Alma Mater Studiorum - Università di Bologna

Istituto Nazionale per la Fauna Selvatica

DOTTORATO DI RICERCA

Biodiversità ed Evoluzione

Ciclo XX

Settore scientifico disciplinare di afferenza: BIO/05 ZOOLOGIA

Non-invasive genetics and wolf (*Canis lupus*) population size estimation in the Northern Italian Apennines

Presentata da: **Dott. ROMOLO CANIGLIA**

Coordinatore Dottorato:
Prof. GIOVANNI CRISTOFOLINI

Relatori:
Prof. BARBARA MANTOVANI
Prof. ETTORE RANDI

Esame finale anno 2008

Alma Mater Studiorum - Università di Bologna

Istituto Nazionale per la Fauna Selvatica

DOTTORATO DI RICERCA

Biodiversità ed Evoluzione

Ciclo XX

Settore scientifico disciplinare di afferenza: BIO/05 ZOOLOGIA

Non-invasive genetics and wolf (*Canis lupus*) population size estimation in the Northern Italian Apennines.

Presentata da: Dott. ROMOLO CANIGLIA

Coordinatore Dottorato:
Prof. GIOVANNI CRISTOFOLINI

Relatori:
Prof. BARBARA MANTOVANI

Prof. ETTORE RANDI

Esame finale anno 2008

INDEX

CHAPTER FIRST: INTRODUCTION

1.1. INTRODUCTION TO THE SPECIES

1.1.1 Systematics

The wolf (*Canis lupus* L., 1758, ord. Carnivora, fam. Canidae) is the second largest predator in Europe, after the brown bear**.** It looks like a large German Shepherd dog. Since the species has a large distribution area and lives in a variety of habitats, its phenotypic variation (body size, color, and weight) is remarkably high (Mech, 1970; Boitani, 1995; 2000). On the basis of this variability (external morphology and skull characteristics), Sokolov & Rosolino (1985) identified 9 subspecies of *Canis lupus* in the Eurasian area (fig 1.1).

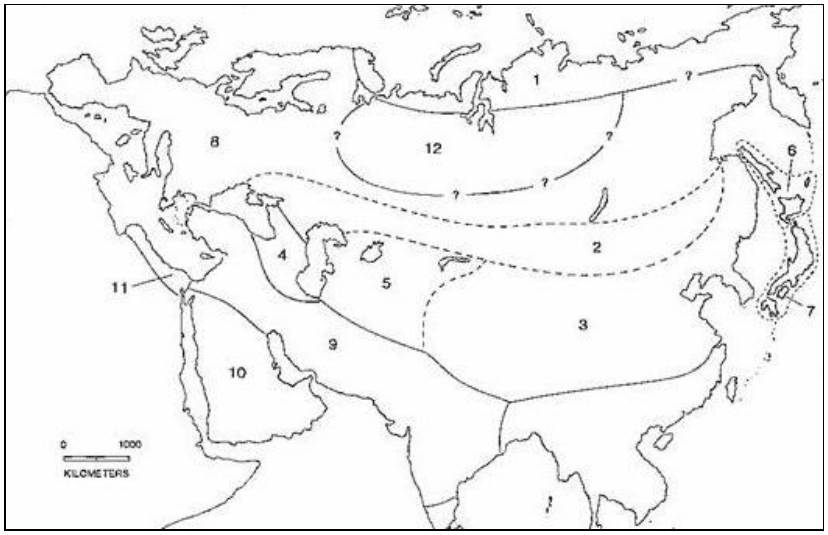

Fig. 1.1: Eurasian distribution of the different wolf subspecies (Sokolov & Rossolino 1985). 1) *C. lupus albus*, 2) *C. lupus campestris*, 3) *C. lupus chaco*, 4) *C. lupus cubanensis*, 5) *C. lupus desertorum,* 6) *C. lupus hattai*, 7) *C. lupus hodophilax*, 8) *C. lupus lupus*, 9) *C. lupus pallipes,* 10) *C. lupus arabs* (sometimes considered synonym of *pallipes*) 11*) C. lupus lupaster* (sometimes considered synonym of *lupus*), 12) *C. lupus communis*.

In this classification the Italian wolf population belongs to the subspecies *Canis lupus lupus*. Anyway, as the Italian wolf presents particular phenotypic characteristics such as a typical gray-brownish coat and a black stripe on the frontal part of the anterior legs, Altobello (1921) had already proposed for it the status of subspecies *Canis lupus italicus*. But as Altobello's description was based only on few phenotypic characteristics, it was rejected.

More recently, however, new taxonomic methods based both on morphometric studies (Nowak & Federoff, 2002) and genetic analyses (Randi *et al*., 2000; Randi & Lucchini, 2002), have suggested that the Italian wolf population seems to be differentiated enough from the other European ones to support Altobello's classification. The distinctive morphologic and genetic traits of the current Italian wolf population could be due to geographical barriers to its (wolf) dispersal. In fact Quaternary glacial/interglacial cycles affected the distributions of plant and animal communities and species, which contracted into southern refugia and expanded re-colonizing de-glaciated regions (Hewitt, 1996; 2000). The ice caps covering the Alps and the wide expansion of the Pò River, which cut the alluvial plains throughout the Holocene, might have isolated wolves in central-southern Apennines since the Last Glacial Maximum (ca. 18.000 years before present). Alternatively, deforestation, which was already widespread in the fifteenth century in northern Italy,

and direct human persecution, might have limited the rate of gene flow among wolves in the Apennines and any other population in Europe during the last few centuries (Lucchini *et al*., 2004).

1.1.2 Habitat and diet

Wolves live in the most diverse types of habitat and their broad distribution ranges show the species' adaptability to the most extreme habitat conditions (Mech, 1970; 1989). The wolf's habitat has been described as everywhere where humans do not kill the species and where food resourches are sufficient. Where wolves live depend on wild ungulate prey, their habitat is that of their prey (Mech & Boitani, 2003). Habitat quality should then be interpreted in terms of human disturbance (wolves are rarely found where human density is above 30-40 persons/km^2; Thiel, 1985), prey densities and range size. In general, large forest areas are particularly suitable for wolves in Europe, although wolves are not primarily a forest species (Boitani, 2000).

In Italy, like in the rest of Europe, the species usually lives in mountain and surmounting forested areas with lower human densities and less extensive agricultural utilisation, opportunistically eating what is most available in its habitat: wild boars (*Sus scrofa*), roe deers (*Capreolus capreolus*), red deers (*Cervus elaphus*), fallow deers (*Dama dama*) and small vertebrates, invertebrates, vegetables and carcasses (Ciucci *et al.*, 1996; Meriggi *et al*., 1996; Pezzo *et al*., 2003; Peterson & Ciucci, 2003).

1.1.3 Social behaviour and reproduction

Wolves live in social units (packs) consisting from 2 till as much as 36 animals which co-operate in hunting, reproducing and defending their territories. A pack is fundamentally a family unit that originates when a pair establishes a territory and reproduces. It is generally made up by a mating pair, its yearling pups and by a few other adults which are generally the offspring of the previous years remaining with the pack for a year or more, when new pups are born (Mech, 1970; Rothman & Mech, 1979).

Among pack members there is a strong hierarchy that regulates internal stability and the dynamics of the pack: individuals at higher dominance level coordinate every kind of pack activity and have most of the privileges in feeding and reproducing. Only the alpha female can breed preventing the other females from breeding (by aggressive behaviours and sometimes by violent fights) even if they can help it to bring up its pups (Olson,1938; Murie, 1944;Young & Goldman, 1944).

When a mating member disappears, it can be substituted by another wolf of the same pack or by a wolf coming from other packs or from other territories (Meier *et al*., 1995).

A wolf is sexually active when it is two years old. In a thriving population a wolf pair can produce pups every year (Fritts & Mech, 1981; Mech & Hertel, 1983; Mech, 1995d). The breeding can happen from January to April, it depends on the latitude (Rosenzweig, 1968), oestrus lasts 5-7 days once a year, the parturition occurs after a two month gestation period and litter size varies from 1 to 11 pups (Mech, 1970; 1981; Mech & Hertel, 1983; Stahler *et al*., 2002). Generally only the dominant pair breeds producing only one litter per pack but data from North America show that when food supplies are flush, some maturing wolves, rather than replacing a pack breeder, may breed in addition to their pack's established breeders while remaining in their natal pack (multiple breeding) and more than two litters (extra litters) can be produced within it (Harrington *et al.,*1982; Ballard *et al.,* 1987; Meier *et al.*, 1995; Mech *et al.*,1998).

When food is scarce adults stop provisioning young wolves and sexual competition and aggression might become the factor triggering dispersal (Mech, 1995c; d). In order to look for new territories where they can settle and found new packs of their own, in fact, some wolves disperse from

their pack as young as 5 months old (Fuller, 1989b), whereas others may remain with the pack for up to 3 years (Gese & Mech, 1991) or occasionally longer (Ballard *et al.*, 1987). Some individuals can also disperse when they lose their status and are rejected by a pack (Mech, 1995c; d).

The distances a wolf can disperse reflect the type of dispersal, from merely moving to an adjacent territory to substantial dispersal distances. Dispersal distances of several hundred kilometres are common, and movements over 1000 km have been documented (Fritts & Mech, 1981; Ballard *et al.*, 1983; Fritts, 1983; Mech *et al.*, 1995; Wabakken *et al.*, 2001). The data suggest the younger the disperser is, the farther it disperses (Wydeven *et al.*, 1995) and that the record dispersal lengths of males and females tend to be about the same (Ballard *et al.*, 1983; Peterson, Woolington & Bailey, 1984).

Wolves generally are highly territorial (Mech, 1973; 1944a; Mech *et al.*, 1998) and each pack territory could be considered a mini-ecosystem (Haber, 1997) whose size (from 80 to 2.500 km^2 in North America and from 100 to 500 km^2 in Europe) depends on the pack size (Mech, 1970; Peterson, Woolington & Bailey, 1984), on prey density (Walters *et al.*, 1981), on landscape, geographical and morphology features (Peterson, Woolington & Bailey, 1984; Peterson & Page, 1988), and on human disturbance. The immediate territory limits of neighbouring packs may partially overlap (Peters & Mech, 1975b; Peterson & Page, 1988) in a kind of buffer zone between packs (Mech, 1977d) but territory boundaries are rarely trespassed and when this occurs, it may lead to violent aggressions and intra-specific mortality (Peters & Mech, 1975b; Zimen, 1976; Harrington & Mech, 1979).

The internal pack cohesion, the conservation of social structure and the territory use and defence depend on the communication ability of wolves. In fact they have developed a complex communication system based on looks, face expression (Schenkel, 1947; Zimen, 1981), vocal signals (Mech, 1970; 1988a; Harrington & Mech, 1979; Schassburger, 1978; 1978; 1993; Coscia *et al.*, 1991; Coscia, 1995) and olfactory communication (Montagna & Parks, 1948; Parks, 1950; Aoki & Wada, 1951; Block *et al.*, 1981; Brown & Johnston, 1983; Mech, 2001b).

Wolves may use feces, with or without streaks of anal sac secretions (Peters & Mech, 1975b; Vilà *et al.*, 1994), and urine (Eisenberg & Kleiman, 1972; Johnson, 1973; Gosling, 1982; Doty, 1986) in territorial marking and as a response to unfamiliar or frightening surroundings (Kleiman, 1996).

1.1.4 Distribution and population numbers

After man, wolves (*Canis lupus*) are the terrestrial mammals with the largest distribution area in recent historical times because they are highly adaptable and widely distributed in ecosystems ranging from Arctic tundra to Arabian deserts in the Old and New World (Mech, 1970) (Fig. 1.2).

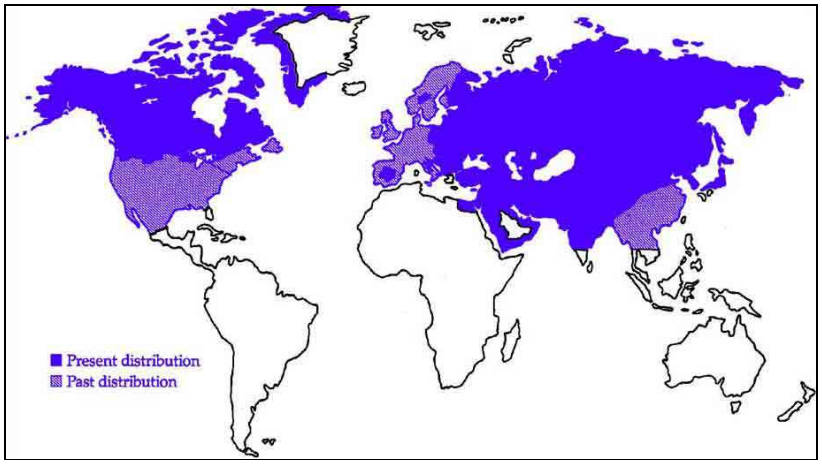

Fig. 1.2: historical and present World wolf distribution

The present European distribution of the species is greatly reduced if compared to the past one. Extermination efforts by man caused the species extinction in many countries through hunting, destruction of its habitats and the decrease of its natural prey (Delibes, 1990). Originally found through out Europe, at the end of the 18th century, wolves were still present in all European countries with the exception of Great Britain and Ireland. During the 19th century, and especially in the years following the Second World War, wolves were exterminated from all central and northern European countries. Now the largest European wolf populations live in Romania, Russia, Bulgaria, Poland, Balkan area. Three smaller sub-populations can be identified in the Iberian peninsula, in Scandinavia and in Italy/France: they appear to be relatively isolated from other wolf populations and are expected to remain distinct for long time (Boitani, 1999; 2000; 2003) (Fig. 1.3).

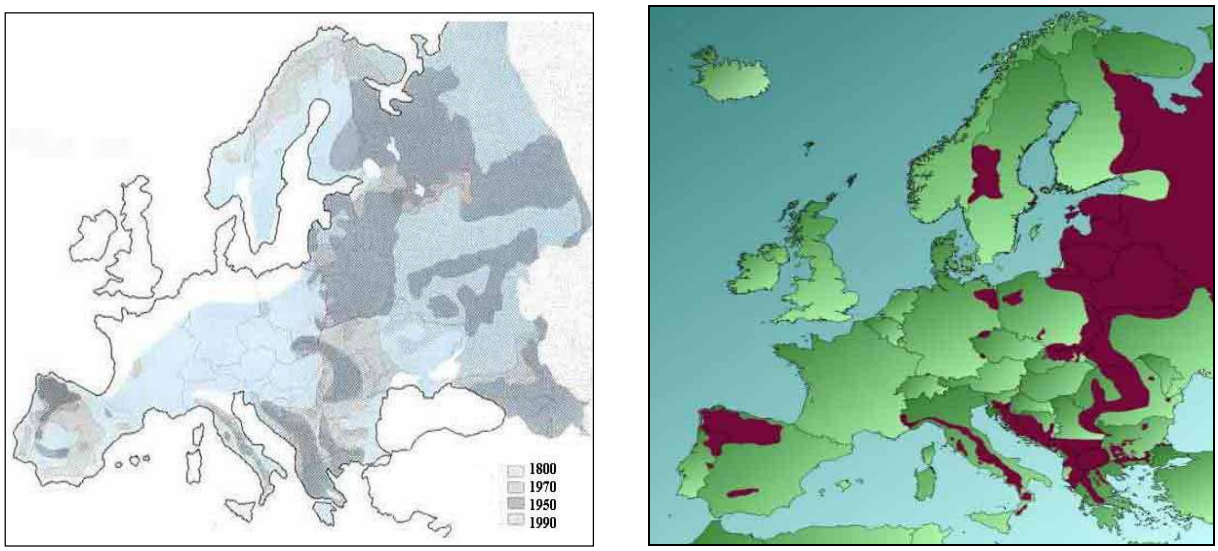

Fig. 1.3: variations of the European wolf distribution from 1800 to 1990 and its current distribution in Europe.

1.1.5 Italian situation and recent re-expansion

Italian wolf population had a continuous distribution from Alps to Sicily until the beginning of the twentieth century, but human activities rapidly reduced it so much that wolves disappeared from the Alps in the 1920s and drastically declined in the two decades after World War II.

During the seventies it approximately consisted of only about 100 individuals surviving isolated in small areas along central and southern Italian Apennines (Zimen & Boitani, 1975). Towards eighties, after that wolf hunting was stopped (1971) and the species was legally protected (1976), it was possible to stop the wolf's decline and its distribution naturally increased.

A census in 1983 suggested the presence of about 220 wolves (Boitani, 1984; Ciucci & Boitani, 1991) estimated an annual population increase of 7% from 1973 to 1988, leading them to argue that the current approximate Italian wolf population should now number about 600-800 individuals (Boitani, 1992; 2003) distributed along the whole Apennine ridge from which they started a natural re-colonization process of previously inhabited areas of their historical range.

From the 1980s onward wolves expanded, crossed the north-western Apennines and reached the south-western Alps in 1992 (Breitenmoser, 1998; Corsi *et al.*, 1999; Poulle *et al.*, 1999) until France and Switzerland (Lucchini *et al.*, 2002; 2004; Boitani, 2003; Valière *et al.*, 2003; Fabbri *et al.*, 2007) (Fig. 1.4).

This quick natural increasing and re-expansion of the Italian wolf population could be due to its great dispersal ability, to the depopulation of the countries, to the mountain area protection and

to the return of the wild ungulates (Boitani, 1986; 1992; Ciucci *et al*., 1997; Ciucci & Boitani, 1999a, b, c).

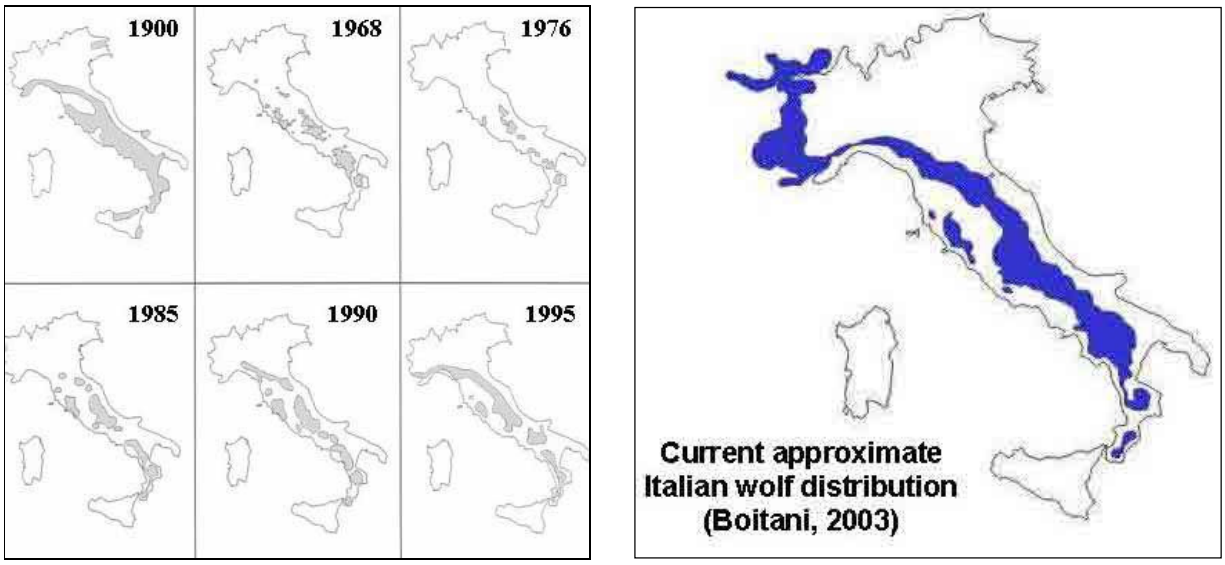

Fig. 1.4: variations of the Italian wolf distribution pattern from 1900 to 1995 and its current approximate distribution in Italy (Boitani, 2003).

1.1.6 Conservation status and recent conservation measures

The conservation of natural wolf populations represents a priority in several European countries, where the species is endangered or has been severely threatened, in the recent past.

At international level the wolf is included in several conservation agreements. The 1996 Red List of the IUCN-World Conservation Union classifies the wolf as vulnerable. The IUCN has also approved a Manifesto of Wolf Conservation, initially drafted in 1973 and later revised to incorporate the changes in wolf population status, public attitudes and management techniques. CITES (Convention on International Trade in Endangered Species of the Wild Fauna and Flora (3.3.1973)) lists the wolf in Appendix II (potentially endangered species), with the exception of Bhutan, Pakistan, India and Nepal where it is listed in Appendix I (species in danger of extinction).

The wolf is also included in Appendix II (strictly protected species) of the Bern Convention (Convention on the Conservation of European Wildlife and Natural Habitats, 19.9.1979). The Standing Committee of the Bern Convention adopted an articulate Recommendation on the protection of the wolf in Europe (Rec. No. 17/1989).

The EC Habitats Directive (92/43 of 21.5.1992) (European Union members only) also lists the wolf in Appendix II (needs habitat conservation) with the exception of the populations in Spain north of the river Duero, the populations in Greece north of the 39° longitude and the populations in Finland. The wolf is moreover listed in Appendix IV (fully protected) with the exception of the populations in Spain north of the river Duero, the populations in Greece north of the 39° longitude and the populations in Finland in areas of reindeer management.

The European Parliament has approved (24.1.1989) a resolution (Doc. A2-0377/88, Ser. A) which calls for immediate steps in favour of wolf conservation in all European States, adopts the IUCN Wolf Manifesto and invites the European Commission to expand and provide financial means to support wolf conservation (Promberger and Schröder, 1993).

In Italy the wolf is a strictly protected species, with law implementation fully delegated from the Ministry of Environment and the Ministry of Agriculture to the Regional Authorities, also responsible for compensation of damage caused by wolf on livestock, thus procedures and amount

of compensation varies across regions. Legal protection of the Italian wolf population started on the 23rd July 1971 when a ministry decree stopped its hunting and was completed in 1976 when the species was given fully protected status, a process stimulated by *WWF International*, that funded a long-term project including a public educational campaign, scientific works and management solutions to protect wolves. The EC Habitats Directive (92/43 of 21.5.1992) lists the Italian wolf in Appendix II (needs habitat conservation) and D.P.R. 357 of 8.11.1997 of Habitats Directive in Appendix IV (fully protected) (Boitani, 2000; 2003; Genovesi, 2002).

1.1.7 Threats, limiting factors and obstacles to conservation

The wolf represents one of the most important conservation and management priorities of our Country. The Italian wolf population has a particular role because it represents one of the few surviving populations in southern Europe after the past persecutions. The quick natural re-expansion of the population after the past decline and the re-colonization of part of its historical range caused some conservation problems for the species, which is still considered as a nuisance in many areas of the peninsula (Boitani & Ciucci, 1993). The main limiting factors to the Italian wolf conservation are:

- Poaching: even if the Italian wolf population is a protected one, poaching is widespread and is probably the single most important mortality factor for the Italian wolves, threatening their survival or recovery. In the last few decades, intense illegal killing (an estimated 15-20% and more of the total) has occurred in Italy in spite of the legal protection established in 1971 (Guberti & Francisci, 1991; Boitani & Ciucci, 1993). Poaching is mainly originated from conflicts between wolves and farmers, because of depredation and damage on livestock (Fico *et al*., 1993; Cozza *et al*., 1996; Ciucci *et al.*, 1997b; Ciucci & Boitani, 1998b; Duchamp *et al*., 2004), and between wolves and hunters, because of the competition for wild ungulates (Boitani, 1982; 1992; 1995, Meriggi *et al*., 1991; Meriggi & Lovari, 1996). Wolf killing is often the accidental result of other hunting and poaching practices (snares for and shooting of wild boar) (Boitani & Fabbri, 1983; Boitani & Ciucci, 1993; Francisci & Guberti, 1993; Boitani, 2000; Genovesi, 2002).

- Habitat quality and food availability: although wolves may survive in the most diverse types of habitat, there seem to be at least a significant correlation between wolf presence and two limiting environmental factors: vegetation cover in which to hide from human sight, and availability of some food resources. Wolves are rarely found where human density is above 30-40 persons/km^2. This would suggest that wolf presence or its diffusion in new areas needs natural habitats populated by wild ungulate prey without human disturbance (Thiel, 1985; Mech, 1989, Corsi *et al*., 1999).

- **Small numbers, low densities and demographic fluctuation:** wolves normally live at low densities (1-3/100 km^2), more rarely at higher densities, and this contributes to making them more vulnerable to ill-planned harvest schemes (Mech, 1970; 1973; Peterson & Page, 1988). Little is known of population dynamics in European contexts, but numerical fluctuations are frequent and they often annihilate the entire local population. These fluctuations are caused or favoured by excessive hunting or poaching. If wolf populations fluctuate too widely, their survival probability will be significantly lower, and their dispersal and re-colonisation rates will also be lower (Ciucci & Boitani, 1998a).

- Feral dogs and wolf hybrids: wolf-like canids form a monophyletic clade of closely related species within the dog family Canidae (Wayne *et al*., 1997). Recent studies of mtDNA supported the single origin hypothesis of the domestication of the dog and suggest that the initial domestication occurred in the eastern part of Asia during the late Pleistocene (ca. 10.000 years ago) (Savolainen *et al.*, 2002). Due to their close relationship, wolves and dogs can successfully hybridize in captivity and in the wild when they co-occur (Wayne *et al*., 1995; Vilà & Wayne, 1999). Risk of natural hybridization may be higher in areas where a species is locally rare and in sympatry with another overabundant species (i.e., wolves (*Canis lupus)* and coyotes (*C. latrans)* in

Minnesota and eastern Canada; Lehman *et al*., 1991), or where wild canids are in contact with feral and free-ranging domestic dogs, as it was for some wolf populations in Europe (Butler, 1994; Rhymer & Simberloff, 1996). Hybridization has the potential to produce morphological, physiological and behavioral changes in captive and wild-living canids (Mengel, 1971; Thurber & Peterson, 1991; Lariviere & Crête, 1993), and eventually led to the origin of a new taxon, as the red wolf (*C. rufus*; Wayne *et al*., 1995). Therefore, hybridization and introgression of domestic genes can diffuse diseases and threaten the integrity of the gene pool of wild canids (Boitani, 1984; Gotelli *et al*., 1994).

In Italy, during the wolf population bottleneck the number of feral and free-ranging dogs in rural areas increased dramatically, thus raising the risk of hybridization (Boitani & Fabbri 1983). Nowadays, the genetic integrity of wolf's gene pool might be seriously compromised by recurrent hybridization (Boitani, 2003). Despite a substantial demographic recovery, Italian wolves are still largely outnumbered by feral and free-ranging domestic dogs, which are estimated to be more than 1 million (Corsi *et al*. 1999; Genovesi & Dupré, 2000). Anyway, genetic studies did not show any evidence of introgression of dog mitochondrial DNA (mtDNA) into Italian wolves (Randi *et al*., 2000), although a few cases of dog-wolf hybrids were already observed in nature (Boitani, 1982); and detected by DNA analyses (Randi *et al.,* 2000). Hybridization has been studied as well as using hypervariable unlinked and linkage canine microsatellite *loci* suggesting that in Italy it is an uncommon process, strictly directional, in fact wolves and free-ranging dogs sporadically hybridise, but Apennine and Alpine wolf populations do not show substantial dog gene introgression (Randi & Lucchini, 2002; Verardi *et al.*, 2006, Fabbri *et al*., 2007).

1.2 INTRODUCTION TO CONSERVATION GENETICS

1.2.1 DNA structure and function

Every individual, with the exception of identical twins, is genetically unique because he possesses a unique patrimony of genetic information (DNA) organized in the chromosomes that are contained in a cell nucleus (nuclear DNA), and in mitochondria, organelles present in cell cytoplasm (mitochondrial DNA or mtDNA). Each DNA molecule takes the form of a double helix built by four nucleotides - the chemical building blocks (Adenine-A, Thymine-T; Guanine-G and Cytosine-C). The structure of the double helix consists of two ribbon-like entities that are entwined around each other and held together by crossbars composed of two bases that have strong affinities for each other. The bases within each chain are bound together by a pentose sugar and phosphate ion, while the opposing strands are held together by weak hydrogen bonds that are relatively easy to break by heating. The linear order in which these four nucleotides follow each other in the double helix of the DNA is called a nucleotide sequence. This very simple structure is extremely stable and allows the DNA to act as a template for protein synthesis and replication (Watson & Crick, 1953).

1.2.2 Mitochondrial DNA

Vertebrate mitochondrial DNA is a circular double helix made up of 15.000-20.000 nucleotides, depending on the species (Hartl & Clark, 1993). It is replicated, independently from cell and DNA nuclear replication, each time the mitochondria divide. During the gametogenesis, the content of cytoplasm and, therefore, the number of mitochondria contained in the gametes significantly change. Mitochondria are provide entirely by cell eggs, therefore during fertilization is the egg cell of the mother that transmits all the mitochondria to the zygotes. Hence mtDNA is haploid and does not recombine. The different types of mtDNA that are originated from mutations and that are present in populations are called "mitochondrial haplotypes".

1.2.3 Nuclear DNA

The genome of vertebrates and many other living organisms is largely made up of coding and non-coding DNA sequences. The first ones are organized in functional domains and are necessary to regulate the protein synthesis consisting of a first phase of transcription of DNA into messenger RNA followed by a phase of translation of the messenger RNA into protein. On the contrary the second ones exist in families of repeated sequences. These tandemly repetitive sequences, commonly known as "satellite DNAs" are classified into three groups:

- Satellite DNA: highly repetitive sequences with very long repeat lengths (up to 5.000.000 nucleotides), usually associated with centromeres.
- Minisatellite DNA: present in hundreds or thousands of *loci* in eukaryotic genomes. These tandem repeats often contain a repeat of more than 10 nucleotides and are present in multiple pairs that produce clusters of 500-30.000 nucleotides. Profiling of these minisatellite *loci* is done using *multi-locus* probes-MLP or *single-locus* probes-SLP to obtain DNA fingerprinting.
- Microsatellite DNA: present in many thousands of *loci* in eukaryotic genomes. They are made up of very short repeats, from 2 to 8 nucleotides, repeated only few times that produce clusters of a few dozen or few hundred nucleotides at every *locus*. Microsatellites are used extensively in forensic genetics and are profiled through PCR.
- Single Nucleotide Polymorphisms (SNPs Single Nucleotide Polymorphisms, widespread in many species' genomes (coding and non-coding regions), and they evolve in a manner well described by simple mutation models, such as the infinite sites model (Vignal *et al*., 2002). These polymorphisms are base substitutions, insertions, or deletions that occur at single positions in the genome (Budowle, 2004). For such a base position with sequence alternatives in genomic DNA to be considered as an SNP, it is considered that the least frequent allele should have a frequency of 1% or greater. Although in principle, at each position of a sequence stretch, any of the four possible nucleotide bases can be present, most SNPs are usually biallelic in practice. One of the reasons for this, is the low frequency of single nucleotide substitutions at the origin of SNPs, estimated to being between 1×10^{-9} and 5×10^{-9} per nucleotide and per year at neutral positions in mammals (Li *et al*., 1981; Martinez-Arias *et al*., 2001).

1.2.4 Genetic mutations and polymorphisms

Mutations generate variability in individuals and populations because they modify DNA sequences and produce the basis on which natural selection can act. Different mutational processes exist and they mainly depend on the structure and function of involved DNA:

- Nucleotide substitution: is the substitution of a nucleotide with another at a certain point in the DNA strand.
- Insertion or deletion of a single nucleotide or series of nucleotides.
- Crossing-over and recombination: crossing-over can be symmetrical, which produces exchanges of corresponding sequences and genetic recombination between two chromosomes, or asymmetrical, which occurs between tandemly repeat DNA that do not precisely align themselves and gives rise to the deletion of a DNA fragment from a chromatid and its insertion on another one.
- DNA slippage: can occur during tandemly repeated DNA replication when the single strand nascent DNA can pair in another point of the DNA template.
- Gene conversion: produces the transfer of a DNA sequence from one allele to another one**.**

1.2.5 Genetic markers

A variable gene, present with two or more variables of the same nucleotide sequence, is defined as polymorphic. Gene coding polymorphisms can generate protein and phenotype polymorphisms which can be used as markers in the identification of samples in forensic science.

Genetic markers are the main tools used to study the genetic variability within and among populations, in fact they allow to estimate which alleles are present inside them (Avise, 1994; Muller & Wolfenbarger, 1999; Parker *et al*., 1995; Sunnucks, 2000).

A genetic marker can be represented by any variable and in hereditable characteristics in populations, determined by genes and not by environment. The main characteristics of a molecular marker are: polymorphism, expression stability during environmental, ontogeny and morphologic changes, well identifiable and amplifiable, Mendelian heredity, expression codominance, many species application. Many kinds of markers exist:

- Visible polymorphisms: phenotype characters with few distinctive variants (*morfi*) not environmental influenced. They are not very common in the eukaryotic genome.
- Molecular markers: macromolecules (proteins, RNA, DNA) which can be separated through electrophoresis in agarose gel within an electric field with a migration speed depending on their weigh and electric charge and visible under ultraviolet light. Alloenzymes belong to these markers (Murphy *et al*., 1996).
- DNA markers: they allow to isolate genetic variability in DNA fragments with different dimensions and weighs and to separate them within electrophoresis gel. Many kinds of markers belong to them:

RFLP: restriction enzymes and restriction fragments length polymorphisms analysis (Jeffreys *et al*., 1985).

RAPD: random amplified polymorphic DNA (Williams, 1990).

AFLP: amplified fragment length polymorphisms (Vos *et al.,* 1995).

VNTRS: variable number of tandem repeats. They are non-coding regions characterized by tandemly repeated sequences. Each repeat can be made up from 10 to 64 nucleotides (minisatellites) or from 2 to 9 nucleotides (microsatellites).

- **SNPs:** Single Nucleotide Polymorphisms. They are hypothesized to become the marker of choice in evolutionary, ecological and conservation studies as genomic sequence information accumulates. As a biallelici marker, SNPs are innately less variable than microsatellites but SNPs are the most prevalent form of genetic variation and hence there is a substantial increase in the number of loci available (Brumfield *et al*. 2003). Furthermore, the simpler mutational dynamics of SNPs lends the advantage of a lowered rate of homoplasy, and, importantly, there is a capacity for rapid, large scale and cost-effective genotyping (Syvänen, 2001; Vignal *et al*., 2002; Brumfield *et al*., 2003; Chen & Sullivan 2003; Schlötterer, 2004).

1.2.6 Conservation Biology and Conservation Genetics

Conservation Biology is a multidisciplinary applied field drawing on ecology, wildlife biology, resource biology, evolutionary, taxonomy, molecular biology, population and conservation genetics.

The aim of *population genetics* is to describe the genetic composition of populations and to understand the causes that determine changes (evolutionary forces). Every species is made up of many evolutionary units, the populations, that contain a certain quantity of genetic variability on which evolution can act. Genetic variability in populations is described through allele frequencies at each *locus* that can vary in the course of generations due to mutations, natural selection, migration or genetic drift.

Conservation genetics is the application of genetic techniques and analysis methods to preserve species and dynamics entities capable of coping with environmental change. It deals with the genetic factors that affect extinction risk and genetic management regimes required to minimise these risks. There are 11 major genetic issues in conservation genetics (Frankham *et al.,* 2002):

- The deleterious effects of inbreeding on reproduction and survival (inbreeding depression).
- Loss of genetic diversity and ability to evolve in response to environmental change.
- Fragmentation of population and reduction in gene flow.
- Random processes (genetic drift) overriding natural selection as the main evolutionary process.
- Accumulation and loss (purging) of deleterious mutations.
- Resolving taxonomic uncertainties.
- Defining management units within species.
- Use of molecular genetic analyses in forensics.
- Use of molecular genetic analyses to understand aspects of species biology (mating, dispersal and migration patterns, reproduction systems) important for conservation.

