in the neighbouring province overlapping the wider distribution area of the H'mong chicken. Therefore, what is the reliability of the true existence of these two breeds? They have the same phenotypic characteristics, the same distribution area. One could expect to obtain results similar to the Ha Giang province (i.e. no genetic differentiation).

Such observations have already been reported in other species such as cattle in Namibia (Nortier *et al.* 2002) and in buffalo by Zhang *et al.* (2007), who first showed that Chinese buffalo clustered in four groups which did not correspond to the previously assumed nomenclature, and also that the White Guizhou breed was not genetically different from the Grey Guizhou breed.

In the context of a global network of breeds such as the DAD-IS database from FAO, countries have reported local breeds. However it seems that there are as many breeds as there are local names. During recent decades, there has been an increase of breed records (Scherf, 2000), but information on breed characteristics was not available. Huyen *et al.* (2006) reported that, in Vietnam, among the 14 local breeds of indigenous pigs, five breeds are vulnerable, two are in a critical state, and three are facing extinction. However, how were these 14 local breeds defined and by whom? There could be less than 14 breeds with similar populations sharing

different local names, but there could also be more. Therefore, breed extinction could also be under- or overestimated. Nevertheless, as urgent need to conserve local breeds has been recognised, competing countries may be pushed into a frantic "local breed race" in order to obtain conservation investments. This could have a negative effect, from a genetic point of view, as two conservation projects of two populations with different names but no genetic differentiation would be less efficient than a bigger project that considers both populations as only one population. From a social point of view, taking into account the cultural value of a livestock population or the economic market demand is much relevant. As a breed is simultaneously defined from a genetic and a social point of view, conservation decisions should consider both points of view.

7.4 Conservation policies and recommendations

These potential animal genetics reservoirs are not only endangered because of the risk of their being replaced by "exotic" breeds or machines but also because of their relatively high inbreeding values. The need to conserve such local breeds becomes therefore obvious. However, the establishment of conservation policies, and practices that are adapted for genetics as well as for the communities who rear the animals, is not so easy to implement.

As the swamp buffaloes in the Ha Giang province are a unique genetic population the main objective of a conservation project will be to decrease the rate of inbreeding or, if possible, to stabilize the current inbreeding. Such an objective could be accomplished by simply choosing breeding animals originating from distant districts.

This kind of practice could not be applied to cattle as two subpopulations were identified. Therefore, it would be necessary to identify if the two subpopulations are valuable for conservation or only one would be significantly relevant. In order to be able to answer such question, subpopulations of the Ha Giang cattle have to be genetically compared to other Vietnamese breeds and more especially with the Yellow cattle population from the bordering province of Cao Ban. Such comparison will confirm and estimate admixture. If admixture significantly occurred, the cattle population from the northeastern districts will show low differentiation with the Yellow cattle from Cao Ban. As a consequence, conserving the South-

Western cattle population would have priority. Also, the absence of preference in a specific phenotype alleviates the constraints of the selection of reproductive bulls and the importance of the district market in the animal exchanges naturally highlights this as a possible location for the implementation of a possible AI or mating centre.

If a phenotype with strong cultural value suffice for being considered as a breed, therefore conservation project of the black phenotype H'mong chicken should be implemented in the three communes where this phenotype was found. One solution could be to establish a nucleus for improvement of black phenotypes rate. The chicks with all the black characteristics would have two purposes: the first one as future reproductive animals for the nucleus; the remaining black chicks would be raised and managed by householders trough the establishment of cooperative or breeder associations. Also for sustainable development, market network should be analysed and developed at the same time.

7.5 Conclusion

Small-scale farming is the predominant form of farming in many regions of the world and it is often in areas where ecological constraints are heavy. The cost of adjusting the production environment to the conditions needed by high-performance breeds may exceed that of improving locally adapted breed and so is an efficient way to alleviate poverty and increase resilience of farming systems. However, genetic improvement programmes in developing countries have failed because of inappropriate strategies. Because of a lack of information on genetic resources, most livestock improvement programmes have been unable to target the most appropriate genetic resources. This could be the case for the Ha Giang cattle population as it could be considered as only one genetic population whereas two subpopulations are actually reared in the province. Also, conservation and use of livestock genetic resources *in situ* are dependent on a suitable enabling policy environment and so equal attention should be given to genetic aspects and the social and economic context. Highlighting constraints such as lack of fodder, or proximity as the main factor for the choice of the mating bull will allow to implement appropriate community-based management projects that through the breed use will ensure the *in situ* conservation of the breed.

The Ha Giang province is inhabited by many ethnic communities which are distributed across the sino-indochinese mountains. Therefore, guidelines observed from the farming system analysis might be representative and could be used in the sino-indochinese mountains region. Moreover we demonstrated the importance of a fine scale while discontinuous sampling surveys could lead to artefactual subpopulations. Following our results, further studies should be carry on the Yellow cattle from Cao Ban province in order to quantify admixture rate between H'mong and Yellow cattle, but also on the chicken mtDNA to confirm gene flow between wild and domestic populations.

REFERENCES

REFERENCES

Abadie M. (1924) Les races du Haut-Tonkin, de Phong-Tho a Lang-Son .Eds: P., Sté. d'Editions Géogr., Maritimes et Coloniales.

Akishinonomiya F., Miyake T., Sumi S., Takada M., Ohno S., Kondo N. (1994) One subspecies of the red jungle fowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *P. Natl. Acad. Sci. USA*, **91,** 12505-12509.

Akishinonomiya F., Miyake T., Takada M., Shingu R., Endo T., Gojobori T., Kondo N., Onho S. (1996) Monophyletic origin and unique dispersal patterns of domestic fowls. *P. Natl. Acad. Sci. USA*, **93,** 6792-6795.

Amano T., Nozawa K., Namikawa T., Hasnath M.A., Mostafa K.G., Faruque M.O. (1984) Blood proteins polymorphisms of water buffaloes in Bangladesh. Genetic studies on Breed Differentiation of the Native Domestic Animals in Bangladesh.25-42, Tokyo University of Agriculture, Japan.

