Conservation genetics of the lion

The impact of new technologies

Klaas Vrieling
Institute of Biology
Leiden University
The Netherlands

In cooperation with
Hans de longh and
Laura Bertola

Conservation genetics

Apply genetic methods to the conservation and restoration of biodiversity.

Conservation genetics

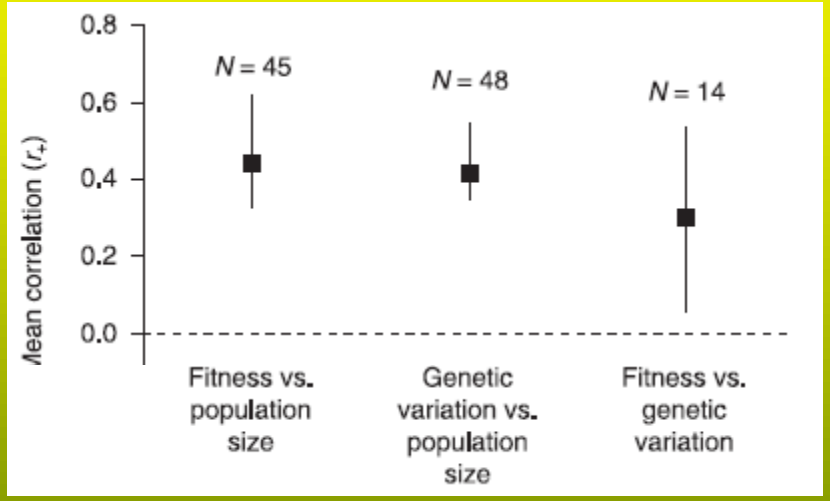
Goals:

- **Resolve taxonomic uncertainties**
- **Measure genetic diversity (evolutionary potential)**
- **Measure inbreeding**
- **Defining management units**
- **Guide successful breeding in captivity**

Conservation genetics

Small populations:

1) Have an increased homozygosity and an increased frequency of deleterious alleles leading to an average reduction of individual fitness



Leimu et al. 2006
Based on studies of plants only

Saccheri et al. 1998
Field study showing that populations with low heterozygosity have a higher incidence of extinction in the *Glanville fritillary*