Deleterious effects on fitness that sometimes occur as a result of outcrossing (outbreeding depression).

1.2.7 Non-invasive genetics

Endangered populations are complicated to study due to their low densities and limited observations (Dalèn *et al*., 2004). Conservation and management of wildlife populations require information on parameters such as population size, demography, gene flow, and population structure but these parameters are difficult to obtain for species that are rare or elusive such as carnivores (Creel *et al*., 2003).

Recent developments in molecular genetics have created new methods such as Non-invasive Genetics or Non-invasive Genetic Sampling (NGS), that have found many applications in ecology, and can resolve some problems of Conservation Biology. They allow populations to be studied and censused (Frantz *et al.*, 2003; Broquet *et al*., 2007) analysing DNA extracted from biological traces such as hairs (Goossens *et al*., 1998; Flagstad *et al*., 1999; Woods *et al*., 1999; Sloane *et al*., 2000), faeces (Taberlet *et al*., 1996, 1999; Gagneux *et al*., 1997; Kohn & Wayne 1997; Kohn *et al*., 1999) and less direct sources of cells (urine and blood traces on snow (Valiere & Taberlet, 2000), sloughed skins (Amos *et al*., 1992; Bricker *et al*., 1996), chewed food material containing buccal cells (Sugiyama *et al*., 1993; Takenaka *et al*., 1993), and bird feathers (Smith *et al*., 1992; Segelbacher, 2002) or egg shells (Pearce *et al*., 1997)).

Non-invasive genetic sampling was introduced about 15 years ago (Taberlet & Bouvet, 1991; Taberlet & Bouvet, 1992; Hoss *et al*., 1992) and consists in a set of field, laboratory and analytical techniques that allow to study the biology of natural populations analysing DNA extracted from biological traces left by individuals and then collected without having (even) to observe, disturb or capture them (Kohn & Wayne, 1997). Conservation biologists in particular have shown a deep interest in these techniques, which are now routinely used in forensic genetics and for investigating the biology and the genetic diversity of elusive, rare and/or endangered species avoiding any risks to impact their survival, their recapture rates or their population dynamics (Kohn & Wayne, 1997; Piggott & Taylor, 2003).

The chief molecular tools used in NGS are mitochondrial DNA (mtDNA) sequencing (Höss *et al*., 1992; Farrell *et al*., 2000) and above all microsatellite *multilocus* genotyping (Palsbøll 1999; Taberlet & Luikart, 1999). They allow to correctly assign the belonging species, to characterise the genetic identity of individuals and their molecular sexing. Many mammal conservation genetic studies using NGS have been recently published, providing abundant information on population parameters, identification, conservation and management strategies of elusive, rare and endangered species (Tikel, Blair & Marsh, 1996; Reed *et al*., 1997; Kohn *et al*., 1999; Bayes *et al*., 2000; Ernest

et al., 2000; Lucchini *et al*., 2002; Waits, 2004; Boulanger *et al*., 2004; McKelvey & Schwartz, 2004; Fabbri *et al*., 2007).

1.2.8 Potential problems of NGS

However, NGS methods might present numerous potential problems which generally tend to limit the efficiency of this approach (Taberlet *et al*., 1996; Broquet *et al*., 2007). Non-invasively collected samples usually provide DNA extracts characterized by low target DNA concentration, low target DNA quality (Taberlet *et al*., 1999), contaminations by alien DNA and various molecules that can disturb or inhibit the polymerase chain reaction (PCR) (Monteiro *et al.*, 1997; Nievergelt *et al*., 2002; Roon *et al*., 2003; Broquet *et al*., 2007), making it unreliable (Gerloff *et al*., 1995; Taberlet *et al*., 1996; Gagneux *et al*., 1997a).

As amplification success and genotyping errors can be sensible to template DNA concentration and composition (Gerloff *et al*., 1995; Wasser *et al*,. 1997; Goossens *et al*., 1998; Morin *et al*., 2001), microsatellite genotypes from non-invasive samples can be affected by errors (Taberlet *et al*., 1996; 1999; Gagneux *et al*., 1997 Smith *et al.*, 2000) such as allelic dropout (ADO) which is the stochastic failure of one allele to amplify for heterozygous individuals, producing false homozygotes (Navidi *et al*., 1992; Taberlet *et al*., 1996; Goossens *et al*., 1998; Constable *et al*., 2001) and false alleles ('misprinting') which are artefacts of amplification products generated during the first steps of PCR that can be misinterpreted as true alleles (Taberlet *et al*., 1996; Goossens *et al*., 1998; Bradley & Vigilant, 2002).

Microsatellite genotypes are commonly used for individual identification, parentage, relatedness, and population genetics (Taberlet *et al*., 1997; Constable *et al*., 2001; Garnier *et al*., 2001). So those genotyping errors affect both the allele frequency estimates and the accurate discrimination of different genotypes. False estimates of allele frequency can create an artificial excess of homozygotes (Taberlet *et al*., 1996; Gagneux *et al*., 1997a), a false departure from Hardy–Weinberg equilibrium (Xu *et al*., 2002), an overestimation of inbreeding rate (Gomes *et al*., 1999; Taberlet *et al*., 1999) or unreliable inferences about population substructures (Miller *et al*., 2002). Erroneous genotypes can distort or overestimate population size estimates (Creel *et al*., 2003; McKelvey & Schwartz, 2004), individual identification (Taberlet & Luikart, 1999; Paetkau, 2003) and parentage analysis (Miller *et al*., 2002).

1.2.9 Possible solutions to NGS problems

Many authors have recognized the complexities of non-invasive genotyping, and have developed methods to address these problems (Taberlet *et al.,* 1996, 1999; Gagneux *et al*., 1997; Morin *et al*., 2001; Miller *et al*., 2002). Contaminations among samples could be avoided using dedicated rooms for extraction and amplification of low-DNA-content samples, while amplification from alien DNA could be avoided by using specific primers (Bradley & Vigilant, 2002).

Numerous quality control protocols have been developed, including the adoption of multiple tube approaches where the same DNA samples are amplified independently several times *per locus* (Navidi *et al*., 1992; Taberlet *et al*., 1996; Lucchini *et al*., 2002; Frantz *et al*., 2003; Fabbri *et al.*, 2007), comparison of genotypes obtained with those from matched blood or tissue (Wasser *et al*., 1997; Kohn *et al*., 1999; Ernest *et al*., 2000; Sloane *et al.,* 2000; Parsons, 2001; Fernando *et al.,* 2003), strategic re-amplification at *loci* likely to harbour errors (Miller *et al*., 2002) and that present one or two mismatches among couples of individuals very similar (Palsboll *et al*., 1997; Woods *et al*., 1999; Paetkau, 2003), pre-screening of samples for DNA quantity (Morin *et al* . 2001; Segelbacher, 2002) and the use of pilot studies (Taberlet & Luikart, 1999) and simulations (Taberlet *et al*., 1996; Valiere *et al*,. 2002). Anyway all these methods can involve a large extra experimental

effort (Brzustowicz *et al*., 1993; Ghosh *et al*., 1997; Ewen *et al.*, 2000), increasing the consumables, costs and time required (Morin *et al.*, 2001).

It is therefore cheaper to conduct statistical tests on already available data. Commonly, the Hardy-Weinberg equilibrium test (Gomes *et al.,* 1999) is checked to reveal the homozygous excess resulting from either null alleles or allelic dropout.

1.3 STATISTICAL METHODS

Even though many reasonable statistic approaches are available to analyse the genetic structure of populations and to estimate the absolute and effective population sizes, most of them, used in this study are based on ***F*** and ***Bayesian Statistics***.

In population genetics, ***F*-statistics** (also known as fixation indices) describe the level of heterozygosity in a population; more specifically the degree of a reduction in homozygosity when compared to Hardy-Weinberg expectation. Such changes can be caused by the **Wahlund effect** (the reduction of heterozygosity in a population caused by subpopulation structure), **inbreeding**, **natural selection** or any combination of these.

The concept of *F*-statistics was developed during the 1920s by the American geneticist Sewall Wright who was interested in inbreeding in cattle, but its applications deeply increased after the 1960s when the advent of molecular genetics allowed heterozygosity in populations to be reliably measured.

F-statistics measure the correlation between genes drawn at different levels of a (hierarchically) subdivided population. This correlation is influenced by several evolutionary forces, such as mutation and migration, but it was originally designed to measure how far populations had gone in the process of fixation owing to genetic drift.

The different F-statistic measures, F_{IS}, F_{ST}, and F_{IT}, are related to the amounts of heterozygosity at various levels of population structure. Together, they are called F-statistics, are derived from *F*, the inbreeding coefficient, and look at different levels of population structure: $\mathbf{F_{IT}}$ is the inbreeding coefficient of an individual (**I**) relative to the total (**T**) population, as above; $\mathbf{F_{IS}}$ is the inbreeding coefficient of an individual (**I**) relative to the subpopulation (**S**), using the above for subpopulations and averaging them; and $\mathbf{F_{ST}}$ is the effect of subpopulations (**S**) compared to the total population (**T**), and is calculated by solving the equation:

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

In a simple two-allele system with inbreeding, the genotypic frequencies are:

$$p^2 + Fpq \text{ for } \mathbf{AA};\ 2pq(1 - F) \text{ for } \mathbf{Aa};\ \text{and } q^2 + Fpq \text{ for } \mathbf{aa}.$$

The value for **F** is found by solving the equation for **F** using heterozygotes in the above inbred population. This becomes one minus the observed number of heterozygotes in a population divided by its expected number of heterozygotes at Hardy–Weinberg equilibrium:

$$F = 1 - \frac{\mathrm{O}(f(\mathbf{Aa}))}{\mathrm{E}(f(\mathbf{Aa}))} = 1 - \frac{\mathrm{ObservedNumber}(\mathbf{Aa})}{n\mathrm{E}(f(\mathbf{Aa}))},$$

where the expected value at Hardy–Weinberg equilibrium is given by

$$\mathrm{E}\left(\int(\mathbf{Aa})\right) = 2\,p\,q$$

where p and q are the allele frequencies of A and a, respectively. It is also the probability that at any locus, two alleles from the population are identical by descent.

A within subpopulations F-statistic can be estimated from a ratio of the observed to expected heterozygosity where,

$$F_{IS} = \frac{\bar{H}_S \quad H_I}{\bar{H}_S}$$

A reformulation of the definition of **F** would be the ratio of the average number of differences between pairs of chromosomes sampled within diploid individuals with the average number obtained when sampling chromosomes randomly from the population (excluding the grouping per individual). One can modify this definition and consider a grouping per sub-population instead of per individual. Population geneticists have used that idea to measure the degree of structure in a population.

Unfortunately, there is a large number of definitions for F_{ST}, causing some confusion in the scientific literature. A common definition is the following:

$$F_{ST} = \frac{\mathrm{var}(p)}{p\,(1-p)}$$

where the variance of *p* is computed across sub-populations (Wright, 1951; 1965; 1969; 1978; Weir & Cockerham, 1984; Slatkin, 1991; Weir & Hill, 2002).

Bayesian Statistic is based on Bayes' theorem (also known as Bayes' rule or Bayes' law), set out by Thomas Bayes (1702-1761), an English clergyman in 1764. It is a result in probability theory relates the conditional and marginal probability distributions of random variables. In some interpretations of probability, Bayes' theorem tells how to update or revise beliefs in light of new evidence "*a posteriori*", according to which, the probability *a posteriori* of an event (given by evidence) can be obtained combining the observations (probability conditional or *likelihood*) with the subjective degree of belief (*a priori*) about the same event based on experiences or theories independent from data. Bayesian probability is an interpretation of the probability calculus where the concept of probability can be defined as the degree to which a person (or community) believes that a proposition is true. The probability of an event *A* conditional on another event *B* is generally different from the probability of *B* conditional on *A*. However, there is a definite relationship between the two, and Bayes' theorem is the statement of that relationship.

Some researchers consider the scientific method as an application of Bayesian probabilist inference because they claim Bayes' Theorem is explicitly or implicitly used to update the strength of prior scientific beliefs in the truth of hypotheses in the light of new information from observation or experiment. This is said to be done by the use of Bayes' Theorem to calculate a posterior probability using that evidence and is justified by the Principle of Conditionalisation that P'(h) = P(h/e), where P'(h) is the posterior probability of the hypothesis 'h' in the light of the evidence 'e', but which principle is denied by some. Adjusting original beliefs could mean (coming closer to) accepting or rejecting the original hypotheses.

Since the 1950s, Bayesian theory and Bayesian probability have been widely applied and it has recently been shown that Bayes' Rule and the Principle of Maximum Entropy (MaxEnt) are completely compatible and can be seen as special cases of the Method of Maximum (relative) Entropy (ME). This method reproduces every aspect of orthodox Bayesian inference methods. In addition this new method opens the door to tackling problems that could not be addressed by either

the MaxEnt or orthodox Bayesian methods individually (Lindley, 1990; West & Harrison, 1989; O'Hagan, 1994; Sivia, 1996; Pritchard *et al*., 2000; Tijms, 2004).

The main differences between ***F* (or *frequency*)** and ***Bayesian Statistics*** lie in the definition, interpretations and in the effective calculus of probabilities (Press, 1972), in fact:

- F statistics assigns probabilities to random events according to their frequencies of occurrence or to subsets of populations as proportions of the whole and allows to compare the test hypothesis to a model/hypothesis (the "null" hypothesis). The probability p of an event H depends on the number of times (n) the event occurs on the total number of tests (N). The probability p of H corresponds therefore to its frequency:

$$pH = n(H)/N.$$

- Bayesian statistics assigns probabilities to propositions that are uncertain; conditions on the data actually observed, and is therefore able to assign posterior probabilities to any number of hypotheses directly. The requirement to assign probabilities to the parameters of models representing each hypothesis is the cost of this more direct approach. The probability p is an estimation of *likelihood* that that the event H occurs. We can have convictions (subjective) or information (objective, even though not exactly quantifiable) that an event may more or less occur frequently. Posterior probability of an event H corresponds on the probability that the event H occurs given the evidence E:

$$Pr\,(H) = Pr\,(H/E).$$

1.4 GENETIC APPLICATION IN WOLF STUDIES

During the last tirthy years many studies about the Italian wolf population have been carried out to investigate its biology, distribution range, size, and its possible coexistence with people (Boitani & Zimen, 1975; Boitani & Ciucci, 1992; 1993; Boitani, 1995; 2000; 2003). Most of these studies have particularly involved the populations living in the southern, central, and northern Apennine Mountains (Francisci & Mattioli, 1992; Meriggi *et al*., 1991; Mattioli *et al*., 1995; Randi *et al.*, 1993; 1995; 2000), but after that the species naturally increased and started a natural re-colonization of western Alps until France and Switzerland, many projects and research programs have been planned and carried out also in these newly colonized areas (Lucchini *et al*., 2002; Valière *et al.,* 2003; Fabbri *et al*., 2007).

This recent and quick natural re-expansion of wolves in areas where they had been exterminated caused some problems about their management determining the need to create an Action Plan to ensure their conservation and their coexistence with people in Italy (Boitani, 2000; 2003), but the protection and conservation of so interesting predators needs a continual monitoring of their biology, presence and distribution (Boitani, 2000; 2003).

As wolves are shy and elusive predators, with a great dispersal ability and adaptable to every kind of environmental conditions (Mech, 1970), it is very difficult to study them using the only classical field research methods such as diet analysis (Guberti *et al*., 1993; Ciucci *et al*., 1996) snow-tracking (Ciucci & Boitani,1999a;b;c), wolf-howling and radio-tracking (Ciucci *et al.*, 1997).

For these reasons projects based on modern molecular techniques, and in particular non-invasive genetic sample studies are getting more and more applied and useful to monitor the presence, distribution and colonization events of the species (Lucchini *et al*., 2002; Valière *et al.,* 2003; Fabbri *et al*., 2007). The first genetics studies of population variability about the Italian

wolves examined variation in allozymes (Randi *et al.*, 1993; Lorenzini & Fico, 1995). Following studies utilized nucleotide indirect sequence variation in mitochondrial DNA analysing RFLP to estimate nucleotide sequence variation within populations and their relationships (Randi *et al.*, 1995). More recently mtDNA sequencing by PCR have been applied to wolflike canids analysing a non-coding hypervariable region of the mitochondrial genome (control region) to allow a more precise reconstruction of historical demographic events such as colonization and gene flow, bottlenecks and hybridization (Randi *et al.*, 2000; Scandura *et al.*, 2001).

Recent developments in molecular genetics have created new methods that involve microsatellite *loci* (tandem repeats of two to six nucleotide sequence) (Bruford & Wayne, 1993; Hancock, 1999) to quantify components of variations within and among populations and to study individual relatedness within social groups (Smith *et al*., 1997; Bossart & Prowell, 1998).

Moreover microsatellite studies allow wolf populations to be censused through non-invasive DNA sampling, in fact the intestinal lining cells shed contained in faeces can represent an alternative source of DNA particularly useful to characterise the genetic identity of individuals (*multilocus* genotype) (Palsbøll, 1999; Taberlet & Luikart, 1999) and to provide abundant information on population parameters, home ranges, genetic variation and phylogenetic relationships in a free ranging and elusive mammal species such as the wolf (Tikel, Blair & Marsh, 1996; Reed *et al*., 1997; Kohn & Wayne, 1997; Lucchini *et al*., 2002; Valière *et al.,* 2003; Fabbri *et al*., 2007).

For these reasons this wolf study was almost completely carried out using non-invasive genetic analysis techniques based on microsatellite *loci* genotyping.

1.5 AIMS OF THE THESIS

The main purposes of this conservation genetic study are:

- to monitor the presence and the distribution of wolves living in the northern Apennines (Emilia Romagna) analysing DNA extracted from non-invasive samples, mainly from presumed wolf scats collected in the study area;
- to create an useful dataset containing both genetic information and field data and a specific regional digital cartography about all the collected wolf data to estimate the minimum number of reproducing individuals, mapping pack localizations and carry out some preliminary hypotheses on pack dynamics;
- to investigate whether wolves and feral or free-ranging domestic dogs eventually hybridize with a substantial dog gene introgression in the wild population;
- to investigate whether there are feral or free-ranging domestic dogs that usually frequent the same areas in which wolf packs established their territories and home ranges;
- to use genetic data as capture-mark-recapture ones to obtain a reliable population size estimation necessary for wolf conservation and management. In fact non-invasively detected *multilocus* genotypes, if individuals are sampled sufficiently often to estimate re-sighting probabilities (Otis *et al*., 1978; Seber, 1982), can be used for censusing also populations whose individuals are difficult to locate like wolves;
- to develop new genotyping methods faster and more reliable than microsatellite *loci* genotyping, in fact microsatellite genotyping from non-invasive samples can be error prone due to allelic dropout and false alleles. SNPs genotyping could represent a near future application in non-invasive genetics as a promising and innovative faster and more reliable method to analyse low quality and quantity DNA samples like non-invasive ones.

Emilia Romagna Region represents a very important study area to explain the ongoing expansion process of the Italian wolf population because it acts as a natural narrow ecological corridor along the ridge of the north-western Apennines linking the central-northern Apennine Mountains with

western Alps. For these reasons I based my study on about 4000 presumed wolf scat samples, collected non-invasively, from 2000 to 2007, in the populations living in the Apennine ridge of Emilia Romagna Region and which contributed to the ongoing Alpine recolonization.

DNA samples were extracted using different extraction methods and genotyped by PCR at 6 autosomal microsatellite *loci*. All the samples analysed were mapped by GIS to obtain spatiotemporal locations of the individual genotypes and wolf pack hypotheses within the study areas, in comparison with observations that are being collected during ongoing field research.

At the end, to investigate about the genetic variability within the wolf population living in Emilia Romagna and to examine the gene flow among this and the other populations living in Italy, all the genetic data obtained in this study were compared to the ones collected during past projects.

The knowledge of the genetic status and the continuous monitoring of the presence and distribution of the wolf population living in an area of strong ongoing expansion process with eventual hybridization risk, such as the Emilia Romagna Region, are essential for the conservation and management of the species, both at regional and at national level.

CHAPTER SECOND: MATERIALS AND METHODS

Procedures for single genome analyses depend on the kind of biological samples and vary according to the different analyses carried out. In this study all the genetic analyses were performed at the Italian Wildlife Institute (INFS).

2.1 SAMPLES AND COLLECTION LOCALITIES

In this study more than 3.500 presumed wolf samples were analysed. Most of them were represented by presumed wolf scat samples which were mainly noninvasively collected in the northern Apennines from March 2000 to March 2007, while only some tissue samples were collected from presumed wild-living wolves accidentally or illegally killed during the study period. The study area includes most of the Emilia-Romagna Apennine Ridge (Bologna, Forlì-Cesena, Modena, Parma, Piacenza, Ravenna and Reggio-Emilia Provincial Administrations), 2 National Parks (Foreste Casentinesi, Alto Appennino Tosco-Emiliano), 4 Regional Parks (Corno alle Scale, Frignano, Gigante, Cento Laghi) (Fig. 2.1); all the monitoring project territory is managed by the Italian Forestry Corp (CFS) in collaboration with the Regional Police and Parks' Personnel .

As the study area is a very vast one, a great logistic effort was required to carry out an appropriate sampling activity during the whole project period. Faeces were usually collected by travelling defined systems of transects opportunistically chosen along human trails/roads to optimize the study area monitoring and look for signs of wolf presence. The sampling effort was more intensive where the species had already been reported and where there was a major availability of financial resources (National Park of Foreste Casentinesi and Life Natura 2000 Project Area) (Fig. 2.1). Many samples were collected in winter during snow-tracking sessions. This strategy slows down DNA degradation because of lower temperatures and makes it possible to sample not only the dominant individuals that mark with high frequency and are easy to find during all seasons, but also juveniles that don't usually mark. Moreover, sampling on snow allows associating faeces sampled along each snow track to individuals that could belong with high probability to the same pack.

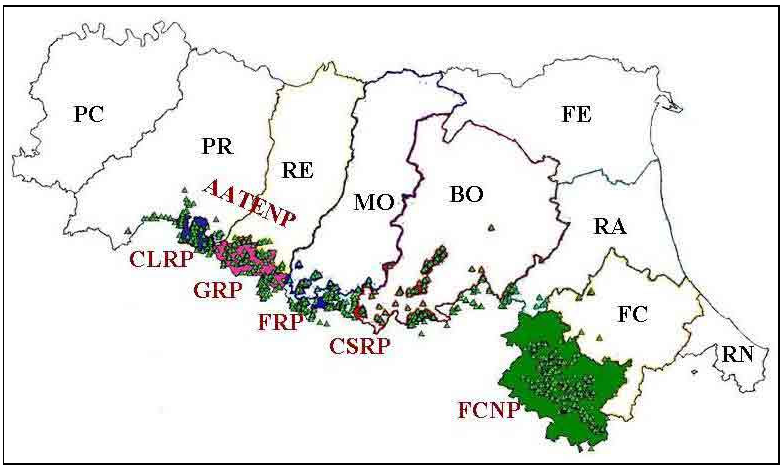

Fig.2.1: Emilia-Romagna Apennine ridge study area including provincial administration of Bologna (BO), Forlì-Cesena (FC), Modena (MO), Parma (PR), Piacenza (PC), Ravenna (RA) and Reggio-Emilia (RE), National Park of Foreste Casentinesi (FCNP) and Alto Appennino Tosco-Emiliano (AATENP), Regional Park of Corno alle Scale (CSRP), Frignano (FRP), Gigante (GRP) and Cento Laghi (CLRP). Green points represent sampling effort and distribution.

2.2 SAMPLE COLLECTION AND THEIR PRESERVATION

The sample collection phase is fundamental to ensure a good success of the following genetic analyses based on PCR techniques because analyses procedures and the quality of the results are dependent on the quality of samples and possible contaminations. For these reasons it necessary to collect and preserve biological samples in the best possible way.

According to the INFS sampling protocol, our field collaborators were asked to avoid the collection of samples older than 2 weeks, favouring the most fresh ones. Samples collected in the field were individually stored and separately enclosed into 50 ml plastic tubes containing 95% ethanol, and preserved at -20 °C until shipped to the laboratory. It is extremely important to preserve the samples in volumes of ethanol at least 3 times greater than the sample weight. Ethanol dehydrates the samples blocking the biochemical reactions that could degrade the DNA. The shipping can be made at room temperature because DNA is stable for several days in ethanol. Once arrived at the laboratory, before any further manipulation, all the samples were deep-frozen at -80 °C for at least 10 days to kill any *Echinococcus* eggs.

Our field collaborators were also asked to compile, during samplings a technical card containing important and useful field information such as sampling localities, sample quality but above all the geographic coordinates necessary to map by gis the spatiotemporal locations of each collected sample, of the individual genotypes and wolf packs within the study areas, allowing comparison with observations collected during ongoing field research.

2.3 DNA MARKERS USED IN THE ANALYSES

2.3.1 Nuclear DNA: Microsatellites

Microsatellites have quickly become of standard usage as genetic markers in DNA fingerprinting. They are nuclear DNA sequences made up of a simple motif of 2-8 nucleotides, that is repeated in tandem for a certain number of times with or without interruptions due to the insertion of other nucleotides or other sequences. Microsatellites have been identified in the genome of all organisms analyzed up to now and are distributed in a more or less random way in chromosomes (Mellersh & Ostrander, 1997). They are not frequent in coding sequences of genes (exons), while they may be present in introns. The composition of microsatellite sequences is variable. In fact the short DNA segments can be made up of mono, di, tri or tetranucleotides (Mellersh & Ostrander, 1997; Stallings *et al.*, 1991; Tautz & Renz, 1984). Microsatellites present very high estimated mutation rates (in vertebrates 10^{-4}-10^{-5} mutations per *locus* for every generation) which determine high levels of polymorphisms, in fact in a single *locus* more than 10 alleles can be present which differ for the number of repeats and therefore for their molecular weight.

Two models have been hypothesized to explain the main mutation mechanisms that could generate microsatellites:

- DNA slippage: it occurs during replication when the nascent DNA separates and reassociates itself temporarily from the DNA template. During replication of non-repetitive sequences the possible disassociation of the sister chromatid does not usually generate mutations because the nascent DNA can reassociated only and exactly in the complementary point of the DNA template. Instead, during tandemly repeat DNA replication, the single strand nascent DNA can pair in another point of the DNA template. When replication continues, the nascent DNA is found to be longer or shorter than the template (Hancock, 1995).
- DNA recombination: it can vary microsatellite length through asymmetrical crossing-over or gene conversion. Asymmetrical crossing-over occurs very frequently between tandemly repeated

DNAs that do not align themselves precisely giving rise to the deletion of a DNA fragment from a chromatid and its insertion into another chromatid. It may occur between two chromatids of the same chromosome or between two different chromosomes. Gene conversion, to answer to DNA damages, produces the unidirectional transfer of a DNA sequence from one allele to another one (Hancock, 1999).

As microsatellites show a high polymorphism rate and a high-quality result reliability, they are considered very popular genetic markers among molecular biologists. In fact these markers are important for map building since the distribution of this sequence repeats within the genome is random and act as landmarks for the organization of the DNA (Mellersh & Ostrander, 1997).

Moreover they find many applications in population genetics, in fact they represent particularly useful tools to study population story and structure, their genetic variability and allow to investigate about the presence and distribution of wild species during non-invasive monitoring projects because they can be used for the identification of individuals and their relationships and they can contribute to the population size estimation (Mellersh & Ostrander, 1997; Reed *et al*., 1997; Kohn & Wayne, 1997; Lucchini *et al*., 2002; Valière *et al.,* 2003; Fabbri *et al*., 2007).

2.3.2 Nuclear DNA: Single Nucleotidic Polimorphisms (SNPs)

As suggested by the acronym, a SNP (single nucleotide polymorphism) marker is just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. For such a base position with sequence alternatives in genomic DNA to be considered as an SNP, it is considered that the least frequent allele should have a frequency of 1% or greater. Although in principle, at each position of a sequence stretch, any of the four possible nucleotide bases can be present, SNPs are usually biallelic in practice. One of the reasons for this, is the low frequency of single nucleotide substitutions at the origin of SNPs, estimated to be between 1×10^{-9} and 5×10^{-9} per nucleotide and per year at neutral positions in mammals (Li *et al*., 1981, Martinez-Arias *et al*., 2001). Therefore, the probability of two independent base changes occurring at a single position is very low. Another reason is due to a bias in mutations, leading to the prevalence of two SNP types. Mutation mechanisms result either in transitions: purine-purine (A↔G) or pyrimidine-pyrimidine (C↔T) exchanges, or transversions: purine-pyrimidine or pyrimidine-purine (A↔C, A↔T, G↔C, G↔T) exchanges. With twice as many possible transversions than transitions, the transitions over transversions ratio, should be 0.5 if mutations are random. However, observed data indicate a clear bias towards the transitions. One probable explanation for this bias is the high spontaneous rate of deamination of 5-methyl cytosine (5mC) to thymidine in the CpG dinucleotides, leading to the generation of higher levels of C↔T SNPs, seen as G↔A SNPs on the reverse strand (Cooper & Krawczak, 1989; Wang *et al.,* 1998). Some authors consider one base pair indels (insertions or deletions) as SNPs, although they certainly occur by a different mechanism.

During the last ten years, the use of molecular markers, revealing polymorphism at the DNA level, has been playing an increasing part in animal genetics studies. Amongst others, the microsatellite DNA marker has been the most widely used, due to its easy use by simple PCR and to the high degree of information provided by its large number of alleles *per locus.*

Despite this, and even though they are only biallelic markers, SNPs are now on the scene and have gained high popularity in animal genetics, in fact, the increasing progress made in the molecular techniques used to produce SNP data, the automation of allele scoring and the development of algorithms for genetic analyses (Abecasis *et al*., 2002) allow to overcome the limitations due to the low heterozygosity of SNPs and to produce an equivalent amount of information as with microsatellites. The very high density of SNPs in genomes usually allows to analyse several of them at a single *locus* of a few hundred base pairs, so that SNPs could represent a more reliable and faster genotyping method because amplifying short sequences and extending

single nucleotides, SNPs genotyping should increase PCR success and reduce the allelic dropout and false allele rates.

2.4 ANALYSIS PROCEDURES

2.4.1 DNA extraction

Extraction process is a crucial step because it must isolate DNA molecules which are present in a sample producing available solutions of DNA without contaminants and must impede further degradations during laboratory procedures. In this study both manual and automated extraction methods to isolate available DNA from scats and tissues were used. Negative (no scat or tissue material added to the extractions) and positive (samples with known genotypes) controls were always used to check possible contaminations during both extraction processes.

2.4.2 Manual extraction

2136 scat samples and 12 tissue samples were manually extracted using a guanidinium thiocyanate and diatomaceus earth (guanidinium-silica) protocol (Gerloff *et al*., 1995). Excremental DNAs were extracted in a separate room only dedicated to low-DNA-content samples to avoid contaminations among them. The used solutions are characterized by the presence of:
TRIS: it maintains a constant pH value that inhibits the activity of enzymes that degrade DNA;
EDTA: it acts as chelants of bivalent calcium and magnesium ions inhibiting the activity of DNase that requires the presence of these ions;
GUS (Guanidinium Thiocyanate): it produces the chemical disintegration of protein structures.

Guanidinium-silica protocol (summary)
Preparation of the samples:
- a piece of tissue (50 mg) or of scat material (80 mg) is cut and transferred into an "eppendorf" test tube of 2.0 ml containing 500/900 µl of ***GUS Lysis Buffer***; flamed sterilized scalpels and forceps are used.

Digestion of the samples:
- in rotation at 57°C overnight.

Collecting DNA:
- centrifuge at room temperature for 10 minutes and collect the supernatant;
- add 500/900 µl of ***GUS Binding Solution*** and in rotation for 1 hour;
- centrifuge at room temperature for 1 minute and eliminate the supernatant.

DNA is now bound to micro-granules of pelleted silica at the bottom of the test tube. Each pellet is washed twice, each time with 500/900 µl of ***GUS Washing Solution*** and then centrifuged at room temperature for 1 minute. The supernatant is eliminated, each pellet is washed again twice, each time with 1ml of **EtOH 70%** and centrifuged at room temperature for 3 minutes. The pellet is dried in open "eppendorf" in a thermostatic multiblock at 56 °C for 10 minute.

The pellet is re-suspended in 200 µl of **TE** for 15 minutes at 56°C. The supernatant with the DNA is transferred in a new "eppendorf" and preserved in freezer at -20°C

2.4.3 Automated extraction

1402 scat samples and 22 tissue samples were extracted in a automated manner by the ***MULTIPROBE IIEX*** robot (Perkin Elmer) and using the QUIAGEN Stool and tissue extraction

kits (QUIAGEN). The robot consists of 2 mechanical hands controlled by an appropriate software which can be set up each time according to the number of samples and to the extraction kind and conditions. This procedure consists of a first manual phase and of a second automated one.

Manual phase:

Preparation of the samples:

- a piece of tissue (50 mg) or of scat material (80 mg) is cut and transferred into an "eppendorf" test tube of 2.0 ml containing 20 µl of ***Proteinase K*** and 180 µl of ***ATL Lysis Buffer*** (previously warmed up at 57°C for 5 minutes); flamed sterilized scalpels and forceps are used.

Digestion of the samples:

- in rotation at 56°C for 30 minutes.

Collecting DNA:

- centrifuge at room temperature for 10 minutes and collect the supernatant;
- transfer the supernatant in a new "eppendorf" and centrifuge at room temperature for other 10 minutes;
- transfer the supernatant in a new appropriate QUIAGEN tube.

Automated phase:

- link the multiblock with QUIAGEN tubes to the robot's platform containing a vacuum pump system to aspirate liquid solutions and a serious of silica-gel filters to trap the DNAs.
- the mechanical hands add 410 µl of ***AL/E Lysis Buffer*** (previously warmed up at 57°C for 5 minutes) to each QUIAGEN tube containing digested sample solutions and the software activates the pup system to isolate the DNA;
- the mechanical hands add 500 µl of ***AW1 Washing Solution*** and the software activates the vacuum for 10 minutes;
- the mechanical hands add 500 µl of ***AW2 Washing Solution*** and the software activates the vacuum for 10 minutes;
- the mechanical hands add 300 µl of ***AE Solution*** (elution solution) to each sample re-suspending the DNA linked to silica filters at room temperature for 1 hour.

The solution with the DNA is transferred in a new "eppendorf" and preserved in freezer at -20°C

2.4.4 DNA amplification

DNA amplification is a necessary procedure to obtain sufficient DNA quantity to carry out molecular analyses. DNA sequences made up of a few dozen or thousands nucleotides and present in a single copy in DNA samples can be amplified effectively up to 10 million times in a few hours.

using Polymerase chain reaction (PCR) (Mullis *et al.*, 1986). PCR occurs by reconstructing the chemical conditions necessary to obtain DNA synthesis in vitro.

First, it is necessary to identify the gene or DNA sequence that one wishes to amplify. The sequence to be amplified is flanked on both side by sequences that must be at least partially known, in fact to start off PCR it is necessary to chemical synthesise a pair of oligonucleotides (20-30 bp) "primers" that are at least partially complementary to the flanking sequences and can bind to flanking regions starting the duplication process of the target sequence. PCR uses single stranded DNA as a template and, by the action of DNA polymerase enzyme, it synthesizes a complementary strand over and over again, until extensive quantities are produced. Every PCR consists of a cycle, repeated many times, made up of the following steps: denaturation of the DNA sample at temperatures up to 90-95°C; binding of the primers to the flanking sequences: it occurs at temperatures which vary from 40°C and 55°C, depending on the length of the primers and their base sequence; extension of the primers through the enzymatic action of a thermoresitant DNA

polymerase (Taq Polymerase) which catalyses the extension of the primers: it occurs at 72°C end ends in the complete replication of both strands of the target sequence (Fig.2.2).