Amano T., Miyakoshi Y., Takada T., Kikkawa Y., Susuki H. (1994) Genetics variants of ribosomal DNA and mitochondrial DNA between swamp and river buffaloes. Anim. Genet., 25, 29-36.

Arranz J.J., Bayon Y., San Primitivo F. (1996a) Genetic variation at five microsatellites loci in four breeds of cattle. J. Agric. Sci. 127, 533-538.

Arranz J.J., Bayon Y., San Primitivo F. (1996b) Comparison of protein markers and microsatellites in differentiation of cattle populations. Anim. Genet. 27,415-419.

Baker C.M.A., Manwell C. (1991) Population genetics, molecular markers and gene conservation of bovine breeds. In: *Cattle Genetic Resources*. (1st ed.), (HICKMAN CG, ed.),Amsterdam: *Elsevier. pp*. 221-304. (NEIMANN-SØRENSEN A, TRIBE DE, eds. *World Animal Science*; vol. **B7**).

Bailey J.F., Richards M., Macaulay VA., Colson I., James IT., Bradley DG., Hedges REM (1996) Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species. *Proc. R. Soc. B,* **263**, 467-1473.

Barker J.S.F., Moore S.S., Hetzel DJS., Evans D., Tan S.G., Byrne K. (1997) Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein coding loci. *Anim. Genet.*, **28,** 103-115 .

Barker J.S.F. (1999) Conservation of livestock diversity. *AGRI*, **25,** 33-45

Belkhir K., Borsa P., Goudet J., Chikhi L., Bonhomme F. (2000) GENETIX, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions: CNRSUPR. 9060. Université Montpellier II, Montpellier, France.

Boichard D., Maignel L., Verrier E. (1996). Analyse généalogique des races bovines laitières françaises. *INRA Productions animales*, **9,** 323-335.

Bowcock A.M., MacHugh D.E., Tomforhde J., Minvh E., Kidd J.R., Cavalli-Forza L.L. (1994) High resolution of human evolutionnary trees with polymorphic microsatellites. *Nature*, **368** , 455-457.

Bradley D.E., MacHugh D.E., Loftus R.T., Sow R.S., Hoste C.H., Cunningham E.P. (1994) Zebu-taurine variation inY chromosomal DNA: a sensitive assay for genetic introgression in West African trypanotolerant cattle populations. *Anim. Genet.,* **25**, 7-12.

Bradley D.E., Cunningham D.G., MacHugh P., Loftus R.T. (1996) Mitochondrial diversity and the origins of African and European cattle. *Proc. Natl. Acad. Sci. USA*, **93**, 5131-5135.

Bruneau, M. (2002). Evolution des etagements ethnopolitiques sans les montagnes sino-indochinoises. *Herodote,* **107,** 88-117.

Bruford M.W., Bradley D.G., Luikart G. (2003) DNA markers reveal the complexity of livestock domestication. Nature Review Genetics, **4,** 900-910

Burton J.A., Hedges S., Mustari A.H. (2005) The taxonomic status, distribution and conservation of the lowland anoa *Bubalus depressicornis* and mountain anoa *Bubabus quarlesi*. *Mammal revue*, **35,** 25-50.

Caballero A., Toro M.A. (2002) Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Gen.,* **3,** 289-299

Cai X., Chen H., Wang S., Xue K., Lei.C. (2007) Polymorphisms of two chromosome microsatellites in Chinese cattle. *Genet. Selec. Evol,*. **38**, 525-534.

Castella J-C., Manh H.P., Kam P.S., Villano L.,Tronche N.R. (2005). Analysis of village accessibility and its impact on land use dynamics in a mountainous province of northern Vietnam. *Applied Geography*, **25,** 308-326.

Cazes P., Chessel D., Doledec S. (1988) L'analyse des correspondences internes d'un tableau partitionné. Son usage en hydrobiology. *Rev. Statistique Appliquée*, **36,** 39-54.

Chen Y.C., Li XH (1989) New evidence of the origin and domestication of the Chinese swamp buffalo (*Bubalus bubalis*). *Buffalo Journal*, **1,** 51-55.

Chessel D., Dufour A.B.,Thioulouse J. (2004) The ade4 package-I :One –table methods, R-News **4**, 5-10.

Choquart L. (1928) Les Marchés de bestiaux et le commerce du bétail au Tonkin : Thèse pour le doctorat vétérinaire (Diplôme d'Etat). - Paris : Ed. "M. Lac", 1928. - 108p. ; 25cm. - (École Nationale Vétérinaire d'Alfort. Année 1928).

Choudhury A. (1994) The decline of the wild water buffalo in northeast India. *Oryx*, **28,** 70-73.

Ciampolini R., Moazami-Goudarzi K., Vaiman D. Dillmann C., Mazzanti E., Foulley JL., Leveziel H., Cianic D. (1995) Individual multilocus genotypes usong microsatellites polymorphisms to permit the analysis of the genetic variability within and between Italian beef cattle breeds. *J. Anim. Sci.*, **73,** 3259-3268.

Cockrill W.R. (1974) The husbandry and health of the domestic buffalo. FAO, Rome

Coquerelle G. (2000) Les poules:diversité génétique visible. Eds: INRA-Editions, Paris

Crooijmans R.P.M.A, Groen F.A., Van Kampen A.J., Van der Beek S., Van der Poel J.J., Groenen M.A. (1996) Microsatellite polymorphism in commercial broiler and layer lines estimated using pooled blood samples. *Poultry Science*, **75,** 904-907.

Culas C., Michaud J. (1997) *A Contribution to the Study of H'Mong (Miao) Migrations and History.* Leiden, KITLV, 1997. 33p. references, 8vo wrs. Bijdragen tot de Taal-, Land- en Volkenkunde, Deel 153.

Custodio C.C., Lepiten M.V., Heaney L.R. (1996) *Bubalus mindorensis. Mammalian Species*, **520,** 1-5.

Del Hoyo J., Elliot A., Sargatal J. (1994) *Handbook of the Birds of the World.* Vol 2. New World Vultures to Guineafowl. Lynx Editions, Barcelona.