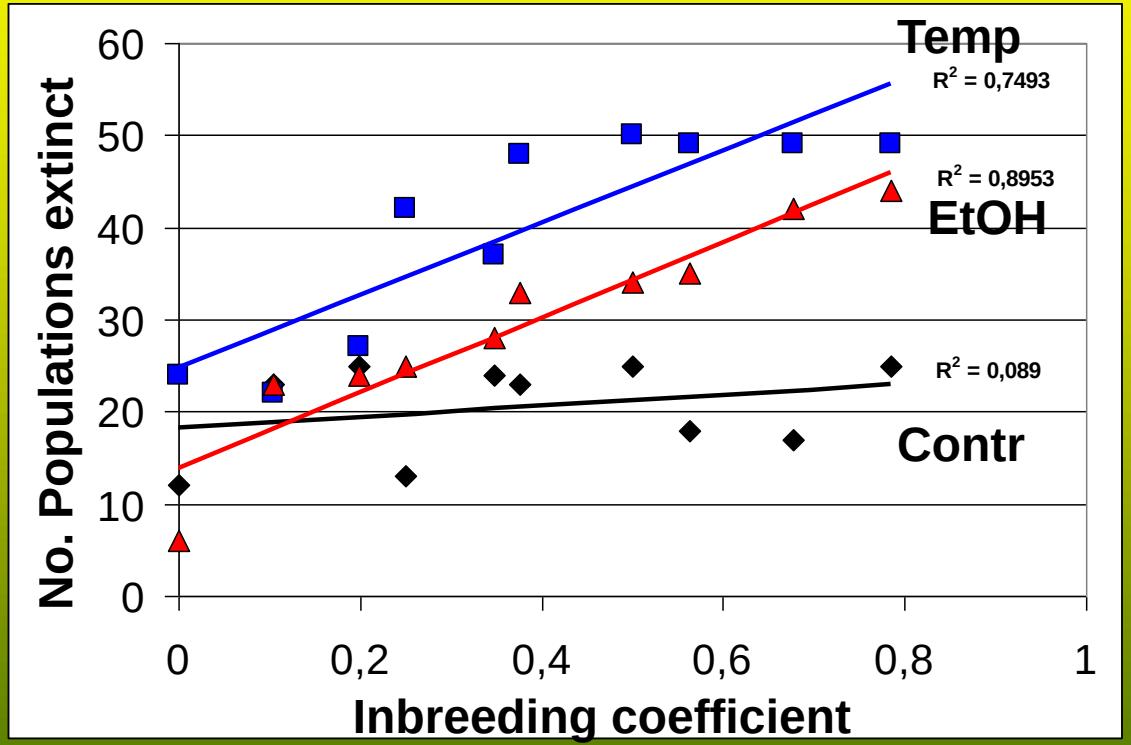
Conservation genetics

Small populations:

2) The loss of genetic variants compromises the evolutionary adaptive potential of the population

Bijlsma et al. 2000

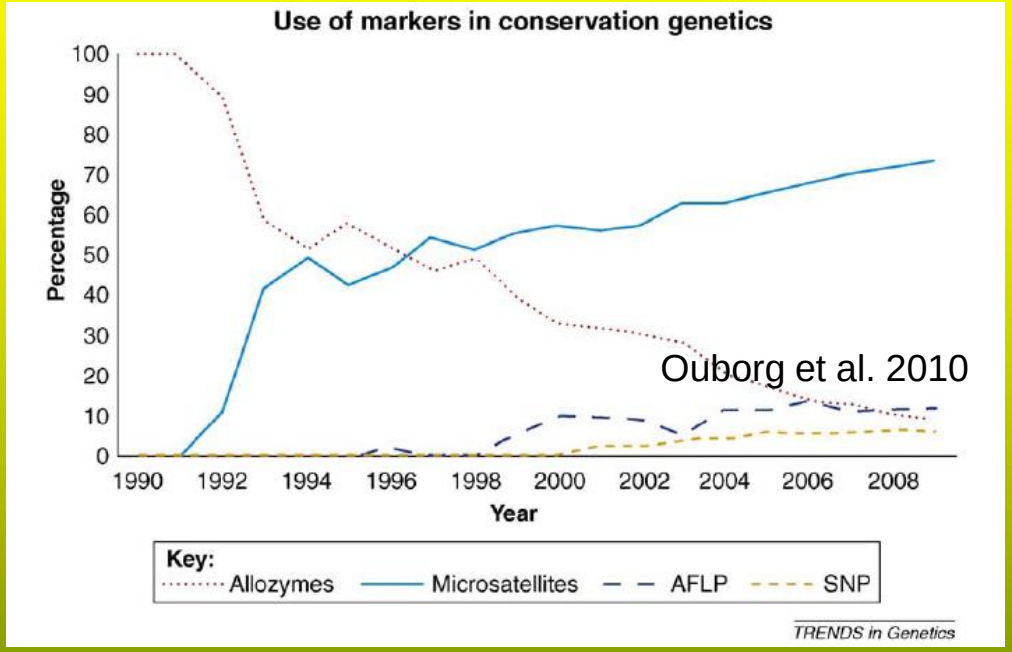
50 populations per treatment
10 males and 10 females
Followed for 8 genartions



Conservation genetics

Small populations:

3) Small population will become increasingly more genetic divergent potentially leading to outbreeding depression



Ouborg et al. 2010

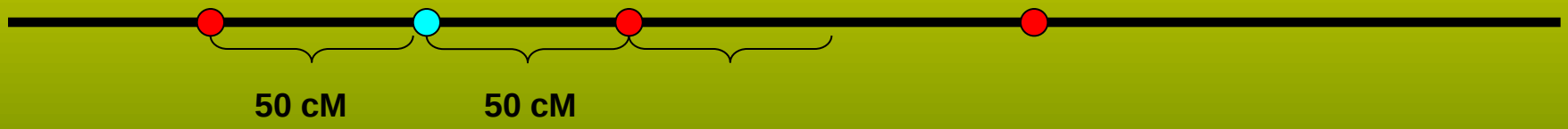
The most common markers in conservation genetics used up to now are microsatellites

How well do genetic markers cover the genome?

