

NGS for ecologists: DNA metabarcoding

François Pompanon & Pierre Taberlet

Laboratoire d'Ecologie Alpine, CNRS UMR 5553

Université Joseph Fourier, Grenoble, France





What determine patterns of species diversity?

**Foodwebs:
niche partitioning,
energy flows in
ecosystems ...**



Barcoding : an old concept ...

Kangethe, E.K., *et al.* **1982**

Identification of the species origin of fresh meat using an enzyme-linked immunosorbent-assay procedure. *Meat Science* 7, 229-240

... and recent developments

Biological identifications through DNA barcodes

**Paul D. N. Hebert^{*}, Alina Cywinska, Shelley L. Ball
and Jeremy R. deWaard**

Different approaches of DNA Barcoding

Sensu stricto

« short DNA sequence from a standardized and agreed-upon position in the genome as a molecular diagnostic for species-level identification »



Sensu lato

« DNA-based taxon identification using diverse techniques »

Review

Cell
PRESS

DNA barcoding for ecologists

Alice Valentini^{1,2}, François Pompanon¹ and Pierre Taberlet¹

DNA metabarcoding

Combination of the barcoding concept
and Next generation sequencing

→ High throughput plant & animal
identification

DNA metabarcoding

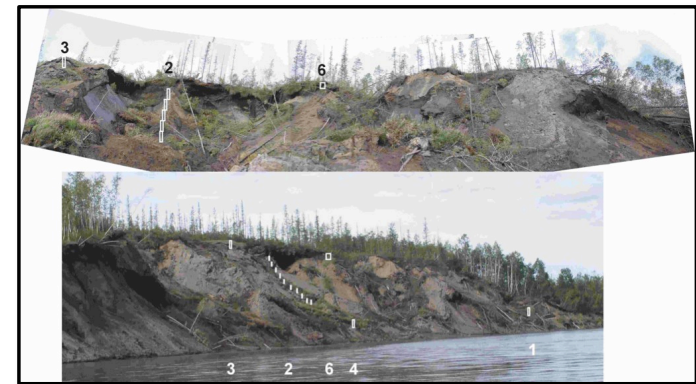
An alternative to traditional methods



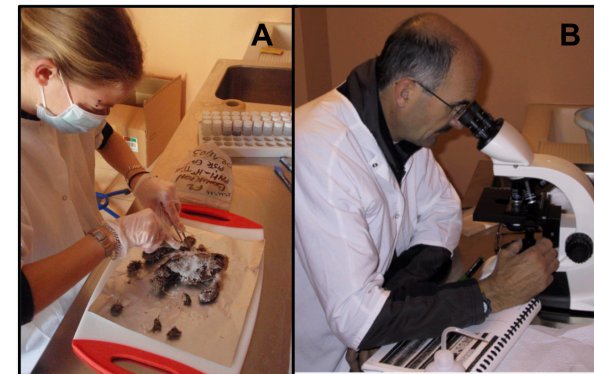
Too many species ...



Too secretive species ...

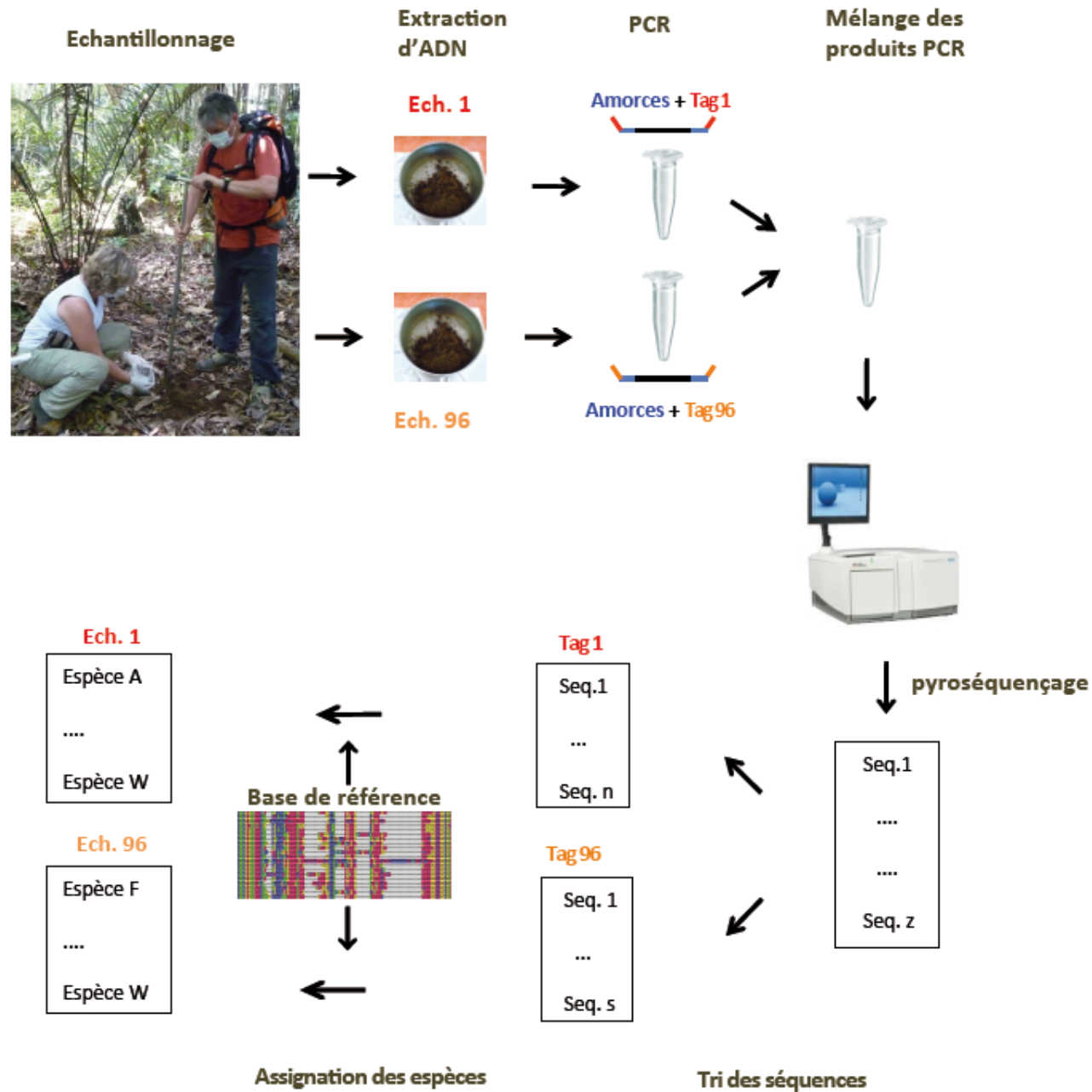


Too old ecosystems ...