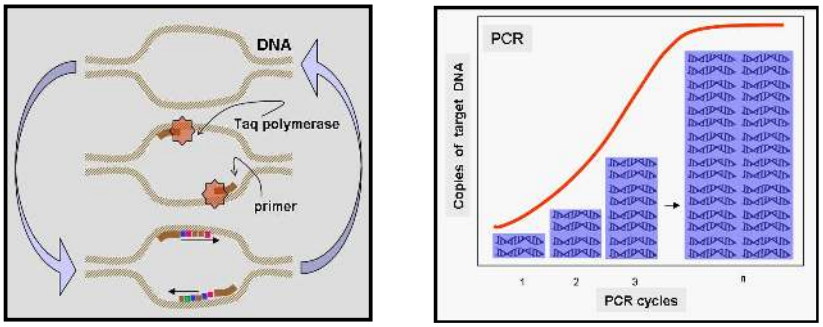

Fig. 2.2: different phases of the Polymerase Chain Reaction and exponential amplification of target DNA.

By the end of the first cycle, every form of the target sequence present in the sample is replicated once, and the thermal cycle of the PCR is repeated a second time and then many other times (20-40) producing an exponential replication of the target sequence because with every successive cycle the synthesised DNA is doubled (Fig. 2.2).

The advantage of using PCR is that the DNA does not have to be in large amounts or even purified to be amplified. It has also been successfully used to amplify ancient DNA (Hofreiter *et al*., 2001).

PCR efficiency depends on the capacity to faithfully amplify the target DNA. If the primers anneal to the target sequence and also to other sequences present in the DNA samples, then the PCR would amplify "aspecific" sequences which would make the analyses and interpretations of the results problematic and even impossible.

2.5 MICROSATELLITE ANALYSESES

As they are very polymorphic and reliable molecular markers, microsatellites genotypes reveal to be very useful tools commonly used to correctly assign the belonging species and probable hybrid detection, to characterise the genetic identity of individuals, their molecular sexing, their parentage and relatedness and to study population genetics (Taberlet *et al*., 1997; Constable *et al*., 2001, Lucchini *et al.,* 2002; Fabbri *et al.,* 2007).

Microsatellite genotyping can be rapidly performed by simple PCR fallowed by electrophoresis gel or automated sequencer analyses. The precision and the reliability of the results increase with the number of microsatellite *loci* which are used.

2.5.1 Microsatellite Amplification

As repeated sequences of microsatellites are flanked by unique sequences, it is possible to design PCR primers (Forward and Reverse) that selectively amplify microsatellite *loci*. Genotyping analysis is done to identify the molecular weigh of the alleles present at each *locus* via electrophoresis.

In this work 12 microsatellite *loci* including 6 dinucleotides (CPH2, CPH4, CPH5, CPH8, CPH12; Fredholm & Wintero, 1995; C09.250; Ostrander *et al*., 1993), and 6 tetranucleotides (FH2004, FH2079, FH2088, FH2096, FH2132 and FH2137; Francisco *et al.,* 1996), were selected, for their polymorphism and reliable scorability in wolves and dogs, among 18 canine microsatellites previously used in a study about Italian wolves (Randi & Lucchini, 2002) and used for all the analyses. 6 microsatellites (CPH2, CPH8, FH2004, FH2088, FH2096, FH2137) were amplified by

PCR to identify the individual genotypes, and the other 6 PCR-amplified *loci* (CPH4, CPH5, CPH12, FH2079, FH2132 and C09.250) were added them to improve estimates of kinship, to clarify doubts about similar genotypes with a few differences, to confirm or not possible hybrids wolf-dog and to carry out preliminary Italian wolf population studies (Randi & Lucchini, 2002; Lucchini *et al*., 2002; Fabbri *et al*., 2007).

As non-invasively collected samples usually provide low target DNA concentration and low target DNA quality (Taberlet *et al*., 1999), to delete those lacking enough DNA to complete the genotyping and to impede possible problems during further laboratory procedures, all the DNA samples were initially screened using a multiple-tube approach (Taberlet *et al*., 1996; Gagneux *et al*. 1997; Lucchini *et al*., 2002). The screening consisted in amplifying each sample four times at 2 microsatellite *loci* (FH2096 and FH2137) chosen, for their high PCR success and their low dropout and false allele rates, among the first 6 microsatellite used for the individual identification (Lucchini *et al.*, 2002; Fabbri *et al*., 2007).

Only the samples with positive PCRs major than 50 % pass the screening and they are amplified four times at the other 4 microsatellite *loci*, always using a multiple-tubes approach by which the samples heterozygote at least in 2 replicates or homozygote at least in 4 replicates at a given *locus* were scored as reliable at that *locus* and genotypes were recorded; while all the other heterozygote, homozygote and uncertain genotypes (due to failure of one amplification or to allelic dropout) were additionally replicated four times. All samples that could be not reliably typed at all *loci* after 8 amplifications were discarded.

Microsatellites were PCR-amplified (Randi & Lucchini, 2002; Lucchini *et al*., 2002; Fabbri *et al*., 2007) separately in 10 µl of volume, using 2 µl of DNA solution, 1 µl of PCR Buffer 10X (1,5 mM of $MgCl_2$), 2 µg of BSA (Bovine Serum Albumin), 0,4 µl of dideossinucleotides (dATP, dCTP, dTTP, dGTP) 2,5mM, 0,15 µl of each primer 10 µM, 0,25 units of Taq and 4,25 µl of PCR water.

Cycling conditions were optimized for each primer pair and for tissue or scat samples, the number of cycles varied from 30 to 45, starting from the following general PCR program:

94°C x 2'→(94°C x 15''→ 55°C x 30''→ 72°C x 30'') for 30-45 cycles → 72°C x 10' → 4°C x 10' → 15°C

2.5.2 Sex identification

All the DNA samples, after genotyping were submitted to a reliable DNA-based sex identification or molecular sexing (Lucchini *et al.*, 2002) amplifying by PCR ZFX/ZFY(zinc-finger protein) sequences (Garcia-Muro *et al*., 1997). As universal primers used for ZFX/ZFY amplifications (P1-5EZ and P2-3EZ; Aasen & Medrano, 1990) are conserved in vertebrates and might amplify DNAs from wolf prey, ZFX/ZFY canid specific primers were used. They had been previously designed by Lucchini *et al.*(2002) who detected a sex-specific RFLP pattern by digestion of PCR product with10 units of ***TAQ* I** restriction enzyme that can cut a fragment only on the Y chromosome. The ZFX/ZFXY product, in fact, includes one ***TAQ* I** restriction site that produces two fragments of different lengths which can be separated by electrophoresis and then observed by automated sequencer. The electrophoretic pattern shows two visible bands in females (an uncut maternal ZFX fragment and a digested paternal ZFY fragment), but only one band in females (generated by both uncut ZFX fragments).

2.5.3 Analysis of microsatellites in automated capillary sequencers

Microsatellite analysis consists in separating the different alleles (the alleles differ for the number of repetitions of the repeat) by electrophoresis in a denaturing gel which clearly separates

the 2 alleles present at the heterozygous *loci*. In automated capillary sequencers the electrophoresis does not require the gel preparation because they can automatically inject it in a serious of capillaries through which fragment migration takes place. Electrophoresis is programmed through a particular computer software that activates and controls all operation performed by the automated sequencer. The capillary sequencer does not use radioactive markers but fluorescent marker systems (*fluorescent dyes*) that are incorporated in the DNA during PCR amplification or sequencing, utilising primers labelled with a fluorescent dye or incorporating a labelled nucleotide in the DNA. When the labelled DNA fragment passes a pre-set location the fluorescent dye is picked up by a laser and the emission of fluorescence is detected and measured by the software that analyses the results of electrophoresis and convert the weights of the different alleles (the alleles differ for the number of repats) in an image file and in an electropherogram in which the molecular weights of the alleles is precisely determined by the use of internal standards.

Homozygous sample at a given *locus* present a single band (that appears as a single peak in an electropherogram) while heterozygous samples present 2 bands (that appear as 2 different peaks in an electropherogram) (Fig. 2.3).

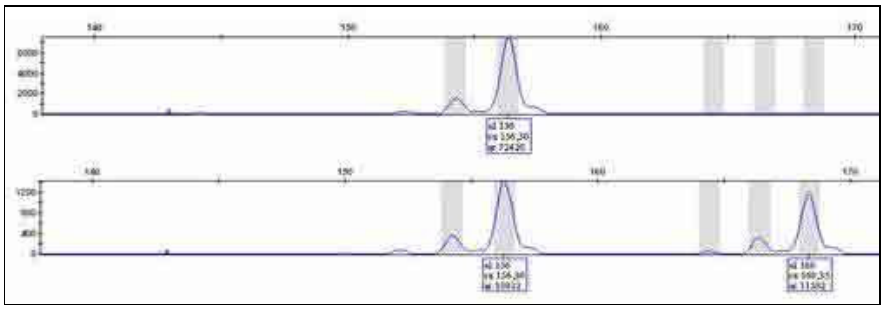

Fig. 2.3: example of electropherograms, the single peak stands for a homozygous sample at a given *locus*, the double peaks indicate a heterozygous sample at a given *locus*.

In automated sequencers it is possible to analyse several microsatellite *loci* in the same capillary column simultaneously. The analysis of multiple *loci* can be done via multiplex PCR or via electrophoresis of mixtures of single PCR (electrophoresis multiplex). In multiplex systems (both PCR and electrophoresis systems) it is necessary to choose microsatellite *loci* that produce clean and clear signals (electropherograms). As in the automatic analysis of microsatellite one of the two PCR primers is labelled with a fluorescent dye, in multiplex systems it is necessary to label primers at different *loci* with different colours. Three colours (yellow, green and blue) are currently used to label the primers while a fourth colour (red) is used to label the standard molecular weight. Microsatellite whose alleles have different molecular weights can be combined in multiplex systems and PCR products are separated in different areas of the gel or capillary and the identification of alleles is facilitated by reading the coloured signals that do not overlap.

2.5. 4 Probability of Identity and selection of the microsatellite *loci*

When using microsatellite *loci* to establish a genetic profile, it is possible for different individuals of the same population to have identical profiles if an insufficient number of *loci* has been used. Mills *et al.* (2000) and Waits *et al.* (2001) showed that, in order to be useful in population size estimations, genetic profiles should consist of enough microsatellite *loci* to distinguish between individuals with 99% certainty. Estimating the required number of *loci* can be achieved by computing probability of identity (PID) which is the probability that 2 individuals, randomly chosen within the same population, have the same *multilocus* genotype and therefore the proportion in a population of individuals with the same *multilocus* genotype (Paetkau & Strobeck

1994). If in a population PID is not zero, some individuals cannot be detected (*shadow effect*) leading to a population size underestimation (Mills *et al*., 2000). Where there is the potential for relatives to be present in the sample, it is best to use an estimate of PID among siblings (PID_{Sibs}: Evett & Weir 1998; Woods *et al*. 1999; Waits *et al*. 2001). The overall PID_{Sibs} is the upper limit of the possible ranges of PID in a population and thus provides the most conservative number of loci required to resolve all individuals, including relatives.

The probability of identity (PID) and the expected PID among full sib dyads (PID_{sib}), were estimated in a set of 100 Italian wolves (Randi & Lucchini, 2002; Lucchini *et al*., 2002) using the software **GIMLET v. 1.3.2**. (Valière, 2002 http://pbil.univ-lyon1.fr/software/Gimlet/gimlet.htm) (Table 2.1).

Locus	PID	PID_{cor}	PID_{sibs}	*Locus*	PID	PID_{cor}	PID_{sibs}
CPH2	2,35 E-1	2,39 E-1	5,22 E-1	**FH2079**	1,87 E-1	1,89 E-1	4,68 E-1
CPH4	3,54 E-1	3,59 E-1	6,22 E-1	**FH2088**	1,68 E-1	1,71 E-1	4,57 E-1
CPH5	2,07 E-1	2,10 E-1	4,86 E-1	**FH2096**	1,86 E-1	1,88 E-1	4,66 E-1
CPH8	9,86 E-1	1,01 E-1	3,97 E-1	**FH2132**	6,47 E-2	6,77 E-2	3,70 E-1
CPH12	4,55 E-1	4,60 E-1	6,93 E-1	**FH2137**	6,52 E-2	6,80 E-2	3,68 E-1
U9.250	1,45 E-1	1,48 E-1	4,45 E-1				
FH2004	1,37 E-1	1,40 E-1	4,39 E-1	**Total**	**3,75E-10**	**4,88 E-10**	**1,15 E-4**

Table 2.1: Probability of Identity for each *locus* estimated in a set of 100 Italian wolves using Gimlet v. 1.3.2. (Valière, 2002). Pid is the probability of identity for individuals randomly chosen within the same population, Pid cor is the probability of identity corrected for small population size, Pid sibs is the probability of identity corrected for siblings. Pid sibs < Pid cor < Pid. Each total probability was computed by multiplying single *locus* probabilities, assuming that *loci* are independent, as suggested by the microsatellite linkage map of the domestic dog (Neff *et al*., 1999).

As wolves in a pack are known to be partially related, sharing alleles which are identical by descent (Mech, 1970; Lehman *et al*., 1992; Wayne *et al*., 1995), and field observations suggested that about 100 wolves were present in the whole study area, it was necessary to achieve PID_{sibs} values < 0,01, meaning that 1 wolf in 100 siblings was expected to share, by chance, an identical genotype with another wolf.

As showed in Fig. 2.4, the minimum number of microsatellite *loci* necessary to obtain such PID values corresponds to 6 microsatellites that produce a PID_{sibs} of $7{,}11 \times 10^{-3}$ allowing detection of unique genotypes also if related individuals were sampled.

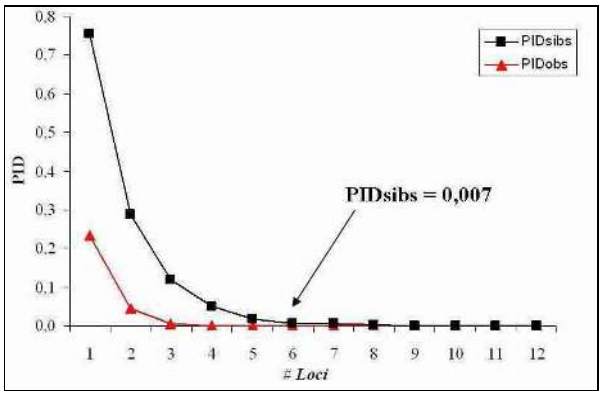

Fig. 2.4: Values of probability of identity observed (PID_{obs}) and corrected for siblings (PID_{sibs}), related to the number of microsatellites typed in a sample of 100 Italian wolves, computed using Gimlet v. 1.3.2. (Valière, 2002). The first 6 *loci* were used for individual genotyping and the additional 6 microsatellites were included for kinship analysis and to clarify uncertain genotyping. The arrow indicates the PIDsib value obtained using six *loci*.

2.6 MICROSATELLITE DATA ELABORATION

The software used to manually or automatically correct the results of the automated analyses is ***GeneMapper v.3.0*** of the Applied Biosystems (ABI). When the electrophoresis ends every allele may be made up of a single band (that appear as a single peak in an electropherogram) or of a main band plus a serious of secondary bands that represent aspecific amplification products. After defining the variation range of molecular weight and of the main peak of the electropherogram as well as the colour of the *locus*, the software allows to identify the signal produced by the main band and assign the respective molecular weight. The program uses an algorithm to filter that information which ignores the secondary signals and assigns the correct molecular weight to the principle signal of the allele. The final result can be visualized as a correct electropherogram, and the data, that contains the values of the molecular weight assigned to each allele, can be exported to database Microsoft Excel-type format , or to input formats of various data elaboration software.

2.6.1 Data reliability: *RelioType*

In this study *multilocus* genotypes were detected using a multiple-tube approach by which the same DNA samples were amplified independently several times *per locus* and the results of each replicate were compared. In this way it was possible to detect eventual dropouts or false amplification. The necessary number of replicates to obtain a reliable *multilocus* genotype was computed using the software ***RelioType*** **(Miller, Joyce & Waits, 2002).** It is a program for assessing how reliable an observed multilocus genotype is and for directing further replication if it is not sufficiently reliable. It is based on the model developed by Miller, Joyce and Waits (2002). The program requires two input files: a first file with allele counts from the population which the program converts into allele frequencies and a second file containing the genotyping data. The software calculate for each *multilocus* genotype of the second input file a probability of reliability using the allele frequencies contained in the first input file. The estimation of reliability assumes that false alleles do not exist in the data set, which is a clearly unrealistic assumption. One simple way to catch false alleles is to require that all alleles are observed multiple times.

2.6.2 *Multilocus* genotype comparison: *Gimlet*

When using the multi-tube approach, it is useful to easily construct consensus genotypes and to rapidly calculate the error rates. ***Gimlet v. 1.3.2*** (**G**enetic **I**dentification with **M**ulti**L**ocus **T**ags) (Valière, 2002) is a software dedicated for geneticists who work on individual identification using molecular tags in diploid species. This software allows to easily construct consensus genotypes from a set of PCR repetitions for each samples choosing the alleles that appeared the most at each *locus* (an allele is retained in the consensus if its score is above a threshold set by the user), and to rapidly calculate the error rates (allelic dropouts (ADO) and the false alleles (FA)) comparing the repeated genotypes and the consensus. The program can be also used to compare the different genotypes to reference ones already analysed. A genotype is identified when its *multilocus* genotype matches completely with a reference genotype. Moreover the program possesses also an option for pooling several genotypes that match themselves. The regrouping is conducted as an identification where all genotypes are potential reference genotypes

In both cases the software indicates the pairs of genotypes where only one allele (for one or two *loci*) or two alleles (for one *locus*) is (are) different between the genotypes. In this way it is possible to re-check these genotypes by re-looking at the electropherograms or by repeating PCRs

at doubt *loci* because, considering very low PID, it is improbable to find 2 identical genotypes differing only for 1 allele on 12.

2.6.3 Species detection: *Structure*

It is possible to discriminate wolves from dogs and detect hybrids using *multilocus* genotypes. In fact they can present the same alleles but with different frequencies, or it can occur that particular alleles ("private alleles") are fixed only in wolves, dogs or hybrids (Randi& Lucchini, 2002; Lucchini *et al*., 2004; Verardi *et al*., 2006; Randi, 2007).

In this study, population assignment and hybrids detection were performed using a Bayesian clustering procedure implemented in ***Structure v. 2.1*** (Pritchard *et al*., 2000; http://pritch.bsd.uchicago.edu; Falush *et al,*. 2003).

The program *structure* implements a model-based clustering method which uses *multilocus* genotype data, consisting of unlinked markers, to infer population structure and to assign individuals to populations. The model assumes that there are *K* populations (where *K* may be unknown), each of which is characterized by a set of allele frequencies at each *locus*. Individuals in the sample are assigned (probabilistically) to populations, or jointly to two or more populations if their genotypes indicate that they are admixed (or hybrids). This method can be used to detect the presence of cryptic population structure and to perform assignment testing. Pritchard *et al.*'s model assumes Hardy-Weinberg (HWE) and linkage (LE) equilibrium among the unlinked *loci*. Departures from HWE and LE lead the population to be split into subpopulations, to which individuals are assigned. The number of contributing populations can be estimated and, for a given number of populations, their gene frequencies and the admixture proportions for each individual are all jointly estimated. In this way the sampled population is subdivided into a number of different subpopulations that effectively cluster the individuals. Then, individuals of *a-priori* known or unknown origin may be assigned probabilistically to the subpopulations.

The model does not assume a particular mutation process, and it can be applied to most of the commonly used genetic markers including microsatellites, SNPs and RFLPs, provided that they are unlinked. In this study I performed the analyses using the "admixture" model which assumes that each individual may have ancestry in more than one parental population and that allele frequencies of a K population can be obtained independently from the others. As wolves and dogs are genetically detectable this approach is very useful to detect F1, F2 hybrids and first-generation backcrosses, in fact using a clustering threshold $Q > 0.90$, all the Italian wolf population individuals should assign to a cluster and all the dogs to another cluster, while hybrids should present an admixture clustering.

2.6.4 Genetic population study: *GeneAlex*

GeneAlex v. 6.0 (Peakall & Smouse, 2005; 2006) is a software provided as an Excel add-in, with a compiled module and an associated menu, particularly useful to study population genetics and produce output files which can be directly used in other elaboration software.

In this study GeneAlex was used to estimate allele frequency by *locus* and population, observed (H_O) and expected unbiased (H_E) heterozygosities, mean number of alleles per *locus* (N_A), number of private alleles (N_P) per population (i.e. the number of alleles unique to a single population in the data set) and to compute the HWE and Chi-square testing procedures. GeneAlex was also used to perform the **AMOVA** (analysis of molecular variance), which was used to assess the level of global and pairwise population differentiation based on φ_{pt}, an analogue of $\boldsymbol{F_{ST}}$, which estimates the proportion of the genotypic variance among populations, relative to the total variance.

The software was also utilized for assignment tests and Principal Coordinate Analysis (PCA), in fact for each sample the expected genotype frequency at each *locus* is calculated and log-transformed to give a log likelihood value which is calculated even for each population, using the allele frequencies of the respective population. A sample is assigned to the population with the highest log likelihood.

Genetic distance and assignment tests allow, through Principal Coordinate Analysis (PCA), to detect the different considered populations despite a Cartesian axe system not linked to a geographic reference system. The software, in fact, synthesizes all variability of the populations, expressed by many variables, in 2 or 3 variability axes around which the analyses and the further assignments occur.

To simplify the graphic visualization of Structure and GeneAlex results the program ***Genetix v.4.2*** (Belkhir *et al.,* 2001; http://www.University-montp2.fr/-genetix/genetix.htm) was used. It can describe in three dimensions all the variability analyzed in GeneAlex by Principal Coordinate Analysis and the different Structure clusterings.

2.6.5 Data mapping: *ArcView GIS*

Sample mapping localization was obtained by the software ***ArcView GIS*** (ESRI) that is a geographical information system able to organizer, control, analyse and update spatial and multidimensional data source using geographic coordinates. The software allows to create a series of informative themes which represent distinct sets of geographic features in a particular geographic data source simplifying their graphical representation. A collection of themes creates a an interactive map (view) that lets you display, explore, query and analyze geographic data in ArcView.

2.7 POPULATION SIZE ESTIMATION

Because hair and fecal samples can be collected without capture, non-invasive sampling methods have great promise for population estimation in fact genotypes can be used to estimate population size in several ways (Sloane *et al*., 2000). Most directly, the number of distinct genotypes is an estimate of the minimum population size, which can be identified by the asymptote of a curve relating the number of distinct genotypes to the number of samples (Kohn *et al*., 1999). If individuals are sufficiently sampled, mark-resight methods of estimating population size can also be applied to genetic data to estimate resighting probabilities (Otis *et al.*, 1978; Seber, 1982).

In recent years, the use of non-invasive genetic sampling and individual *multilocus* genetic profiles for capture-recapture studies has rapidly increased and the method has been applied to a diverse array of taxa to assess population size (Lukacs, 2005; Lukacs & Burnham, 2005)

Abundance is just one parameter that can be estimated from capture-recapture data. Many models have been recently developed to estimate survival, emigration rates, movement or transition rates, fecundity, and population growth (Nichols; 1992). In natural closed and open populations, variation in capture probability (behavioural responses to capture, variation over time with constant trapability for all individuals) and individual heterogeneity in capture probability (the variation among individuals in their probability of being detected) are the most difficult problem facing the estimation of animal abundance and of the other biological parameters (Otis *et al*., 1978).

Anyway, a large number of models and software exists for a wide range of capture-recapture analyses. Most capture-recapture theory builds off a reparameterization of a multinomial model (Burnham, 1991); therefore, software can be designed to analyse a wide variety of capture-recapture data within a common framework.

In this study the *multilocus* genotype identifications from wolf scats sampled over the Emilia Romagna Apennines from spring 2000 to winter 2007 were used as capture-recapture data. The first detection of a genotype is alike marking, while further detections are alike recapture with the approximation that genotyping errors are assumed to be negligible as a result of careful lab procedures. Due to the open nature and the long time span of this wolf population project, these genetic data were analysed using the open population multistate and multievent models (Lebreton & Pradel, 2002; Pradel, 2005) incorporating mixture heterogeneity (Pledger *et al*., 2003) to detect the main biological parameters necessary to obtain a reliable population size estimation. The programs utilized to built and test the possible statistic models essential to compute the population size estimation were U-Care and E-Surge.

2.7.1 ***U-Care*** (Choquet *et al.,* 2005) is a computer program that deals with the first steps of the analyses of capture-recapture data, the preparation of the datasets and the assessment of the fit of a general model: the Cormack-Jolly-Seber (CJS) model which has both capture **(*P*)** and survival **(Φ)** probabilities dependent on time **[Φ(t)*P*(t)]**. Goodness of fit test (GOF) of the CJS model is generally done to explore the fit of the CJS to the data, but also to identify a general model that fits the data from which to start an eventual further model selection (Lebreton *et al*., 1992).

Goodness of fit (GOF) of the general model was evaluated in a series of tests: test 3.SR, 3.SM, 2.Ct and 2.Cl (Burnham *et al*., 1987; Pradel, 1993; Pradel *et al*., 1997). Among them the most informative ones are the specific tests for behavioural patterns or recapture heterogeneity: **test 3.SR**, which is a specific directional test for temporary emigration and transience (Pradel *et al* ,. 1997), relevant to detect whether there was heterogeneity in survival between individuals, depending on whether or not they had been captured previously (it is significant, for example, when there is an age effect on survival or because of the presence of transients), and test **2.Ct**, which is a specific test trap for trap dependence which is significant if there is either an immediate trap effect on recapture probability or if there is non-random temporal emigration (trap-happiness or trap-shyness, Pradel, 1993).

If the model does not fit the data, there is a clear signal of heterogeneity of capture and it necessary to carry out a new model selection to find more complex heterogeneous models using specific software such as E-Surge (Choquet *et al.,* 2007) and find the best one. On the contrary, if the model fits the data, it is necessary to run other homogeneous models to find the best one.

2.7.2 ***E-Surge v. 1.1.1*** (Multi**E**vent **G**eneralized **Sur**vival **E**stimation) (Choquet *et al.*, 2007**,** http://ftp.cefe.cnrs.fr/biom/soft-cr/) is a program for fitting multievent models to capture-recapture(CR) data (Pradel, 2005). Multievent models are an extension of multistate models in which observations do not necessarily correspond to states but are defined.

Because the observations in multievent models do not necessarily correspond to individual states, they can handle state uncertainty, the software can provide a general framework for problem such as: heterogeneity of capture (Pledger *et al*., 2003), determination of the sex when sex is not available (Nichols *et al.,* 2005) and memory model (Pradel, 2005).
Several programs exist for CR analysis (MSSURVIV, Hines, 1994; MARK, White & Burnham, 1999; M-SURGE, Choquet *et al.,* 2004) but E-Surge is the first general one for multievent models which has powerful capabilities for maximum likelihood estimation of complex age and time-dependent models with linear constraints among parameters, in a generalized linear model fashion.

Multievent models assumes that individuals move independently among a finite set **E** of states over a finite number **K** of sampling occasions and that successive states obey a Markov chain. They are defined in terms of three kinds of parameters: initial state probabilities **π**, transition probabilities **Φ**, and encounter probabilities ***b***. The matrices associated to these parameters together define the general model (GM) under which an umbrella model (the most general model) retained by Goodness of Fit can be fitted.

In M-Surge and Mark transition probabilities are defined either directly, or in terms of survival and transition conditional on survival. In E-Surge a pattern generator **GEPAT** (for **GE**nerator of **PAT**tern of elementary matrices) makes it possible to generate the GM using the elementary matrices (transition and encounter matrices) and initial state vector under which the UM is defined. This feature or model is called DES for Decomposition in Elementary Steps.

Model-building in E-Surge (as in M-Surge and Mark) proceeds by imposing linear constraints on the parameters of the umbrella model, in the spirit of generalized linear models. The vector $\boldsymbol{\theta}$ of "biological parameters" (parameters of direct interest to the biologist e.g., $\boldsymbol{\theta} = (\boldsymbol{\pi}, \boldsymbol{\Phi}, \boldsymbol{b})$, organized as a vector) is expressed as a linear transformation of a vector $\boldsymbol{\beta}$ of "mathematical parameters" which can be expressed in a "matrix of constraints" (Matrix $\boldsymbol{X}$). In general, it expresses hypotheses about the dependence of the parameters on stage (of departure or arrival), age (since first capture), time, group, and/or covariates. The design matrix is built by the program **GEMACO** (**GE**nerator of **MA**trices of **CO**nstraints) based on a powerful language similar to those used in general statistical software packages such as SAS, S-Plus, Genstat or GLIM (for instance, the formula t+g generates a model with additive effects of time and group) avoiding tedious and error-prone matrix manipulations.

2.8 SINGLE NUCLEOTIDE POLYMORPHYSM ANALYSESES

Although ascertainment bias is a problem for some applications, single nucleotide polymorphisms (SNPs) can often generate equivalent statistical power as providing broader genome coverage and higher quality data than can either microsatellites or mtDNA (Morin *et al*., 2004). They represent an efficient and cost-effective genetic tool which can be used as novel genetic markers for common questions in population genetics: forensic identity testing (Andreàsson *et al.*, 2002) include: Y chromosome SNPs for lineage-based studies of highly degraded DNA samples using autosomal SNPs (Budowle *et al*., 2004) assessing biogeographical origin (Frudakis *et al.*, 2003), individual identification and individual assignment to a population (Seddon *et al*., 2005; Holm Andersen & Fabbri *et al.,* 2006). There are two principle steps to the use SNP markers: *locus* discovery (ascertainment) and genotyping. SNP discovery is the process of finding the polymorphic sites in the genome of the species and populations of interest (Morin *et al*., 2004).

Genotyping is a laboratory procedure that identifies the alleles presented in a given sample. To accomplish this goal, genotyping biochemistry must be highly specific. A biochemical reaction identifies one and only one allele at a time. Since multiple reactions can occur simultaneously at multiple templates and target *loci* of a sample, collectively the same biochemical reaction can identify multiple alleles and multiple *loci*. Popular SNP genotyping technologies currently available are based on one or more properties of these enzymes and processes: DNA polymerases; DNA ligases; and hybridization. A genotyping protocol is normally has two parts: biochemical reactions to form allele-specific products (allele discrimination) and detection procedures to identify the products (Chen & Sallivan, 2003).

2.8.1 SNP characterization in the Italian wolf population

Noninvasive DNA analyses are often prone to genotyping errors (false alleles and allelic dropouts) due to DNA degradation. Thus, the use of single nucleotide polymorphisms (SNPs), which requires amplification of much shorter DNA sequences may allow more efficient genotyping of noninvasive samples (Seddon *et al*., 2005).

This study contributed to characterize canine single nucleotide polymorphisms (SNPs) in the endangered Italian wolf (*Canis lupus*) population, which were discovered by resequencing sequence-tagged-site (STS) DNA sequences that were known to contain SNPs in domestic dogs

(see Holm Andersen & Fabbri *et al.,* 2006). DNA fragments, extracted from 14 Italian wolf samples collected in north and central Italy, were amplified by polymerase chain reaction (PCR) using 76 primer pairs for SNPs containing dog STS sequences (Guyon *et al*., 2003).

PCRs were carried out in 15 µl volumes using the following touchdown program: 8min 95°C, followed by 20 cycles of 30 sec 94°C, 30 sec 63°C decreasing of 0.5°C per cycle, 1 min 72°C and 15 cycles of 30 sec 94°C, 30 sec 53°C, 1 min 72°C, and a final extension of 2 min 72°C.

On the base of this first SNP characterization in the Italian wolf population 15 new reliable primer sets (1C06/138, 38K22/150, 96B17/422, 182J12/119, 182M20/250, 189H18/294, 218J14/81, 309N24/298, 310M20/207, 310M20/332, 120D19/347, 133N13/219, 148L07/169, 168J14/149, 182B11/138, Holm Andersen & Fabbri *et al.,* 2006) for analysing SNPs using Pyrosequencing technology (Ronaghi *et al*., 1998) were designed. The SNPs found by Pyrosequencing were verified by comparing the results with the sequences.

Afterwards these 15 SNP *loci* allowed to design 6 new primer sets (182B11/138; 309N24/298; 1C06/138; 38K22/150; I96B17/422; F310M20/207), that could be used also by other 2 SnP genotyping methods, SNaPshot and RealTime PCR. 43 non-invasive DNA samples of different qualities were amplified, according to a multiple tube approach (Taberlet *et al*., 1996; Gagneux *et al*. 1997; Lucchini *et al*., 2002), 3 times for each of these primer sets using all the 3 technologies. The results obtained were compare to finally establish which was the best one, from success, cost and time points of views, to use for the identification of the individuals and their assignment to the belonging populations.

2.8.2 Pyrosequencing analyses

Pyrosequencing™ AB (http://www.pyrosequencing.com) is a non-electrophoretic real-time DNA sequencing method which uses an enzyme-cascade system, consisting of four enzymes and specific substrates, to produce light whenever a nucleotide is incorporated to form a base pair with the complementary base in a DNA template strand. The amount of light is proportional to the number of incorporated nucleotides (Ronaghi, 2001; Ronaghi *et al.,* 1996; 1998; Berg *et al.,* 2002). The pyrosequencing technology is suitable for both scoring and discovery of SNPs. Because not all enzymes involved are thermostable, heat cannot be used to denature the templates so the assay is run at room temperature. Either the forward or the reverse primer must be biotinylated for later immobilization of PCR products using the Vacuum Prep Tool (Biotage) to obtain single-strand DNA. It consists of a hand-grip with 96 replaceable filter probes. The hand-grip is connected to a vacuum source by an extendable hose with an on/off control switch. The sample preparation tool streamlines the preparation of single-stranded DNA prior to sequence primer annealing. To begin sample preparation, streptavidin coated sepharose beads are added to the PCR plate containing the DNA template with one strand 5'-biotinylated and this is mixed for 10 minutes. The PCR product with beads attached is picked up by the tool from the PCR plate and, from a separate trough, 70% ethanol is aspirated through the filter probes. This step positions the 5'-biotinylated strands attached to streptavidin coated sepharose beads at the end of all 96 filter probes. The sample preparation tool is then placed into a trough of wash buffer and the strands are rinsed by aspiration. The single-stranded templates are then transferred to a previously prepared PSQ HS 96 plate containing annealing buffer and primer. After the primer is annealed, this plate is then placed into the PSQ HS 96A System for analysis. The method designs extension primer a few bases upstream from the polymorphic site. The chosen SNP sequences are entered into the SNP Software or imported from external sample databases The software automatically recommends the most effective order for the dispensation of nucleotides. The theoretical sequence results are displayed as bar graphs (Fig. 2.5).
During primer extension dNTPs are added one by one as dictated by the target sequences. If the incoming base matches the template the base would be added to the extending primer and a pyrophosphate would be produced. The pyrophosphate then triggers the synthesis of ATP, which in

turn is used by a luciferase to produce a chemiluminescence signal. The amount of light is proportional to the number of incorporated nucleotides. This light signal is detected, the base registered (seen as a peak in the resulting Pyrogram™) and the next nucleotide added. If the base added does not match the template, the primer would not be extended, the dNTP is then degraded into dNMP (deoxyribonucleoside monophosphate) by an apyrase and no light will be generated. Since DNA polymerases do not use dNMP, it would not interfere with subsequent reactions. DNA polymerases are faster than the apyrases, so the polymerases always process the incoming nucleotide first, but this process is conditioned on the sequences of the template. If the nucleotide matches the template, it will be incorporated onto the primer by the polymerases, otherwise, it will be left to the apyrases, which are phosphodiesterases that do not discriminate among the four bases and do not produce pyrophosphate.