Lion genome has 38 chromosomes and is 3.475.000.000 bp long.
The total length of the lion genome in mapping units is 4633 cM.

Suppose genetic variation is measured with 20 microsatellites

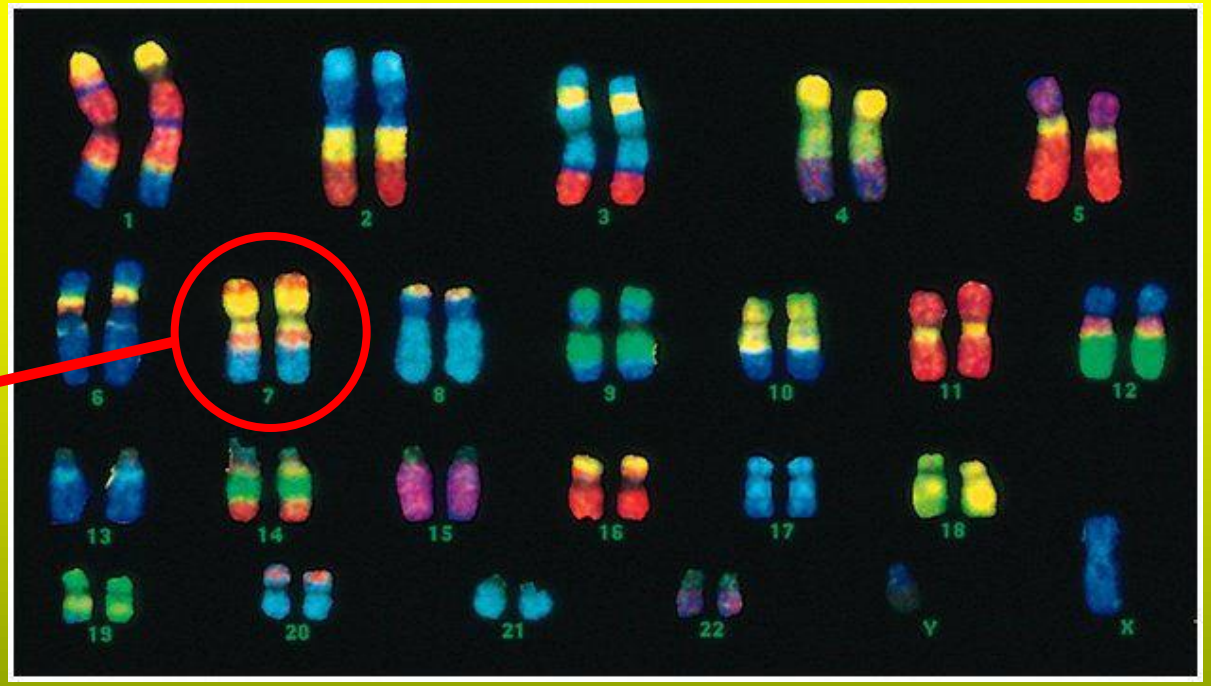
If they are ideally distributed they cover 2000 cM or 42% of the lion genome



Microsatellites are only found in non-coding parts of the DNA and mutation rates of microsatellites are 100 to 1000 higher than of SNPs (or genes)

What is a SNP?

The human nuclear genome 22 pairs of chromosomes and 1 pair of sex chromosomes



2 homologous chromosomes

--AACTTGCAAATTGAACTCTGA---
--TTGAACGTTTAACTTGAGACT---

One homologous chromosome

--AACTTGCAAGTTGAACTCTGA---
--TTGAACGTTCAACTTGAGACT---

The other matching homologous chromosome

What is a SNP?