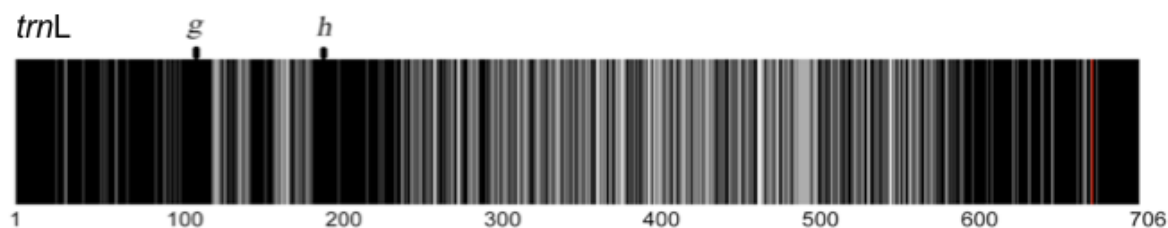
Too many time...

The metabarcoding approach



Constraints of working with degraded DNA

A barcode should be		taxonomy	ecology
Discriminant	Low variation within species & high variation between species	✓ ✓	✓
Standardized	One region for as many taxa as possible	✓ ✓	✓
Phylogenetically informative	Assigning unknown species to a taxon (genus, family, etc.)	✓ ✓	✓
Robust	Highly conserved priming sites, and highly reliable DNA amplifications	✓	✓ ✓
Short	Allowing amplification from as many substrates as possible (even degraded)	✓	✓ ✓



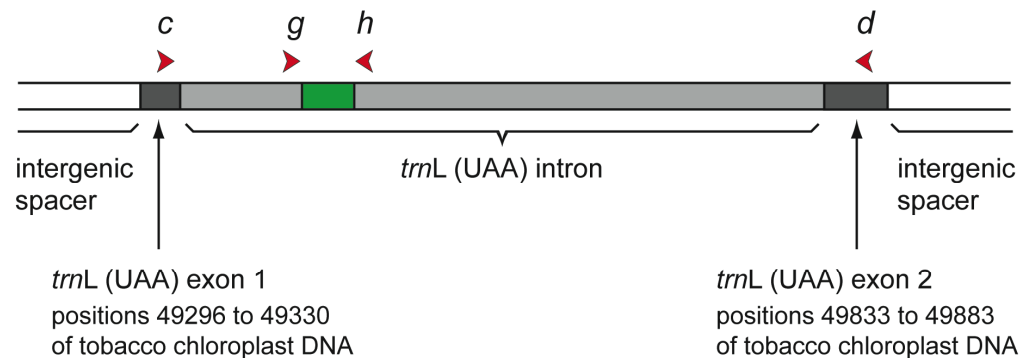
sampling



DNA amplification

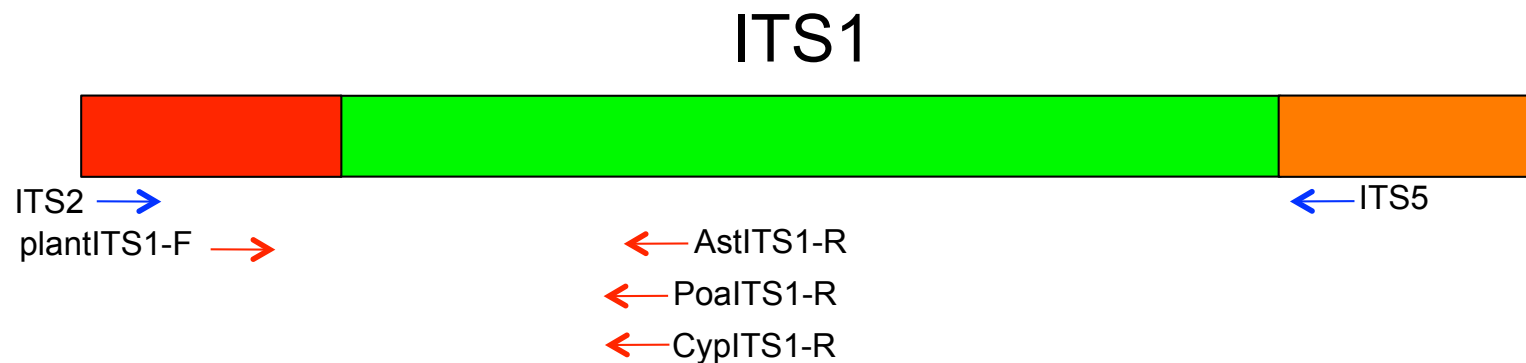
- Using universal primers amplifying a very short DNA fragment (< 100 bp)
- Possibility of using many primer pairs targeting different taxonomic groups

The chloroplast
trnL(UAA) intron



New ITS barcode primers for Asteraceae, Poaceae, and Cyperaceae

- Asteraceae
Length: < 90bp; 68% Asteraceae with a maximum of 2 mismatches
- Poaceae
Length: 54bp- 88bp; 98.5% Poaceae with a maximum of 2 mismatches
- Cyperaceae
 - Length: 46bp - 78bp; 93% Cyperaceae with a maximum of 2 mismatches



sequencing



400 bp fragments
 10^6 reads /run

GS FLX Roche Diagnostic®



100 bp fragments
(2x100 paired-ends)

$2 \cdot 10^9$ reads / run

HighSeq 2000 Illumina®



taxonomic assignment

- Reference database (extracted from public databases / built on purpose)
- Difficult with very short barcodes (identification of families, genera, or species)
- Possibility to work with a low resolution or without taxonomic assignment (Ecological clades / MOTUs)



Search

[Login](#) | [Help/Guide](#) | [About Trac](#) | [Preferences](#)

[Wiki](#) | [Timeline](#) | [Roadmap](#) | [Browse Source](#) | [View Tickets](#) | [New Ticket](#) | [Search](#)

[Start Page](#) | [Index](#) | [History](#) | [Last Change](#)

OBITools

OBITools is a set of python scripts developed to simplify manipulation of sequence files in our labs. They were mainly designed to help us during the analysis of Next Generation Sequencer outputs (454 or solexa). These scripts rely on a Python library of the same name "["obitools"](#)".

The library and the scripts evolved with our needs. They mainly provide facilities to manipulate sequence files in [Fasta format](#). Other formats are supported as genbank or embl.

New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trnL* approach

ALICE VALENTINI,^{*,†} CHRISTIAN MIQUEL,^{*} MUHAMMAD ALI NAWAZ,^{‡,§} EVA BELLEMAIN,^{*} ERIC COISSAC,^{*} FRANÇOIS POMPANON,^{*} LUDOVIC GIELLY,^{*} CORINNE CRUAUD,[¶] GIUSEPPE NASCETTI,[†] PATRICK WINCKER,[¶] JON E. SWENSON,^{‡**} and PIERRE TABERLET^{*}