When the run is completed, the genotype is determined by comparison of the peak heights of the SNP positions, with the theoretical results predicted by the SNP Software (Fig. 2.5).

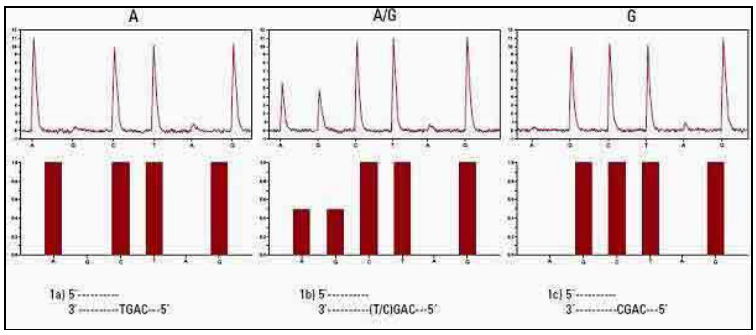

Fig. 2.5: theoretical results, automatic genotyping and assessment for a multiple SNP sample.

In this study unlinked SNPs in 15 canine STS sequences were genotyped designing 15 different primes (1C06/138, 38K22/150, 96B17/422, 182J12/119, 182M20/250, 189H18/294, 218J14/81, 309N24/298, 310M20/207, 310M20/332, 120D19/347, 133N13/219, 148L07/169, 168J14/149, 182B11/138) suitable for Pyrosequencing by the assay design software version 1.0.6 (Biotage). Those primers were used to analyse SNPs in 14 Italian wolves by Pyrosequencing™ technology (Biotage), using the PSQ 96MA system. Then 6 (182B11/138; 309N24/298; 1C06/138; 38K22/150; I96B17/422; F310M20/207) of these 15 primers were used to genotype 43 non invasive DNA samples for a method comparison with results obtained through SNaPshot and RealTime technologies. All the PCRs were performed in 25 µl of volume using 4 µl of DNA solution, 2,5 µl of PCR Buffer 10X (1,5 mM of $MgCl_2$), 2,5 µg of BSA (Bovine Serum Albumin), 1 µl of dideossinucleotides (dATP, dCTP, dTTP, dGTP) 2,5 mM, 0,5 µl of each primer 10 µM, 0,125 units of Taq and 14µl of PCR water. PCRs were performed with 45 cycles and with an annealing temperature of 55° C (Holm Andersen & Fabbri *et al.,* 2006) using the following program:

94°C x 2'→(94°C x 15''→ 55°C x 30''→ 72°C x 30'') for 45 cycles → 72°C x 10' → 4°C x 10' → 15°C

The biotinylated PCR products were immobilised to streptavidin-coated beads (Streptavidin Sepharose HP, Amersham Bioscience) following the standard protocol for the PSQ 96 Sample Preparation Kit (Pyrosequencing™ AB).

The DNA strands were separated in 50 ml denaturation solution for 1 min. The immobilised template was rinsed twice with 150 ml washing buffer, resuspended in a reaction mix (containing 43,7 ml of annealing buffer and 1,3 ml of sequencing primer 10µM, 588µl of enzyme and 588µl of substrate) and transferred to a PSQ 96 plate to complete the genotyping (Holm Andersen & Fabbri *et al.,* 2006).

2.8.3 SNaPshot analyses

SNaPshot™ (ABI) is a solution-based assay that uses the single nucleotide primer extension assay (Syvanen *et al.*, 1990; Syvanen, 1999; Budowle *et al.*, 2004). The method is based on the use of three primers for the analysis of each SNP: an external forward, an external reverse and an internal primer consisting of a few nucleotides until the one which precedes the mutation. The first two primers are necessary for a first amplification of the fragment containing the SNP, while the SNP extension primer is used during a second amplification to detect the polymorphism. During this minisequencing PCR the SNP extension primer is annealed to the denatured template amplicon and is extended at the SNP site by the incorporation of one of the four fluorescently labelled terminator ddNTPs. The primer cannot be extended further, because only ddNTPs are in the extension reaction. The extended SNP primer is subjected to capillary or slab-gel electrophoresis. The particular incorporated nucleotide is identified by the different labelled fluorescent tag as in Sanger sequencing. The specific SNP *locus* (or in actuality the extended SNP primer) in a multiplex assay is identified by its mobility during electrophoresis. The mobility can be modified by incorporating varying-length polynucleotide tails or by incorporating mobility modifiers at the 5' end of the SNP primer.

In this study, 6 of the 15 primer sets for analysing SNPs through Pyrosequencing technology used for the Italian wolf characterization, were applied to design 6 new primer sets (182B11/138; 309N24/298; 1C06/138; 38K22/150; I96B17/422; F310M20/207) for SNaPshot technology. These 6 primers were used to genotype 43 non invasive DNA samples for a method comparison with results obtained through Pyrosequencing and RealTime technologies.

SNPs were PCR-amplified separately in 10 µl of volume, using 2 µl of DNA solution, 1 µl of PCR Buffer 10X (1,5 mM of $MgCl_2$), 1 µg of BSA (Bovine Serum Albumin), 0,4 µl of dideossinucleotides (dATP, dCTP, dTTP, dGTP) 2,5 mM, 0,25 µl of each primer 10 µM, 0,25 units of Taq and 5,05 µl of PCR water.

Cycling conditions were optimized for each primer pair and for tissue or scat samples with the number of cycles varied from 30 to 45, starting from the following general PCR program:

94°C x 2'→(94°C x 30''→ 55°C x 30''→ 72°C x 45'') for 30-45 cycles →
72°C x 10' → 4°C x 10' → 15°C

As the excess dNTPs and PCR primers interfere with primer extension, to use the PCR products as templates for the extension reaction, they were purified using Exo-Sap (Amersham) which is a clean-up step to remove excess dNTPs and PCR primers left over from the reaction necessary for all primer extension-based methods. It can be done enzymatically with shrimp alkaline phosphatase (SAP) and *E. coli* exonuclease I (Exo I), which inactivate dNTPs and PCR primers respectively, or physically via gel filtrations.

In this study it was enzymatically carried out adding 1 µl of Exo/Sap mix to the PCR products and using the following thermocycling program:

37°C x 30' → 80°C x 15' → 4°C x 10' → 15°C

The extension reaction were carried out in 10 µl of volume, using 1 µl of PCR product, 1 µl of SNaPshot Reaction Mix, 0,2 µl of the extension primer 10 µM, 7,8 µl of PCR water using the following thermocycling program:

(96°C x 10''→ 55°C x 5''→ 60°C x 30'') for 25 cycles →4°C x 10' →15°C

SNaPshot Reaction Mix contains a reaction buffer, the enzyme Taq polymerase and the four dideossinucleotides (ddATP, ddGTP, ddCTP, ddTTP) labelled with different colours: A=green; C= black; G=blue; T=red. Dideossinucleotides are modified bases which posses an OH in 3'-position and avoid the formation of a phosphodiesteric link with another deossinucleotide so that the incorporation of one of them stop the extension generating fragments consisting of the primer and the SNP at that *locus*.

2.8.4 SNaPshot data elaboration

One µl of each purified minisequencing PCR product was resuspended in a denaturation solution (Formammide) and analysed by electrophoresis on an AB Prism 3130 Genetic Analyser with a 36 cm capillary array, POP4 polymer. GeneScan-120 LIZTM, labelled with a colour that was not used to mark the nucleotides, was used as internal size standard.

sequencing results are saved in the form of electropherograms and visualized in the form of peaks (Fig.2.6) because during the electrophoresis, when a fluorescent dye is picked up by a laser the data produce a luminous emission that is registered as a peak. The height of the peak indicates the intensity of the emission and the colour indicates the colour of the fluorescent dye.

In this study the data were analysed using GeneScan Analysis software v. 3.7 and GeneMapper Analysis software v 3.0 (Applied Biosystems). *GeneScan* automatically analyses sequencing data which derive directly from the electrophoretic runs, sorting the peaks into bins according to sizes by comparison to the internal size standard. *GeneMapper* allows to visualize and manually correct the electropherograms. Peaks above 100 relative fluorescence units can be considered positive signals and a SNP type was assigned (Fig.2.6).

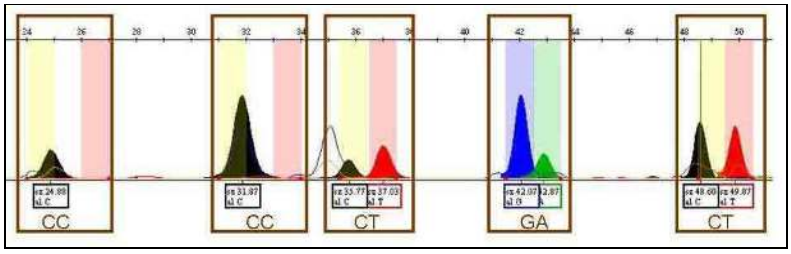

Figura 2.6: Example of minisequencing SNaPshot results. Data were analyzed through GeneScan Analysis software v. 3.7 (Applied Biosystems) and visualized by GeneMapper Analysis software v 3.0 (Applied Biosystems).

2.8.5 RealTime PCR analyses

RealTime PCR consists in a SNP genotyping method based on DNA Polymerases, versatile enzymes that have multiple functions. Their major function is to replicate DNA during cell division, but they also have 5' and 3' exonuclease activities in order to repair errors and remove RNA primers used in DNA replication. These activities form the basis of a mutation detection system (pyrophosphorolysis-activated polymerization) and the TaqMan assay (Liu & Sommer, 2000; 2002). The Invader assay was developed from a special structure specific endonuclease activity of some archael bacteria (Lyamichev *et al.*, 1993). The TaqMAn 5' nuclease assay is an elegant assay that exploits the 5' nuclease activity of DNA polymerases. It is a closed tube, single-step assay, and can score genotypes in real time or at the end of reaction. It combines target DNA amplification with allele discrimination in a single reaction (Livak & Goodsaid, 1997).

In the 5' nuclease first polymerase chain reaction assays, as first described by Holland *et al.* (1991), a hybridization probe included in the PCR is cleaved by the 5' nuclease activity of Taq DNA polymerase only if the probe target is being amplified. Using a fluorescent probe, first

synthesized by Lee *et al.* (1993), enables cleavage of the probe to be detected without post-PCR processing. The flurogenetic probe consists of an oligonucleotide labelled with both a fluorescent reporter dye and a fluorescent quencher. In the intact probe, proximity of the quencher causes Fluorescence Resonance Energy Transfer (FRET) and thus reduces the fluorescence from the reporter dye (Förster, 1948). Cleavage of the fluorogenetic probe during the PCR assay liberates the reporter dye causing an increase in fluorescence intensity.

Livak *et al.* (1995a) discovered that probes wit a reporter dye on the 5' end and a quencher on the 3' end can be used in the 5' nuclease assay, greatly simplifying the design of fluorogenetic probes.

Fluorogenetic probes and the 5'nuclease assay can be used for allelic discrimination through the hybridization of two doubly labelled allele-specific fluorescence probes to the target polymorphisms (Livak, 1999) (Fig. 2.7).

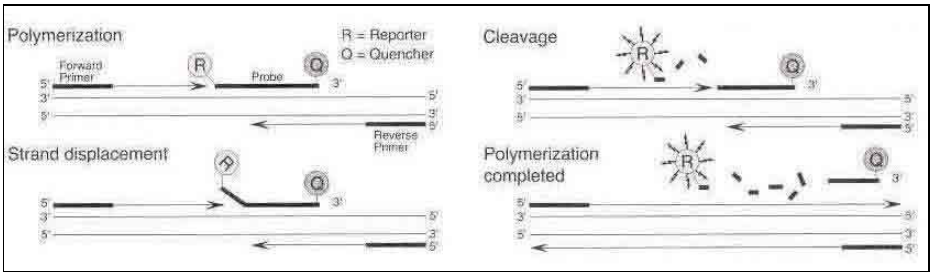

Figura 2.7: Fluorogenic 5' nuclease chemistry. (1) Forward and reverse primers are extended with *Taq* polymerase as in a traditional PCR reaction. A probe with two fluorescent dyes attached anneals to the gene sequence between the two primers. (2) As the polymerase extends the primer, the probe is displaced. (3) An inherent nuclease activity in the polymerase cleaves the reporter dye from the probe. (4) After release of the reporter dye from the quencher, a fluorescent signal is generated.

Allelic discrimination assay can detect single-base nucleotide mutations and polymorphisms. These assays for a bi-allelic system require to include in the PCR assay two separate probes, specific for each allele, that differ only by one base mismatch. The probes can be distinguished because they are labelled with different fluorescent dyes (FAM™ dye and VIC™ dye). A fully hybridized probe remains bound during stand displacement, resulting in efficient probe cleavage and release of the reporter dye. A mismatch between probe and target greatly reduces the efficiency of probe hybridization and cleavage. Thus, substantial increase in FAM or VIC dye fluorescence indicates homozygosity for the FAM- or VIC-specific allele, while an increase in both signals indicates heterozygosity (Fig. 2.9). Three factors contribute to the allelic discrimination. First, the mismatch has a disruptive effect. A mismatched probe will have a lower melting temperature (T_m) than a perfectly matched probe. Proper choice of an annealing/ extension temperature in the PCR will favour hybridization of an exact-match probe over a mismatched one. Second, as the assay is performed under competitive conditions with both probes present in the same reaction tube, mismatched probes are prevented from binding. Third, the 5' end of the probe must start to be displaced before cleavage occurs. The 5' nuclease activity of Taq DNA Polymerase actually recognize a forked structure with a displaced with a 5' strand of at least 1 to 3 nucleotides (Lyamichev *et al.*, 1993). Once a probe starts to be displaced, complete dissociation occurs faster with a mismatch than with an exact match so there is less time for cleavage to occur with a mismatched probe. Thus, the presence of a mismatch promotes dissociation rather than cleavage of the probe.

In this study 6 of the 15 primers used by Holm Andersen & Fabbri *et al.* (2006) for the Italian wolf characterization through Pyrosequencing technology, were applied to design 6 primer sets (182B11/138; 309N24/298; 1C06/138; 38K22/150; I96B17/422; F310M20/207) for RealTime PCR fluorescent technologies. RealTime was initially utilized to discriminate between the alternative

alleles of a polymorphism and genotype 43 non invasive DNA samples. This allowed to compare results through different methods (RealTime PCR, Pyrosequencing and SNaPshot technologies). The PCR was performed on the PRISM® 7500 Fast Real Time PCR System (Applied Biosystems) using the 5' nuclease assay with a dual labelled fluorogenic "TaqMan" probe. All oligonucleotides were designed using the Pimer Express® Software according to the parameters recommended in the guidelines by Applied Biosystems. These parameters include a T_m for the probe that is 10°C higher than the primers, primer T_ms between 58°C and 60°C, amplicon size between 50 and 150 bases, absence of 5' Gs, and primer length (Livak *et al*. 1995a).

PCR was carried out in a total volume of 5 µl reaction solution composed of: 2,5 µl of Mix TaqMan® FAST PCR Master Mix, 0,175 µl of a solution containing Forward and Reverse primers and the TaqMan probe, 0,5 µl of BSA 0,2%, 0,825 of PCR water and 1 µl of DNA. The master mix containes AmpliTaq Gold DNA Polymerase (active only after incubation at elevated temperatures).

Each PCR was run for 45 cycles (95°C, 20 sec, denaturation, followed by 45 cycles of 95°C, 5 sec and 60°C, 30 sec, amplification). Initial treatment of the raw data was carried out using the Applied Biosystems SDS software. Contamination was minimised by preparing reaction mixtures in a dedicated clean room with reagents aliquoted into single use volumes.

2.8.6 RealTime data elaboration

The 7500 Fast RealTime PCR Sequence detection system (Applied Biosystems) provides an accurate method for determination of levels of specific DNA sequences in tissue and scat samples. It is based on detection of a fluorescent signal produced proportionally during amplification of a PCR product. It consists of a 96-well thermal cycler connected to a laser and charge-coupled device (CCD) optics system. An optical fiber inserted through a lens is positioned over each well, and laser light is directed through the fiber to excite the fluorochrome in the PCR solution. Emissions are sent through the fiber to the CCD camera, where they are analyzed by the software's algorithms. Collected data are subsequently sent to the computer. Emissions are measured every 7 seconds. The sensitivity of detection allows acquisition of data when PCR amplification is still in the exponential phase. This is determined by identifying the cycle number at which the reporter dye emission intensities rises above background noise; this cycle number is called the threshold cycle (C_t). The C_t is determined at the most exponential phase of the reaction and is more reliable than end-point measurements of accumulated PCR products used by traditional PCR methods. The C_t is inversely proportional to the copy number of the target template; the higher the template concentration, the lower the threshold cycle measured.

Real-time monitoring of the release of fluorescence several times during each cycle allows collection of abundant data. After 40 cycles, data are processed by the software within a few seconds. Data can be viewed in an "amplification window" in the analysis program (Fig.2.8).

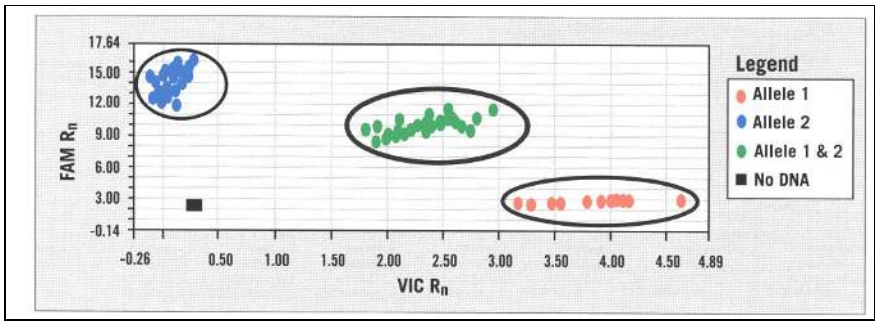

Fig. 2.8: visualization of the total allelic discrimination. Blue dots represent homozygotes for FAM, red ones represent heterozygotes for VIC, green ones stand for heterozygote individuals.

This allows the operator to check the fluorescence from each reaction at each cycle. The linearity of the fluorescence response for each sample at each cycle and the baseline can be checked for each tube. An occasional problem tube can easily be identified and the data point discarded, or the amplification curve may indicate that a different baseline should be chosen for the experiment to generate more accurate C_ts

Fluorescence spectra are collected after the run, and using multicomponent analysis, the software extracts the contribution of each component dye to the observed spectrum. Homozygotes for FAM show an increase in the FAM signal but no increase in the VIC signal, and homozygotes for the VIC probe show an increase in that signal. Heterozygotes show intermediate increases of FAM and VIC signals. All three groups are clearly distinguishable, and the sensitivity is similar to that for the quantitative PCR application (Fig.2.9).

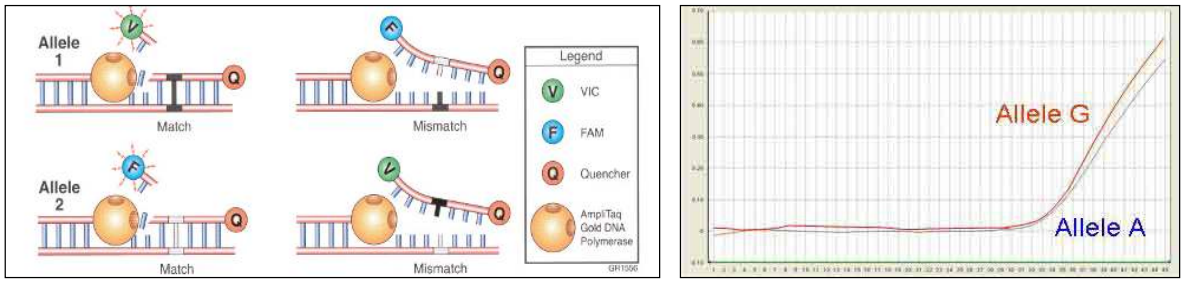

Fig. 2.9: visualization of the results for a single sample that present substantial increase in both FAM and VIC dye fluorescence indicates heterozygosity

CHAPTER THIRD: RESULTS

A total set of 3538 scat samples (including 1664 samples collected from January 2000 to December 2004 and 1874 samples collected from January 2005 to March 2007) and 34 tissue or blood samples (including 7 samples collected from January 2000 to December 2004 and 27 samples collected from January 2005 to June 2007) were analyzed in this study (Table 3.1).

SAMPLING PERIOD	SAMPLE TYPE	BO	FC	MO	PC	PNFC	PR	RA	RE	REGIONAL TOTAL
January 2000 - December 2000	Scats	1	0	27	0	0	0	0	0	28
	Tissues	1	0	0	0	0	0	0	0	1
January 2001 - December 2001	Scats	53	0	77	0	0	0	0	0	130
	Tissues	2	0	0	0	0	0	0	0	2
January 2002 - December 2002	Scats	97	7	75	0	54	48	2	107	390
	Tissues	0	0	0	0	0	0	0	1	1
January 2003 - December 2003	Scats	80	10	88	0	125	78	1	143	525
	Tissues	0	0	0	0	0	0	0	1	1
January 2004 - December 2004	Scats	156	13	66	0	260	29	14	53	591
	Tissues	0	0	0	0	1	0	0	**1**	2
January 2005 - December 2005	Scats	174	10	36	0	208	0	18	26	472
	Tissues	**1**	0	1	0	2	0	0	1	5
January 2006 - December 2006	Scats	231	7	54	9	422	148	39	65	975
	Tissues	1	0	3	1	5	0	0	1	11
January 2007 - March 2007	Scats	157	14	24	2	168	51	11	0	427
	Tissues	8	0	**1**	1	**1**	0	0	0	11
TOTAL	**Scats**	**949**	**61**	**447**	**11**	**1237**	**354**	**85**	**394**	**3538**
	Tissues/Blood	**13**	**0**	**5**	**2**	**9**	**0**	**0**	**5**	**34**

Table 3.1: presumed wolf samples collected and analyzed during the whole study period in the Emilia-Romagna Apennine Ridge. Blood samples are indicated in red. BO stands for Bologna provincial Administration (including samples collected in the Corno alle Scale regional park), FC for Forlì-Cesena provincial Administration, MO (including samples collected in the Frignano regional park) for Modena provincial Administration (including samples collected in Frignano regional park), PC for Piacenza provincial Administration, PNFC for National Park of Foreste Casentinesi; PR for Parma provincial Administration (including samples collected in Cento Laghi regional park), RA for Ravenna provincial Administration and RE for Reggio-Emilia provincial Administration (including samples collected in Gigante regional park).

3.1 MICROSATELLITE ANALYSES

All samples collected until December 2004 (including 136 scat samples and 12 tissue samples) were manually extracted using a guanidinium thiocyanate and diatomaceus earth (guanidinium-silica) protocol (Gerloff *et al*., 1995), while all the samples collected from January 2005 to March 2007 (including1402 scat samples and 22 tissue samples) were extracted in an automated manner by the ***MULTIPROBE IIEX*** robot (Perkin Elmer) and using the QUIAGEN Stool and tissue extraction kits (QUIAGEN).

3.2 SCAT MICROSATELLITE ANALYSES

All the 3538 scat samples were initially submitted to a preliminary quality screening test by PCR at 2 microsatellite *loci,* using a multiple tube approach (Taberlet *et al*., 1996; Gagneux *et al*.

1997; Lucchini *et al*., 2002). 2123 samples (60 % of total screened samples) resulted to posses enough DNA to complete the analyses so they were amplified at the other 4 microsatellite *loci*, necessary for the individual genotype identification, and submitted to a molecular sexing, always using a multiple tube approach, followed by a reliability analysis (Miller *et al*., 2002). 1293 samples (37% of total analyzed samples or 61% of screening test positive samples) obtained a complete and reliable genotyping suggesting that there was no statistical significant divergence (P = 0.7734, χ^2 test) between genotyping success of samples collected during winter (November-April) and summer (May-October) periods (Table3.2).

	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**TOT. PROJECT**
Winter Genotyping success	35%	43%	47%	28%	39%	44%	37%	39%	**39%**
Summer Genotyping success	34%	43%	45%	18%	40%	33%	35%	37%	**36%**
Total success	33%	43%	46%	23%	40%	39%	36%	38%	**37%**

Table 3.2: winter and summer genotyping success obtained for the whole project period. The genotyping value are comparable suggesting a no significant genotyping difference between winter and summer sampling.

The 1293 reliable genotyped scat samples were then submitted to the regrouping procedure carried out by the software Gimlet v.1.3.2 (Valière, 2002) allowing to assign them to 378 different unique individual genotypes. These unique genotypes, detected through genotyping at 6 microsatellite *loci*, were analyzed using an admixture model implemented in ***Structure v. 2.1*** (Pritchard *et al*., 2000; Falush *et al,*. 2003) to establish their belonging population considering their microsatellite allele frequencies that are sharply different between wolves and dogs (Randi & Lucchini, 2002). The program uses *multilocus* genotype data to infer population structure and to assign individuals to populations.

The model assumes Hardy-Weinberg (HWE) and linkage (LE) equilibrium among the unlinked *loci*, and that there are *K* populations (where *K* may be unknown), each of which is characterized by a set of allele frequencies at each *locus*. Individuals in the sample are assigned (probabilistically) to populations, or jointly to two or more populations if their genotypes indicate that they are admixed (or hybrids). Departures from HWE and LE lead the population to be split into subpopulations, to which individuals are assigned.

The program starts with a series of simulations to randomly assign the individuals, computing each time the reliability of these clusterings through a *likelihood* value estimation. Clustering occurs through *Markov Chain* and *Monte Carlo* algorithms that are able to maximize results, collecting only the permutations with high *likelihood* values. As the first simulations are usually not reliable and are considered as *burnings*, they are deleted from the results interpretation that are based only on the following permutations. In this study 130000 simulations were used (30000 as *burnings*) to carried out the assignment, and in this way 290 different Italian wolves ($Q_W > 90\%$), 75 domestic dogs ($Q_D > 90\%$),) and 13 uncertain assignment individuals were detected.

In order to better visualize the clustering so obtained, program Genetix v.4.2 (Belkhir *et al.,* 2001) was used (Fig.3.1).

To resolve these uncertain assignments all the unique genotypes were amplified at other 6 microsatellite *loci* and the Structure assignment test was repeated clarifying the assignments. Using 12 microsatellite *loci*, in fact, all the certain attribution, previously obtained using 6 *loci*, were confirmed through a better assignment probability, 10 of the 13 uncertain assignment individuals finally resulted to belong to the Italian wolf population while the other 3 uncertain individuals, that presented an admixture clustering resulted to be real hybrids, probably of second generation (Fig. 3.2).

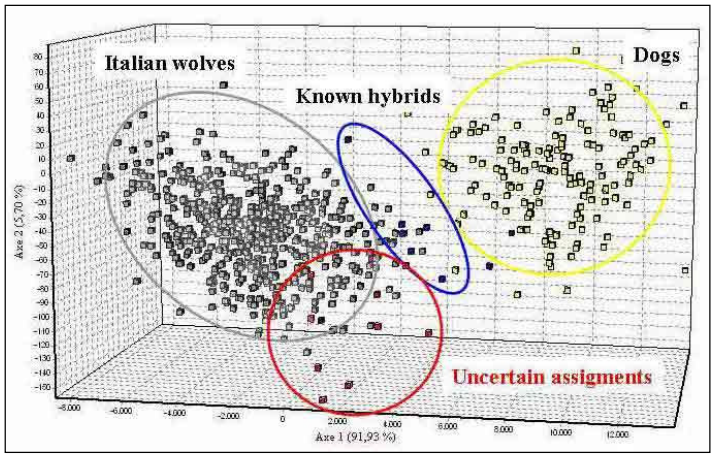

Fig. 3.1: Genetix graphic visualization of the samples genotyped at 6 microsatellite *loci*. Gray dots stand for Italian wolves, yellow dots for domestic dogs, blue for certain known hybrids and red dots for the uncertain assignments individuals.

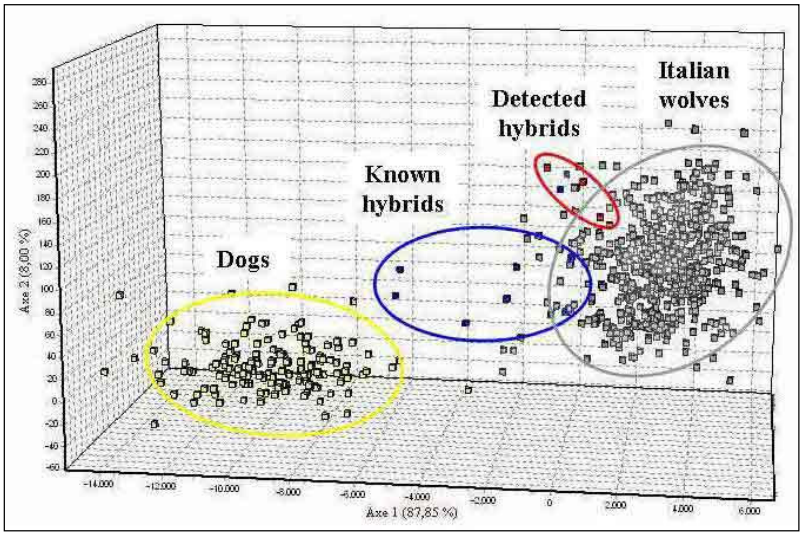

Fig. 3.2: Genetix graphic visualization of the samples genotyped at 12 microsatellite *loci*. Gray dots stand for Italian wolves, yellow dots for domestic dogs, blue for certain known hybrids and red dots for the 3 detected certain hybrids.

3.2.1 Wolf genotype Analyse

Positive PCR, dropout and false allele rates per *locus* in all the genotyped samples were estimated using the software Gimlet v. 1.3.2 (Valière, 2002).

All the 6 *loci* showed high rates of positive PCRs, ranging from 66 % (CPH8) to 98 % (FH2096), in fact they had been selected among the whole 12 microsatellite set because of their very high amplification success. Allelic dropout rates per *locus* varied from 10,1 % (CPH2) to 35,2 % (CPH8) and false allele rates per *locus* varied from 0 % (CPH2) to 8,1 % (FH2137) (Table3.3).

	CPH2	CPH8	FH2004	FH2088	FH2096	FH2137	Mean
Positive PCR	92,0%	66,0%	84,0%	94,0%	98,0%	86,0%	**87,0%**
ADO	10,1%	35,2%	24,8%	15,8%	13,9%	26,7%	**21,1%**
FA	0,0%	2,1%	4,4%	2,9%	3,0%	8,1%	**3,4%**

Table 3.3: rates of positive PCRs, allelic dropouts (ADO) and false alleles (FA) observed using replicated PCRs of 6 microsatellite *loci* in genotyped excremental DNA samples. The 6 *loci* considered are the *loci* used for the individual identification.

All FA rates showed a pattern that increased at the ADO increase, only *locus* CPH8 showed an inverse situation, in fact it presented the highest ADO rate and one of the lowest FA rates (Fig 3.3).

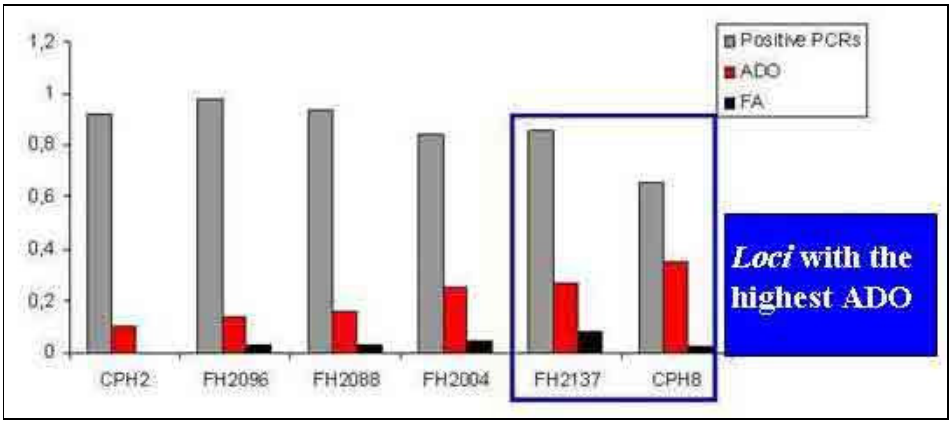

Fig. 3.3: rates of positive PCRs, allelic dropout and false alleles in genotyped samples showed that *loci* FH2137 and CPH8 presented the highest dropout rate.

The fact that *locus* CPH8 presented the highest ADO rate and the lowest amplification success could be due to the length of its amplified DNA sequence and to the high molecular weight of its alleles.Even if individual genotypes were detected using a multiple-tube approach and a quality screening test, as scats contain degraded DNA and noninvasive data could be error-prone, they were submitted to a further reliability check up and to a mismatch analysis that revealed that some wolf genotypes, sampled only once, resulted reliable but with high ADO rates, and presented 1 or 2 mismatches with other wolf genotypes sampled several times (Fig. 3.4). After that these individuals were regenotyped repeating PCRs at suspected *loci*, 31 of them (10%) were deleted because they revealed to have the same genotypes from which they mismatched for one or two alleles.

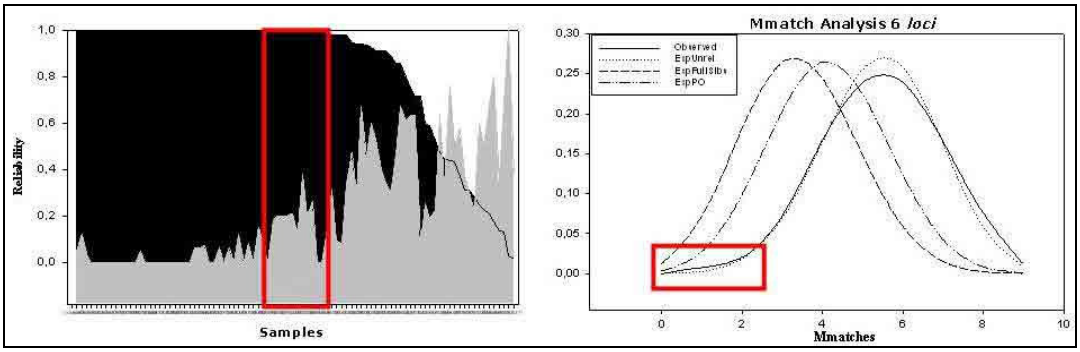

Fig. 3.4: reliability versus dropout and mismatch analyses carried out for the unique genotypes. Some genotypes with high reliability presented high ADO rates and differed from other genotypes only for one or two alleles.

This further check up, thus, allowed to detect during the whole study period, 347 reliable individual genotypes (corresponding to 269 wolves, 3 hybrids and 75 domestic dogs). 272 of them (corresponding to 269 wolves and 3 hybrids) were used to carry out all the necessary population genetic elaborations.

The 272 different individual genotypes were not sampled with the same frequency during the whole project period. In fact, among the wolf and hybrids individual genotypes, 106 (39%) were sampled only once, 98 (36%) from only twice to 5 times and the remaining 68 (25%) for more than 6 times (Fig. 3.5).

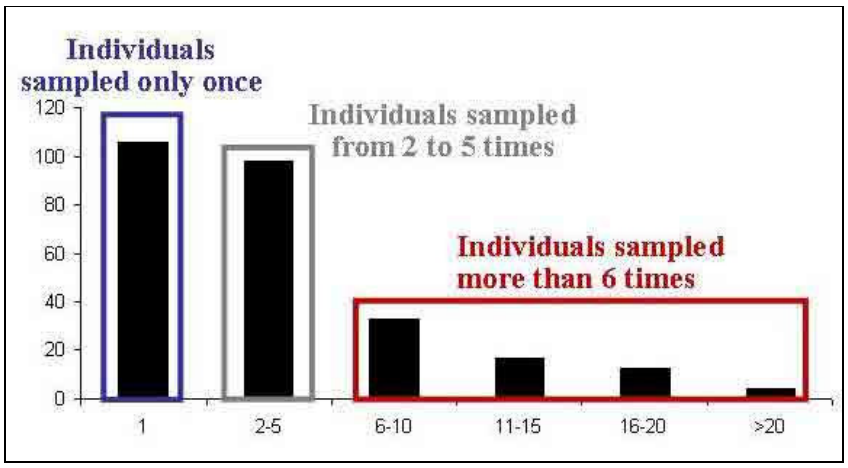

Fig. 3.5: proportion of individuals sampled from once to more than 20 times during the study period.