Pointmutation

Homozygotes

Heterozygotes

```
Genotype 1   GCTGCCAATTTTCG
             GCTGCCAATTTTCG

Genotype 2   GCTGCTAATTTTCG
             GCTGCTAATTTTCG

Genotype 3   GCTGCCAATTTTCG
             GCTGCTAATTTTCG

Genotype 4   GCTGCCAATTTTCG
             GCTGCCAATTTTCG

Genotype 5   GCTGCTAATTTTCG
             GCTGCCAATTTTCG
```

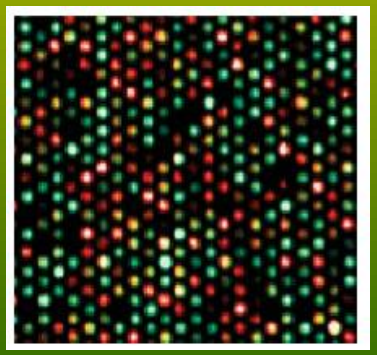
It is estimated that human genome contains 10 millions SNPs.
On average every 350 bp contains 1 SNP

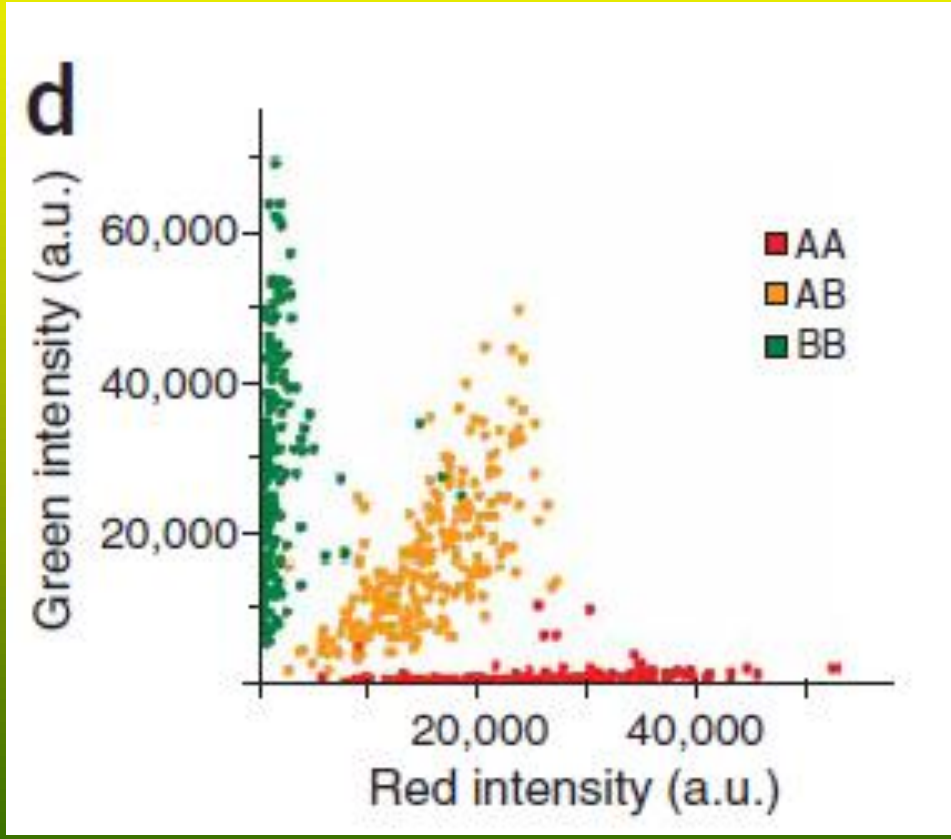
SNP detection

An array of methods is available many of them having limitations in accuracy, throughput or simultaneous processing (multiplexing).

Microarray based techniques

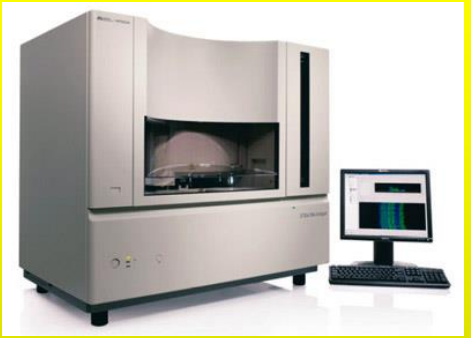
Infinium array allows simulteaneous Detection of 6000 SNPs





SNP discovery: Next generation sequencing

For more than 15 years sequencing traditionally has been carried out with the Sanger method.