Mollusks

Deroceras reticulatum

Arion rufus

Helix aspera



Insects

Gonfophocerippus rufus

Chorthippus biguttulus



Birds

Tetrao urogallus aquitanicus

Tetrao urogallus major



Mammals

Ursus arctos

Marmota caudata

Marmot and bear diets



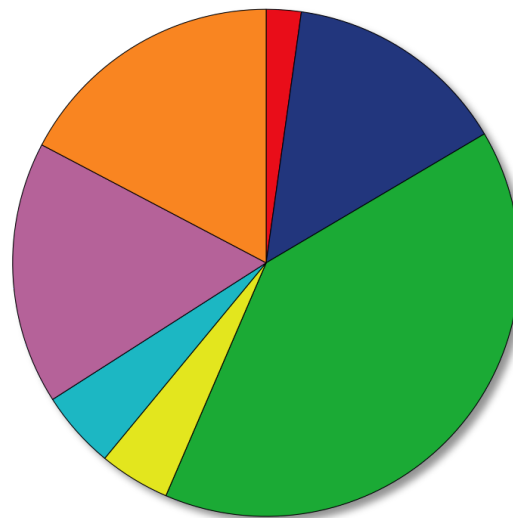
Deosai National Park, Pakistan



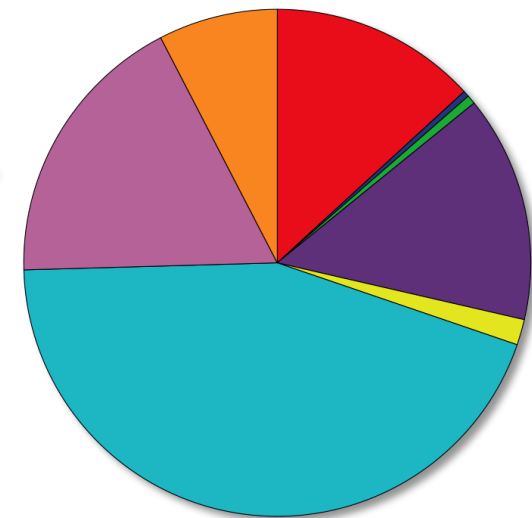
Golden marmot



Brown bear



- Apiaceae
- Asteraceae
- Caryophyllaceae
- Cyperaceae
- Fabaceae
- Poaceae
- Polygonaceae
- Others