Moreover, if the time interval during which genotypes were sampled several times is considered, 183 individuals (67%) were observed for a period shorter than one year, while only 89 individuals (33%) were observed during a period longer than one year (Fig. 3.6).

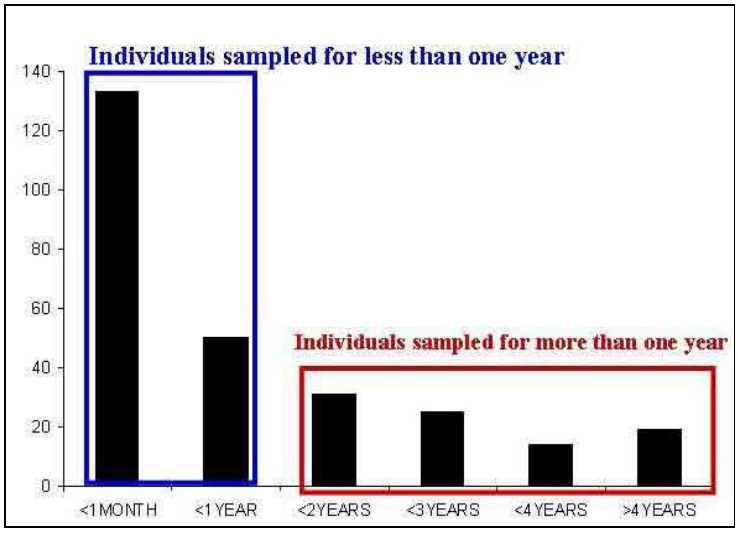

Fig. 3.6: proportion of individuals sampled for different time intervals. Most of genotypes were sampled for period minor than one year, only 33 % of the genotypes were sampled for period major than 1 year.

All the genotyped samples were also successfully sexed through a molecular sexing method proposed by Lucchini *et al.* (2002). Samples analyzed allowed to detect 153 males and 119 females with a sex ratio among detected individuals greater than one (1,29M:1,00F). Sex ratio values remained almost constant during the whole study period with values ranging from 1,00 to 1,43 (Table3.4).

During the study period the number of samples analyzed considerably increased ranging from a minimum value of 28 during the first study year (2000) to a maximum value of 940 during year 2006 allowing to detect from 4 to 114 different individuals per year (Fig.3.7, Table 3.4). The percentage of wolf genotypes detected per year remained almost constant ranging from 13% to 19%, with a mean value of 14,62%.

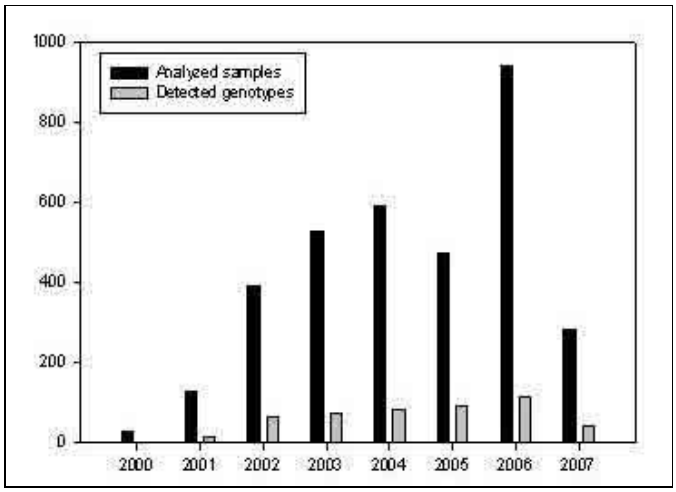

Fig. 3.7: number of samples analyzed and genotypes detected during the different years of the study period. Samples analyzed during 2007 include only samples collected until June.

Year	Analyzed Samples	Detected Genotypes	Genotyping percentage	Males	Female	Sex Ratio
2000	28	4	14%	2	2	1,00M:1,00F
2001	130	17	13%	10	7	1,43M:1,00F
2002	390	63	16%	32	31	1,03M:1,00F
2003	525	74	13%	40	34	1,18M:1,00F
2004	591	85	14%	37	48	0,77M:1,00F
2005	472	92	19%	51	41	1,24M:1,00F
2006	940	114	13%	64	50	1,28M:1,00F
2007	427	67	15%	35	32	1,09M:1,00F

Table 3.4: number of samples analyzed and genotypes detected during the different years of the study period. Samples analyzed during 2007 include only samples collected until June. Genotyping percentage and sex ratio (M/F) remain constant during the whole study period.

3.2.2 Population genetic analyses and microsatellite variability

The 269 distinct wolf genotypes noninvasively detected in this study were also used to estimate some of the chief parameters of Population Genetics using the programs GeneAlex v. 6.0 (Peakall & Smouse, 2005; 2006) and Genetix v.4.2 (Belkhir *et al.,* 2001).

All *loci* were polymorphic in the Emilia-Romagna wolf population, showing high values of heterozygosity ($\mathbf{H_O}$ = 0,59-0,74; $\mathbf{H_E}$ = 0,62-0,80) and a mean number of alleles per *locus* of 6,5 ranging from 3 (FH2096) to 12 (FH2137) (Table 3.5).

	CPH2	CPH8	FH2004	FH2088	FH2096	FH2137	Mean value
$\mathbf{N_A}$	6	6	6	6	3	12	6,5
$\mathbf{H_O}$	0,59	0,60	0,63	0,64	0,66	0,74	0,65
$\mathbf{H_E}$	0,62	0,72	0,63	0,65	0,65	0,80	0,68

Table 3.5: genetic diversity in Emilia-Romagna wolves genotyped at the 6 unlinked microsatellite *loci* used for the individual identification. $\mathbf{H_O}$ = observed heterozygosity, $\mathbf{H_E}$ = expected heterozygosity, $\mathbf{N_A}$ = number of alleles.

Emilia-Romagna wolves showed a slight deficit of heterozygotes (significantly positive $\mathbf{F_{IS}}$ values, Table 3.6) determining that *loci* were not completely in Hardy-Weinberg equilibrium.

Results of single-*locus* Hardy-Weinberg equilibrium tests showed that departures from HWE were contributed only by *loci* FH2137 (χ^2_{66} = 157,75; P < 0,005) and CPH8 (χ^2_{15} = 34,74 P < 0,005), which were both significant. This could be due to the fact that FH2137 and CPH8 *loci* presented the highest dropout rates and FH2137 even the highest false allele rate as showed in Table 3.3.

Population	$\mathbf{H_O}$	$\mathbf{H_E}$	$\mathbf{N_{Amean}}$	$\mathbf{N_P}$	$\mathbf{F_{IS}}$	P	$\mathbf{F_{ST}}$
Emilia-Romagna population	0,65 (0,03)	0,68 (0,03)	6,5 (0,75)	7	0.011**	0,010	0,09 (P = 0,01)

Table 3.6: genetic diversity in Emilia-Romagna wolves genotyped at 12 unlinked microsatellite *loci.* $\mathbf{H_O}$ = observed heterozygosity, $\mathbf{H_E}$ = expected heterozygosity, $\mathbf{N_{Amean}}$ = mean number of alleles per *locus* (direct count), $\mathbf{N_P}$ = number of private alleles. Standard errors in parentheses. Departures from Hardy–Weinberg equilibrium were assessed for the population from average multilocus $\mathbf{F_{IS}}$ values (the average individual inbreeding coefficient within the population), **P** = probability to obtain $\mathbf{F_{IS}}$ values lower (for negative $\mathbf{F_{IS}}$), or higher (for positive $\mathbf{F_{IS}}$) than observed after 1000 random permutations of alleles in each population determined using Genetix (**$\mathbf{P}$ < 0.01; *$\mathbf{P}$ < 0.05). $\mathbf{F_{ST}}$ = effect of subpopulations or individual diversity coefficient within the population.

Differentiation between Emilia-Romagna wolf population and the other Italian wolf populations (Alp, Central and Northern Apennine wolf populations) was assessed also by Analysis of MOlecular VAriance (AMOVA). A significant average *multilocus* $\mathbf{F_{ST}}$ = 0.09 (**P** = 0,01; computed from AMOVA) indicated that genetic diversity was significantly partitioned among the four wolf groups.

Moreover, Emilia-Romagna wolf population presented 7 alleles never found in the rest of the Italian wolf population (private alleles) (Table 3.6).

3.2.3 Individual genotype mapping

During sample collection, field collaborators compiled a technical card associated to the corresponding sample and containing important and useful field information such as sampling localities, sample quality but above all the geographic coordinates, necessary to map by gis the spatiotemporal locations of each collected sample. After completing the genetic analyses the mapped samples were utilized to realize thematic maps, through the software **ArcView GIS** (ESRI), about detected *multilocus* genotypes in the study area.

3.2.4 Mapping pack localizations

As all analyzed samples were characterized by the geographical coordinates of their sampling localities, mapping the detected *multilocus* genotypes by **ArcView GIS** (ESRI), it was possible to have an idea of the areas with high wolf density and thus to carry out some preliminary hypotheses about the different probable packs living in the study area.

In fact, circumscribing a perimeter around the localities in which the different samples belonging to the same individual were collected, it was possible to detect the territory where each individual was considerably stable in the time. Overlapping these individual areas and comparing the sampling periods of the individuals observed in them, it was possible to detect the different plausible packs and to make some preliminary hypotheses about their localizations, structure dynamics and interactions. Using this scheme, and integrating the data so obtained with the wolf-howling ones, 22 different possible wolf packs were identified in the study area, localized along the whole Emilia-Romagna Apennine Ridge: 1 in the area between Parma provincial administration and Cento Laghi regional park, 2 in the area between Reggio-Emilia provincial administration and

Gigante regional park, 3 in the area of Alto Appennino Tosco-Emiliano national park, 3 in the area between Modena provincial administration and Cento Laghi regional park; 6 in the area between Bologna provincial administration and Corno alle Scale regional park and 8 in the area between Foreste Casentinesi national park and Forlì-Cesena provincial administration (Fig. 3.8).

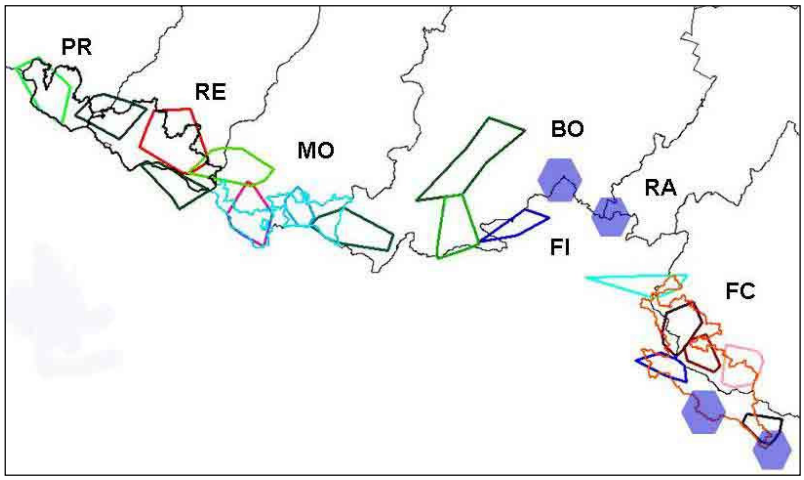

Fig. 3.8: probable mapping pack localizations in the study area. The polygons indicate the probable pack localizations detected mapping the *multilocus* genotypes, the blue hexagons indicate the probable pack localizations detected through wolf-howling.

3.2.5 Dispersal events

Mapped individual genotypes were utilized also to monitor eventual dispersal events, that can be detected when the same individual is sampled after a certain period, in an area different from the one in which it was sampled the first time. In this study, considering air line movings of at least 25 kilometres (corresponding to longer real distances on the territory) as dispersal events, it was possible to detect 17 presumed dispersal events through noninvasive genotyped samples (Table 3.7).

Genotype	First sampling year	Provincial administration	Successive sampling year	Provincial administration	Km	Direction
WBO10M	2001	BO	2002	PR	114	SE-NW
WBO14F	2001	BO	2005	FCPN	74	NW-SE
WBO16M	2002	BO	2006	RE	66	SE-NW
WFO15M	2002	FCPN	2006	PR	142	SE-NW
WFO25M	2002	FCPN	2006	BO	52	SE-NW
WRE4M	2002	RE	2003	MO	27	NW-SE
WRE6M	2002	RE	2003	BO	76	NW-SE
WFO47M	2003	FCPN	2005	RA	25	SE-NW
WPR3M	2003	MO	2004	PR	53	SE-NW
WBO27F	2004	BO	2006	FI	32	NE-SW
WBO38M	2004	BO	2006	RA	34	SE-NW
WFO46F	2004	FCPN	2006	RE	124	SE-NW
WFO61M	2004	FCPN	2006	BO	66	SE-NW
WBO44M	2005	BO	2007	PR	105	SE-NW
WFI12M	2005	FI	2007	BO	85	SE-NW
WFO77M	2005	FCPN	2007	BO	67	SE-NW
WMO46M	2005	MO	2007	PR	64	SE-NW

Table 3.7: 17 dispersal events detected during the study period. Most of dispersals were male biased and occurred in direction SudEst-NordWest.

Most of these presumed dispersal events (13) occurred in direction Southeastern-Northwestern, while only 4 of them in direction Northwestern-Southeastern; moreover most of the dispersal events (14) were male biased while only 3 were female biased (Table 3.7).

3.2.6 Mapping hybrid localizations

Scat samples that resulted to be wolf-x-dog hybrids were 11, corresponding to 3 different individuals (2 males and 1 female). The genetic analyses were based on 12 microsatellite *loci* genotyping (Randi *et al*., 2000; Randi & Lucchini, 2002). The 2 male hybrids were both sampled only once, the first on December 2002 in Gigante Regional Park during a snow tracking sampling activity, the second on January 2005 in an area between Gigante regional Park and Reggio-Emilia provincial administration. On the contrary the female hybrid was noninvasively sampled 9 times, from December 2002 to June 2006, in an area between Bologna and Firenze provincial administrations. On December 2006, in the same area, a female canid carcass, that had the typical Italian wolf coat colour pattern, was found. It was genotyped at 12 microsatellite *loci* and the genetic analyses proved a complete match with the female hybrid *multilocus* genotype.

All the hybrid samples detected during this study were mapped by **ArcView GIS** (ESRI) (Fig 3.9) to investigate whether they frequented the same areas in which some wolf packs established their territories and home ranges and whether other animals sampled in these areas were characterized by traces of hybridization and dog gene introgression.

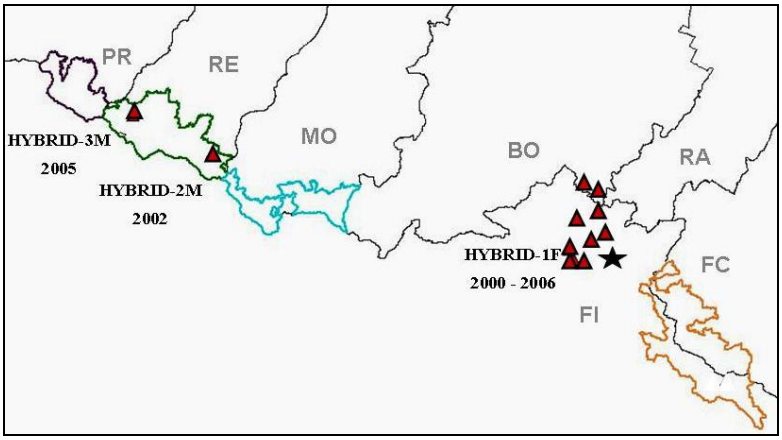

Fig. 3.9: mapping hybrid localizations in the study area. The red triangles are referred to the non-invasively hybrid samplings while the black star is referred to the hybrid carcass sampling.

3.2.7 Mapping free-ranging domestic dog localizations

Scat samples that resulted to be domestics dogs were 78, corresponding to 75 different individuals. In fact 72 dog individuals were sampled once, while only 3 individuals (1 male and 1 female in Bologna provincial administration and 1 female in Forlì-Cesena provincial administration) were sampled twice. All these samples were mapped by **ArcView GIS** (ESRI) (Fig 3.10) to examine whether, in the study area, feral or free-ranging domestic dogs could stably live in the same territories frequented by wolves interacting with them, and to understand whether they could be considered as wolf food competitors.

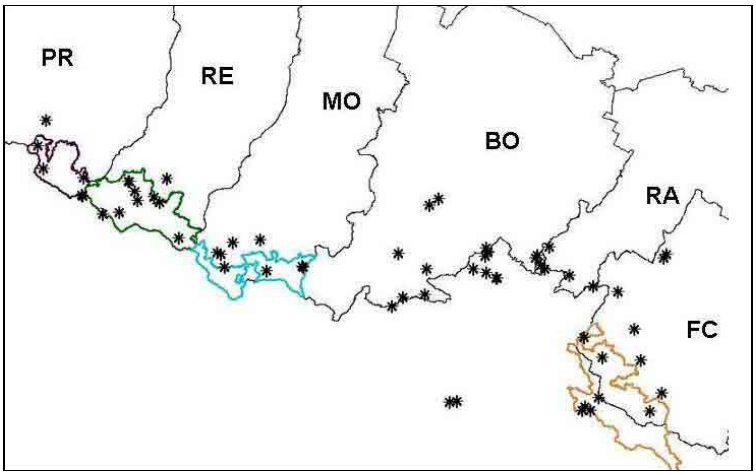

Fig. 3.10:mapping feral or free-ranging domestic dog localizations in the study area.

3.3 TISSUE AND BLOOD SAMPLE ANALYSES

During this study also 34 invasive samples (30 carcasses and 4 blood samples) were genetically analyzed. They were genotyped using the same 12 microsatellite *loci* used for the individual identification and for the population assignment of non-invasive samples. For 5 carcasses it was not possible to obtain any kind of information because of their low quality DNA content. The remaining 25 carcasses were successfully genotyped and correctly assigned allowing to detect a domestic dog and 24 Italian wolves. Among the wolves, 5 carcasses resulted to be corresponding to 5 different genotypes previously noninvasively sampled (1 of them matched with the female hybrid previously cited), and 19 resulted to be new individuals never previously sampled. Two of the 4 blood samples were found during snow tracking, while the other 2 were collected from 2 live-trapped individuals. 1 of the 2 live-trapping samples and one of 2 blood traces on snow resulted to be corresponding to 2 *multilocus* genotypes already noninvasively sampled different times and probably already alive. The other 2 blood samples resulted to be new individuals never previously sampled (Table 3.8).

Sample	Provincial admnistration	Sampling year	Sample kynd	Population assignement	Genotype	Match
W555	BO	2000	CARCASS	WOLF	NEVER SAMPLED	NO
W719	BO	2001	CARCASS	WOLF	NEVER SAMPLED	NO
W720	BO	2001	CARCASS	WOLF	OLD	YES
W667	RE	2002	CARCASS	WOLF	NEVER SAMPLED	NO
W730	RE	2003	CARCASS	WOLF	OLD	YES
W774	BO	2005	BLOOD	WOLF	OLD	YES
W975	RE	2004	BLOOD	WOLF	NEVER SAMPLED	NO
W777	PN	2004	CARCASS	WOLF	NEVER SAMPLED	NO
W892	PN	2005	CARCASS	WOLF	NEVER SAMPLED	NO
W893	MO	2005	CARCASS	WOLF	NEVER SAMPLED	NO
W903	FC	2005	CARCASS	NOT DETECTED		
W904	MO	2006	CARCASS	WOLF	NEVER SAMPLED	NO
W905	MO	2006	CARCASS	WOLF	OLD	YES
W906	MO	2006	CARCASS	WOLF	OLD	YES
W911	RE	2006	CARCASS	DOG	NEVER SAMPLED	NO
W912	PC	2006	CARCASS	WOLF	NEVER SAMPLED	NO
W914	PN	2006	CARCASS	NOT DETECTED		
W915	PN	2006	CARCASS	NOT DETECTED		

Sample	Provincial admnistration	Sampling year	Sample kynd	Population assignement	Genotype	Match
W915a	PN	2006	CARCASS	WOLF	NEVER SAMPLED	NO
W919	PN	2006	CARCASS	NOT DETECTED		
W929	BO	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W931	PN	2006	CARCASS	WOLF	NEVER SAMPLED	NO
W933	BO	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W934	BO	2006	CARCASS	WOLF	NEVER SAMPLED	NO
W936	FI	2006	CARCASS	HYBRID	OLD	HYBRID1F
W937	FI	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W938	FI	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W940	FC	2007	BLOOD	WOLF	NEVER SAMPLED	NO
W942	MO	2007	BLOOD	WOLF	OLD	YES
W943	BO	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W945	BO	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W955	BO	2007	CARCASS	WOLF	NEVER SAMPLED	NO
W958	PC	2007	CARCASS	NOT DETECTED		
W974	BO	2007	CARCASS	WOLF	NEVER SAMPLED	NO

Table 3.8: tissue and blood samples analyzed during the study period. NOT DETECTED is referred to those samples for which it was not possible to complete the genetic analyses because of their low quality DNA content. NEVER SAMPLED is referred to those samples never sampled before.

3.4 POPULATION SIZE ESTIMATION

Multilocus genotypes, obtained from non-invasively collected samples, can be used to estimate population size in several ways (Sloane *et al*., 2000). If individuals are sufficiently sampled, mark-resight methods of estimating population size can also be applied to genetic data to estimate resighting probabilities (Otis *et al.*, 1978; Seber, 1982).

In this study analyses about Emilia-Romagna wolf population size estimation were carried out with capture-recapture models for open populations that allow to estimate apparent survival **(Φ)** accounting for the probability of recapture **(*P*)** (Lebreton *et al.,* 1992). All the 272 *multilocus* genotypes detected from wolf scats sampled over the Emilia Romagna during the whole study period, from eastern 2000 to winter 2007,: were used as capture-recapture data, thus the study period was subdivided in 28 "3 month period ***occasions*"**. The first detection of a genotype was considered as marking, while further detections were considered as recaptures with the approximation that eventual genotyping errors were assumed to be insignificant as a result of careful lab procedures.

3.4.1 Goodness of Fit analyses

The goodness-of-fit tests (GOF) of the Cormack-Jolly-Seber (CJS) model which has both capture **(*P*)** and survival **(Φ)** probabilities dependent on time **[Φ(t)*P*(t)]** represent the first step of the analyses of capture-recapture data. These preliminary tests were performed using the program **U-Care** (Choquet *et al*., 2005) and allowed to explore the fit of the CJS to the Emilia-Romagna wolf population data and also to identify a general model that fits the data from which to start model selection. When the GOF tests were run on the full parameter Cormack-Jolly-Seber model, the overall GOF for Emilia-Romagna dataset was not significant (χ^2_{167} = 175,56; P = 0.052). However, both **Test 3.SR** (χ^2_{22} = 44,12; P = 1.68 x 10^{-6}) and **Test 2.Ct** (χ^2_{23} = 69,10; P = 3.42 x 10^{-3}) were strongly significant. The significant positive **Z-Statistics** for transience (Z= 6,15; $P_{\text{one-sided test for transience}}$ = 3.75 x 10^{-10}) suggested that more wolves than expected under the CJS model were seen only once (occurrence of transient individuals in the marked population).

Furthermore, significant negative **Z-Signed Statistics** for trap-dependence (Z = -5,11; $P_{two\text{-}sided\ test\ for\ transience} = 3.22 \times 10^{-7}$) suggested there was evidence for a significant trap-happiness effect on capture probabilities (probability of first capture < probability of recapture) (Table 3.9).

	Df	χ^2	P_{χ^2}	Z-statistics	P_Z	Significant	Meaning
TEST 3.SR	**22**	44,12	1.68×10^{-6}	+ 6,15	3.75×10^{-10}	YES	Transience excess
TEST 2.Ct	23	69,10	3.42×10^{-3}	- 5,11	3.22×10^{-7}	YES	Trap-happiness
Global Test	167	175,56	0.052			NO	

Table 3.9: results about GOF tests obtained using the software U-Care (Choquet *et al*., 2005).

Even if the GOF global test was not significant, the fact that both the Test **3.SR** (**Z-Statistics** for transient > 0) and **Test 2.Ct** (**Z-Signed Statistic**s for trap-dependence < 0) were strongly significant showed that Emilia-Romagna wolf population data presented a clear signal of permanent heterogeneity of capture among individuals and that the CJS model did not completely fit the data suggesting that more complex models were needed.

3.4.2 New model selection

The GOF tests showed that Emilia-Romagna data seemed ideal for using multievent models (Lebreton & Pradel, 2002; Pradel, 2005) incorporating mixture heterogeneity (Pledger *et al*., 2003) to detect the main biological parameters necessary to obtain a reliable population size estimation. Multievent models are models in which events (observation) do not match with states (physiological, geographical). Models with heterogeneity of capture are models in which individuals are in 2 classes of capturability: a proportion $\boldsymbol{\pi}$ of the N individuals that have low capture probabilities ($\mathbf{P_L}$) and a proportion ($\mathbf{1-\pi}$) of the N individuals that have high capture probabilities ($\mathbf{P_H}$). Total capture probability **P** is a mixture of the 2 groups: $\mathbf{P = \pi \times P_L + (1-\pi) \times P_H}$. This probability is necessary to obtain a reliable population size estimation: **Nj = Cj/Pj**, where **Cj** is the number of counted individuals and **Pj** is the estimated detection probability at the capture occasion **j**.

Multievent model building and selection were performed using the software **E-Surge v. 1.1.1** (Multi**E**vent **G**eneralized **Sur**vival **E**stimation) (Choquet *et al.*, 2007), in fact it is a program for building and fitting multievent models to capture-recapture (CR) data by Maximum Likelihood.

3.4.3 Model buildings

18 different multievent models with heterogeneity of capture and characterized by 2 events ("seen"; "not seen") and 3 states (belongs to class with low capture probabilities, belongs to class with high capture probabilities, Dead) were built using the software E-Surge (Table 3.10). The definition of each model and its parameterization were obtained using the software's tool GEPAT (GEnerator of PATtern matrices). Constrained models were built using a model description language interpreted by the software's tool GEMACO (generator of constrained or design matrices), avoiding tedious and error-prone matrix manipulations. In the Model Definition Language (MDL) of GEMACO classical effects, such as time (**t**), age (**a**) and group (**g**), were widely used to explain variability in the data and to build the matrices of constraint. These effects were represented by reserved Keywords and synonyms that facilitated the writing models. More complex models were built combining the classical effects through two main effects, **from** (departure from a previous state) and **to** (arrival to a current or next state), and two operators **dot product (.)** and **sum (+)**.

Analyzed Model	AICc	ΔAICc	AICc weight	np	Deviance (PI,PHI,B)
[Φ(*ct*)*p*1(*S*+*H*)*p*2(*S*+*H*)] Best model	2028,386	0	0,597625926	7	2014,386
[Φ(*ct*)*p*1(*S*)*p*2(*S*)]	2029,180	0,794	0,401804245	10	2009,180
[Φ(*ct*)*p*1(*t*+*g*)*p*2(*t*+*g*)]	2042,878	14,492	0,00042612	56	1930,878
[Φ(*ct*)*p*1(*t*)*p*2(*t*)]	2045,056	16,670	0,000143412	53	1939,056
[Φ(*ct*)*p*1(*t*.*g*)*p*2(*g*)]	2059,046	30,660	1,3143E-07	56	1947,046
[Φ(*ct*)*p*1(*g*)*p*2(*t*.*g*)]	2059,046	30,660	1,3143E-07	56	1947,046
[Φ(*ct*)*p*1(*t*.*g*)*p*2(*t*)]	2063,115	34,729	1,71839E-08	78	1907,115
[Φ(*ct*)*p*1(*t*)*p*2(*t*.*g*)]	2063,115	34,729	1,71839E-08	78	1907,115
[Φ(*ct*)*p*1(*ct*)*p*2(*ct*)]	2078,347	49,961	8,46323E-12	4	2070,347
[Φ(*ct*)*p*1(*t*+*g*)*p*2(*t*.*g*)]	2078,915	50,529	6,37084E-12	79	1920,915
[Φ(*ct*)*p*1(*t*.*g*)*p*2(*t*+*g*)]	2078,915	50,529	6,37084E-12	79	1920,915
[Φ(*ct*)*p*1(*g*)*p*2(*g*)]	2079,891	51,505	3,91076E-12	6	2067,891
[Φ(*g*)*p*1(*ct*)*p*2(*ct*)]	2080,160	51,774	3,4186E-12	5	2070,160
[Φ(*t*)*p*1(*ct*)*p*2(*ct*)]	2090,515	62,129	1,92881E-14	30	2030,515
[Φ(*t*+*g*)*p*1(*ct*)*p*2(*ct*)]	2091,580	63,194	1,13247E-14	31	2029,580
[Φ(*ct*)*p*1(*t*.*g*)*p*2(*t*.*g*)]	2096,990	68,604	7,57286E-16	100	1896,990
[Φ(*g*.*t*)*p*1(*ct*)*p*2(*ct*)]	2134,789	106,403	4,6915E-24	55	2024,789
[Φ(*ct*)*p*1(*f*+*s*)*p*2(*f*+*s*)]	2553,899	525,513	4,5997E-115	8	2537,899
[Φ(*ct*)*p*1(*s*)=*p*2(*s*)] HOMOGENEITY	2103,767	75,381	2,55654E-17	5	2093,767

Table 3.10: analyses of capture-recapture Emilia-Romagna wolf data using different heterogeneous models. **AICc** = Akaike's information criterion, **ΔAICc** = differences between that model and the model with the lowest AICc, **np** = number of parameters; **Deviance**= minimum relative deviance; **Φ** = survival probability; ***p*1** = low recapture probability; ***p*2** = high recapture probability; ***ct***= constant; ***t*** = time effect; ***g*** = group effect; ***S*** = seasonal effects , ***H*** = heterogeneity effect, . = product effect; + = plus effect, ***f*** = from effect. A homogeneous model was run to carry out a comparison with the best heterogeneous one.

3.4.4 Best Model selection

As Pledger's models are based on the Likelihood theory, model selection was undertaken on the basis of Akaike's information criterion (**AICc**) (Lebreton *et al.*, 1992), where the AIC is equal to the deviance from the model plus two times the number of estimable parameters (**dev/ĉ + 2** x **np**) and where the model with the lowest AIC is considered to be the most parsimonious. A difference in AIC of two or more units is generally accepted to indicate a significant difference in model fit (Lebreton *et al.*, 1992). The best model (AICc = 2028,386) resulted to be the one built considering a constant survival probability (**Φ(*ct*)**) and low and high recapture probabilities depending on the additional effects of heterogeneity and season (***p*1(*S*+*H*)*p*2(*S*+*H*)**) (Table 3.10). The heterogeneity effect considers the heterogeneity of capture among individuals while the season one considers the same capture probability for the same seasons during the different project years. The biological parameters (**π, P_L, P_H** and **C**), obtained running this model and calculated for each considered season, were submitted to 1000 bootstrap replicates to reduce the gap between simulation and approximation. Using the formula **Nj = Cj/Pj**, for each bootstrap replicate Nj was calculated considering 27 "3 month period ***recapture occasions***" (the first one was considered as marking and thus fixed). The mean of the 1000 bootstrap replicates for all the 27 "3 month period ***recapture occasions***" **Nj** allowed to obtain a reliable parameterization of the selected model and thus a reliable Emilia-Romagna wolf population size estimation.

Bootstrap techniques permitted also to achieve the 95% confidence intervals discarding the lowest 2.5% estimates and the highest 2.5% estimates (Table 3.11).

HETEROGENEITY							
Occasion	**Cj/Pj=Nj**	**Lower-CI**	**Upper-CI**	**Occasion**	**Cj/Pj=Nj**	**Lower-CI**	**Upper-CI**
Spring 2000	Fixed first occasion			**Autumn 2003**	154,11	108,45	203,56
Summer 2000	8,56	0,00	26,82	**Winter 2004**	178,09	130,70	235,63
Autumn 2000	4,75	0,00	15,39	**Spring 2004**	101,55	60,18	148,78
Winter 2001	29,11	11,36	53,24	**Summer 2004**	86,47	38,27	137,76
Spring 2001	22,96	5,16	49,11	**Autumn 2004**	172,59	125,09	229,56
Summer 2001	61,71	22,20	107,77	**Winter 2005**	184,46	129,21	244,39
Autumn 2001	22,91	4,72	46,98	**Spring 2005**	162,89	111,65	213,09
Winter -2002	69,11	39,49	104,53	**Summer 2005**	148,39	86,00	228,06
Spring 2002	67,48	32,90	109,36	**Autumn 2005**	150,57	100,67	207,85
Summer 2002	138,88	82,96	208,07	**Winter 2006**	250,57	181,91	322,69
Autumn 2002	213,29	154,60	275,21	**Spring 2006**	236,81	181,87	292,68
Winter 2003	161,68	118,90	211,36	**Summer 2006**	149,37	86,70	229,84
Spring 2003	72,73	35,60	114,78	**Autumn 2006**	183,54	128,34	247,80
Summer 2003	156,63	91,87	227,01	**Winter 2007**	157,28	113,68	206,92

Table 3.11: values of estimation of Emilia-Romagna wolf population size (Nj) computed for the 27 different capture occasions and their relative lower and upper confidence intervals (CI).

Emilia-Romagna wolf population size estimation presented a clear seasonal trend during the whole study period, with a total mean value of 123,95 (95%CI 80,83-174,01) individuals, ranging from a minimum of 4,75 (95%CI 0,0-15,38), during autumn 2000, to a maximum of 250,57 (95%CI 181,91-322,68), during winter 2006 (Table 3.11; Fig. 3.11). During the first study years (2000, 2001) the population size seemed to be characterized by a regular size increasing during winter and summer periods and a regular contraction during spring and autumn periods. During year 2002 population size increased in winter, remained constant in spring and widely increased during both summer and autumn. During year 2003 population size decreased in winter and spring, reincreased in summer and remained constant in autumn period. During year 2004 population size increased in winter, decreased in both spring and summer and re-increased in autumn. During years 2005 and 2006 population size presented a similar pattern characterized by a great increasing in winter, a high decreasing in spring and summer and a little re-increasing during autumn. Year 2007, on the contrary, represented a population size decreasing during winter period

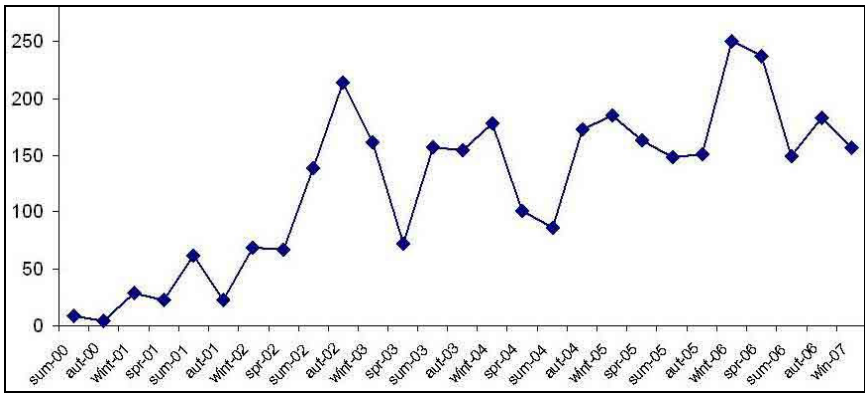

Fig. 3.11: seasonal pattern of Emilia-Romagna wolf population size estimated through the best heterogeneous model

3.4.5 Heterogeneity *versus* Homogeneity

To investigate the effective reliability of the mixture heterogeneity model used for the Emilia-Romagna wolf population estimation, it was compared to a homogeneous model built considering a constant survival probability (**Φ(*ct*)**) and a homogeneous capture probability among individuals depending on the seasonal effects (***p*1(*S*)=*p*2(*S*)**) (Table 3.10). Even in this case the seasonal effect considers the same capture probability for the same seasons during the different project years.