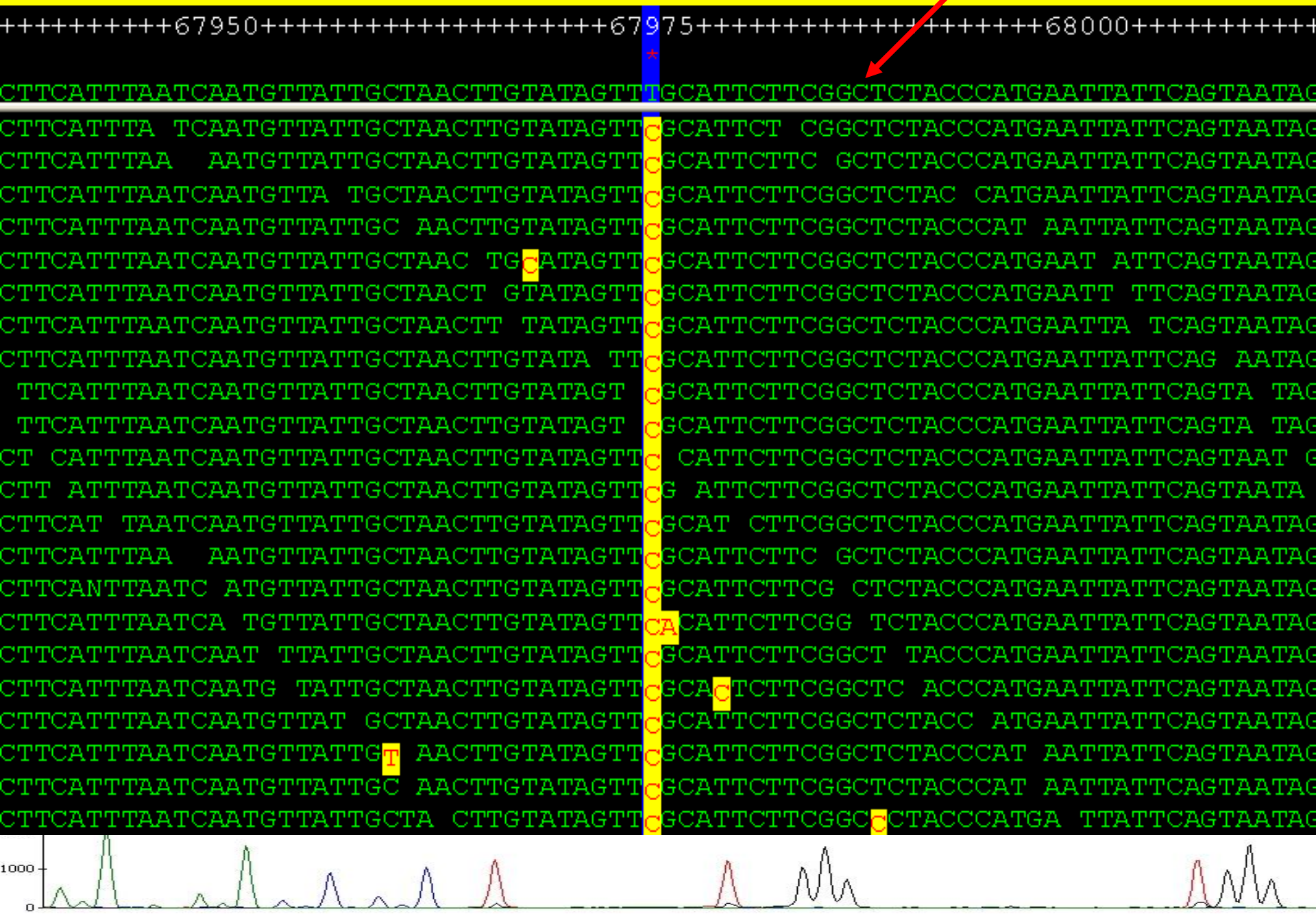
One day throughput is 1,5 Mb of sequence

At minimum 10 persons are needed to prepare the runs to reach this capacity.

What are the new solutions?