Douarche E .(1902) Les bovins de l'Indochine Impr. Typo-litographique F-H Schneider

Dray S., Dufour A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22(4) http://www.jstatsoft.org/v22/i04/.

Eding J.H., Meuwissen T.H.E .(2001) Marker based estimates of between and within population kinships for the conservation of genetic diversity. *J. Anim. Breed. Genet.,* **118**, 141-159.

Eding J.H., Meuwissen T.H.E. (2003) Linear methods to estimate kinships from genetic marker data for the construction of core sets in genetic conservation schemes. *J. Anim. Breed. Genet.,* **120**, 289-302.

Epstein H., Mason I.L. (1984) Cattle. In: *Evolution of domesticated animals* (1st ed.). (Mason I.L, ed.) London: Longman. p.p. 6-27.

European Cattle Genetic Diversity Consortium (2006) Marker assisted conservation of European cattle breeds: an evaluation. *Anim. Genet.,* **37**, 475-481

Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE, a simulation study. *Mol. Ecol.,* **34,** 2611-2620.

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

FAO (1977) The Water Buffalo. *Ed: Food and Agricultural Organization,* Rome, ISBN: 92-5-100108-1.

FAO, 2000. World Watch list for Domestic Animal Diversity. FAO, Rome, Italy.

Felius M.(1995) Cattle breeds – an encyclopaedia. *Ed:* Misset uitgeverij, Doetinchem NL.

Fischer . & Ulbrich F. (1968) Chromosomes of the Murrah buffalo and its crossbreeds with the Asiatic swamp buffalo (*Bubalus bubalis*). *Z. Tiersucht. Zuchtungsbiol.,* **84,** 110-114.

Frankham R., Ballou J.D., Briscoe D.A. (2002) In: Introduction to conservation genetics, Cambridge University Press

Freeman A.R., Meghen CM., MacHugh D.E., Loftus R.T., Achukwi M.D., Bado A., Sauveroche B., Bradley D.G (2004) Admixture and diversity in West African cattle populations. *Mol. Ecol.,* **13**, 3477-3487.

Freeman A.R., Hoggart CJ, Hanotte O., Bradley DG (2006) Assessing the relative ages of admixture in the bovine hybrid zone of Africa and the Near East using X chromosome haplotype mosaicism. *Genetics,* **173**, 1503-1510.

Granevitze Z., Hillel J., Chen G.H., Cuc N.T.K., Feldam M., Eding H.,Weigend S. (2007) Genetic diversity within chicken populations from different continents and management histories. *Anim. Genet*., doi:10111/j1365-2052.2007.01650.

Grooves C.P. (1981) Systematic relationship in the Bovini (Artiodactila, Bovidae). *Sonderdruck aus Z. F. Zool. Systematik u. Evolutionforshung,* **19,** 264-278.

Grigson C. (1980) The craniology and relationships of four species of *Bos*. 5. *Bos indicus L. J. Arc.h Sci*., **7**,3-32

Gupta P., Singh L., Ray-Chudhuri S.P. (1974), Chromosomes of Indian breeds of cattle. *Nucleus*, **17**, 129-132.

Hall S.J.G. (1991) Body dimensions of the Nierian Cattle, sheep and goats. *Anim. Prod.*, **53,** 66-69.

Hanotte O., Tawah C.L., Bradley D.G., Okomo M., Verjee Y., Oghieng J., Rege J.E.O. (2000). Geographic distribution and frequency of a taurine *Bos taurus* and indicine *Bos indicus* Y specific allele amongst sub-Saharan African cattle breeds. *Mol. Ecol.,* **9**, 387-396.

Hanotte O., Bradley D.G., Ochieng J.W., Verjee Y., Hill E.W., Rege J.E. (2002) African pastoralism: genetic imprints of origins and migrations. *Science*, **296,** 336-339.

Hassanin A, Ropiquet A (2004). Molecular phylogeny (Bovidae, Bovinae) and the taxonomic status of the Kouprey, *Bos sauveli* Urbain 1937. *Mol. Phylo. Evo.,* **33,** 896-907.

Heinen J.T. (1992) Phenotypic and behavioural characteristics used to identify wild buffalo *Bubalus bubalis* from feral backcrosses in Nepal. *Bombay Natural History Society,* **99,** 173-183.

Heinen J.T. & Singh G.R. (2001) A census and some management in implications for wild buffalo (*Bubalus bubalis*) in Nepal. *Biol. Conser.,***101**,391-394.

Hill M.O., Smith J.E. (1976) Principal component analysis of taxonomic data with multi-state discrete characters. *Taxon*, **25,** 249-255.

Hoang C.C.(1978) Survey on Buffalo breed characteristics in number of Southern provinces of Vietnam (1976-1978) In proceedings (Vietnamese).The First Agricultural Science Conference of IVth Agricultural University. Volume II.

Hoffmann I (2003) Biodiversity management in West African pastoral and agropastoral systems. A case of study from northwest Nigeria. Available at www.fao.org/DOCREP/005/Y4586E/y4586e03.htm#Po_o

Hoffmann I., Marsan P.A., Barker J.S.F., Cothran E.G., Hanotte O., Lenstra J.A., Milan D., Weigend S., Simianer H. (2004) New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a join ISAG/FAO grouping. In: *Proceedings of 29th International Conference on Animal Genetics*, Tokyo, Japan.

Huyen L.T.T., Roessler R., Lemke U., Valle Zarate A.(2006) Impact of the use of exotic compared to local pig breeds on socio-economic development and biodiversity in Vietnam. In A. Valle Zarate, K. Musavaya & C. Schafer, Eds. Gene flow in animal genetic resources: a study on status, impacts and trends, pp 459-508. FAO, GTZ, BMZ.

ILRI (1996) *ILRI 1995:Building a global research institutes*. International Livestock Research Institute, Nairobi, Kenya.

IUCN available at www.iucnredlist.org.