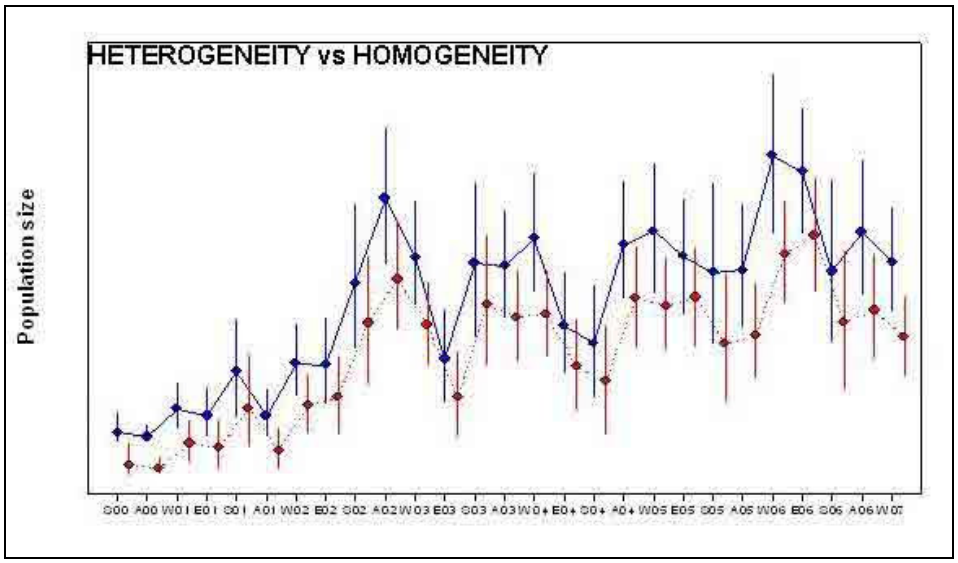

Fig. 3.12: Best heterogeneous model (blue continues line) compared to a homogeneous one (red dot line). In both Heterogeneous and Homogeneous models points represent the 27 different capture occasions (the first one is fixed) and vertical bars represent confidence intervals.

Even if the homogenous model showed confidence intervals lower about 1,5 times than the heterogeneous one, it produced a population size estimation about 1,5 times lower than the one obtained by the heterogeneous model (Table 3.12, Fig 3.12).

Lower CI	Upper CI	Pop Size Estimation
CI-Het/CI-Hom	CI-Het/CI-Hom	N-Het/N-Hom
1,586042472	1,623206587	1,57323988

Table 3.12: relative merits of the CJS (homogeneity) and mixture heterogeneity model. In the heterogeneous model lower and upper confidence intervals, and population size estimation are about 1,5 times major than in the homogeneous model.

3.5 SINGLE NUCLEOTIDES POLYMORPHISM (SNP) ANALYSES

Noninvasive DNA analyses are often prone to genotyping errors (false alleles and allelic dropouts) due to DNA degradation. Thus, the use of single nucleotide polymorphisms (SNPs), which requires amplification of much shorter DNA sequences may allow more efficient genotyping of noninvasive samples (Seddon *et al*., 2005; Holm Andersen & Fabbri *et al.,* 2006).

Canine Single Nucleotide Polymorphisms (SNPs) characterization in the Italian wolf population, started resequencing sequence-tagged-site (STS) DNA sequences that were known to contain SNPs in domestic dogs (see Holm Andersen & Fabbri *et al.,* 2006). DNA fragments, extracted from 14 Italian wolf samples, collected in north and central Italy, were amplified by PCR using 76 primer pairs for SNPs containing dog STS sequences (Guyon *et al*., 2003).

49 (64%) of these 76 tested primer pairs reliably amplified and so their PCR products were purified and sequenced in both directions. Sequence analysis and alignment allowed to detect 59 different SNPs in the Italian wolf population.

On the basis of this first SNP characterization in the Italian wolf population 21 new primer sets (see Holm Andersen & Fabbri *et al.,* 2006) for analysing SNPs using Pyrosequencing technology (Ronaghi *et al*., 1998) were designed. 15 (72%) of these 21 tested primer pairs reliably amplified and allowed to detect 59 different SNPs which were verified by comparing the results with the sequences and resulted to be the same ones previously found in the Italian wolf population (Table 3.13).

Method	Tested primers	Reliable primers	Detected SNPs
Traditional sequencing	76	49 (64%)	59
Pyrosequencing	21	15 (72%)	59

Table 3.13: first Italian wolf SNP characterization using both the traditional sequencing method and Pyrosequencing technology.

3.5.1 SNP analysis method comparison

These 15 SNP *loci* were successively used to design 6 new primer sets (182B11/138; 309N24/298; 1C06/138; 38K22/150; I96B17/422; F310M20/207), that could be used not only through Pyrosequencing technology but also by other 2 genotyping methods, SNaPshot and RealTime PCR.

43 non-invasive samples, extracted using a guanidinium-silica protocol (Gerloff *et al*., 1995) and previously genotyped by PCR at 12 microsatellite *loci*, were amplified, following a multiple tube approach (Taberlet *et al*., 1996; Gagneux *et al*. 1997; Lucchini *et al*., 2002), 3 times for these 6 SNP primer sets by the 3 different genotyping techniques: Pyrosequencing, SNaPShot and Real Time PCR to estimate which was the best method for a future use in individual genotyping of low content DNA samples and for their assignment to the belonging populations.

The 43 samples were chosen on the base of their different amplification success during their previous microsatellite genotyping at 12 *loci*, in fact PCR success was 90% in 14of them, 60% in 14 of them and minor than 20% in 15 of them.

Real Time PCR resulted to be the best method, in fact even if it did not present the best PCR success (66%) it was the faster method with the lowest ADO rate (4,09 %) calculated on the total of amplifications and the lowest ADO rate (3,3 %) calculated in samples with positive PCRs greater than 50%. Moreover RealTime method produced reliable results with the lowest number of necessary replicates, max 2, per sample per *locus* (Table 3.14). In the future, if ADO rate should become lower, reliable results might be obtained also through only 1 replicate per sample per *locus*.

Method	Positive PCR %	Total ADO %	ADO in Positive PCRs > 50%	Necessary Replicates
Pyrosequencing	62	14,00	9,3 (N=28)	> = 3
SNaPshot	80	10,09	8,2 (N=34)	> = 3
RT-PCR	66	4,09	3,3 (N=30)	Max 2

Table 3.14: results obtained amplifying 43 noninvasive samples 3 times at 6 SNP *loci* using 3 different SNP analysis methods: Pyrosequencing, SNaPshot and RealTime PCR. RealTime method resulted to be the one with the lowest ADO rate and the minor number of necessary replicates to obtain reliable results.

3.5.2 SNPs *versus* Microsatellites

As RealTime PCR resulted to be the best among the 3 SNP genotyping methods, it was used to try to create an efficient and reliable laboratory protocol for the individual identification and population assignment in low content DNA samples based on *multilocus* SNP genotyping.

The 6 RealTime primer sets used for the 3 different method comparison were then used to characterize 30 Italian wolf tissue DNA samples and 30 dog tissue DNA samples to test their real discrimination power. The analyses showed that only 4 of them (182B11/138; 309N24/298; 1C06/138; 38K22/150;) were reliably polymorphic both in the Italian wolves and between wolves and dogs so other 5 primer sets (168J14/149; 218J14/81; 372M9/32; BLA22/199; BLB52/368) suitable for RealTime were designed.

The application reliably and the performance of these 9 RealTime primer sets was compared to the application reliably and the performance of the 6 microsatellite *loci* used for the individual identification and population assignment.

For this reason 28 scat DNA samples, that passed the preliminary microsatellite quality screening test (consisting in amplifying each sample four times at 2 of the 6 microsatellite *loci* used for the individual discrimination) were genotyped using both genotyping methods. The comparison started considering a multiple tube approach, based on 2 initial replicates per sample per *locus* for RealTime genotyping method and 4 initial replicate per sample per *locus* for usual microsatellite genotyping method, followed by a reliability analyses (Miller *et al*., 2002).

Microsatellite genotyping showed a discrete PCR success (62%), high allelic dropout (15,10 %) and high false allele (2,60 %) rates (Table 3.15).

Moreover, using microsatellites, after the first 4 replicates per sample per *locus*, only 15 samples (54 %) resulted reliable while 13 (46 %) needed 4 further additional PCRs at different *loci*. At the end 3 samples (11 %) were deleted because not reliable while 25 samples (89 %) were reliable and successfully genotyped allowing to detect, by the software structure, 18 different Italian wolves, 1 domestic dog and 1 uncertain assignment individual (Table 3.15).

Method	Positive PCR %	ADO %	FA %	Genotyping % after 4-2 replicates	Final Genotyping %	Different wolves	Different dogs	Uncertain individuals
Microsatellite genotyping	62	15,10	2,60	54 (N=15)	89 (N=25)	18	1	1
RT-PCR SNP genotyping	86	1,09	0,0	79 (N=22)	96 (N=27)	17	1	1

Table 3.15: results about 28 scat DNA samples, that passed the preliminary microsatellite quality screening test, and that were genotyped using both 9 RealTime SNP *loci* and 6 microsatellite *loci*.

SNP genotyping showed a high PCR success (86%), very low allelic dropout rates (1,09 %) while false allele were not present (Table 3.14).

Moreover, using SNPs, after the first 2 replicates per sample per *locus*, 22 samples (79 %) resulted reliable while only 6 (21 %) needed 2 further additional PCRs at different *loci*. At the end, only one sample (4%) was deleted, while 27 samples (97 %) resulted reliable and successfully genotyped allowing to detect, by the software structure, 17 different Italian wolves, 1 domestic dog and 1 uncertain assignment individual (Table 3.14).

A comparison between the samples genotyped through both methods, showed they both allowed to detect the dog and the uncertain assignment individual. Microsatellite genotyping permitted to detect 18 different wolves, while SNP genotyping was not able to distinguish all the individuals detecting only 16 different wolves. This suggested that an insufficient number of SNP *loci* was probably used and that Probability of Identity (PID) values were not low enough, so that some different individuals of the same population might have identical profiles and thus not detected (*shadow effect*) leading to a population size underestimation (Mills *et al*., 2000).

As Emilia-Romagna wolf population could consist of about 100-200 individuals, genotyping primer sets used for the individual identification, to ensure a reliable detection of unique genotypes, also if related individuals were sampled, should produce PID_{sibs} values minor than 0,01.

For this reason a set of 30 Italian wolves were genotyped using both microsatellite and SNP genotyping methods and the probability of identity (PID), the probability of identity corrected for small population size (PIDcor) and the expected PID among full sib dyads (PID_{sibs}) values were estimated using the software **GIMLET v. 1.3.2**. (Valière, 2002) (Table 3.16).

SNPs				Microsatellites			
Locus	PID	PIDcor	PIDsibs	*Locus*	PID	PIDcor	PIDsibs
1C06/138	4,18E-01	4,30E-01	6,51E-01	**FH2004N**	1,37E-01	1,40E-01	4,40E-01
38K22/150	3,77E-01	3,84E-01	6,05E-01	**FH2088N**	1,68E-01	1,71E-01	4,57E-01
96B17/422	4,73E-01	4,89E-01	7,00E-01	**FH2096N**	1,84E-01	1,89E-01	4,66E-01
120D19/347	8,21E-01	8,31E-01	9,12E-01	**FH2137N**	6,52E-02	6,82E-02	3,69E-01
133N13/219	5,00E-01	5,17E-01	7,21E-01	**CPH2**	2,36E-01	2,40E-01	5,22E-01
168J14/149	3,73E-01	3,78E-01	5,98E-01	**CPH8**	9,86E-01	1,86E-01	3,98E-01
182B11/138	3,70E-01	3,75E-01	5,94E-01				
218J14/81	4,51E-01	4,66E-01	6,82E-01				
309N24/298	4,25E-01	4,38E-01	6,58E-01				
Total	8,07E-04	1,01E-03	2,89E-02	**Total**	6,43E-05	1,38E-05	7,19E-03

Table 3.16: Probability of Identity for each SNP and microsatellite *locus* estimated in a set of 30 Italian wolves using Gimlet v. 1.3.2. (Valière, 2002). **Pid** is the probability of identity for individuals randomly chosen within the same population, $\mathbf{Pid_{cor}}$ is the probability of identity corrected for small population size, $\mathbf{Pid_{sibs}}$ is the probability of identity corrected for siblings. Pid sibs < Pid cor < Pid. Each total probability was computed by multiplying single *locus* probabilities, assuming that *loci* were independent, as suggested by the microsatellite linkage map of the domestic dog (Neff *et al*., 1999).

PID values did not result to be low enough to discriminate among individuals (table 3.15), in fact the 9 RealTime SNP primer sets produced, respectively, a PID_{cor} of $1{,}006 \times 10^{-3}$ and a PID_{sibs} of $2{,}89 \times 10^{-2}$ (meaning that 2,8 wolves in 100 siblings are expected to share, by chance, an identical genotype with another wolf), values not comparable with PID values obtained from genotyping the same individuals by the 6 microsatellite *loci*.

For this reason the probability of identity (PID), the probability of identity corrected for small population size (PID_{cor}) and the expected PID among full sib dyads (PID_{sibs}) values were re-estimated in the same set of 30 Italian wolves, previously genotyped, using the same data related to the 9 RealTime SNP primer sets to which data related to the 2 microsatellites, usually used for the quality screening test (FH2096 and FH2137) were added (Table 3.16).

This time both PID_{cor} and PID_{sibs} values significantly reduced, in fact already adding only *locus* FH2096 to the 9 SNPs, PID_{cor} decreased from $1{,}006 \times 10^{-3}$ to $1{,}90 \times 10^{-4}$ and PID_{sibs} from $2{,}89 \times 10^{-2}$ to $1{,}35 \times 10^{-2}$. When *locus* FH2137 too was added, PID values resulted to be low enough to discriminate not only between individuals, but also among partially related or full sib dyads, in fact PID_{cor} decreased to $1{,}29 \times 10^{-5}$ and PID_{sibs} to $4{,}97 \times 10^{-3}$ (Table 3.17, Fig.3.13), values even lower than the ones obtained from genotyping individuals with 6 microsatellite *loci*.

Considered *loci*	PID	PIDcor	PIDsibs
Only 9 SNPs	8,10E-04	1,01E-03	2,89E-02
9 SNPs+FH2096	1,50E-04	1,90E-04	1,35E-02
9 SNPs+FH2096+FH2137	9,70E-06	1,29E-05	4,97E-03

Table 3.17: Pid (probability of identity for individuals randomly chosen within the same population), $\mathbf{Pid_{cor}}$ (probability of identity corrected for small population size) and $\mathbf{Pid_{sibs}}$ (probability of identity corrected for siblings) estimated in a set of 30 Italian wolves using Gimlet v. 1.3.2. (Valière, 2002), considering 9 SNP *loci* and 2 microsatellite *loci*. Each total probability was computed by multiplying single *locus* probabilities, assuming that *loci* were independent.

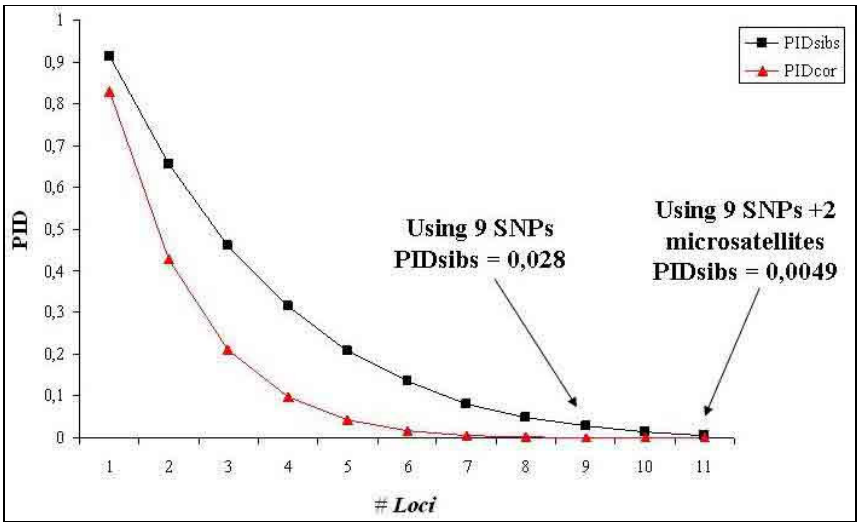

Fig. 3.13: PID_{cor} and PID_{sibs} values estimated using 9 SNPs and 2 microsatellite *loci*. Both values significantly reduced at the increasing of the number of applied *loci*. Using only the 9 RealTime SNPs, PID values were not low enough to discriminate among individuals, but adding the 2 microsatellite *loci* used for the quality screening test, PID values considerably decreased allowing a reliable discrimination also among related individuals.

CHAPTER FOURTH: DISCUSSION

As the Italian wolf population represents one of the few surviving populations in southern Europe after the past persecutions, it symbolizes one of the most important Italian conservation and management priorities. Italian wolf population had a continuous distribution from Alps to Sicily until the beginning of the twentieth century, but persecution, deforestation and a decrease of its natural preys reduced it so much that wolves disappeared from the Alps in the 1920s and continue to drastically decline until the seventies when it approximately consisted of only about 100 individuals surviving isolated in small fragmented areas in central southern Apennines (Zimen & Boitani, 1975; Delibes,1990).

Towards the eighties, Italian wolf population naturally increased and expanded along the Apennine ridge with a partial recolonization of its historical range (Boitani, 1992) owing to a more effective legal protection and to substantial changes in the ecology of mountain areas (decrease of human density and increase of wild ungulates). Wolves crossed the north-western Apennines and reached the south-western Alps in 1992 (Breitenmoser, 1998; Corsi *et al*., 1999; Poulle *et al*., 1999) and reappeared again in the central Italian Alps in 2000 (Lucchini *et al*., 2002; Fabbri *et al.,* 2007).

Nowadays the Italian wolf population consists of more than 600-800 individuals (Boitani 2003), but its quick natural re-expansion after the past decline and its return in areas from which it was eradicated caused some conservation problems for the species which is still considered as a nuisance in many areas of the peninsula (Boitani & Ciucci, 1993; Duchamp *et al*., 2004). Wolves, in fact, often establish packs near urban areas, where dumps offer them easy food source raising social conflicts with breeders, because of depredation and damage on livestock, and with hunters, because of the competition for wild ungulates.

Moreover, despite a substantial demographic recovery, wolves are still largely outnumbered by feral or free-ranging domestic dogs, which are estimated to be more than 1 million and widespread, particularly in the central-southern Italian Apennine (Genovesi & Dupré, 2000). As a consequence of such striking disparity in population size, risk of recurrent and extensive introgressive hybridizations might seriously rise, threatening the genetic integrity of wolf's gene pool (Boitani, 2003; Verardi *et al*., 2006; Randi, 2007).

All these management problems related to the recent reexpansion and colonization events of the Italian wolves determined the need to ensure their conservation and their coexistence with people in Italy through a continuous and careful monitoring of their biology, presence and distribution (Boitani, 2000; 2003).

As wolves are shy and elusive predators, with a great dispersal ability and adaptable to every kind of environmental conditions (Mech, 1970), it is very difficult to study them using only classical field research methods such as diet analysis (Guberti *et al*., 1993; Ciucci *et al*., 1996) snow-tracking (Ciucci & Boitani,1999a;b;c), wolf-howling and radio-tracking (Ciucci *et al.*, 1997). For this reason noninvasive genetics, through the analyses of DNA extracted from biological traces left by individuals and then collected without having (even) to observe, disturb or capture animals (Kohn & Wayne, 1997) seems to be particularly suitable to study the biology and the genetic diversity of an elusive, rare and/or endangered species such as the wolf, avoiding any risks to impact on its survival, its recapture rates or its population dynamics (Kohn & Wayne, 1997; Piggott and Taylor, 2003).

Many wolf conservation genetic studies using non-invasive genetic sampling have been recently carried out allowing to characterize the genetic identity of individuals and their molecular sexing, and thus to provide abundant information on population parameters, presence, distribution, colonization events, conservation and management strategies of the species in the different study areas (Lucchini *et al*., 2002; Valière *et al.,* 2003; Fabbri *et al*., 2007). Moreover, in some studies non-invasively detected *multilocus* data were also used to estimate wolf population size (Creel *et al.*, 2003) in fact, if individuals are sufficiently sampled to estimate re-sighting probabilities (Otis *et*

al., 1978; Seber, 1982), noninvasive genotypes can be used as capture-recapture data and allow population size estimations also for populations whose individuals are difficult to locate like wolves.

4.1 WOLF PRESENCE AND DISTRIBUTION IN EMILIA-ROMAGNA

One of the goals of this study was to monitor the presence and the distribution of the Italian wolf population living in the Apennine ridge of Emilia Romagna Region through the analysis of DNA extracted from non-invasively collected presumed wolf scat samples. Emilia Romagna Region represents, in fact, a very important study area to explain the ongoing expansion process of the Italian wolf population because it acts as a natural narrow ecological corridor along the ridge of the north-western Apennines linking the central-northern Apennine Mountains with the western Alps.

As the study area is a very vast one, the achievement of this large scale wolf monitoring required a great technical and logistic organization and a wide financial effort.

Emilia-Romagna Region strongly supported this project in fact all its provincial administrations (Bologna, Forlì-Cesena, Modena, Parma, Piacenza, Ravenna and Reggio-Emilia) and its national (Foreste Casentinesi and Alto Appennino Tosco-Emiliano) and regional parks (Corno alle Scale, Frignano, Gigante, Cento Laghi) took part in the project ensuring not only an appropriate sampling activity but also an useful professional collaboration, sharing all their field experience and knowledge, during the whole project period.

The genetic analyses of all the noninvasively collected presumed wolf scat samples, necessary for this monitoring project, were performed at the Genetic Laboratory of the Italian Wildlife Institute (INFS) using an already existing noninvasive genetic multiple-tube protocol, based on microsatellite *loci* genotyping and molecular sexing (Lucchini *et al*., 2002), followed by a statistical reliability analysis (Miller *et al*. 2002). In the current study, the efficiency of this protocol was authenticated because the microsatellite *loci* utilized allowed individual identification and population assignment, thus confirming previous comparable analyses (Lucchini *et al*., 2002; Fabbri *et al.*, 2007). Moreover, during this study this protocol was improved and optimized making analyses faster and adding further reliability check up procedures that allowed to discover and remove a certain number of false or "ghost" genotypes due to allelic dropouts and false alleles.

A preliminary microsatellite quality screening allowed to discard about 40% of DNA samples that were too degraded for molecular sexing and to complete microsatellite genotyping, avoiding unnecessary money and time wasting and possible error sources.

The six microsatellite *loci* used for the individual discrimination performed well enough to successfully genotype 1293 samples (37% of total analyzed samples) showing that the presumed "age" of samples and sampling "season" did not significantly affect the performances of DNA analyses in fact samples collected in summer performed generally as well as samples collected in winter (Table3.2). This values could appear low if compared to the ones obtained by Lucchini *et al* (2002) in a similar wolf noninvasive monitoring project in the western Alps in which they obtained a final genotyping rate of 51%. But they analyzed only 130 scat samples and as sampling activities were performed in a much smaller area, they could focus on collecting fresh scats in winter along wolf snow tracks.

The six microsatellite *loci* used for the individual discrimination had a great success also in the belonging population assignment of the individuals, allowing to well distinguish wolf, dog and also wolf-dog hybrid genotypes in fact only in a few cases samples really needed the adding of other further 6 microsatellite *loci* to clarify their assignment. Thus, among the 1293 scat samples which were successfully genotyped it was possible, using Structure v. 2.1 (Pritchard *et al.,* 2000) to identify 269 distinct wolf genotypes, which presumably correspond to at least 269 different wolf individuals, 3 hybrids and 75 domestic dogs.

An estimation of total and per *locus* positive PCR, dropout and false allele rates in all the genotyped samples (Table3.3), using the software Gimlet v. 1.3.2 (Valière, 2002), showed that all the 6 microsatellite *loci*, used for the individual identification and population assignment, presented very high amplification success, with a mean success value of 87%. Total allelic dropout and false allele rates across PCRs resulted to be, respectively, of 21,1% and 3,4 %, values comparable with the ones obtained in other wolf noninvasive genetic studies (Lucchini *et al*., 2002; Fabbri et *al.*, 2007). A large per *locus* variance in dropout rates was observed indicating that some *loci* were amplified less efficiently than others, probably because of the length of their amplified DNA sequence and the high molecular weight of their alleles, using excremental DNA samples (Lucchini *et al*., 2002).

Allelic dropouts produced 31 false genotypes (10% of total genotypes) that were discovered using the multiple-tube protocol of Lucchini *et al*. (2002) reinforced with a further reliability check up and mismatch analyses for all the noninvasive genotyped samples (Fig. 3.4).

Inefficient *loci* could be replaced, in future studies, with other *loci*, or even with new molecular markers such as SNPs which should be equally informative, but less prone to allelic dropouts.

Recapture rates among the different detected individuals were very variable: a few individuals (25%) were collected several times, some genotypes (36%) were sampled from twice to five times and the remaining ones (39%) only once (Fig. 3.5). Genotypes sampled only once were mainly from scats collected in winter along wolf snow tracks. Following a pack travel route in the snow it is possible to collect scats of each individual aside from the individual marking behaviour. Instead, scats collected in summer along human trails/roads, as wolf marking behaviour affects the patterns of defecation on trails (Peters & Mech, 1975; Vilà *et al*., 1994; Kohn *et al*., 1999), were likely samples from dominant individuals that frequently mark the territory. Other genotypes sampled only once might be juveniles that disperse to look for new territories where they can found new packs.

Moreover, considering the time interval during which genotypes were sampled several times, most of them (67%) were observed for a period shorter than one year, while a few individuals (33%) were observed during a period longer than one year (Fig. 3.6), suggesting that only a few individuals, among the sampled ones, can be considered stable in the study area while the others could be considered a portion of the population that does not influence the effective population size.

All the genotyped samples were also successfully sexed through a molecular sexing method proposed by Lucchini *et al.* (2002). Samples analyzed allowed to detect 153 males and 119 females with a sex ratio (M/F) among detected individuals greater than one (1,29M:1,00F).

Sex ratios among detected individuals during all the study period years remained almost constant during the whole study period with values ranging from 1,00 to 1,43 (Table3.3) suggesting that males were often more abundant than females.

However these data should not be used to estimate the real sex ratio of Emilia-Romagna wolf population, because the sampling probabilities of the different genotypes were very variable.

Analysing individuals sampled for a period greater than 1 year, 45 males and 44 females were detected suggesting that sex ratio M/F among more stable animals on the study territory could probably be 1,00M:1,00F confirming the trend according to which there is not a skewed ratio toward males in wolf populations (Mech, 1970; 1975). This was already showed in other Italian wolf population noninvasive studies by Lucchini *et al*. (2002) and Fabbri *et al*. (2007).

During the whole study period the number of collected and analyzed samples considerably increased, allowing to obtain more and more useful biological information about the wolf presence and distribution in the study area (Fig. 3.7; Table 3.4). During the first 2 project years, 2000 and 2001, 28 and 120 presumed wolf scat samples, respectively, were analyzed allowing to detect, respectively, 4 and17 different wolf genotypes.

Anyway the following 5 project years (2002, 2003, 2004, 2005 and 2006) revealed to be the most productive ones both for the number of analyzed sample and for the monitored area size.

During these years, in fact, 2918 samples were analyzed, allowing to detect, respectively, 63, 74, 85, 92, and 114 different individual genotypes showing a mean annual individual increasing of 9.75 individuals, corresponding to 3,6 % of the total detected genotypes.

These data can give only an idea of the number of individuals frequenting the study territory during the project years, and thus they cannot be confused with census data, because the estimation of population size needs accurate marc-recapture analyses to be guessed.

A microsatellite variability estimation carried out on the detected unique genotypes (Table 3.5; Table 3.6) showed that all the 6 used microsatellite *loci* were polymorphic in the Emilia-Romagna wolf population presenting from 3 to 12 alleles per *locus*, and even a total of 7 alleles never found in the rest of the Italian wolf population (private alleles). Even if heterozygosity values ($\mathbf{H_O}$ = 0,67; $\mathbf{H_E}$ = 0,68) resulted to be high, in the comparison between the genetic variability of the Emilia-Romagna wolf population and of the Alpine, Central and Southern Apennine wolf populations, the Analysis of Molecular Variance (AMOVA) indicated that genetic diversity was significantly partitioned among the four wolf groups suggesting that the ongoing population expansion process is sustained by limited gene flow, and that formerly isolated populations have not completely admixed yet (see Fabbri *et al.,* 2007).

The studied wolf population resulted to be not completely in Hardy-Weinberg equilibrium due to deficit of heterozygotes (the probability to obtain by chance a value of $\mathbf{F_{IS}}$ greater than the observed was $\mathbf{P}$ = 0.010), likely because of inbreeding in local patches and presence of geographical substructuring along the Apennines (see also Randi & Lucchini, 2002).

4.2 ANALYZED AND INDIVIDUAL GENOTYPE MAPPING

Wolf noninvasive genetic monitoring project contributed to develop an efficient data collection strategy and a standardized procedure to coordinate all the wolf monitoring activities. This professional reorganization allowed to collect a wide series of data about the wolf presence and its distribution in Emilia-Romagna and to build an useful common Regional dataset containing all the information about each collected sample (both field and genetic ones) and necessary to preserve and share all the produced results.

All that allowed also to create a specific Emilia-Romagna digital cartography about all the collected wolf data. During sample collection in fact, field collaborators compiled, for each sample, a technical card containing useful field and biological information such as sample quality, sampling localities and the relative geographical coordinates, necessary to map their spatiotemporal locations.

Mapping by the software **ArcView GIS** (ESRI) the exact localities of the detected *multilocus* genotypes collected several times it was possible to have an idea of the areas with high wolf density and where some individuals were stable in the time. Integrating then the data so obtained with the biological and wolf-howling information it was possible also to carry out some preliminary hypotheses about the different probable packs living in the study area. In this way 22 different probable wolf packs were identified in the study area, localized along the whole Emilia-Romagna Apennine Ridge (Fig. 3.8).

Monitoring the wolf presence and the distribution of the 22 probable wolf packs it was possible to observe that wolves seems to prefer the most meridional areas of the region that present high altitudes and correspond to the most natural territories, populated by wild ungulate preys with the lowest human density and disturbance. This confirmed previous results by Corsi *et al*. (1999).

In some zones within Bologna and Forlì-Cesena provincial administrations, though, wolf scat samples were collected also at lower altitudes, very close to urban areas where dumps could represent easier food sources for the wolves.

Through the regional scale sampling, the mapped *multilocus* genotypes were also used to follow their spatiotemporal moving detecting 17 presumed dispersal events (Table 3.5). The mean

air line distance of these movings was of about 68 kilometres (corresponding to longer real distances on the territory).

Fabbri *et al.* (2007) studies on the recolonization genetics of the Italian Alps demonstrated that wolf migration were unidirectional from the Apennines to the Alps and male-biased. Data obtained in this study seem to confirm this trend in fact among the 17 presumed dispersal events, 13 occurred in direction Southeastern-Northwestern and 14 resulted to be male-biased (Table 3.7).

During the whole project period 3 certain wolf x dog hybrids (2 males and 1 female) were also genetically detected (Fig. 3.9). Their genetic structure suggested they were not F1 hybrids but probably F2 hybrids, originated through the backcross of an F1 wolf-dog hybrid with another wolf. The 2 males were both sampled only once while the female was noninvasively sampled 9 times, from December 2002 to June 2006, between Bologna and Firenze provincial administrations, and one time through the finding of its corresponding carcass on December 2006, in the same area. The carcass was also submitted to an accurate veterinary analysis that showed it was illegally killed, that it did not present detectable morphological signals of hybridization but above all that it never mated.

The 3 detected hybrids represented just the 1,1% compared with the 269 detected wolf genotypes, and were sporadically collected, always far from the territories of the probable wolf packs. The only one collected several times, over a period of 5 years never mated, suggesting a clear signal of how the hybrids could meet some behavioural and ecological reproductive barriers.

All these consideration let us believe that also in Emilia-Romagna hybridization is an uncommon process, strictly directional and that Emilia-Romagna wolf population do not show substantial dog gene introgression, confirming what showed by many other genetic studies about hybridization between wild wolves and free-ranging domestic dogs in Italy (Randi *et al.*, 2000; Randi & Lucchini, 2002; Lucchini *et al.*, 2002; Lucchini *et al.*, 2004; Verardi *et al.*, 2006, Fabbri *et al.*, 2007; Randi, 2007).

During the project, 75 different domestics dog individuals were also detected (Fig. 3.10). The data suggested that in the study area no feral or free-ranging domestic dogs stably live in the same territories frequented by the wolves interacting with them in fact most of them were collected once while only 3 of them were collected twice but in areas very close to urban zones.

Dog samples were never found near depredated wild ungulates or livestock carcasses suggesting that in the study area feral or free-ranging dogs cannot be considered as wolf food competitors.

4.3 POPULATION SIZE ESTIMATION

Multilocus genotypes sampled during the whole study period cannot be considered as census data because they don't consider neither the different sample frequencies of each animal nor the monitoring heterogeneity among the different zones within the study area. For this reason another crucial goal of this study was to use genetic data as capture-mark-recapture ones to obtain a reliable population size estimation necessary for Emilia-Romagna wolf conservation and management.

Multilocus genotypes obtained analysing noninvasive samples can be used to estimate population size in several ways (Sloane *et al.*, 2000). In recent years, their use for capture-recapture studies rapidly increased and the method was already applied to a diverse array of taxa to assess population size (Lukacs, 2005; Lukacs & Burnham, 2005). In natural populations the estimation of biological parameters such as survival, migration rates, movement or transition rates, fecundity, and population growth can be problematic because of variation in capture probability (behavioural responses to capture, variation over time with constant trapability for all individuals) and individual heterogeneity in capture probability (the variation among individuals in their probability of being detected) (Otis *et al.*, 1978). Anyway, a large number of models and software exists for a wide range of capture-recapture analyses.

Due to the open nature and the long time span of this wolf population project, these genetic data were analysed using the open population multistate and multievent models (Lebreton & Pradel, 2002; Pradel, 2005) incorporating mixture heterogeneity (Pledger *et al*., 2003) to detect the main biological parameters necessary to obtain a reliable population size estimation.

Capture-recapture analyses were carried our considering all the 272 Emilia Romagna *multilocus* genotypes detected during the whole study period as capture-recapture data, in fact the first detection of a genotype was considered as marking, while further detections were considered as recaptures.

A preliminary analysis, using the program **U-Care** (Choquet *et al*., 2005), to explore the fit of a general model (the Cormack-Jolly-Seber model) in which both capture **(*P*)** and survival **(Φ)** probabilities dependent on time, to the Emilia-Romagna *multilocus* genotypes, showed that Emilia-Romagna wolf population data presented a clear signal of permanent heterogeneity of capture among individuals suggesting that more complex models were needed.