454 GS FLX (Roche)	**384 Mb/day**
SOLiD (Applied Biosystems)	**670 MB/day**
Solexa (Illumina)	**1440 Mb/day**
Heliscope tSMS (Helicos Biosciences)	**24000 Mb/day**

Next generation sequencing



SNP discovery for lion

Three types of DNA:

1) Nuclear DNA (autosomes) : inherited biparentally and recombining

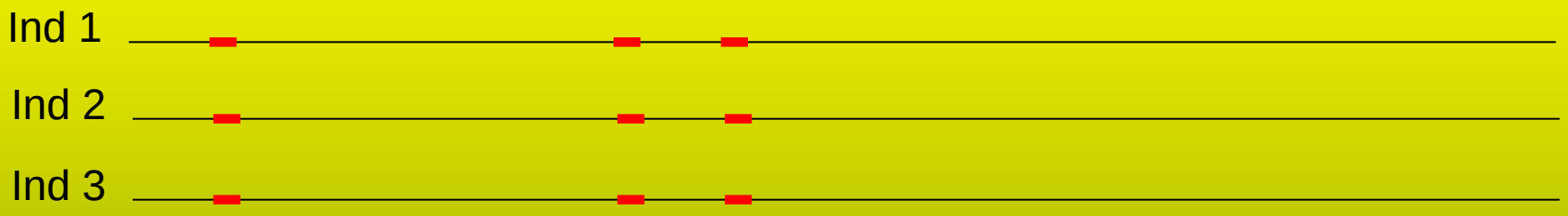
2) Nuclear DNA, Y chromosome: Inherited only through the male line as a single locus (no recombination)

3) Mitochondrial DNA : Inherited only through the female line as a single locus (no recombination)

SNP discovery for lion (nuclear DNA and Y chromosome)

1) Extract DNA of 10 individuals spread over the distribution range

2) Select less than 1% of the same DNA of each individual



3) Give DNA of each individual an unique tag. Pool all samples

4) Sequence the fragments on Illumina platform (30.000.000 reads of 250 basepair) and assemble reads and score SNPs

5) Circa 30.000 SNPs expected

* Parts of DNA always missing in female samples but always present in male samples are very likely are parts of the Y chromosome. Between 200 and 400 Y chromosoom SNPs expected

SNP discovery for lion (mitochondrial DNA)

1) Amplify total mtDNA (17.000 bp) with long range PCR for 10 individuals (1 or 2 primer pairs)

2) Tag PCR products per individual and sequence all in one lane on Illumina platform (30.000.000 reads of 36 base pair)

3) Use mtDNA of closely related species (Leopard) as reference

4) Assemble read of each individual and find SNPs (circa 300 SNPs expected)

Power of Illumina: 30.000.000 reads of 36 bp = 1.080.000.000 bp
Mt DNA is 17.000 bp long
Each base is therefore 63.529 sequenced