Ji C.L., Chen G.H., Wang M.Q., Weigend S. (2005) Genetic structure and diversity of 12 Chinese indigenous chicken breeds. In: The role of biotechnology for characterization and conservations of crops, forestry, animal and fishery genetic resources. Villa Gualino, Turin, Italy. Eds: FAO. 5-7 March 2005.

Jordana J., Alexandrino P., Beja-Pereira, Bessa I., Cañon, Carretero Y., Dunner S., Laloë D., Moazami-Goudarzi K., Sanchez A., Ferrand N. (2003) Genetic structure of eighteen local south beef cattle breeds by comparative F-statistics analysis. *J. Anim. Breed Genet.,* **120,** 73-87.

Kamalzadeh A., Koops W.J., Van Bruchem J. (1998) Feed quality restriction and compensatory growth in growing sheep, modeling changes in body dimensions. *Livest. Prod. Sci*, **53,** 57-67.

Kantanen J., Olsaker I., Holm H.E., Lien S., Vilkki J., Brusgaard K., Eythorsdottir E., Daneli B., Adalesteinsson S. (2000) Genetic diversity and populations structure of 20 North European cattle breeds. *J. Hered.*, **91**, 446-457.

Kierstein G., Vallinoto M., Silva A., Scnhneider M.P., Iannuzzi L., Brenig B. (2004) Analysis of mitochondrial D-loop region casts new light on domestic water buffalo (Bubalus bubalis) phylogeny. *Mol. Phylo. Evol.*, **30,** 308-324.

Kikkawa Y., Yonekawa H., Suzuki H., Amano T. (1997) Analysis of the genetic diversity of domestic water buffaloes and anoas based on variations in the mitochondrial gene for the cytochrome b. *Anim. Genet.*, **28,** 195-201.

Köhler-Rollefson I. (1997) Indigenous practices of animal genetic resource management and their relevance for the conservation of domestic animal diversity in domestic countries. *J. Anim. Breed. Genet.,* **114,** 231-238.

Kumar P., Freeman A.R., Loftus R.T., Gaillard C., Fuller D.Q., Bradley D.G. (2003) Admixture analysis of South Asian cattle. *Heredity*, **91**, 43-50

Kumar S., Gupta J., Kumar N., Dikshit K., Navani N., Jain P., Nagarajan M.(2006) Genetic variation and relationships among eight Indian riverine buffalo breeds. *Mol. Ecol.,* **15,** 593-600.

Kumar S., Nagarajan M., Sandhu J.S., Kumar N. Behl V., Nishanth (2007a) Mitochondrial DNA analyses of Indian water buffalo support a distinct genetic origin of river and swamp buffalo. *Anim. Genet.*, **38,** 227-232.

Kumar S., Nagarajan M., Sandhu J.S., Kumar N., Behl V. (2007b) Phylogeography and domestication of Indian River buffalo. *BMC genetics*, **7,** 186-193.

Lai S-J., Liu Y-P., Liu Y-X., Yao Y-G .(2006) Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Mol. Phylogenet. Evol.,* **38**, 146-154

Laloë D., Moazami-Goudarzi K., Chessel D (2002) Contribution of individual markers to the analysis of the relationships among breeds by correspondence analysis. 7th World Congress on Genetics Applied to Livestock Production, August 19-23, 2002.Communication 26-06. Vol 33, p 493-496. Montpellier, France.

Lau C.H., Drinkwater R.D., Yusoff K., Tan S.G., Hetzel D.J.S., Barker J.S.F. (1998) Genetic diversity of the Asian water buffalo (*Bubalus bubalis*): Mitochondrial DNA D-loop and cytochrome b sequence variation. *Anim*. *Genet*., **29,** 253-264.

Lauvergne J., Souvenir Zafindrajaona P. (1992) Comparaison de deux populations de Zébu Malgache par les distances génétiques biométriques. *Rev. Elev. Med, Vet.,* **45,** 167-174.

Laval G., SanCristobal M., Chevalet C. (2002) Measuring genetic distances in various short-term evolution methods. *Genet. Sel. Evol.*, **34,** 481-507.

Le Louet G. (1931) La production animale en Indochine. *Receuil de Medecine Vétérinaire exotique Ecole d'Alfort*, **4,** 144-157.

Lei C.Z., Chen H., Zhang H.C., Cai X., Liu R.Y., Luo L.Y. Wang C.F., Zhang W., Ge Q.L., Zhang R.F., Lan X.Y., Sun W.B. (2006) Origin and phylogeographical structure of Chinese cattle. *Anim. Genet.*, **37**, 579-582.

Lei C.Z., Zhang W., Chen H., Lu F., Yang X.Y., Zhang H.C., Liu Z.G., Yao L.B., Lu Z.F. , Zhao Z.L. (2007) Independent maternal origin of Chinese swamp buffalo (Bubalus bubalis). *Anim. Genet.*, **38,** 97-102.

Liu Y.P., Wu G.S., Yao Y.G., Miao Y.W., Luikard G., Baig M., Beja-Pereira A., Ding Z.L., Palanichamy M., Zhang Y.P. (2006) Multiple maternal origins of chickens: Out of the Asian Jungles. *Mol. Phylogenet. Evol.,* **38,** 12-19.

Loftus R.T., MacHugh D.E., Bradley D.G., Sharp P.M., Cunningham E.P. (1992) Mitochondrial DNA and inferredrelationships between European, African and Asian cattle. Anim. Genet., **23**, 65

Loftus R.T., MacHugh D.E., Bradley D.G., Sharp P.M., Cunningham E.P. (1994a) Evidence for two independent domestications of cattle. *Proc. Natl. Acad. Sci.* USA **91**, 2757-2761.

Loftus R.T., MacHugh D.E., Ngere L.O., Balain D.S., Badi A.M., Bradley D.G., Cunningham E.P. (1994 b) Bos Mitochondrial genetic variation in European, African and Indian cattle populations. *Anim. Genet.*, **25**, 265-271.

Lunet et Lajonquiere. *Ethnographie des territoires militaires*. 1904. Impr. F-H Schneider.

Macgregor R. (1939) The domestic buffalo. Thesis presented to the Royal College of Veterinary Surgeons (quoted by I.L. Mason in Cockrill (1974)), 1-47.