A further analysis, performed by the software E-Surge v. 1.1.1 (Choquet *et al.*, 2007), allowed to build 18 different multievent models with heterogeneity of capture and characterized by 2 events ("seen"; "not seen") and 3 states (belongs to class with low capture probabilities, belongs to class with high capture probabilities, Dead) (Table 3.10).

Among them the best model resulted to be the one built considering a constant survival probability among individuals and 2 classes (low and high) of recapture probability both depending on the additional effects of heterogeneity and season.

This model permitted to resolve the problem related to the heterogeneity of capture among individuals hypothesizing, both for individuals with low capture probability and for individuals with high capture probability, the same capture probability for the same seasons during the different project years.

It allowed to obtain the total capture probability and the estimated detection probability (number of counted individuals at each capture occasion), parameters necessary for population size estimation.

Emilia-Romagna wolf population size estimation showed a clear seasonal trend during the whole study period, suggesting a true and significant increase over that period.

Emilia-Romagna wolf population size was characterized by a total mean value of 123,95 (95%CI 80,83-174,01) individuals, with fluctuations through each year, ranging from a minimum of 4,75 (95%CI 0,0-15,38), during autumn 2000, to a maximum of 250,57 (95%CI 181,91-322,68), during winter 2006 (Table 3.11; Fig. 4.1).

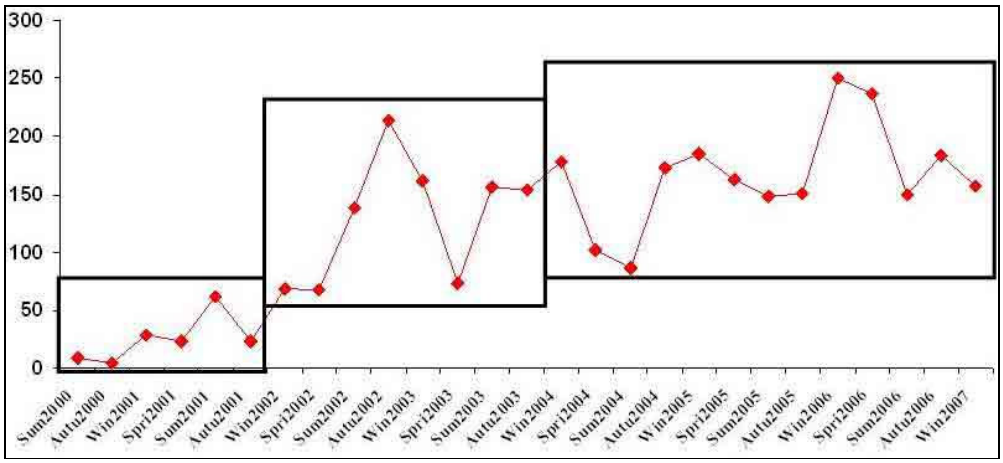

Fig. 4.1: seasonal pattern of Emilia-Romagna wolf population size estimated through the best heterogeneous model. It shows the presence of 3 main patterns of fluctuations during the whole study period.

During the first 2 project years Emilia-Romagna wolf population presented the largest sizes during summer periods, and the smallest ones during autumn periods. This could be due to the fact that during summer, after pups births the greater annual incremental to wolf population generally occurs, while during autumn and winter pup and adult mortality typically peaks reducing overall population size (Mech, 1970; 1973; 1982b). During the third and fourth project years population size presented anomalous trends in fact the year 2002 was characterized by the smallest value in winter and spring and the largest values in summer and autumn, while the year 2003 was characterized by the smaller value in spring and similar values in summer, autumn and winter.

During all the other project years population size followed a regular trend characterized by the largest values during winter and the smallest ones during summer periods. This could be due to the fact that during summer, according to what Mech (1973; 1982b) describes about observed packs, members are more often together during winter and more often travel alone during summer.

Comparing the population size estimation *per* year with the number of detected genotypes *per* year it was possible to observe that the mean number of individuals estimated *per* year was about two times wider than the number of detected genotypes (Table 4.1).

Year	Detected genotypes	Population size	Population size / Number of genotypes
2000	4	6	1,50
2001	17	34	2,01
2002	63	122	1,94
2003	74	136	1,84
2004	85	134	1,58
2005	92	161	1,75
2006	114	205	1,80
2007	67	157	2,34
Mean	64,5	119	1,85

Table 4.1: values per year of detected genotypes, population size mean and ratio between population size and number of detected genotypes. Population size mean values per year are about two times wider than the correspondent detected genotypes per year.

It could be due to the fact that most of samples were collected near mark sites where it is more probable to find dominant individuals that mark with high frequency the territory during all seasons than juveniles that don't usually mark.

On the contrary, estimation of population size was carried out using a multievent model able to consider not only the individuals with a high capture probability but also the individuals characterized by a lower capture probability.

The efficiency of this model was tested by comparing it with a homogeneous one characterized by a homogeneous capture probability among individuals (Fig 4.2).

The homogenous model showed confidence intervals lower about 1,5 times minor than the heterogeneous one, but it produced also an underestimation of the population size that resulted to be about 1,5 times minor than the one obtained by the heterogeneous model (Table 3.12). For this reason Emilia-Romagna wolf population size estimation, obtained using the best heterogeneous model, resulted to be the most reliable one.

The method used in this study for Emilia-Romagna wolf population size estimation seemed to be more reliable than the method based on accumulation or rarefaction curves used by Creel *et al.* (2003) to estimate population size in Yellowstone wolves.

Capture-recapture analyses, in fact, present several advantages over accumulation curves. Accumulation curves do not account for the sampling design used to obtain the data (Cam *et al.*, 2002) and are just designed to approximate the appearance of the data not the process that generates the data, whereas capture-recapture models directly estimate detection probability.

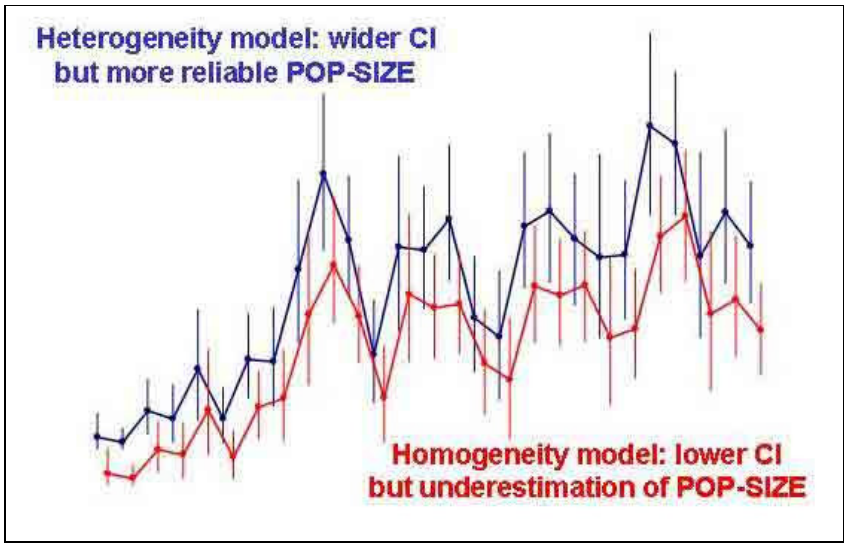

Fig. 4.2: Best heterogeneous model (blue line) compared to a homogeneous one (red line). In both Heterogeneous and Homogeneous models points represent the 27 different capture occasions (the first one is fixed) and vertical bars represent confidence intervals. Heterogeneous model produced wider confidence intervals but allows a more reliable population size estimation.

Accumulation curves do not efficiently use the data collected, they only use the first detection of an individual, whereas capture-recapture methods can use all detections.

Finally, accumulation curves cannot account for variation in detection probability. Detection probability is known to vary widely in many situations across time, space and individuals. Estimation methods, on the contrary, need to be able to account for these differences in order to appropriately estimate abundance.

4.4 SINGLE NUCLEOTIDE POLYMORPHISMS (SNP)

As noninvasive DNA analyses are often prone to genotyping errors (allelic dropouts and false alleles) due to DNA degradation, the last crucial goal of this work was to develop new genotyping methods faster and more reliable than microsatellite *loci*, able to powerfully analyse also low quality and quantity DNA samples like non-invasive ones. In this study the possible application of SNPs in the Emilia-Romagna wolf population noninvasive genetic monitoring was investigated in fact their ability to amplify much shorter DNA fragments could make SNPs of particular use for population monitoring, where faecal and other noninvasive samples are routinely used (Seddon *et al*., 2005).

This study contributed to characterize 59 canine single nucleotide polymorphisms (SNPs) in the Italian wolf population, which were discovered by resequencing sequence-tagged-site (STS) DNA sequences that were known to contain SNPs in domestic dogs (see Holm Andersen & Fabbri *et al.,* 2006). On the bases of this first SNP characterization in the Italian wolf population new primer sets (see Holm Andersen & Fabbri *et al.,* 2006) for analysing SNPs using Pyrosequencing technology (Ronaghi *et al*., 1998) were designed. A comparison of the results obtained by Pyrosequencing technology with sequence data showed the efficiency of this method that permitted to detect the same 59 SNPs previously found in the Italian wolf population trough classical sequencing.

Despite the long procedures for SNP discovery in nonmodel organisms, as SNPs are the most prevalent form of genetic variation, and thus characterized by a substantial increase in the number of *loci* available (Brumfield *et al* . 2003) and by a capacity for rapid, large scale and cost-effective

genotyping (Vignal *et al*., 2002; Chen & Sullivan, 2003), their application seemed to be particularly suitable in ecological and conservation studies based on noninvasive genetic population monitoring.

Because of the large number of samples usually analyzed in the Emilia-Romagna wolf population noninvasive genetic monitoring project, this study also tried to investigate which was the most reliable, faster but also less expensive SNP genotyping method, among the large variety of available ones, that could allow a near future replacing of noninvasive sample microsatellite genotyping. For this reason, in this study 3 different SNP analysis methods, Pyrosequencing, SNaPShot and Real Time PCR, were compared to evaluate which was the best one, from success, reliability, time and cost points of views, to use for future noninvasive applications. Thus, 43 noninvasive DNA samples, of different qualities, were amplified, according to a multiple tube approach (Taberlet *et al*., 1996; Gagneux *et al*. 1997; Lucchini *et al*., 2002), 3 times for the same 6 primer sets suitable for each of the 3 tested technologies (Table 3.14).

Analysing scat samples, Pyrosequencing method resulted to be the worst one. Pyrosequencing genotyping, because of its particular procedures, needed a lot of time and too laborious passages to prepare samples rising the contamination risk among samples. Moreover, it produced the lowest PCR success and the highest allelic dropout rates, with a necessary number of replicates equal or greater than 3 to resolve uncertainties due to the high values of allelic dropouts.

Intermediate results were produced by the SNaPshot genotyping system. Even if it achieved to use already existing equipment in the forensic laboratory and produced the highest PCR success, it showed allelic dropout rates minor than Pyrosequencing but still high, and it also needed a necessary number of replicates equal or greater than 3 to resolve uncertainties due to its high allelic dropout rates.

RealTime PCR method resulted to be the best one because it revealed to be not only the faster among the tested methods but also the most reliable presenting an intermediate PCR success, the lowest allelic dropout rate and the minor number of necessary replicates to obtain reliable results, in fact in most of the cases only one replicate was sufficient to produce clear data and only in a few occasions the method needed an additional replicate to resolve some uncertainties.

As individual identification by *multilocus* genotypes and their assignment to the true belonging populations are central themes for many noninvasive genetic studies, after establishing that RealTime PCR was the most suitable method to use for low content DNA sample analyses, its application reliability and its performance to detect unique genotypes and their corrected assignment to a population were tested.

Thus, 28 wolf scat DNA samples were genotyped using both 9 RealTime SNP *loci*, polymorphic in the Italian wolves and even between wolves and dogs, and the same 6 microsatellite *loci* usually applied for the individual identification and population assignment. SNP genotyping revealed to be much more efficient than microsatellite one, producing a greater final genotyping percentage (Table 3.15). Moreover, SNP genotypes were wanting in false alleles and showed much higher PCR success and much lower allelic dropouts than microsatellites. Using SNPs most of the reliable data were obtained through maximum 2 replicates per *locus*, while using microsatellites almost the 50% of reliable data were obtained after 8 replicates per *locus*. The 2 methods were equally efficient in the assignment of the genotypes to their belonging populations, in fact they clearly allowed to distinguish wolves, dogs and even wolf-dog hybrids. Anyway the 9 used SNPs, even if very variable among the Italian wolf population, producing PID values that were not low enough to perfectly discriminate among the different individuals, revealed to be less efficient than the 6 microsatellites for the individual identification (table 3.15; Fig 4.3).

As SNPs are bi-allelic markers characterized by lower heterozygosity rates and lower information contents than microsatellites, a greater number of SNP *loci* is needed to distinguish individuals, number that depends on the size of the population, the genetic diversity of the population, the polymorphism of the SNP *loci* and the desired probability level.

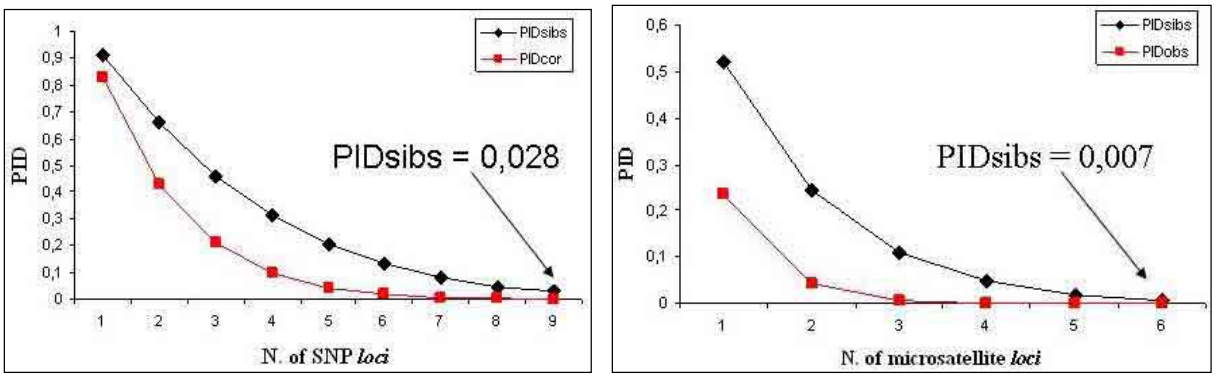

Fig. 4.3: comparison of the PID values obtained genotyping a set of 30 Italian wolves both through 9 SNP *loci* and 6 microsatellite *loci*. PID values obtained using SNPs (on the left) were not low enough to distinguish among different individuals.

This study suggested the possibility to combine microsatellite and SNP genotyping methods to obtain a faster, economic and more reliable genotyping tool, completely efficient not only to distinguish wolves, dogs and even hybrids but also able to perfectly discriminate among the different individuals.

Many combinations of the two types of markers were performed to find the most suitable, and at the end the best solution resulted to be a genotyping method based on the simultaneous use of the 2 microsatellite *loci* applied for the preliminary quality screening test with the addiction of all the 9 RealTime SNP *loci*. In this way it was possible to obtain PID values low enough to discriminate not only between individuals, but also among partially related or full sib dyads (Fig 4.4). Moreover in this way it was possible also to combine the efficiency of a preliminary quality screening test based on microsatellites with the genotyping reliability and rapidity based on the 9 SNPs. In fact the preliminary quality screening test based on the 2 microsatellites should continue to be applied to discard of the low quality-quantity DNA content samples, while the 9 SNPs should continue to be used to complete the genotyping of the good quality samples that passed the previous screening, ensuring high rates of final genotyping, limited false alleles and allelic dropout values (Table 4.2).

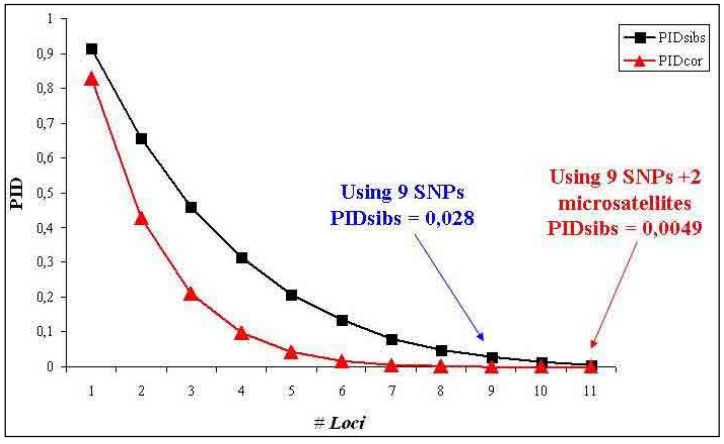

Fig. 4.4: PID_{cor} and PID_{sibs} values estimated using 9 SNPs with the addiction of the 2 microsatellite *loci* used for the preliminary quality screening test. In this way PID values considerably decreased allowing a reliable discrimination not only between individuals, but also among partially related or full sib dyads.

Considered *loci*	PID	PIDcor	PIDsibs	Positive PCRs	ADO	FA
9 SNPs	8,10E-04	1,01E-03	2,89E-02	86%	1,09%	0
FH2096	1,84E-01	1,89E-01	4,66E-01	58%	3,28%	1,6%
FH2137	6,52E-02	6,82E-02	3,69E-01	59%	14,40%	0
Total	**9,70E-06**	**1,29E-05**	**4,97E-03**	**68%**	**6,26%**	**0,53%**

Table 4.2: values of PID, PIDcor, PIDsibs, positive PCRs, allelic dropouts and false alleles considering separately 9 SNPs, *locus* 2096 and *locus* 2137, and corresponding total values.

CONCLUSIONS

This study showed that to use of non invasive genetic sampling (NGS) represents a powerful tool to study endangered species when its data are efficiently supported by additional ecological and field information, confirming that noninvasive genetic sampling methods can provide several issues that could not be addressed in any other way.

In this study the screening of a limited number of genetic markers, such as microsatellite *loci*, produced information reliably useful to monitor the presence and distribution of wolves living along the Emilia-Romagna Apennine Ridge, identify species, eventual hybridization processes, origin of local populations and individuals, estimate the number of reproducing individuals, dispersal events, mapping pack localizations and carry out some preliminary hypothesis on the fine structure of wolf packs and their dynamics.

Moreover the intensive sampling collection, the open nature and the long time span of this wolf population project allowed to use noninvasive genetic data as capture-recapture data highlighting that they can provide reliable population size estimations (avoiding the need to genotype every individual from faeces) when supported by appropriate statistical techniques and complex mathematical models built considering not only the noninvasive data characteristics and sampling type but also the studied species biology and its environmental context.

For the future it would be suitable a more intensive and homogeneous sample collection among the whole study area to significantly reduce the heterogeneity of capture among individuals, but randomized across the entire wolf pack range and concentrated in short period (a few months), before and after the reproductive periods, to better interpret the fluctuations of the studied population size through the time.

Anyway this study confirmed also that microsatellite genotypes from non-invasive samples can be affected by errors such as allelic dropout and false alleles due to their low quality-quantity DNA content. For this reason NGS studies can be considered significantly reliable, only applying rigorous error-checking protocols able to minimize error rates and laboratory efforts.

SNPs genotyping could represent a near future application in non-invasive genetics as a promising and innovative faster and more reliable method to analyse low quality and quantity DNA samples like non-invasive ones. In fact, if it is available a sufficient number of SNPs, polymorphic both in the Italian wolves and even between wolves and dogs, able to ensure a reliable individual identification and a certain population assignment, using appropriate laboratory protocols and efficient techniques, SNP *loci* could completely replace microsatellite l*oci* in noninvasive genetic studies.

ACKNOWLEDGEMENTS

I greatly thank everybody who allowed me the achievement of this PhD Thesis in a pleasant way: my Professor and Supervisor Ettore Randi who made possible the achievement of one of my career aims: the PhD; Alberto Santini who taught me most of my knowledge about Non-invasive Genetics;

Nadia Micci and Elena Fabbri for their many suggestions helpful to resolve any kind of technical laboratory problems and to organize this thesis; Andrea Grill for the supervision of the English style of my thesis; Prof Barbara Mantovani for her final supervision and suggestions.

I greatly thank also:

- Federica who deeply changed my life giving it the most beautiful sense by her love, personality and sweetness;
- My parents and my sister who strongly supported me and always helped me to face each difficulties;
- All my friends Francesca, Claudia, Alessandra, Chiara, Cecilia, Severino, Adriano, Vittorio, Max, Cristiano, Alic and Rita who contributed to make pleasant the laboratory experience; and free time moments;
- Jean-Dominique Lebreton who gave me the chance to spend a period at Montpellier allowing me to analyze my data with the hope to produce a great PhD thesis, who considered me as a son during my whole French period;
- Olivier Gimenez who taught me everything I know about the population size estimation and the heterogeneity models, revealing an excellent and very patient teacher and a good friend.

Many thanks also to everybody who took care of the sample and field information collection: Ministero dell'Ambiente e della Tutela del Territorio (Direzione Protezione della Natura), Rome; Italian Forestry Corp, the Regione Emilia-Romagna (in particular M. L. Bruschini, E. Valbonesi, M. L. Zanni), the directors, personnel and CTS of the Parco Nazionale delle Foreste Casentinesi, Monte Falterona e Campigna (N. Agostini, M. D'Amico, M. Mencucci, Mauro Fabbri and Claudio Monti), Parco Nazionale dell'Alto Appennino Tosco-Emiliano (W. Reggioni and Francesca Moretti), Regional Park of Corno alle Scale (Stefano Sozzi), Regional Park of Frignano (Riccardo Fontana), Regional Park of Gigante (Willy Regioni) and Regional Park of Cento Laghi, police technical staff and the personnel of provincial administrations of Bologna (Dari Marelli, Gianpaolo Paladini and Lorenzo Rigacci), Forlì-Cesena (Remo Valmori), Modena (Riccardo Fontana), Parma (Mario Andreani), Piacenza (Enrico Merli), Ravenna (Maurizio Pascucci) and Reggio-Emilia (Willy Regioni).

REFERENCES

Aasen, E. & Medrano J.F., 1990. Amplification of the *ZFY* and *ZFX* genes for sex identification in humans, cattle, sheep and goat. Bio/Technology **8**: 1279-1281.

Abecasis, G.R., Cherny, S.S., Cookson, W.O.& Cardon, L.R., 2002. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees, Nat. Genet. **30:** 97-101.

Altobello, G., 1921. Fauna dell'Abruzzo e del Molise. Vol **IV**:1-6. Campobasso.

Amos, W., Whitehead, H., Ferrari, M.J., Glockner-Ferrari, D.A., Payne, R. & Gordon, J., 1992. Restrictable DNA from sloughed cetacean skin: its potential for use in population analysis. Mar. Mamm. Sci. **8:** 275-283.

Andreàsson, H., Asp, A., Alderborn, A., Gyllensten, U. &Allen, M., 2002. Mitochondrial sequence analysis for forensic identification using pyrosequencing technology, Biotechniques **32:** 124-133.

Aoki, T. & Wada, M., 1951. Functional activity of the sweat glands in the hairy skin of the dog. Science **114:** 123-124.

ArcView® GIS 3.2, a desktop mapping program produced by ESRI (Environmental Systems Research Institute, Inc.).

Avise, J.C., 1994. Molecular Markers, Natural History and Evolution. Nature, **352:** 192.

Ballard W.B., Farnell R. & Stephenson R.O., 1983. Long distance movement by grey wolf, (*Canis lupus)*. Can. Field-Nat. **97:** 333.

Ballard W.B., Whitman, J.S. & Gardner, C.L., 1987. Ecology of an exploited wolf population in south-central Alaska. Wildlife Monograph , no. 98. The Wildlife Society, Bethesda, MD. 54 pp.

Bayes, M., Smith, K., Alberts, S., Altmann, J. & Bruford, M., 2000. Testing the reliability of microsatellite typing from faecal DNA in the savannah baboon. Conservation Genetics **1:** 172-175.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F., 2001. GENETIX 4.02: Logiciel sous Windows TM pour la génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France. Available at www.univ-montp2.fr/-genetix/genetix.htm.

Berg, L.M., Sanders, R. & Alderborn, A., 2002. Pyrosequencinge technology and the need for versatile solutions in molecular clinical research. Expert. Rev. Mol. Diagn. **2:** 361-369.

Block, M.L., Volpe, L.C. & Hayes, M.J., 1981. Saliva has a chemical cue in the development of social behaviour. Science **211:** 1062-1064.

Boitani, L., 1982. Wolf management in intensively used areas of Italy. In: Harrington, F.H. & Paquet, P.C. (eds.), Wolves of the world, Noyes Publications, New Jersey: 158-172.

Boitani, L. 1983. Wolf and dog competition in Italy. Acta Zoologica Fennica **174:** 259-264.

Boitani, L., 1984. Genetic consideration on wolf conservation in Italy. Boll. Zool. **51(3):** 367-373.

Boitani, L., 1986. Dalla parte del lupo. Milano: L'airone di G. Mondatore Associati Spa.

Boitani, L., 1992. Wolf research and conservation in Italy. Biological Conservation **61:** 125-132.

Boitani, L., 1995. Ecological and cultural diversities in the evolution of wolf-human relationship. In: Carbyn, L. N., Fritts, S. H., Seip, D. R. (eds.), Ecology and conservation of wolves in a changing world, Canadian Circumpolar Institute, Occasional pubblication n. **35**: Edmonton, Canada: 3-11.

Boitani L., 1999. Final Draft Action Plan for the Conservation of Wolves (*Canis lupus*) in Europe. Council of Europe, Strasbourg.

Boitani, L., 2000. Action Plan for the Conservation of Wolves in Europe (*Canis lupus*). Nature and environment, n. 113.

Boitani, L., 2003. Wolf conservation and recovery. In Wolves: 317-340. Mech, L.D. & Boitani, L. (Eds). Chicago and London: University of Chicago Press.

Boitani, L. & Ciucci, P., 1992. Stato delle conoscenze del lupo (*Canis lupus*) in Italia: prospettive di ricerca e conservazione. Atti del Convegno Nazionale dalla Parte del Lupo, pp 15-26.

Boitani, L. & Ciucci, P., 1993. Wolves in Italy*:* critical issues for their conservation. In: Promberger, C. & Shroeder, W. (eds.), Wolves in Europe, Status and perspectives, Oberammergau, Germania, WGM: 75-90.

Boitani, L. & Ciucci, P., 2000. Piano d'Azione per la conservazione del lupo (*Canis lupus*)

nelle Alpi. WWF Italia.

Boitani, L. & Fabbri, M. L 1983. Strategia nazionale di conservazione per il lupo (*Canis lupus*).

Ric. Biol. Selvaggina **72:** 1-31.

Boitani, L. & Zimen, E., 1975. Number and distribution of wolves in Italy. Z. Fur Saugetierkunde **40:** 102-112.

Bossart, J.L. & Prowell, D.P., 1998. Genetic estimates of population structure and gene flow. Limitations, lessons and new directions. Trends Ecol. Evol. **13:** 202-206.

Boulanger, C.A., Wagner, K.U. & Smith, G.H., 2004. Parity-induced mouse mammary cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression. Oncogene doi: 10.1038_sj.onc. 1208185.

Bradley, B.J. & Vigilant, L., 2002. False alleles derived from microbial DNA pose a potential source of error in microsatellite genotyping of DNA from faeces. Molecular Ecology Notes **2:** 602-605.

Breitenmoser U, 1998. Large predators in the Alps: The fall and rise of man's competitors. Biological Conservation **83:** 279-289.

Bricker, J., Bushar, L.M., Reinert, H.K. & Gelbert, L.,1996. Purification of high quality DNA from shed skins. Herpetological Review **27:** 133-134.

Broquet, T., Menard, N. & Petit, E., 2007. Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. Conservation Genetics **8:** 249-260.

Brown, D.S. & Johnston, R.E., 1983. Individual discrimination on the basis of urine in dogs and wolves. Pp 343-346 in Müller-Schwarze, D. & Silverstein, R.M., eds. Chemical signals in vertebrates, 3rd ed. Plenum Press, NY.

Bruford, M.W. & Wayne R.K., 1993. Microsatellite and their application to population genetic studies. Curr. Biol. **3:** 939-943.

Brumfield, R.T., Beerli, P., Nickerson, & D.A., Edwards, S.V., 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology and Evolution **18:** 249-256.

Brzustowicz, L.M., Merette, C., Xie, X. *et al*., 1993. Molecular and statistical approaches to the detection and correction of errors in genotype databases. American Journal of Human Genetics **53:** 1137-1145.

Budowle, B., 2004. SNP typing strategies. Forensic Science International **146:** 139-142.

Budowle, B., Planz, J., Campbell R. &. Eisenberg, A., 2004. SNPs and microarray technology in forensic genetics: development and application to mitochondrial DNA, Forensic Sci. Rev. **16:** 22-36.

Burnham, K.P., 1991. On a unified theory for release-resampling of animal populations. In: Proceedings 1990 Taipei Symposium in Statistics (eds Chao, M.T. & Cheng, P.E.), pp. 11-35. Institute of Statistical Science, Academia Sinica: Taipei, Taiwan.

Burnham, K.P., Anderson, D.R., White, G.C., Brownie, C., & Pollock, K.H., 1987. Design and analysis methods for fish survival experiments based on release-recapture. Am Fish Soc Monogr **5:**1-437

Butler, D., 1994. Bid to protect wolves from genetic pollution. Nature **370:** 497.

Cam, E., Nichols, J.D., Sauer, J.R. & Hines, J.E., 2002. On the estimation of species richness based on the accumulation of previously unrecorded species. Ecography **25:** 102-108.

Chen, X. & Sullivan, P.F., 2003. Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. Pharmacogenomics Journal **3**: 77-96.

Choquet, R., Reboulet, A.M., Pradel, R., Gimenez, O. &. Lebreton, J. D., 2004. M-SURGE: new software specifically designed for multistate recapture models. Animal Biodiversity and Conservation, **27(1):**177-185, 2004.

Choquet, R., Reboulet, A.M., Lebreton, J.D., Gimenez, O. & Pradel, R., 2005. U-CARE 2.2

user's manual (Utilities-Capture-Recapture). Technical report, CEFE, UMR 5175, Montpellier, France, http://ftp.cefe.cnrs.fr/biom/Soft-CR/.

Choquet, R., Rouan, L. & Pradel, R., 2007. "Program E-Surge: a software application for fitting Multievent models." submitted to Environmental and Ecological Statistics

Ciucci, P. & Boitani, L., 1991. Variability assessment of the Italian wolf and guidelines for the management of the wild and a captive population. Ricerche Biologia della Selvaggina 89: 58 pp.

Ciucci, P., Boitani, L., 1998a. Il lupo, elementi di biologia, gestione, ricerca. Istituto Nazionale per la Fauna Selvatica "Alessandro Ghigi", Documenti tecnici **23:** 114 pp.

Ciucci, P., Boitani, L., 1998b. Wolf and dog depredation on livestock in central Italy. Wildlife Society Bullettin **26:** 504-514.

Ciucci, P. & Boitani, L.,1999a. Nine-year dynamics of a wolf pack in the Northern Apennines, Italy. Mammalia **63:** 377-384.

Ciucci, P., Boitani, L.,. 1999b. Monitoraggio del lupo su neve all'interno dei parchi nazionali: una necessità, un'opportunità e una proposta operativa. Biologia e Conservazione della Fauna xx:xxx-xxx. In stampa.

Ciucci, P. & Boitani, L., 1999c. Distribuzione dei segni di marcatura del lupo su neve nell'Appennino Toscoemiliano. In Programma e Riassunti del IV Congresso dei Biologi della Selvaggina, 28-30 ottobre 1999, Bologna: 91.

Ciucci, P., Boitani, L., Francisci, F. & Andreoli, G., 1997. Home range, activity and movements of a wolf pack in central Italy. Journal of Zoology (London) **243:** 803–819.

Ciucci, P., L. Boitani, L., Pellicccioni, E.R., Rocco, M. & Guy, I., 1996. A comparison of scat-analysis methods to assess the diet of the wolf *(Canis lupus)*. Wildlife Biology, **2 (1):** 37-48.

Ciucci, P., Raganella E & Boitani, L., 1997b. Impatto del lupo sul patrimonio zootecnico:

entità dei danni e quadro legislativo. Rapporto Ministero Ambiente, 36 pp

Constable J.L., Ashley M.V., Goodall J., Pusey A. 2001. Noninvasive paternity assignment in Gombe chimpanzees. Mol.Ecol. **10:**1279-1300.

Cooper D.N.& Krawczak M., 1989. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes, Hum. Genet. **83:** 181-188.

Corsi, F., Dupré, E. & L. Boitani, L., 1999. A large-scale model of Wolf distribution in Italy for

conservation planing. Conservation Biology **13:** 150-159.

Coscia, E.M., 1995. Ontogeny of timber wolf vocalizations: acoustic properties and behavioural contexts. PhD dissertation, Dalhousie University, Halifax, NS.

Coscia, E.M., Phillips, D.P. & Fentress, J.C., 1991. Spectral analysis of neonatal wolf (*Canis lupus*) vocalizations. Bioacoustics **3:** 275-293.

Cozza, K., Fico, R., Battistini, M.L.. & Rogers, E., 1996. The damage-conservation interface illustrated by predation on domestic livestock in central Italy. Biological Conservation **78 (3):** 329-336.

Creel, S., Spong, G., Sands, J.L., Rotella, J., Zeigle, J., Joe, L., Murphy, K.M. & Smith, D., 2003. Population size estimation in Yellowstone wolves with error-prone non-invasive microsatellite genotypes. Molecular Ecology **12:** 2003-2009.

Dalén, L., Götherström, A. & Angerbjörn, A., 2004. Identifying species from pieces of faeces.

Conservation Genetics **5:** 109-111.

Delibes, M., 1990. Status and conservation needs of the wolf in the Council of Europe member

States. Nature and Environment Series **47:** 1-46.

Doty, R.L., 1986. Odor-guided behaviour in mammals. Experientia **42:** 257-271.

Duchamp, C., Genevey, V., Favier, F. & Lacour, N., 2004. Le retour de loup dans les Alpes Francaises. Project LIFE99NAT/F/006299. Raport final 1999-2004. Available at www.oncfs.gouv.fr/events/point_faune/ mammifere/2005/loup.php.

Eisenberg; J.F. & Kleiman, D.G., 19972. Olfactory communication in mammals. Annu. Rev. Ecol. Syst. **3:** 1-32.

Ernest, B.H., Penedo, M.C.T., May, B.P., Syvanen, M. & Boyce, W.M., 2000. Molecular tracking of mountain lions in the Yosemite Valley region in California: genetic analysis using microsatellites and faecal DNA. Molecular Ecology **9:** 433-441**.**

Evett, I.W.,& Weir B.S., 1998. Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists. Sinauer, Sunderland, MA.

Ewen, K.R., Bahlo, M., Treloar, S.A. *et al*., 2000. Identification and analysis of error types in high-throughput genotyping. American Journal of Human Genetics **67:** 727-736.

Fabbri, E. *et al*., 2007. From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. Molecular Ecology **16:** 1661-1671.

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.

Farrell, L.E., Romant, J. & Sunquist, M.E., 2000. Dietary separation of sympatric carnivores identified by molecular analysis of scats. Mol. Ecol. **9:** 1583-1590.

Fernando, P., Vidya, T.N.C., Rajapakse, C., Dangolla, A. & Melnick, D.J., 2003. Reliable noninvasive genotyping: fantasy or reality? Journal of Heredity **94:** 115-123.

Fico, R., Morosetti, G. & Giovannini, A., 1993. The impact of predators on livestock in the Abruzzo region of Italy. Rev. sci. tech. Off. int. Epiz., **12 (1):** 39-50.

Flagstad, Ø., Røed, K., Stacy, J.E. & Jakobsen, K.S., 1999. Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. Mol. Ecol. **8:** 879-883.