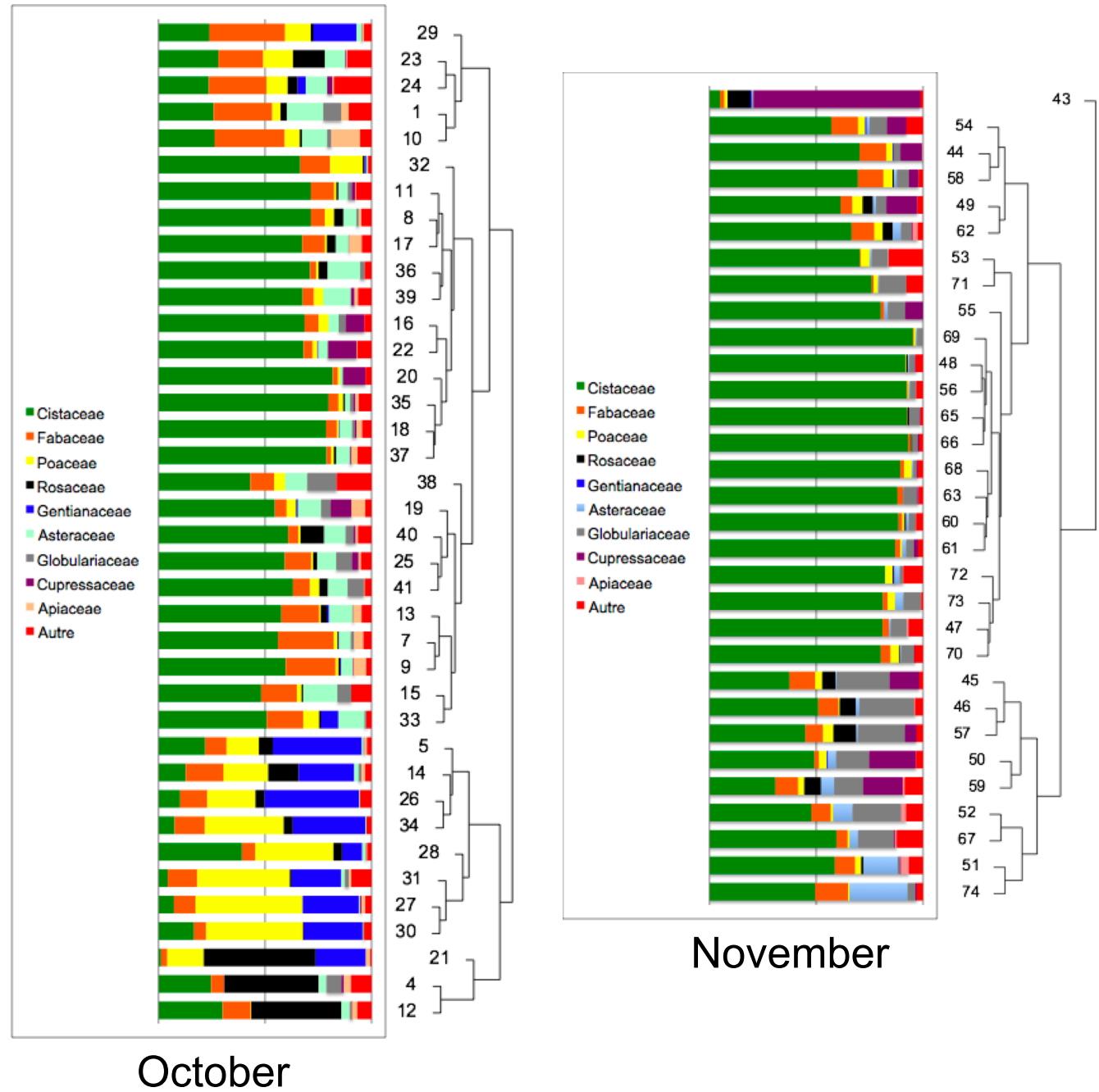




Bauges massif, France

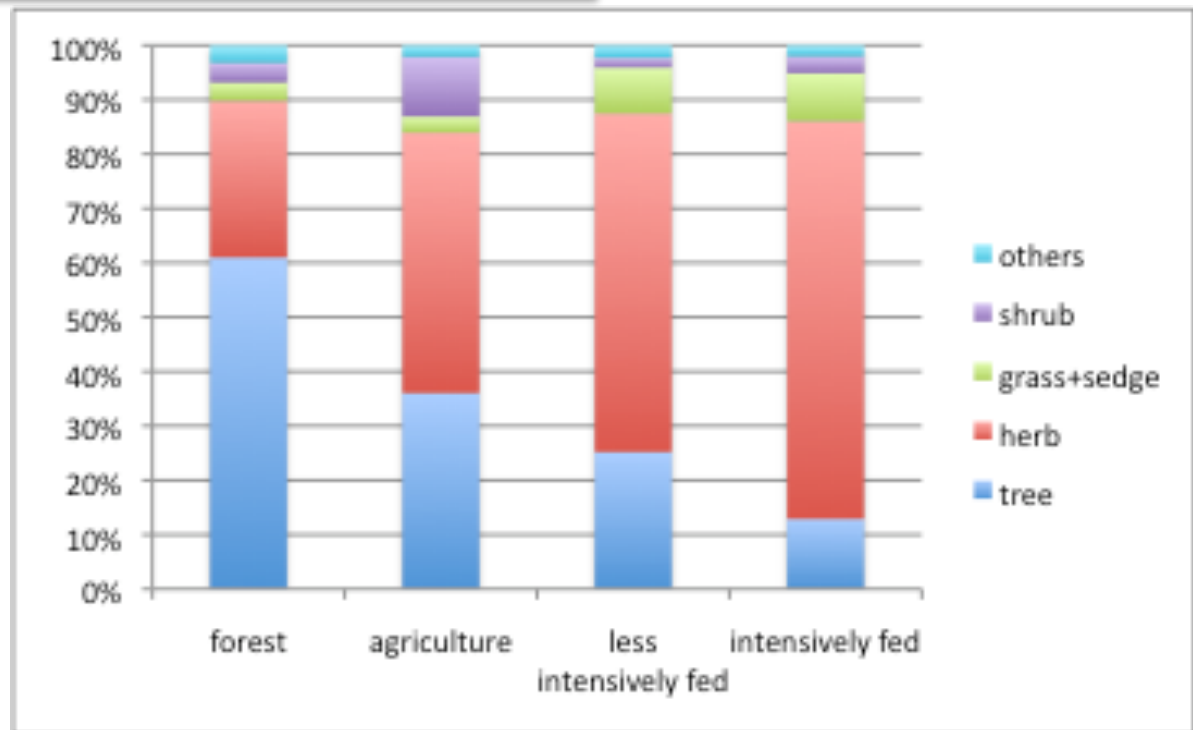


Chamois diet



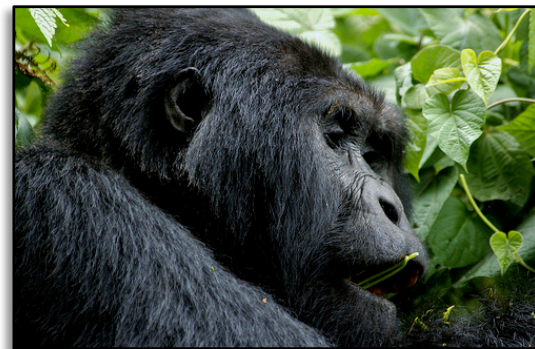


Bison diet in 4 management schemes

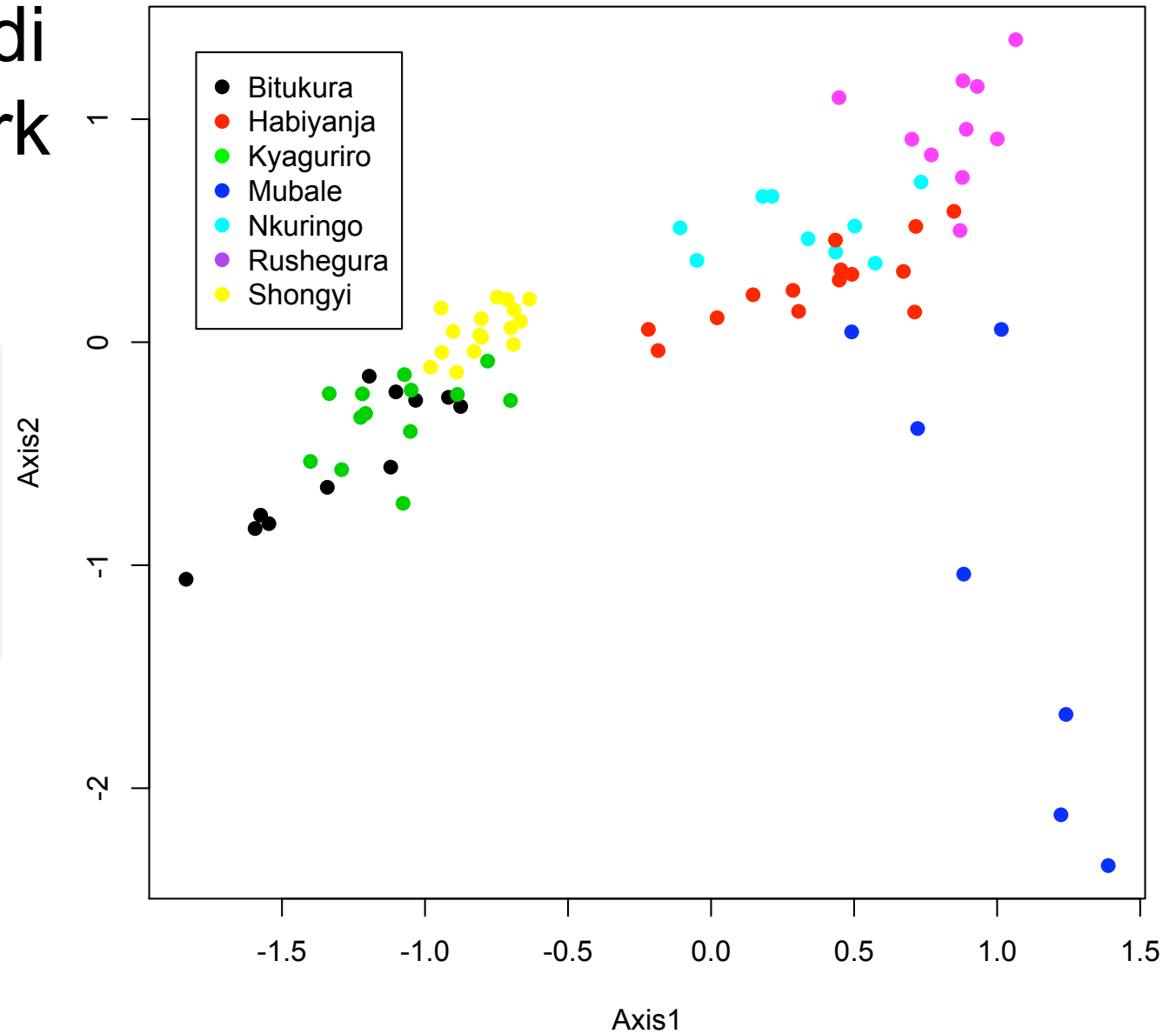


Kowalczyk *et al.* (2011) *Forest Ecology and Management*, **261**, 821-828.

Mountain Gorilla diet in Bwindi National Park (Uganda)



Gorilla diet
Correspondence analysis





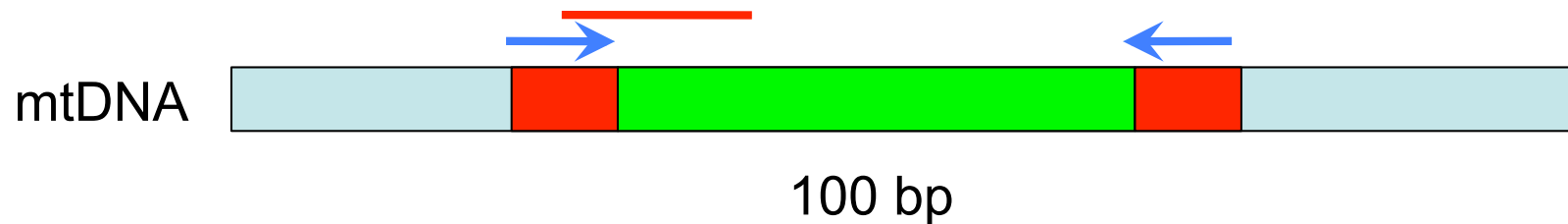
Collaboration with Juliette Ogubi, Innocent Rwego, and Charles Masembe
Makerere University, Kampala, Uganda