MacHugh D.E., Loftus R.T., Bradley D.G., Sharp P.M., Cunningham E.P.(1994) Microsatellite DNA variation within and among European cattle breeds. *Proc. R. Soc. B,* **256**, 25-31.

MacHugh D.E., Shriver M.D., Loftus R.T., Cunningham P. (1997) Microsatellite DNA variation and the evolution ,domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics,* **146,** 1071-1086.

Malécot G.(1948) Les mathématiques de l'hérédité. *Eds:* Masson et cie, Paris, France.

Mannen H., Tsuji S., Loftus R.J., Bradley G.G. (1998) Mitochondrial DNA variation and evolution in Japanese Black Cattle (*Bos taurus*). *Genetics,* **30**, 1169-1175.

Mannen H., Kohno M., Nagat Y., Tsuji S., Bradley D.G., Yeo J.S., Nyamsamba D., Zagdsuren Y., Yokohama M., Komura K., Amano T. (2004) Independent mitochondrial origin and historical genetic differentiation in North Eastern Asia Cattle. *Mol Phylogenet. Evol,*. **32**, 539-544.

Mathias E. (1995) "Recording indigenous knowledge: an overview"pp.19-25 in Scneideri J. (ed). Indigenous knowledge in conservation of crop resources. Proceedings of an international workshop held in Cisarua, Bogor, Indonesia, Jan 30-Feb 3, 1995. International Potato Center, Office for East, Southeast Asian and the Pacific (CIP-ESEAP)/Cntral Reseach Institute for Food Crops (CRIFC), Bogor , Indonesia

Mathias E., Mundy P. (2005) *Herds movements*. Ober-Ramstadt. League for Pastoral Peoples and Endogenous Livestock development

Mattalia S., Barbat A., Danchin-Burge C., Brochard M., Le Mézec P., Minéry S., Jansen G., Van Doormaal B., Verrier E. (2006) La variabilité génétique des principales races bovines laitières françaises : quelles évolutions, quelles comparaisons internationales ? *Rencontres Recherches Ruminants* 13, 239-246

Meadow RH (1989). Osteological evidence for the process of animal domestication. In: *The Walking Larder: Patterns of Domestication, Pastoralism and Predation*. (Clutton-Brock J., ed.) London: *Unwin Hyman*, 80-90.

Meng-Hua L., Tapio I., Vilkki J., Ivanova Z., Kiselyova T., Marzanov N., Cinkulov M., Stojanovic S., Ammosov I., Popov R., Kantanen J. (2007) The genetic structure of cattle populations (*Bos taurus*) in northern Eurasia and the neighbouring Near Eastern regions: implications for breeding strategies and conservation. *Mol. Ecol.,* **16,** 3839-3853.

Morechand G. (1968) Le chamnaisme chez les H.mong. Bulletin de l'Ecole française d'Extrême Orient, Tome LIV, Paris

Moazami-Goudarzi K., Laloë D. Furet J.P., Grosclaude F. (1997) Analysis of genetic relationships between 10 cattle breeds with 17 microsatellites. *Anim. Genet.,* **28,** 338-345.

Moioli B., Georgoudis A., Napolitano F., Catillo G., Giubilei E., Ligda Ch., Hassanane M (2001) Genetic diversity between Italian, Greek and Egyptian buffalo populations. *Livest. Prod. Sci.,* **70,** 203-211.

Moore S.S., Evans D., Byrne K., ., Tan S.G., Vankan D. , Hetzel D.J.(1995) A set of polymorphic DNA microsatellites useful in swamp and river buffalo (*Bubalus bubalis*). *Anim. Genet.,* **26,** 355-359.

Msoffe P.L.M., Mtambo M.M.A., Minga UM, Juul-Madsen HR, Gwakisa P.S. (2005) Genetic structure among te local chicken ecotypes of Tanzania based in microsatellites DNA typing. *Afr. J. Biothec.*, **4,** 768-771.

Muchadeyi F.C., Eding H., Wollny C.B.A., Groeneveld E., Makusa S.M., Shamseldin R., Simianer H., Weigend S. (2007) Absence of population substructuring in Zimbabwe chicken ecotypes inferred using microsatellites analysis. *Anim. Genet.,* **38,** 332-339.

Mwacharo J.M., Drucker A.G. (2005) Production objectives and management strategies of livestock keepers in South-East Kenya: Implications for a breeding programme. *Trop. Anim. Health Prod.*, **37,** 635-652.

Namikawa T. (1981) Geographic distribution of bovine haematoglobin-beta (Hbb) alleles and the phylogenetic analysis of the cattle in Eastern Asia. *Z. Tierz. Zucchtungbiol.,* **98,** 151-159.

Navani N., Jain P.K., Gupta S., Sisodia B.S., Kumar S. (2002) A set of cattle microsatellite DNA markers fro genome analysis of riverine buffalo (*Bubalus bubalis*). *Anim. Genet.,* **33,** 149-154.

Nei M. (1978) Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals. *Genetics*, 89, 583-590.

Nishibori M., Shimogiri T., Hayashi T., Yassue H. (2005) Molecular evidence for hybridization of species in the genus *Gallus* except for *Gallus varius. Animal Genetics,* **36,** 367-75.

Nortier C.L., Els J.F., Kotze A., Van der Bank F.H. (2002). Genetic diversity of indigeneous Sanga cattle in Namibia using microsatellites markers. 7th World Congress on Genetic Applied to Livestock Production, August 19-23, 2002, Montpellier, France, Session 25. Communication 25-12.

Ollivier L., Foulley J.L. (2005) Aggregate Diversity: new approach combining within- and between-breed genetic diversity. *Livest. Prod. Sci.,* **95,** 247-54.

Pyane W.J.A. (1970) Cattle Production in the Tropics (1st ed.) London: Longman.

Payne W.J.A, Hodges J (1997) Tropical Cattle, Origins, Breeds and Breeding Policies. Blackwell Sciences Ltd, Oxford.

Porter V. (1991) Cattle, a handbook to the breeds of the World, A&C blacks : London.