Forster, V.Th., 1948. Zwischenmolekulare energiewanderung und fluoresenz. Annals of Physics (Leipzig) **2:** 55-75.

Francisci, F. & Guberti, V., 1993. Recent trends of wolves in Italy as apparent from kill figures and specimens. In: Promberger , C. & Shroeder, W. (eds.), Wolves in Europe, Status and perspectives, Oberammergau, Germania, WGM: 91-102.

Francisci, F. & Mattioli, S., 1992. Cause della recente ripresa del lupo in Appennino Settentrionale. Atti del Convegno Nazionale dalla Parte del Lupo pp. 34-91.

Francisco, L.V., Langston, A.A., Mellersh, C.S., Neal, C.L.& Ostrander, E.A.,1996. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mammalian Genome **7:** 359-362.

Frankham, R., Ballou, J.D. & Briscoe, D.A. (eds.), 2002. Introduction to conservation genetics. Cambridge University Press, Cambridge, UK.

Frantz, A.C., Pope, L.C., Carpenter, P.J., Roper, T.J., Wilson, G.J., Delahay, R.J. & Burk, T., 2003. Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Molecular Ecology **12:** 1649-1661.

Fredholm, M. & Wintero, A.K., 1995. Variation of short tandem repeats within and between species belonging to the Canidae family. Mammalian Genome **6**: 11-18.

Fritts, S.H., 1983. Record dispersal by a wolf from Minnesota. J. Mamm. **64:**166-167.

Fritts, S.H., Mech, L.D., 1981. Dynamics, movements, and feeling ecology of a newly protected wolf population in northwestern Minnesota. Wildlife Monographs **80:** 1-79.

Frudakis, T., Venkateswarlu, K., Thomas, M.J., Gaskin, Z., Ginjupalli, S., Gunturi, S., Ponnuswamy, V., Natarajan, S. & Nachimuthu, P.K., 2003. A classifier for the SNP-based inference of ancestry, J. Forensic Sci. **48:** 771-782.

Fuller, T.K., 1989b. Population dynamics of wolves in north-central Minnesota. Wildlife Monograph, no. 105. The Wildlife Society, Bethesda, MD. 41 pp.

Gagneux, P., Boesch, C.& Woodruff, D.S., 1997. Microsatellite scoring errors associated with noninvasive genotyping on nuclear DNA amplified from shed hair. Molecular Ecology **6:**861-868.

Gagneux, P., Woodruff, D.S. & Boesch, C., 1997a. Furtive mating in female chimpanzees. Nature **387:** 358-359.

Garcia-Muro, E., Aznar, M.P., Rodellar, C.& Zaragoza, P., 1997. Sex specific PCR/RFLPs in the canine *ZFX/ZFY* loci. Animal Genetics **28**: 156.

Garnier, J.N., Bruford, M.W. & Goossens, B., 2001. Mating system and reproductive skew in the black rhinoceros. Molecular Ecology **10:** 2031-2041.

Genovesi, P.. & Dupré, E., 2000. Strategia nazionale di conservazione del lupo (*Canis lupus*): indagine sulla presenza e la gestione dei cani vaganti in Italia. Biol. Cons. Fauna **104:** 1-36.

Genovesi, P. (Ed.), 2002. National action plan for wolf conservation in Italy (*Canis lupus*). Quad. Cons. Natura **13**. Ministry of Environment-National Wildlife Institute

Gerloff, U. Schlotterer, C. Rassmann, K. *et al.*,1995. Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living Bonobos (*Pan paniscus*). Molecular Ecology, **4**, 515–518.

Gese, E.M. & Mech, L.D., 1991. Dispersal of wolves (*Canis lupus*) in northeastern Minnesota, 1969-1989. Can. J. Zool. **69(12):** 2946-2955.

Ghosh, S., Karanjawala, Z.E., Hauser, E.R. *et al.*, 1997. Methods for precise sizing, automated binning of alleles, and reduction of error rates in large-scale genotyping using fluorescently-labelled.

dinucleotide markers. FUSION (Finland-US Investigation of NIDDM Genetics) study group. Genome Research, **7:** 165-178.

Gomes, I., Collins, A., Lonjou, C. *et al.*, 1999. Hardy-Weinberg quality control. Annals of Human Genetics **63:** 535-538.

Goossens, B., Waits, L.P. & Taberlet, P., 1998. Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. Mol. Ecol. **7:** 1237-1241.

Gosling, L.M., 1982. A reassessment of the function of scent marking in territories. Z.Tierpsychol, **60:** 89-118.

Gotelli, D., Sillero-Zubiri, C., Applebaum, G.D., Roy, M.S., Girman, D.J., García-Moreno, J., Ostrander, E.A. & Wayne; R.K., 1994. Molecular genetics of the most endangered canid: The Ethiopian wolf *Canis simensis*. Mol. Ecol. **3:** 301-312.

Guberti, V. & Francisci, F., 1991. Cause di mortalità di 60 lupi raccolti in Italia dal 1984. In: Spegnesi, M. & Toso, S. (eds.), Atti del Secondo Convegno Nazionale dei Biologi della Selvaggina, Suppl. Ric. Biol. Selvaggina **XIX:** 599-603.

Guberti, V., Stancampiano, L. & Francisci, F., 1993. Intestinal helminth parasite community in wolves (*Canis lupus*) in Italy. Parassitology **35:** 59-65.

Guyon, R., Lorentzen, T.D., Hitte, C. *et al*., 2003. A 1-MB resolution radiation hybrid map of the canine genome. Proceedings of the National Academy of Sciences, USA, **100:** 5296-5301.

Haber, G.C., 1997. Socio-ecological dynamics of wolves and pray in a subarctic ecosystem. PhD dissertation, University of British Columbia, Vancouver. 824 pp.

Hancock, J.M., 1995. The contribution of slippage-like processes to genome evolution. J Mol Evol., 1995 Dec; **41(6):** 1038-47.

Hancock, J.M., 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In "Microsatellites: Evolution and Applications" (Goldstein, D.B. & Schlötterer, C., eds.) (Oxford University Press) pp.1-9.

Harrington, F.H. & Mech, L.D., 1979. Wolf howling and its role in territory maintenance. Behaviour **68:** 207-249.

Harrington F.H., Paquet J. & Fentress J.C.,1982. Monogamy in wolves: a review of the evidence. Pp: 209-222 in Harrigton, F.H. & Paquet, P.C., eds. Wolves of the world: perspectives of behaviour, ecology, and conservation. Noyes, Park Ridge, N.J.

Hartl, D.L. & Clark, A.G., 1993. Principles of Population Genetics (Genetica di popolazione). Eds. Zanichelli, Bologna.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of Linnean Society, **58:** 247-279.

Hewitt, G.M., 2000. The genetic legacy of the Quaternary ice ages. Nature **405:** 907-913.

Hines, J. E., 1994. MSSURVIV user's manual,

Hofreiter, M., Sere, D., Pioneer, H.A.N.& Ouch M, Pääbo, S., 2001. Ancient DNA. Nature **2:**353-359.

Holland, P.M., Abramson, R.D., Watson, R. & Gelfand, D.H., 1991. detection of specific polymerasechain reaction products by utilising the 5' to 3' exonuclease activity of *Thermus acquaticus* DNA polymerase. Proceedings of the National Academy of Sciences **88:** 7276-7280.

Holm Andersen, D., Fabbri, E., Santini, A., Paget, S., Cadieu, E., Galibert, F., André C. & Randi, E., 2006. Characterization of 59 canine single nucleotide polymorphisms in the Italian wolf (*Canis lupus*) population. Molecular Ecology Notes **6:**1184-1187.

Höss, M., Kohn, M., Paääbo, S., Knauer, F. & Schröder, W., 1992. Excrement analysis by PCR. Nature **359:** 199.

Jefferis, A. J., Wilson, V. & Thein, S.L., 1985. Hypervariable ''*minisatellite*'' regions in human DNA. Nature, **314:** 67-73.

Johnson, R.P., 1973. Scent marking in mammals. Anim. Behav. **21:** 521-535.

Kleiman, D.G., 1996. Scent marking in the canidae. Symp. Zool. Soc. Lond. **18:** 167-177.

Kohn, M.K. & Wayne, R,K., 1997. Facts from feces revisited. Trends in Ecological Evolution **12:** 223-227.

Kohn, M., York, E.C., Kamradt, D.A., Haught, G., Sauvajot, R.M.& Wayne R.K., 1999. Estimating population size by genotyping feces. Proceedings of the Royal Society of London, Series B, **266:** 657-663.

Lariviere, S. & Crête, M., 1993. The size of eastern coyotes (*Canis latrans*): a comment. J. Mammal. **74:** 1072-1074.

Lebreton, J.D., Burnham, K.P., Clobert, J. & Anderson, D.R., 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs **62:** 67-118.

Lebreton, J.D. & Pradel, R., 2002. Multistate recapture models: modelling incomplete individual histories. Journal of Applied Statistics **29:** 353-369.

Lee, L.G., Connel, C.R. & Bloch, W., 1993. Allelic discrimination by nick translation PCR with fluorogenic probes. Nucleic Acids Research **21:** 3761-3766.

Lehman, N., Clarckson, P., Mech, L.D., Meyer, T.J. & Wayne R.K., 1992. A study of the genetic relationships within and among wolf packs using DNA fingerprinting and mitochondrial DNA. Behavioral Ecology and Sociobiology **30:** 83-94.

Lehman, N.E., Eisenhawer, A., Hansen, K., Mech, L.D., Petrson, R.O., Gogan, P.J.P. & Wayne, R.K., 1991. Introgression of coyote mitochondrial DNA into sympatric north American gray wolf populations. Evolution **45:** 104-119.

Li, W.H., Gojobori T., Nei, M., 1981. Pseudogenes as a paradigm of neutral evolution, Nature **292:** 237-239.

Lindley, D. V., 1990. The 1988. Wald memorial lecture: The present position in Bayesian Statistics. Statist. Sci. **5:** 44-89.

Liu Q. & Sommer, S.S., 2000.. Pyrophosphorolysis-activated polymerization (PAP): application to allele-specific amplification. Biotechniques **29:** 1072-1080.

Liu Q. & Sommer, S.S., 2002. Pyrophosphorolysis-activatable oligonucleotides may facilitate detection of rare alleles, mutation scanning and analysis of chromatin structures. Nucleic Acids Res **30:** 598-604.

Livak K.J., 1999. Allelic discrimination using fluorogenic probes and the 50 nuclease assay. Genet Anal **14:** 143-149.

Livak, K.J., Flood, S.A.J., Marmaro, J., *et al*., 1995a. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods and Applications **4:** 357-362.

Livak, K.J. & Goodsaid, F., 1997. Allelic discrimination using the 5' nuclease assay. Foster City, CA: PE Applied Biosystems Bullettin.

Lorenzini, R. & Fico, R., 1995. A genetic investigation of enzyme polymorphisms shared by wolf and dog: suggestions for conservation of the wolf in Italy. Acta Theriol. **3 (Suppl.):** 101-110.

Lucchini, V., Fabbri, E., Marucco, F., Ricci, S., Boitani, L. & Randi, E., 2002. Noninvasive molecular tracking of colonizing wolf (*Canis lupus*) packs in the western Italian Alps. Molecular Ecology, **11:** 857-868.

Lucchini, V., Galov, A. & Randi, E., 2004. Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines. Molecular Ecology, **13:** 523-536.

Lukacs, P.M., 2005. Statistical aspects of using genetic markers for individual identification in capture-recapture studies. PhD Thesis, Colorado State University.

Lukacs, P.M. & Burnham, K.P., 2005. Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. Journal of Wildlife Management **69:** 396-403.

Lyamichev, V., Brow, M.A.D.,& Dahlberg, J.E., 1993. Structure-specific endonucleotyc cleavage of nucleic acids by eubacterial DNA polymerases. Science **260:** 778-783.

Martinez-Arias R., Calafell F., Mateu E., Comas D., Andres A., Bertranpetit J., 2001. Sequence variability of a human pseudogene, Genome Res; **11:** 1071-1085.

Mattioli, L., Apollonio, M., Lazzaroni, V. & Centofanti, E. 1995. Wolf food habitats and wild ungulate availability in the Foreste Casentinesi National Park, Italy. Acta Theriologica, **40 (4):** 387-402.

McKelvey, K.S. & Schwartz, M.K., 2004. Genetic errors associated with population estimation using non-invasive molecular tagging: problems and new solutions. Journal of Wildlife Management **68:** 439-448.

Mech, L.D., 1970. The wolf: the ecology and behaviour of an endangered species. New York, NY: Natural History Press.

Mech, L.D., 1973. Wolf numbers in the Superior National Forest of Minnesota. USDA Forest Service Research Paper NC-97. North Central Forest Experiment Station, St. Paul, MN, 10pp.

Mech, L.D., 1975. Disproportionate sex ratios of wolf pups. J. Wildl. Mgmt **39:** 747-750.

Mech, L.D., 1977d. Wolf-pack buffer zones as prey reservoirs. Science **198:** 320-321.

Mech, L.D., 1981. The Wolf: The Ecology and Behaviour of an Endangered Species, 1st ed. Natural History Press, Doubleday, NY.

Mech, L.D., 1982b. The IUCN-SSC Wolf Specialist Group. Pp 327-333 in Harrington, F.H. & Paquet, P.C., eds, Wolf of the world: Perspectives of behaviour, ecology and conservation. Noyes Publications, Park Ridge, NJ.

Mech, L.D., 1988a. The arctic wolf: living with the pack. Voyageur Press, Stillwater, MN.

Mech, L.D., 1989. Wolf population survival in an area of high road density. Am. Midl. Nat. **121:** 387-389.

Mech, L.D., 1994a. Buffer zones of territories of Gray wolves as regions of intraspecific strife. J. Mammal. **75:** 199-202.

Mech, L.D., 1995c. Summer movements and behaviour of an arctic wolf, *Canis lupus*, without pups. Can. Field Nat. **109:** 473-475**.**

Mech, L.D., 1995d. A ten year history of the demography and productivity of an arctic wolf pack. Arctic **48:** 329-332.

Mech, L.D., 2001b. "Standing over" and "Hugging" in wild wolves, *Canis lupus*. Can. Field Nat. 115: 179-181.

Mech, L.D., Adams, L.G., Meier, T.J., Burch, J.W. & Dale, B.W. ,1998. The wolves of Denali. Minneapolis and London: University of Minnesota Press.

Mech, L.D. & Boitani, L. 2003. Wolf social ecology. In Wolves: 1-34. Mech, L. D. & Boitani, L. (Eds). Chicago and London: University of Chicago Press.

Mech, L.D., Fritts, S.H. & Wagner, D., 1995. Minnesota wolf dispersal to Wisconsin and Michigan. Am .Midl. Nat. **133:** 368-370.

Mech, L.D. & Hertel, H.H., 1983. An eight year demography of a Minnesota wolf pack. Acta Zool. Fenn. **174:** 249-250.

Meier, T.J, Burch, J.W., Mech, L.D. & Adams, L.G., 1995. Pack structure dynamics and genetic relatedness among wolf packs in a naturally regulated population. Pp. 293-302 in Carbyn, L.N., Fritts, S.H. & Seip, D.R., eds., Ecology and conservation of wolves in a changing world. Canadian Circumpolar Institute, Edmonton, Alberta.

Mellersh, C.S.& Ostrander, E.A.,1997. The Canine Genome. Adv Veter Med **40:**191-215.

Mengel, R.M., 1971. A study of dog-coyote hybrids and implications concerning hybridization in *Canis*. J. Mammal. **52:** 316-336.

Meriggi, A., Brangi, A., Matteucci, C.& Sacchi, O., 1996. The feeling habits of wolves in relation to large prey availability in northern Italy. Ecography **19:** 1-9.

Meriggi, A. & Lovari, S., 1996. A review of wolf predation in Southern Europe: does the wolf prefer wild prey to livestock? J. Appl. Ecol. **33:** 1561-1571.

Meriggi, A., Paola, R., Brangi, A. & Mattucci, C., 1991. Habitat use diet of the wolf in northen Italy. Acta Theriologica **36 (1-2):** 141-152.

Meriggi, A, Rosa, P., Brangi, A, & Matteucci, C., 1991. Habitat use and diet of the wolf in the northern Italy. Acta Theriol. 36: 141-151.

Miller, C., Joyce, P.& Waits, L.P., 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics **160:** 357-366.

Mills, L.S., Citta, J.J., Lair, K.P., Schwartz, M.K. & Tallmon, D.A., 2000. Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. Ecological Applications **10:** 283-294.

Montagna, W. & Parks, H.F., 1948. A histochemical study of the glands of the anal sac of the dog. Anat. Rec. **100:** 297-318.

Monteiro, L., Bonnemaison, D., Vekris, A., Petry, K.G., Bonnet, J., Vidal, R., Cabrita, J. & Mégraud, F., 1997. Complex polysaccharides as PCR inhibitors in feces: Helicobacter pylori model. J Clin Microbiol **35:** 995-998.

Morin, P.A., Chambers, K.E., Boesch, C. & Vigilant, L., 2001. Quantitative PCR analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Mol. Ecol. **10:** 1835-1844.

Morin P.A., Luikart, G., Wayne, R. K. & the SNP workshop group, 2004. SNPs in ecology, evolution and conservation. Trends in Ecology and Evolution **19:** 208-216.

Muller, U. G. & LaReesa Wolfenbarger, L., 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. **10:** 389-394.

Mullis, K., Erlich, H. , Faloona, F., Horn, G., Saiki, R. & Scharf, S., 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harbor Symp. Quant. Biol., **51:** 263-273.

Murie, A., 1944. The wolves of Mont Mckinley. U.S. National Park Service Fauna Series, no. 5. U.S. Government Printing Office, Washington, D.C. 238 pp.

Murphy, M., Hinman, A. & Levine, A.J., 1996. Wild-type p53 negatively regulates the expression of a microtubule-associated protein. Genes & Dev., **10:** 2971-2980.

Navidi, W., Arnheim, N. & Waterman, M.S., 1992. A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. American Journal of Human Genetics **7:** 347-359.

Neff, M.W., Broman, K.W., Mellersh, C.S. *et al.,* 1999. A second generation linkage map of the domestic dog, *Canis familiaris*. Genetics **151:** 803-820.

Nievergelt, C.M., Mutschler, T., Feistner, A.T.C., & Woodruff ,D.S., 2002. Social system of the Alaotran Gentle Lemur (*Hapalemur griseus alaotrensis*): genetic characterization of group composition and mating system. Am. J. Primatol. **57:** 157-176

Nichols, J.D.,1992. Capture-recapture models. Bioscience **42:** 94-102.

Nowak, R.M. & Federoff, N.E., 2002. The systematic status of the Italian wolf. Acta Theriologica **47**:333-338.

O'Hagan, A., 1994. Bayesian Inference. London: Edward Arnold. Press, S. J. (1972). Applied Multivariate Analysis: using Bayesian and Frequentist Methods of Inference. Melbourne, FL: Krieger.

Olson, F.S., 1938. Organization and range of the pack. Ecology, **19:** 168-170.

Ostrander, E.A., Sprague, G.F. & Rine, J., 1993. Identification and characterization of dinucleotide repeat (CA)*n* markers for genetic mapping in dog. Genomics **16:** 207-213.

Otis, D.L., Burnham, K.P., White, G.C. & Anderson, D.R., 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs, **62:** 1-135.

Paetkau, D. & Strobeck, C. 1994. Microsatellite analysis of genetic variation in black bear populations. Molecular Ecology **3:** 489-495.

Palsbøll, P.J., 1999. Genetic tagging: contemporary molecular ecology. Biology Journal of the Linnean Society **68:** 3-22.

Palsbøll, P.J., Allen, J., Berube, M., Clampham, P.J., Feddersen, T.P., Hammond, P.S., Hudson, R.R., Jørgensen, H., Katona, S., Larsen, F., Lien., J., Mattila, D.K., Sigurjonsson, J., Sears, R., Smith, T., Sponer, R., Stevick, P. & Øin., N., 1997. Genetic tagging of humpback whales. Nature **388:** 767-769.

Parker, Q.A., Phillipps, S. & Morgan, D.H. , 1995. IAU Colloq. 148, ASP Conf. Ser. 84, eds. Chapman, J. M. *et al*. (San Francisco: ASP), p. 96.

Parks, H.F., 1950. Morphological and cytochemical observations on the circumanal glands of the dog. PhD dissertation, Cornell University, Ithaca, NY.

Parsons, K.M., 2001. Reliable microsatellite genotyping of dolphin DNA from faeces. Molecular Ecology Notes **1:** 341-344.

Pearce, J.M., Fields, R.L. & Scribner K.T., 1997. Nest materials as a source of genetic data for avian ecological studies. Journal of Field Ornithology **68:** 471-481.

Paetkau, D., 2003. An empirical exploration of data quality in DNA-based population inventories. Molecular Ecology **12:** 1375-1387.

Peakall, R. , & Smouse, P. E. , 2005. GenAlEx 6: Genetic Analysis in Excel., Population Genetic Software for Teaching and Research. Australian National University, Canberra, Australia. http://www.anu.edu.au/BoZo/GenAlEx.

Peakall, R. , & Smouse, P. E. , 2006. GenAlex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, **6:** 288-295.

Peters, R.P. & Mech, L.D., 1975. Scent-marking in wolves: a field study. American Scientist **63:** 628-637.

Peters, R.P. & Mech, L.D., 1975b. Scent-marking in wolves. A field study. Am. Sci. **63:** 628-637.

Peterson, R.O. & Ciucci, P., 2003. The wolf as a carnivore. In Wolves: 104-109. Peterson, R. O. & Ciucci, P. (Eds). Chicago and London: University of Chicago Press.

Peterson, R.O. & Page, R.E., 1988. The rise and fall of the Isle Royale wolves. J. Mammal. 69: 89-99.

Peterson R.O., Woolington J.D. & Bailey T.N., 1984. Wolves of the Kenai Peninsula, Alaska. *Wildl. Monogr*. **88**:1-52.

Pezzo, F., Parigi, L. & Fico, R., 2003. Food habits of wolves in central Italy based on stomach and intestine analyses. Acta Theriol. **48:** 265-270.

Pledger, S., Pollock, K. H. & Norris, J.L., 2003. Open capture-recapture models with heterogenety: I. cormack-jolly-seber model. Biometrics **59:** 786-794,.

Piggott, M.P. & Taylor, A.C., 2003. Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. Wildlife Res. **30:** 1-13

Poulle, M.L., Lequette, B. & Dahier, T., 1999. Le recolonisation des Alpes par le loup de 1992 à 1998. Bulletin Mensuel de l'Office National de la Chasse **242:** 4-13.

Pradel, R. 1993. Flexibility in survival analysis from recapture data: handling trap-dependence. In Lebreton, J.D. & North, P.M. editors, Marked individuals in the study of bird population, pages 29-37. Birkh Verlag, Basel.

Pradel, R., 2005. Multievent:an extension of multistate capture-recapture models to uncertain states. Biometrics **61:** 442-447.

Pradel, R., Hines, J.E., Lebreton, J.D. & Nichols, J.D., 1997. Capture-recapture survival models taking account of transients. Biometrics **53:** 6072.

Pritchard, J.K., Stephens, M. & Donnelly, P.J., 2000. Inference of population structure using multilocus genotype data. Genetics **155:** 945-959.

Promberger, C.& Schröder, W. (Eds.), 1993. Wolves in Europe: Status and perspectives. Munich Wildlife Society, Ettal, Germany.

Randi, E., 2007. Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology **17:** 285-293.

Randi, E., Francisci, F. & Lucchini, V., 1995. Mitochondrial DNA restriction-fragment-length monomorphism in the Italian wolf (*Canis lupus*). J. Zool. Syst. Evol. Res. **33:** 97-100.

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

Randi, E., Lucchini, V., Christensen, M.F., Mucci, N., Funk, S.M., Dolf, G. & Loeschcke, F., 2000. Mitochondrial DNA variability in Italian and east European wolf: detecting the consequence of small population size and hybridization. Conservation Biology **14**:464-473.

Randi, E., Lucchini, V. & Francisci, F., 1993. Allozyme variability in the Italian wolf (*Canis lupus*) population. Heredity **71:** 516-522.

Reed, J., Tollit, D., Thonpson, P. & Amos, W., 1997. Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal feces. Molecular Ecology **6:** 225-234.

Rhymer, J.M. & Simberloff, D., 1996. Extinction by hybridisation and introgression. Annual Review of Ecology and Systematics, **27:** 29-44.

Ronaghi, M., 2001. Pyrosequencing sheds light on DNA sequencing. Genome Res. **11:** 3-11.

Ronaghi, M., Uhlen, M. & Nyren, P., 1998. A sequencing method based on real-time pyrophosphate. Science **281:** 363-365.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. & Nyren, P., 1996. RealTime DNA sequencing using detection of pyrophosphate release. Anal. Biochem. **242:** 84-89.

Roon, D.A., Waits, L.P. & Kendall, K.C., 2003. A quantitative evaluation of two methods for preserving hair samples. Mol. Ecol. Notes **3:**163-166

Rosenzweig, M.L., 1968. The strategy of body size in mammalian carnivores. Am Midl.Nat. **80:** 299-315.

Rothman, N.J. & Mech, L.D., 1979. Scent-marking in lone wolves and newly formed pairs. Anim. Behav. **27:** 750-760.

Savolainen, P., Zhang, Y., Luo, J., Lundeberg, J. & Leitner, T., 2002. Genetic Evidence for an East Asian Origin of Domestic Dogs. Science **298:** 1610-1613.

Scandura, M., Apollonio, M. & Mattioli, L., 2001. Recent recovery of the Italian wolf population: a genetic investigation using microsatellites. Mammalian Biology **66:** 321-331.

Schassburger, R.M., 1978. The vocal repertoire of the wolf: structure, function, and ontogeny. PhD dissertation, Cornell University, Ithaca, NY. 343 pp.

Schassburger, R.M., 1993. Vocal communication in the timber wolf, *Canis lupus*, Linneaus: structure, motivation, and ontogeny. Advances in Ethology, no. 30. Paul Parey, Berlin. 84 pp.

Schenkel, R., 1947. Ausdrucks-studien an wolfen (expression studies of wolves). Behaviour **1:**81-129 (Translation from German by Harrington, F.H.).

Schlötterer, C., 2004. The evolution of molecular markers-just a matter of fashion? Nature Reviews Genetics **5:** 63-69.

Seber, G.A.F., 1982. Estimation of Animal Abundance. Charles Griffin, London.

Seddon, J.M., Parker, H.G., Ostrander, E.A. & Ellegren, H., 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. Molecular Ecology **14:** 503-511.

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

Sivia,D., 1996. Data Analysis: A Bayesian Tutorial. Oxford: Clarendon Press (1996), pp. 7-8.

Slatkin, M., 1991. Inbreeding coefficients and coalescence times. Genetical Research, Cambridge, **58:** 167-175.

Sloane, M. A., Sunnucks, P., Alpers, D., Behergaray, L. B. & Taylor, A. C., 2000. Highly reliable genetic identification of individual northern hairy-noised wombats from single remotely collected hairs: a feasible censusing method. Mol. Ecol. **9:** 1233-1240.

Smith, D., Meier, T., Geffen, E., Mech, D.L., Burch, J., Adams, L. & Wayne, R., 1997. Is incest common in grey wolf packs? Behavioral Ecology **4:**384-391.

Smith, E.F.G., Arctander, P., Fjeldsa, J. & Amir, O.G., 1992. A new species of shrike (Laniidae: *Lanarius*) from Somalia, verified by DNA sequence data from the only known individual. Ibis 227-235.

Smith, K.L, Alberts, S.C., Bayes, M.K. *et al*., 2000. Cross-species amplification, non-invasive genotyping and no-Mendelian inheritance of human STRPs in Savannah baboons. American Journal of Primatology **51:** 219-227.

Sokolov, V.E. & Rossolino, O.L., 1985. Taxonomy and variability. Pp: 21-50 in Bibikov, D.I. ed. The wolf: history, systematics, morphology, ecology. USSR Acad. Sci. Nauka, Moscow.

Stahler, D.R., Smith, D.W. & Landis, R., 2002. the acceptance of a new breeding male into a wild wolf pack. Can. J.Zool. **80:** 360-365.

Stallings, R.L., Ford, A.F., Nelson, D., Torney, D.C., Hildebrand, C.E.& Moyzis, R., 1991. Evolution and Distribution of (GT)n Repetitive Sequences in Mammalian Genomes. Genomics **10:** 807-815.

Sugiyama, Y., Kawamoto, S., Takenaka, O., Kumizaka, K. & Norikatsu, W., 1993. Paternity discrimination and inter-group relationships of Chimpanzees at Bossou. Primates **34:** 545-552**.**

Sunnucks, P. , 2000. Efficient genetic markers for population biology. Trends Ecol Evol, **15:** 199-203.

Syvanen, A.C., 1999. From gels to chips: mini-sequencing primer extension for analysis of point mutation and single nucleotide polymorphisms, Hum. Mut. **13:** 1-10.

Syvanen, A.C., 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. Nature Genetics Reviews, **2:** 930942.

Syvanen, A.C., Aalto-Setala, K., Harju, L., Kontula, K. &. Soderlund, H.A., 1990. Primer-guided nucleotide incorporation assay in genotyping of apolipoprotein E, Genomics **8:** 684-692.

Taberlet, P. & Bouvet, J., 1991. Single plucked feather as a source of DNA for bird genetic studies. The auk, Vol. 108, no. 4 (Oct 1991), pp.959-960.

Taberlet, P. & Bouvet, J., 1992. Bear conservation genetics. Nature, Lond. **358:** 197.

Taberlet P., Camarra J.J., Griffin S., Uhrès E., Hanotte O., Waits L., Dubois-Paganon C., Burke T. & Bouvet J. 1997. Noninvasive genetic traking of the endangered Pyrenean brown bear population. Molecular Ecology **6**:869-876.

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

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

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

Takenaka, O., Takashi, H., Kawamoto, S., Arakawa, M & Takenaka, A., 1993. Polymorphic microsatellite DNA amplification customised for chimpanzee paternity testing. Primates **34:** 27-35.

Tautz, D.& Renz, M., 1984. Simple Sequences are Ubiquitous Repetitive Components of

Eukaryotic Genomes. Nucleic Acids Res **12:** 4127-4137.

Thiel, R. P., 1985. Relationships between road densities and wolf habitat suitability in Wisconsin. Am.Midl. Nat. **113:** 404-407.

Thurber, J.M. & Peterson; R.O., 1991. Changes in body size associated with range expansion in the coyote (*Canis latrans*). J. Mammal. **72:** 750-755.

Tijms, H., 2004. Understanding probability. Cambridge University Press.

Tikel, D., Blair, D. & Marsh, H., 1996. Marine mammal faeces as a source of DNA. Mol. Ecol. **5:** 456-457.

Valière, N., 2002. GIMLET: a computer program for analysing genetic individual identification data. Molecular Ecology Notes **10:**1046. http://pbil.univ-lyon1.fr/software/Gimlet/gimlet.htm.

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

Valière, N., Fumagalli, L., Gielly, L. *et al.*, 2003. Long distance wolf recolonization of France and Switzerland inferred from noninvasive genetic sampling over a period of 10 years. Animal Conservation **6:** 83-92.

Valieíre, N. & Taberlet, P., 2000. Urine collected in the field as a source of DNA for species and individual identification. Mol. Ecol. **9:** 2150-2152.

Verardi, A., Lucchini, V. & Randi, E., 2006. Detecting introgressive hybridisation between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. Molecular Ecology, **15:** 2845-2855.

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol, **34:** 275-305.

Vilà, C., Urios, V. & Castroviejo, J., 1994. Use of faeces for scent marking in Iberian wolves (*Canis lupus*). Can. J. Zool. **72:** 374-377.

Vilà, C., & Wayne, R. K., 1999. Hybridization between wolves and dogs. Conserv. Biol. 13: 195-198.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., & Zabeau, M., 1995. AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. **23:** 4407-4414.

Wabakken, P., Sand, H., Liberg, O. & Bjärvall, A., 2001. The recovery, distribution, and population dynamics of wolves of the Scandinavian Peninsula, 1978-1998. Can. J. Zool. **79:** 710-725.

Waits, L.P., 2004. Using noninvasive genetic sampling to detect and estimate abundance of rare wildlife species. In: Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters (eds. Thompson, W.L.), pp. 211-228. Island Press, Washington, D.C.

Waits, L.P., Luikart, G.& Taberlet, P., 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology **10:** 249-256.

Walters, C.J., Stocker, M. & Haber, G.C., 1981. Simulation and optimization models for a wolf-ungulate system. Pp. 317-337 in Fowler, C.W. & Smith, T.D., eds; Dynamics of large mammal populations. John Wiley & Sons, New York.

Wang, D.G., Fan J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N.,Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M.S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T.J.& Lander E.S., 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, Science **280:** 1077-1082.

Wasser, S.K., Houston, C.S., Koehler, G.M., Cadd, G.G. & Fain, S.R., 1997. Techniques for application of faecal DNA methods to field studies of Ursids. Molecular Ecology **6:** 1091-1097.

Watson, J. D. & Crick, F. H. C., 1953. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. Nature, **171:** 737-738.

Wayne, R. K., Geffen, E., Girman, D.J., Koepfli, K.P., Lau, L.M.. & Marshal, C.R., 1997. Molecular systematics of the Canidae. Syst. Biol. 46: 622-653.

Wayne, R.K., Lehman, N. & Fuller, T.K., 1995. Conservation genetics of the gray wolf. In: Ecology and Conservation of Wolves in a Changing World (eds. Carbyn, L.N., Fritts, S.H. & Seip, D.R.), 399–407. Canadian Circumpolar Institute, Occasional Publication no. **35**. Edmonton, Alberta.

Weir, B.S. & Cockerham C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution **38:** 1358-1370.

Weir, B.S. & Hill, W.G., 2002. Estimating f-statistics. Annual Review of Genetics **36:**721-750.

West, M. & Harrison, P. J., 1989. Bayesian Forecasting and Dynamic Models. Berlin: Springer.

White, G.C.& Burnham, K.P., 1999. Program MARK: survival estimation from populations of marked animals. Bird Study **46 (suppl.):**120-139.

Williams, J. G. K., Kubelik, A. R., Livak, K. L. & Tingey. S. V., 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., **18:** 6531-6535.

Woods, J.G., Paetkau, D., Lewis, D. *et al.*, 1999. Genetic tagging of free-ranging black and brown bears. Wildlife Society Bulletin **27:** 616-627.

Wright, S., 1951. The genetical structure of populations. Annals of Eugenics **15:** 323-354.

Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution. **19:** 395-420.

Wright, S., 1969. Evolution and the Genetics of Populations, volume 2. The Theory of Gene Frequencies. University of Chicago Press, Chicago, IL, 1969.

Wright, S., 1978. Evolution and the Genetics of Populations. Vol. 4. Variability Within and Among Natural Populations. Univ. of Chicago Press, Chicago.

Wydeven, A.P., Schultz, R.M. & Thiell, R.P., 1995. Monitoring of a gray wolf (*Canis lupus*) population in c Conservation of wolves in a changing world. Canadian Circumpolar Institute, Edmonton, Alberta.

Xu, J., Turner, A., Little, J., Bleecker, E.R. & Meyers, D.A., 2002. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? Human Genetics **111:** 573-574.

Young, S.P. & Goldman E.A., 1944. The wolves of North America. American Wildlife Institute, Washington, D.C.

Zimen, E., 1976. On the regulation of pack size in wolves. Z. Tirpsychol. 40: 300-341.

Zimen, E., 1981. The wolf: a species in danger. Delacorte Press, New York.

Zimen, E. &. Boitani, L., 1975. Number and distribution of wolves in Italy. Zeitschrift für Säugetierkunde **40:** 102-112.