Diet analysis of carnivores

- Technically more difficult
 - How to amplify the prey DNA, without amplifying host DNA when using universal primers?
- using a blocking primer for host DNA during the amplification

Experimental protocol

- Sampling in the field
- DNA extraction
- DNA amplification
 - with mtDNA universal primers 
 - with and without blocking primer 



- DNA sequencing (NGS)

Snow leopard diet

Influence of the blocking primer during DNA amplification



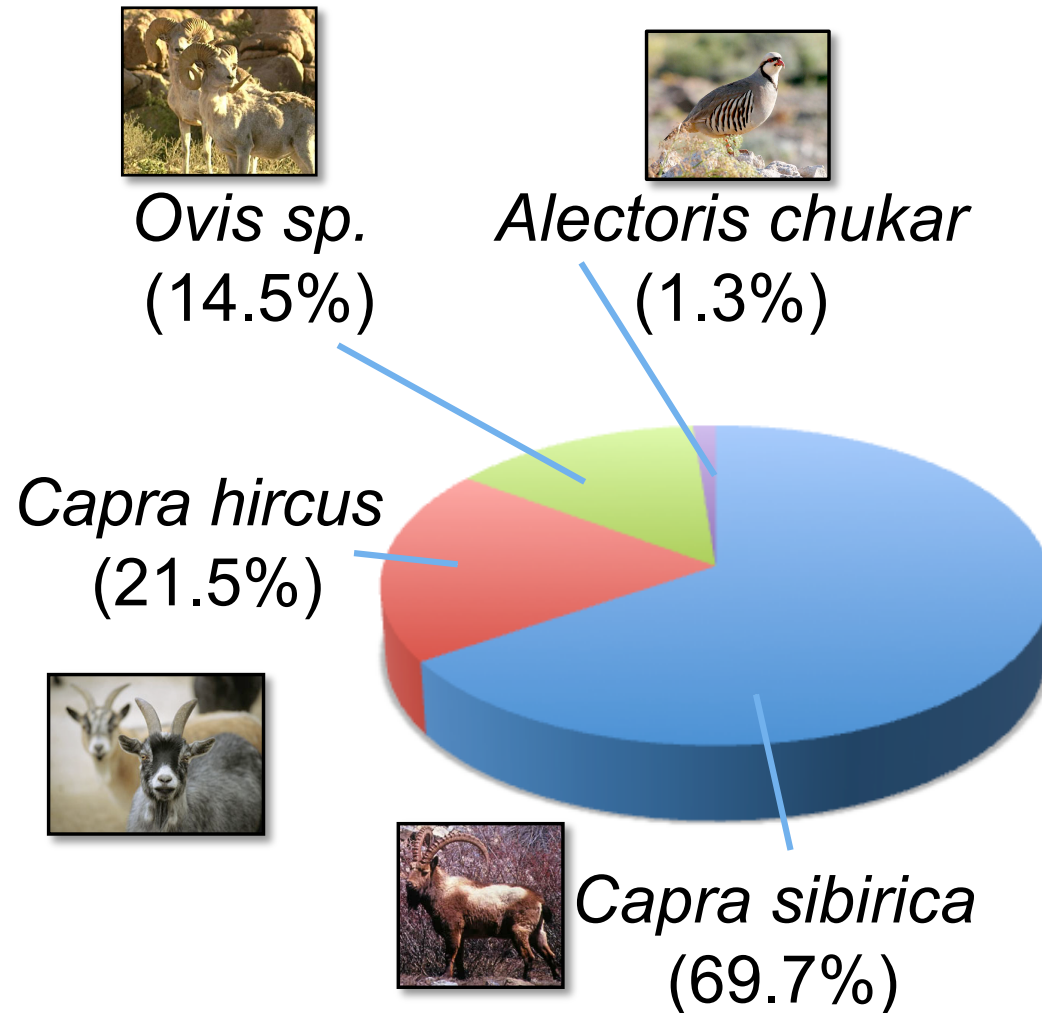
	feces 1		feces 2		feces 3	
<i>Uncia uncia</i>	74%	1%	97%	43%	100%	100%
<i>Capra sibirica</i>	26%	99%	0%	0%	0%	0%
<i>Capra hircus</i>	0%	0%	0%	0%	0%	0%
<i>Ovis sp.</i>	0%	0%	3%	57%	0%	0%