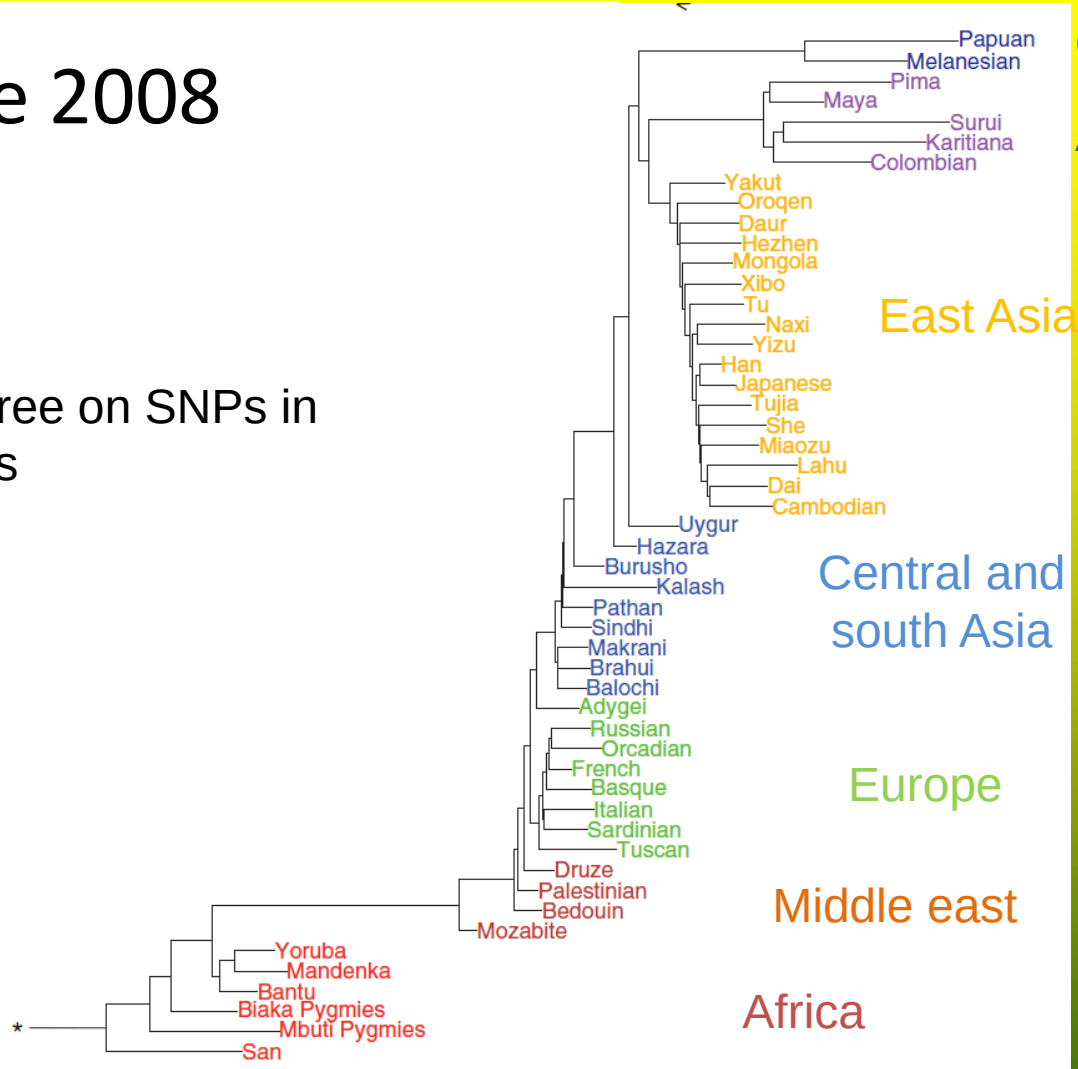
You want to know the sequence of your pet lion? Come to us!

Comparing microsatellites and SNPs for Lions

	Microsatellite	SNP
No. analysed	20	6000
Costs euro	124	120
Hours labour	32	8
Coverage cM/marker	231	0,77