Qu L., Yang N., Li, Xu G. Chen K., Yang H., Zhang L., Wu G., Hou Z., Xu G. (2004) Genetic diversity of 78 Chinese chicken breeds based on microsatellites analysis. Genetics session, paper G7. WPC, Istanbul.

Pritchard J.K., Stephens M., Donnelly P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

R Development Core Team (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rege JEO (2003) "Livestock breeds in traditional animal genetic resource management".pp133-138 in CIP-UPWARD. Conservation and sustainable use of agricultural biodiversity: A source book. International Potato Center and User's Perspectives with Agricultural Research and Development , Los Baños, LagunaNPhilippines.www.eseap.cipotato.org/upward/Publicatios/-Agrobiodiversity/pages%20133138%20(Paper%2016).pdf

Reist-Marti S.B., Simianer H., Gibson J., Hanotte O., Refe J.E.O. (2003) Weitzman's approach and conservation to African cattle breeds. *Conserv. Bio.,***17,** 1299-1311.

Reynolds J., Weir B.S., Cockerham C.C. (1983) Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics,* **105,** 767-779.

Rife D.C .(1962) Colour and horn variations in water buffalo: the inheritance of coat colour, eye colour in the water buffalo of Thailand. *J. Hered.,* **50,** 269-272.

Ritz L.R., Glowatski-Mullis M.L., MacHugh D.E. ,Gaillard C. (2000) Phylogenetic analysis of the tribe Bovini using microsatellites. *Anim. Genet.*, **31,** 178-185.

Rousset F., Raymond M. (1995) GENEPOP (version 3-4): population genetics software for exact test and eucumenicism. *J. Hered.*, **86,** 248-259.

Ruane J. (2000) A framework for prioritizing domestic animal breeds fro conservation purposes at the national level, a Norwegian case study. *Conserv. Bio.,* **40,** 1-10.

Russell N.D., Rios J., Erosa G., Remmenga M.D., Hawkins D.E. (2000) Genetic differentiation amon geographically isolated populations of Criollo cattle and their divergence from other Bos taurus breeds. *J. Anim. Sci.,* **9**, 2314-2322

Salerno A. (1974) The buffaloes of Italy . In: Cockrill WR (Ed.), The husbandry and health of the Domestic Buffalo. FAO, Rome, pp. 737-747.

Savina F.M. (1924) Histoire des Miao. Hong Kong. Societé des Missions Etrangères.

Sokal R.R., Rohlf F.J. (1969) Biometry: the principles and practice of statistics in biological research. Ed: W.H. Freeman and Company, New York.

Scherf B.D. (2000) *World Watch List for Domestic Animal Diversity*, 3rd edn. Food and Agriculture Organization of the United Unions, Rome.

Steglich M., Peters K.J. (2002). Agropastoralists' trait preferences in N'Dama cattle: Participatory methods to assess breeding objectives. 7th World Congress on Genetic Applied to Livestock Production, August 19-23, 2002, Montpellier, France, Session 25.

Developing Sustainable Breeding Strategies in Medium- to Low-Input systems. Communication 25-04

Taberlet P. Valentini A., Resaei H.R., Naderi S., Pompanon F., Negrini R., Ajmone-Marsan P. (2007) Are cattle , sheep and goats endangered species? *Mol. Ecol.,* **17,** 275-284.

Tisdell C. (2003) Socioeconomic values of loss of animal genetic diversity: analysis and assessment. *Ecolog. Econ.,* **45,** 365-376.

Thac N.D.(1979) Results of studies on F1 crossbreed of Murrah and Vietnamese local buffaloes. In proceedings (Vietnamese). Results of research on Science and Technology (1969-1979). Agricultural Publish House. Hanoi : 154-163.

Thu N.V. (1987) Breed characteristics and productivity of local buffaloes in Mekong Delta of Vietnam. Paper presented at the Conference of Animal Production in the South of Vietnam in Sep. 1987.

Thu N.V., Ding N.V.K. Presti T.R., Pearson R.A. (1995) Exploring approaches to Research in Animal Sciences in Vietnam. Australian Center for International Agricultural Research (ACIAR) proceedings n°68, p 261. ACIAR, Hue, Vietnam.

Thu N.V. (1997) A study of performance, physiological parameters and economic efficiency of working buffaloes in the Mekong Delta of Vietnam. *A collection of some current research and development observations*, p165-171.

Troy C., MacHugh D.E., Bailey J.F., Magee D.A., Loftus R.T., Cunningham P., Chamberlain A.T., Sykes B.C.,Bradley D.G. (2001) Genetic evidence for Near-Eastern origins of European cattle. *Nature,* **410**, 1088-1091

Ulbrich F., Fischer H. (1967) The chromosomes of the Asiatic buffalo (*Bubalus bubalis*) and the African buffalo (*Syncerus caffer*). *Z. Tiersucht. Zuchtungsbiol.,* **83**, 219-223.

United State department of Agriculture (USDA) (1996) *Technical Bulletin*, number 8. USDA, New York

Vo Van Su. (2004) At lat : cac giong vat nuoi o Viet Nam. Ed : Nha xuat ban nong nghiep – Hanoi.

Weir B.S., Cockerham C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38,** 1358-1370.

Weitzman M.L. (1993) What to preserve? An application of diversity theory to crane conservation. *Q. J. Eco.,* **108,** 157-183.

Wenping H. (1998) Genetic diversity of Chinese water buffalo. *Buffalo bulletin,* **17,** 30-32.

West B. , Zhou B.X. (1988) Did chickens go north? New evidence for domestication. *J Archaeol. Sci.*, **15,** 515-533.

Wimmers K., Ponsuksili S., Hardge T., Valle-Zarate A., Mathur P.K., Horst P. (2000). Genetic distinctness of African, Asian and South American local chickens. *Animal Genetics,* **31,** 159-165.

Wright S. (1951) The genetical structure of populations. *Ann. Eugen.,* **15,** 323-354.