forward primer + reverse primer

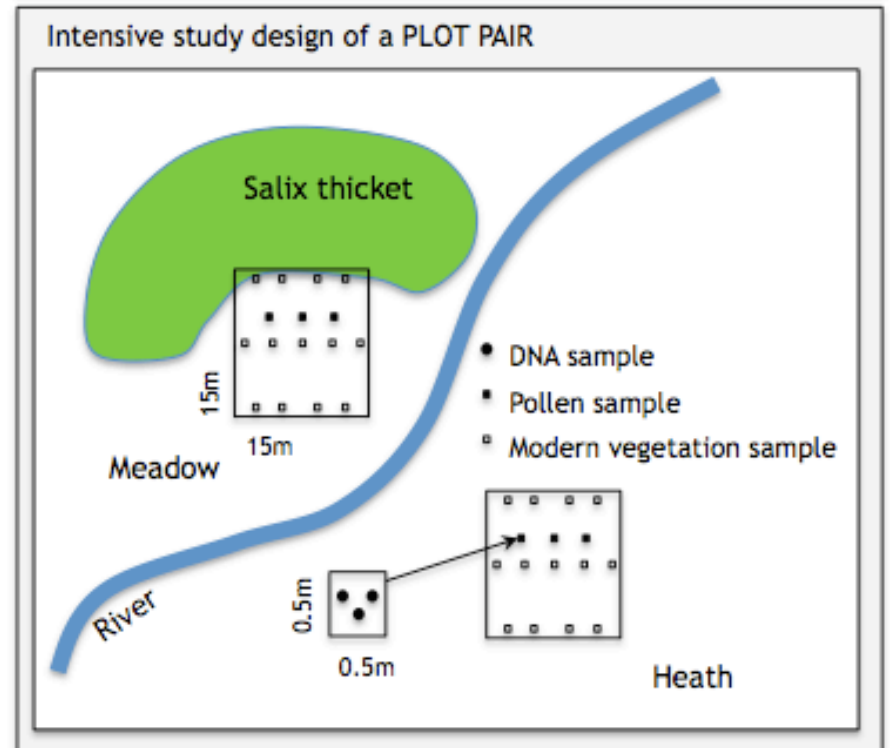
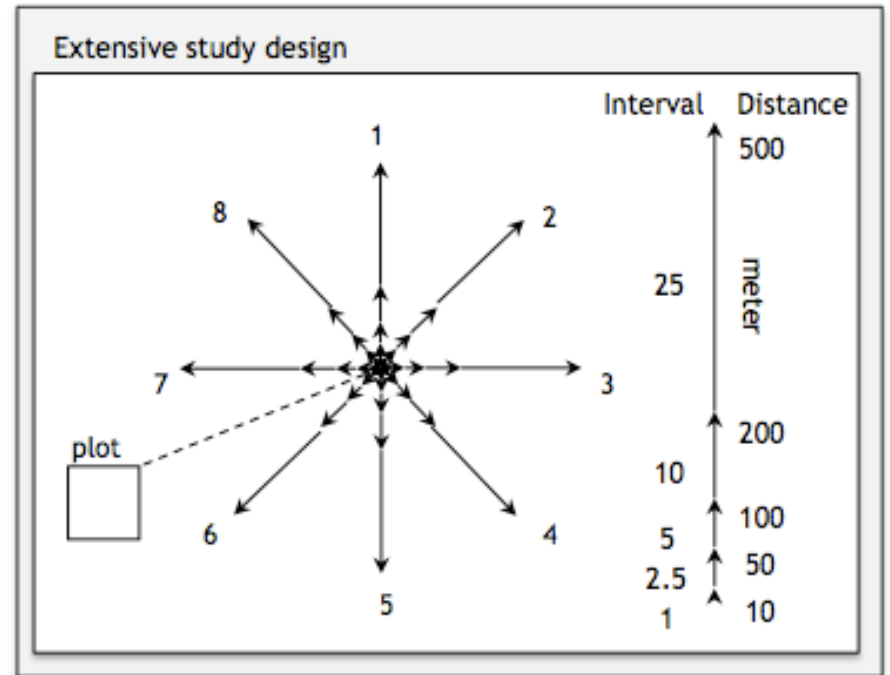
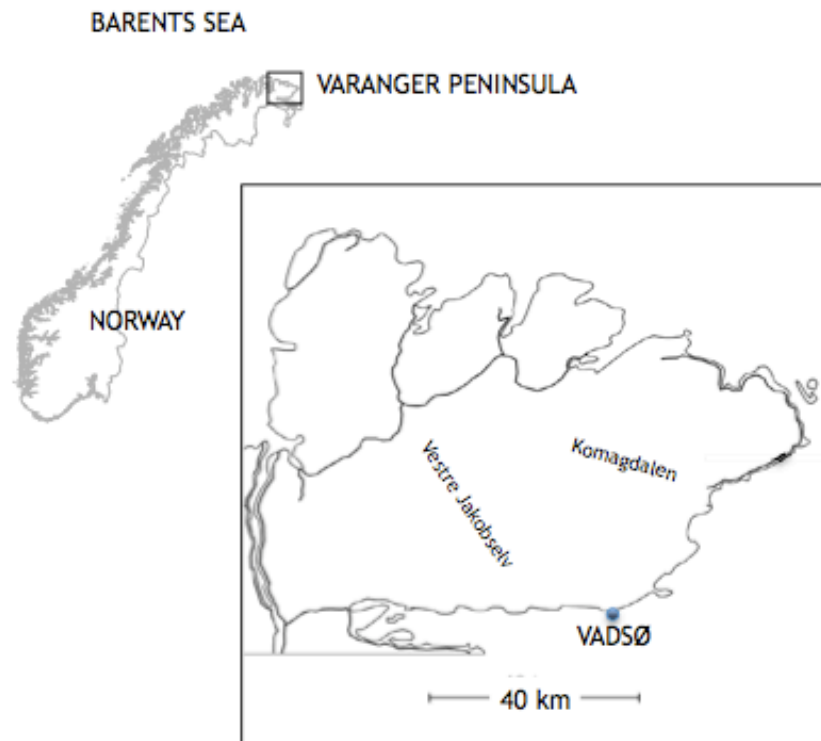
forward primer + reverse primer + blocking primer

Snow leopard diet in Mongolia

70 faeces analyzed



Calibration of soil DNA with aboveground species at Varanger, Northern Norway



Above ground analysis

Avenella flexuosa



Poa sp.



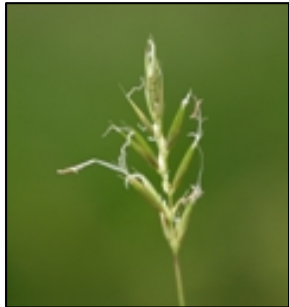
Taraxacum sp.



Carex sp.



Anthoxanthum nipponicum



Alchemilla sp.



Festuca sp.



Viola biflora



Deschampsia sp.



Rumex sp.



Calamagrostis sp.



Salix sp.



Bistorta vivipara



Equisetum sp.