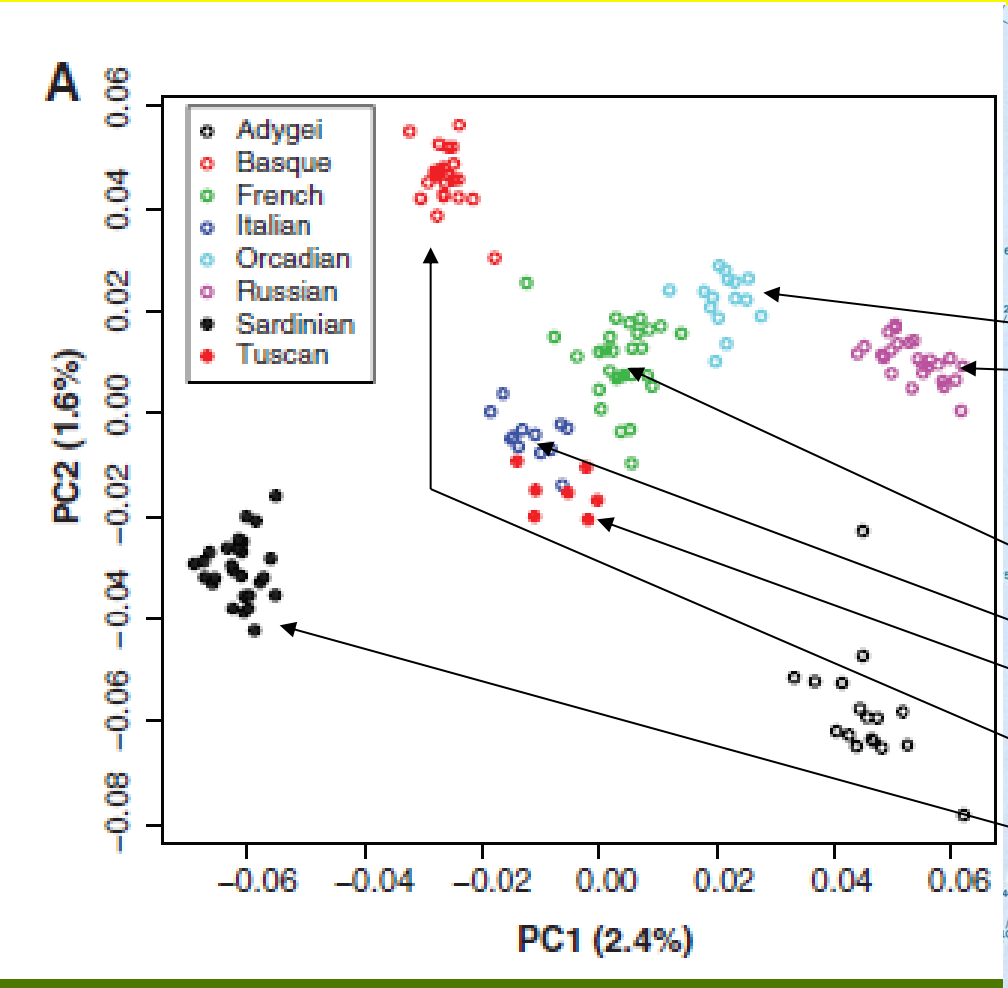
Li et al. Science 2008

Maximum likelihood tree on SNPs in 51 human populations



Li et al. Science 2008

Principal component analysis to separate human populations based on SNPs



How can we use the SNP arrays for lion?

At the species level

We will genotype at least 30 samples (the more the better) from the entire distribution range of the lion (and 1 *Panthera pardus/tigris*)

This will assure delimitation of species, subspecies (and populations)

Indicate the “origin” of the lion (using a phylogenetic analysis)

Estimate the genetic diversity and heterozygosity throughout the distribution range

Find markers for morphological traits (manelessness, skin color, skinfolds, etc)

How can we use the SNP arrays for lion?

At the population level

We will genotype 6 (or more) populations from small to large to relate genetic diversity and homozygosity to population size.
This will create benefits for local lion researchers in terms of genetic information (eg. Etotepe Sogbohossou, Talatu Tende, Pricelia Tumenta , Tuqa Jirmo)

Determine genetic diversity between populations (gene flow)

At the individual level

Reliably estimate the level of genetic diversity and heterozygosity

Determine paternity and kinship
(Avoid inbreeding in zoo's, relationships in prides)

How can we use the SNP arrays for lion?

For an unknown lion sample we can estimate:

1) The “subspecies”

2) Its “genetic” origin (by PCA or phylogenetic analysis, based on nuclear or mtDNA using eg *P. pardus* as an outgroup).
(Important for illegal trading)

3) Reliably estimate the level of genetic diversity and heterozygosity
(important for zoo captive breeding programs)

4) If the sample is a genetically “mixed” individual from different lion populations

5) (Hybridisation with other species)

Problems

SNP discovery

For SNP discovery little problems are foreseen and they should be developed within a year.

SNP genotyping

Is technically feasible and should present no problems for good quality DNA. Success for DNA from scat samples is unsure.

Financially still problematic as more than 1000 chip arrays have to be ordered at once.

So an investment of circa 100.000 euro is needed

Conclusions

Next generation sequencing in combination with high throughput SNP genotyping offers unprecedented opportunities for conservation genetics and the answering of evolutionary and ecological questions in lions (and in any other species)