Yu Y., Nie L., He Z-Q., Wen J-K., Jian C-S., Zhang Y.P. (1999) Mitochondrial DNA variation in cattle of South China : origin and introgression. *Anim. Genet.*, **30**, 245-250.

Zhang G.X., Wang Z.G., Chen W.S., Wu C.X., Han X., Chang H.,Zan L.S. Li R.L., Wang J.H., Song W.T., Xu G.F., Yang H.J., Luo Y.F. (2007) Genetic diversity and population structure of indigeneous yellow cattle breed of China using 30 microsatellite markers. *Anim. Genet.,* **38**, 550-559.

Zhang Y., Sun D., Yu Y., Zhang Y. (2007) Genetic diversity and differentiation of Chinese domestic buffalo based on 30 microsatellites markers. *Anim. Genet.,* **38,** 569-575.

Zhou H., Lamont S.J.(1999) Genetic characterization of biodiversity in highly inbred chicken lines by microsatellites markers. Anim. Genet., **30,** 256-264

ANNEXE 1 FARMER INTERVIEW FORM

Code Commune: Date
Code Village:
GPS: X: Y: Altitude:

Name   Code-Householder:
Nb of person in the family:
Nb of children: Nb going to school:

Pig Breeder: Yes / No   Pig Fattening: Yes / No

Crops:

Crops	Plant (kg seeds)	Produce (kg)	Purchase (Kg)	Sale (Kg)

Livestock

Species	Males	Females	Young	Total
Cattle			/	
Buffalo			/	
Goat			/	
Pig			/	
Horse			/	
Chickens			/	
Duck				

Nutrition:
Cattle: Free Grazing/ Stalling/ Communal Grazing/ Forest leaves/ others:
Buffalo: Free Grazing/ Stalling/ Communal Grazing/ Forest leaves/ others:
Goat: Free Grazing/ Stalling/ Communal Grazing/ Forest leaves/ others:
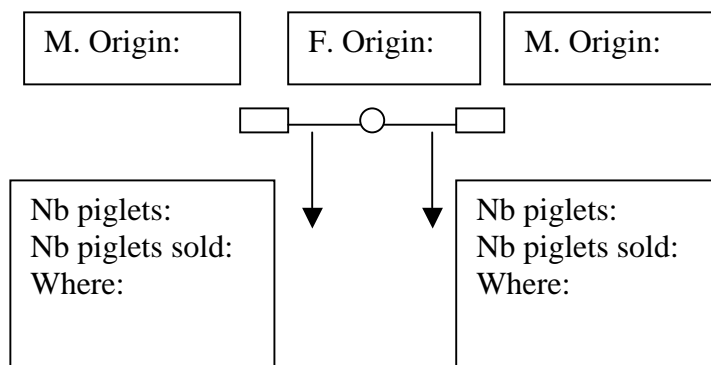Pigs:
Type of fodder:
Quantity of starches;   kg/day/animal
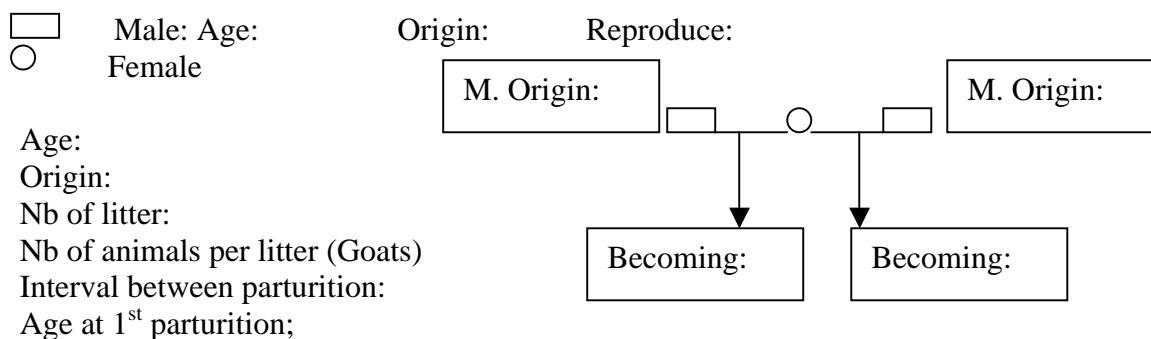Quantity of fodders: kg/day/animal

Pig Husbandry:

	Pig n°1	Pig n°2	Pig n°3	Pig n°4	Pig n°5
Weight Purchase					
Weight Sale					
Current Weight					
Husbandry Time (months)					
Purchase Price					
Sale Price					

Sow: Pedigree done for each sow present in the Farm

M. Origin:	F. Origin:	M. Origin:

Nb piglets:
Nb piglets sold:
Where:

Nb piglets:
Nb piglets sold:
Where:

Cattle/ Buffalo/Goat: Pedigree done for each animal present in the Farm

Male: Age: Origin: Reproduce:

Female

Age:
Origin:
Nb of litter:
Nb of animals per litter (Goats)
Interval between parturition:
Age at 1st parturition;

M. Origin:				M. Origin:

Becoming:

Becoming:

Chicken husbandry: average number over all reproductive hens

	Nb Eggs	Eggs Hatched	Nb of chicks	Nb adults
Laying 1				
Laying 2				
Laying 3				

Overlapping generation: Yes / No

Origin: From Family   From Market   From Forest

Year of the last epidemic:
Nb of survivor:
Nb of new introduced animals: Hens: Cock: From:

ANNEXE 2 PICTURES FROM SURVEYS

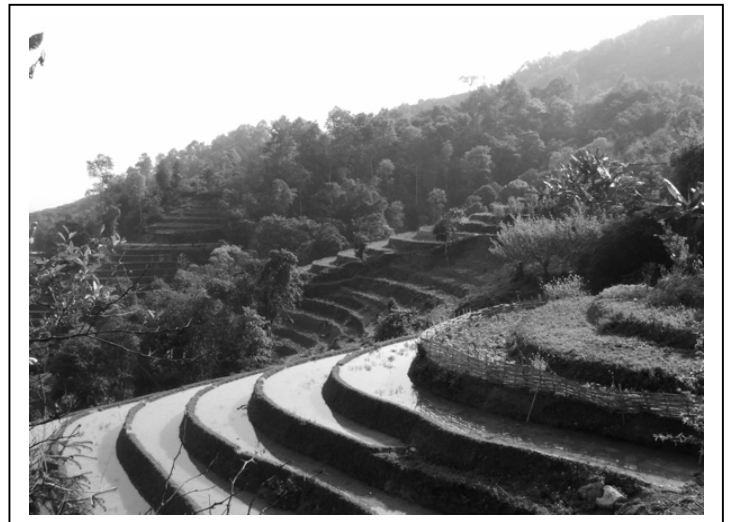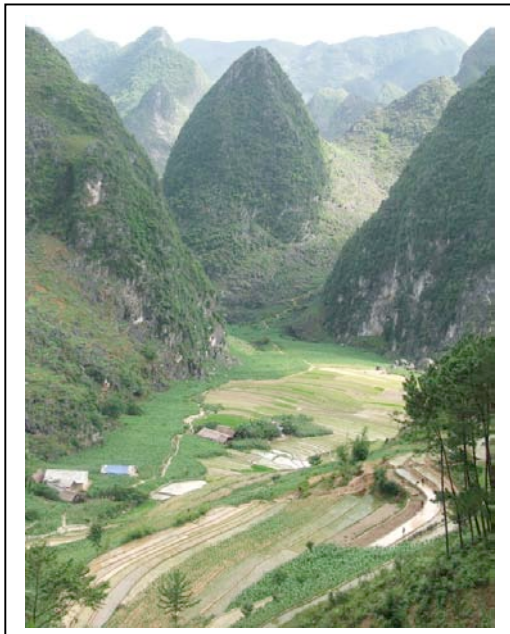Preparation of cam for pigs feeding

Fodder from forest

Canna plant
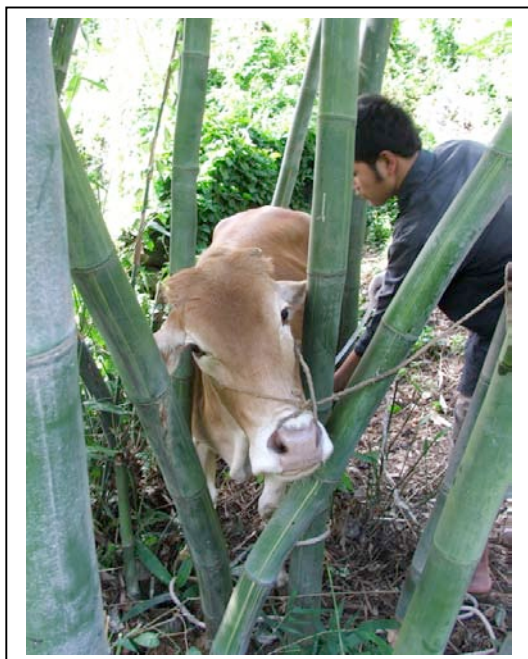

Paddy rice field
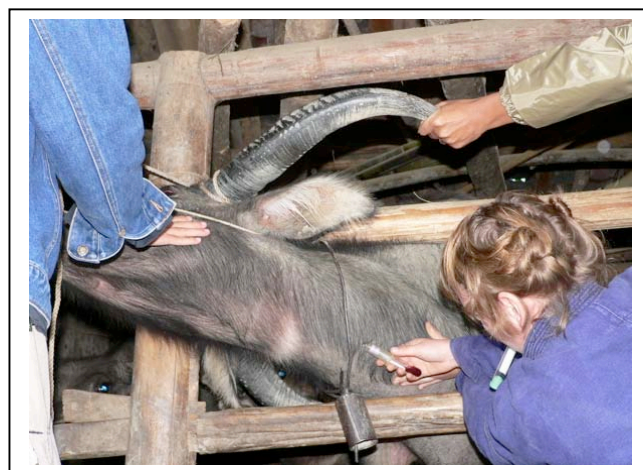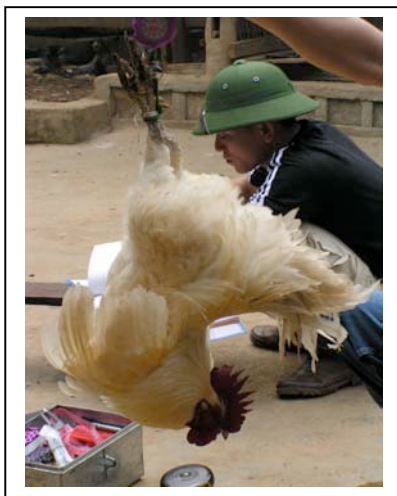

Landscape in Dong –Van district

259

Dao family

H'mong people

ANNEXE 3 PICTURES OF SAMPLINGS

ANNEXE 4. QUALIFICATIONS & TRAININGS

Academic Qualifications

2003-to 2008 **Ph.D.: Doctoral School ABIES-Life Science Engineering.**

Thesis: ***Characterisation of the cattle, buffalo and chicken populations in the northern Vietnamese province of Ha Giang** -BIODIVA Project.* Director: Prof. Dr. Etienne Verrier, Dr. Jean-Charles Maillard.

2002-2003 **M.Sc. agr.: Ecologie des systèmes aquatiques terrestres et lagunaires** - University of Science Toulouse-Montpellier 2. Master Thesis: ***Eutrophisation effect on aquatic microorganism***. Director: Dr Bezhad Mostajir.

Undergraduate courses

November 2003 Trainning course on ACCESS and MAPINFO softwares. (2 weeks) – CIRAD, Montpellier

November 2004 Trainning course on MAPINFO software (1 week)) – NIAH, Hanoi

Mars 2005 Trainning for Cytogenetic (Buffalo karyotypes) (1 week) – INRA/ National Veterinary School of Toulouse "Cytogénétique des populations animals", Toulouse

Communications

August 2006 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil.

December 2007 International workshop on domesticated animal biodiversity of Vietnam in the context of globalization, Ha Noi, Vietnam, December 13-14, 2007