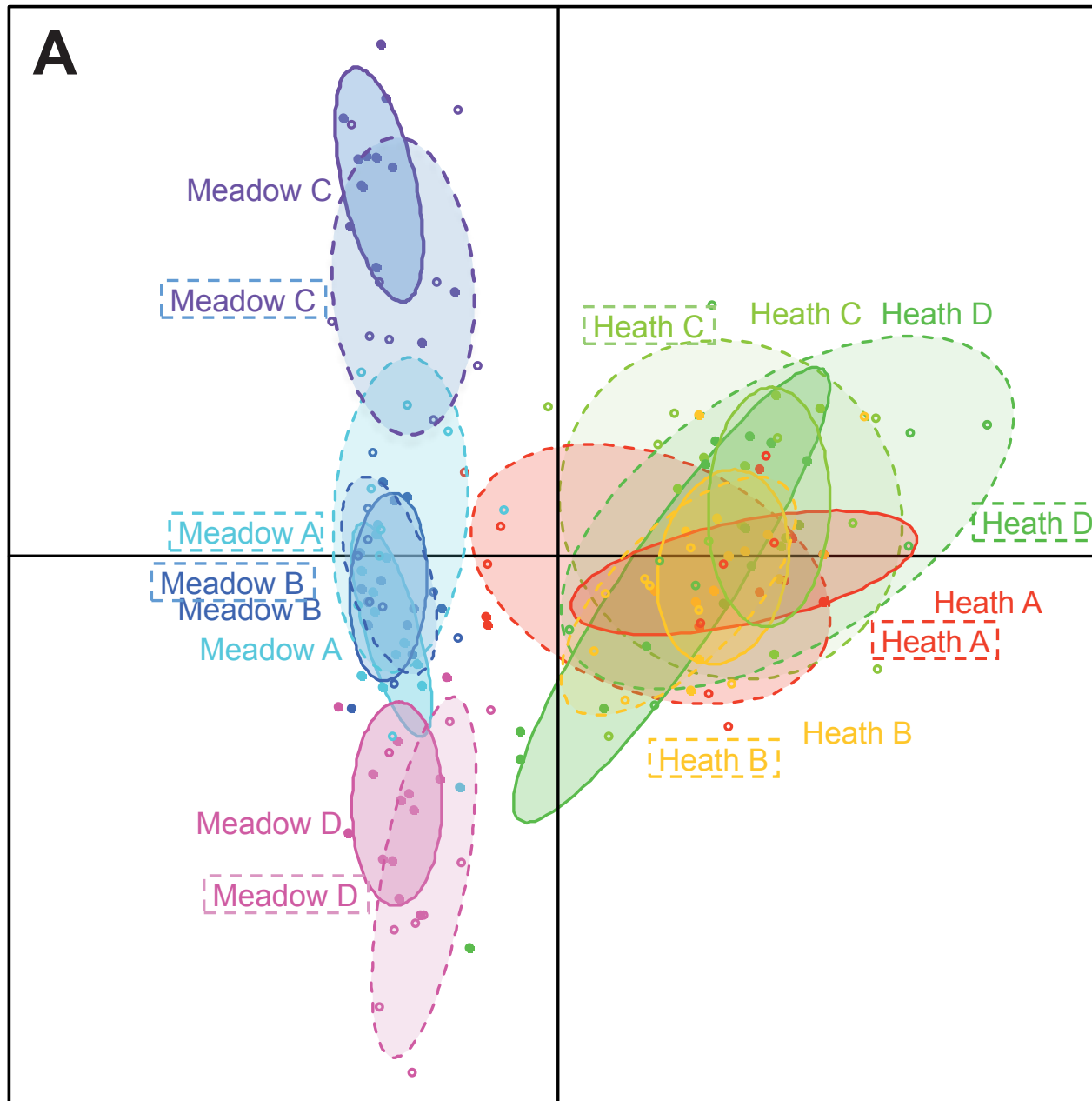
Viola biflora



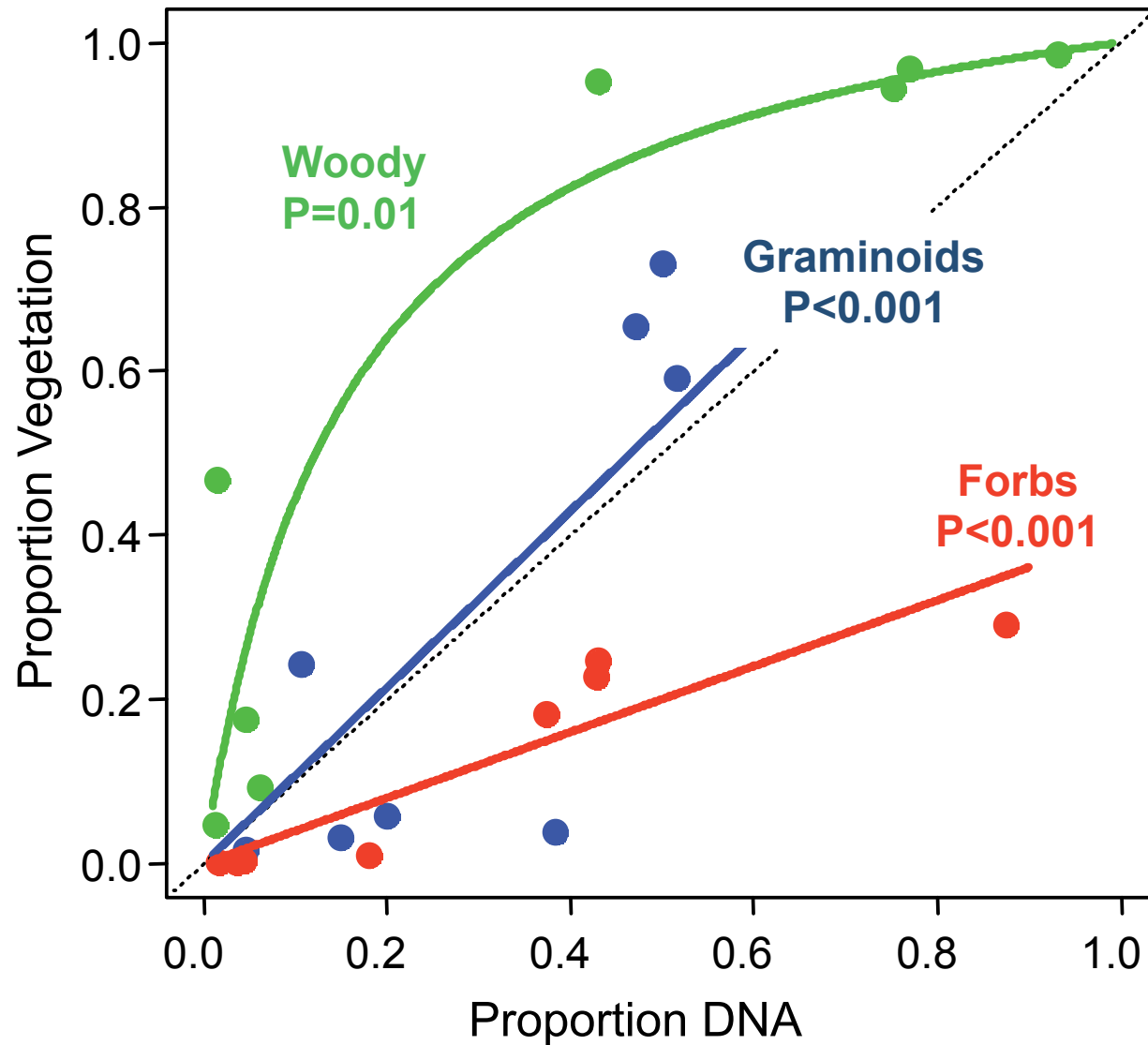
Equisetum sp.



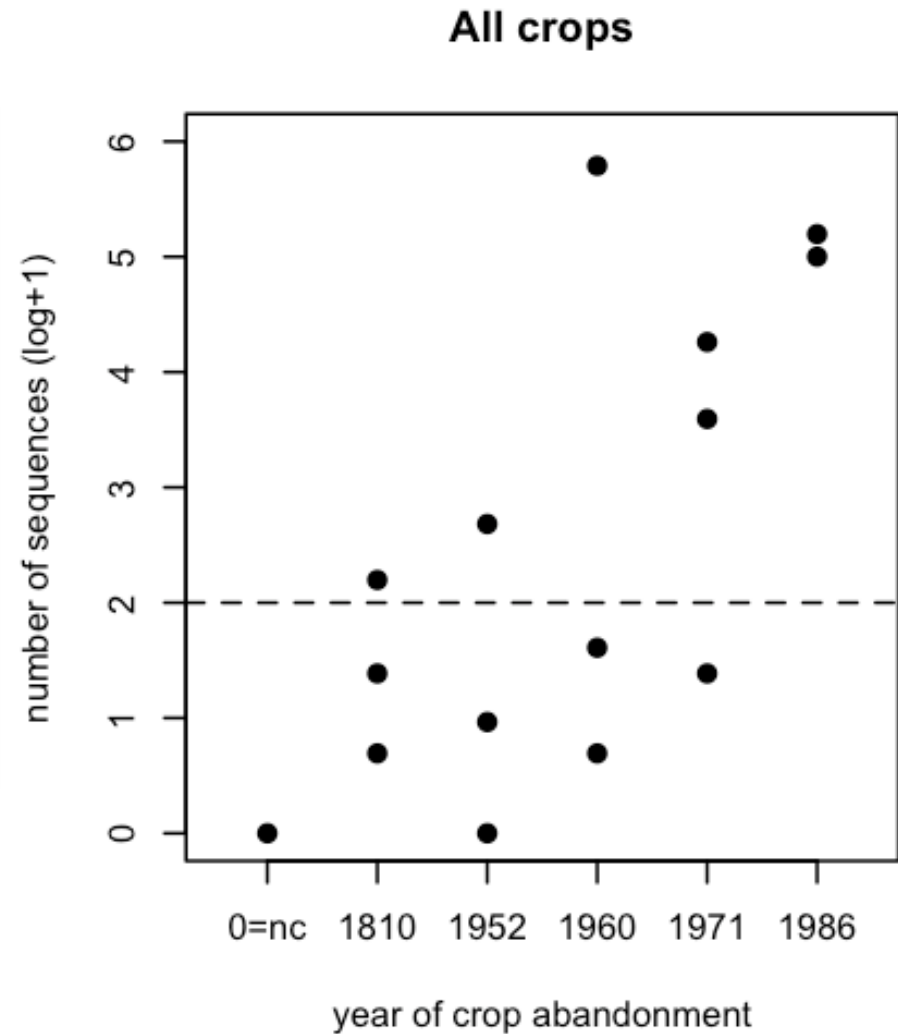
DNA in soil mirrors above ground plant diversity



DNA-based soil analysis: functional groups



How long does a DNA molecule persist in soil?



DNA-based soil analysis: broad applicability



C	Sites	No of core soil samples	No of plots	Sequencing platform	No of sequences analyzed	No of families identified
	Boreal (Varanger Fjord, Norway)	72	8	Roche 454 FLX	176 283	33
	Temperate (French Alps, France)	8	1	Illumina GA IIx	396 054	5
	Tropical (Nouragues Field Station, French Guiana)	49	1	Illumina GA IIx	1 636 455	34

Permafrost analysis

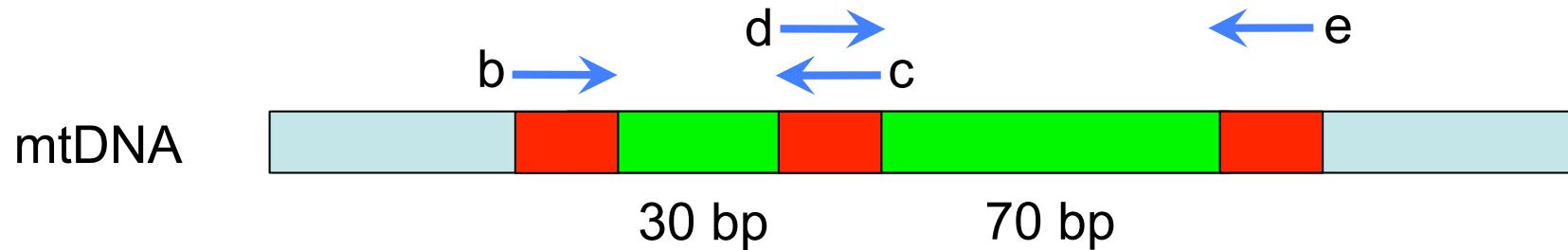


A 22,960 years old
flora

Nb Seq	Family	Genus	Species
937			<i>Bistorta vivipara</i>
482		<i>Equisetum</i>	<i>arvense / fluviatile / sylvaticum</i>
94	Salicaceae	3	<i>Salix sp. / Chosenia arbutifolia / Populus balsamifera</i>
60			<i>Armeria scabra</i>
55			<i>Thymus oxyodontus</i>
43			<i>Lagotis glauca</i>
37	Asteraceae	3	Asteraceae 4
35			<i>Avenella flexuosa</i>
27		<i>Aconogonon</i>	<i>alaskanum / ocreatum / tripterosperrum</i>
26		<i>Rumex</i>	5
19	Asteraceae	2	<i>Packera sp. / Senecio sp.</i>
19	Poaceae		Poaceae 2
16		<i>Ranunculus</i>	<i>acris / subborealis / turneri</i>
15		<i>Festuca</i>	9
13			<i>Hulteniella integrifolia</i>
11			<i>Saxifraga hirculus</i>
9			<i>Trientalis europaea</i>
8	Asteraceae	4	Asteraceae 2
7		<i>Valeriana</i>	<i>capitata / sambucifolia / transjenisensis</i>
6			<i>Myosotis alpestris</i>
6	Asteraceae	5	Asteraceae 3
6		<i>Empetrum</i>	<i>sibiricum / subholarcticum</i>
5			<i>Anthoxanthum nipponicum</i>
5			<i>Crepis chrysantha</i>
4		<i>Saxifraga</i>	5
3		<i>Papaver</i>	19
3	Poaceae	2	<i>Elymus sp. / Leymus sp.</i>
3			<i>Trollius europaeus</i>

Earthworms from soil DNA

- Eight soil samples collected per plot
- Universal short barcodes for earthworms
- Reference database built using samples identified with the standardized COI barcoding approach



Earthworms from soil DNA: results

Species	Barcode	Chartreuse		Grenoble	
		Plot 1	Plot 2	Plot 1	Plot 2
<i>Aporrectodea icterica</i>	catcttaatgaagactaaaacttcactaaa	836954	649677	834031	1359355
<i>Aporrectodea longa</i>	tattttaacaaaaacccaaaaattttcaataaa	2	6	244463	271829
<i>Aporrectodea sp</i>	cattttaataaaaattataaattttactaaa	0	0	236024	236678
<i>Octolasion cyaneum</i>	cattttaatagaagcttactattctaataaa	468462	3823	0	2
Unidentified Oligochaeta	tattttaataaaaatagtaaattttactaaa	334804	96337	0	1
Unidentified Oligochaeta	tattataaatcaattaataattgagcata	0	372828	0	0
<i>Lumbricus terrestris</i>	aattttaataaatataaaaaattttactaaa	0	0	174286	143682
<i>Octolasion tyrtaeum</i>	cattttaatagaaaaataatcctaataaa	306476	0	0	2
Unidentified Oligochaeta	tatcacaatatttatacaataaatattatg	183116	68615	0	0
Unidentified Oligochaeta	tatttttcttatactttagtaaacaaaaa	96924	42148	0	0
<i>Lumbricus castaneus</i>	aattttaataaatataaaaaaattttactaaa	0	0	56	131001
<i>Aporrectodea longa</i>	tattttaacaaaaacccaaaaattttcaataaa	2469	105312	159	145
<i>Allobophora chlorotica</i>	cattttaataaagatataaactttactaaa	0	0	51953	43196
Unidentified Oligochaeta	tattttatttacctaaaacagtaacaaaa	0	0	62901	0
Unidentified Oligochaeta	tatttttcttatactttagtaaataaaaa	592	61802	0	0
Unidentified Oligochaeta	taccttaacaaatattattttcgaag	30571	0	0	0
<i>Aporrectodea caliginosa</i>	tattttaataaaaaaatataaatttttaataa	0	23005	0	0

DNA metabarcoding

- Works extremely well for diet analysis using feces as a source of DNA
 - For herbivores
 - For carnivores
 - Can be adjusted for other type of diet
- Works well for plant biodiversity using soil samples
 - In the Arctic
 - In the temperate region
 - In the tropical region
- Need to be further adjusted for animal biodiversity using soil samples
 - Optimistic for animals with high biomass (earthworms, etc.)
 - Might be problematic for animals with low biomass

Acknowledgements

Thank you for
your attention

Rike Bienert
Kari Anne Bråthen
Christian Brochmann
Anne Krag Brysting
Eric Coissac
Corinne Cruaud
Francesco de Bello
Mary Edwards
Francesco Ficetola
Frédéric Gavory
Ludovic Gielly
James Haile
Christelle Melo de Lima
Christian Miquel
Stéphanie Pellier-Cuit
Delphine Rioux
Wasim Shehzad
Jorn Henrik Sønstebø
Wilfried Thuiller
Alice Valentini
Eske Willerslev
Patrick Wincker
Nigel Yoccoz