Centre for genetic identification

Naturhistoriska riksmuseet

Thomas.Lyrholm@nrm.se

DNA ’barcoding’ and eDNA

Naturhistoriska riksmuseet

DNA ’barcoding’

- Species are genetically unique
- Identification by short, standardised DNA sequences
- Globally accessible reference databases (BOLD) (connected to GenBank)
- Taxonomic vouchers!
- Global coordination: *CBOL Consortium and iBOL* *www.barcodeoflife.org*

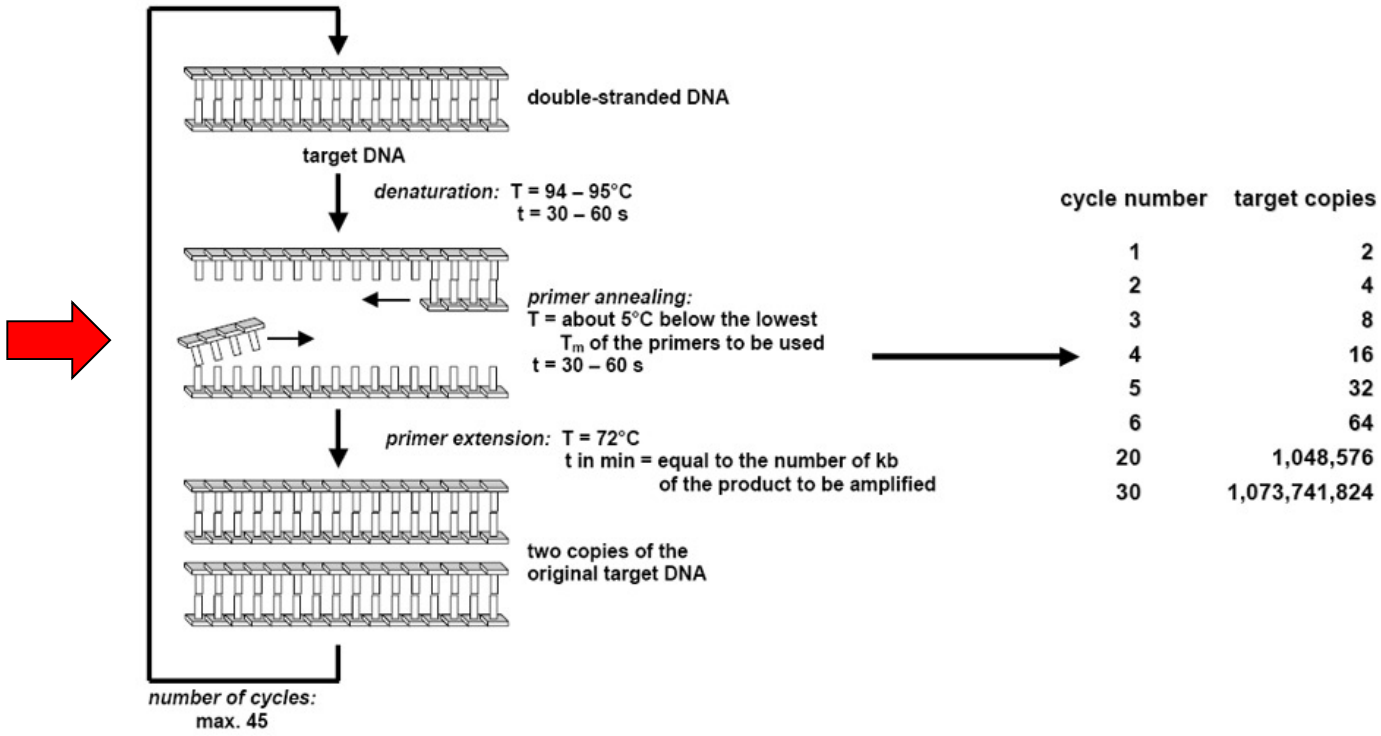
500 000 species (2015)

Naturhistoriska riksmuseet

Taxonomic identification by DNA

- Amplify by PCR and sequence the DNA fragment
- Compare against the taxonomically validated sequence database
- Taxonomic assignments based on sequence similarity
  - ’threshold’ (e.g.. 99% similarity=same species)
  - Tree clustering (e.g. Neighbour joining)

Primers och probes used to copy by PCR and analyze



cycle number	target copies
1	2
2	4
3	8
4	16
5	32
6	64
20	1,048,576
30	1,073,741,824

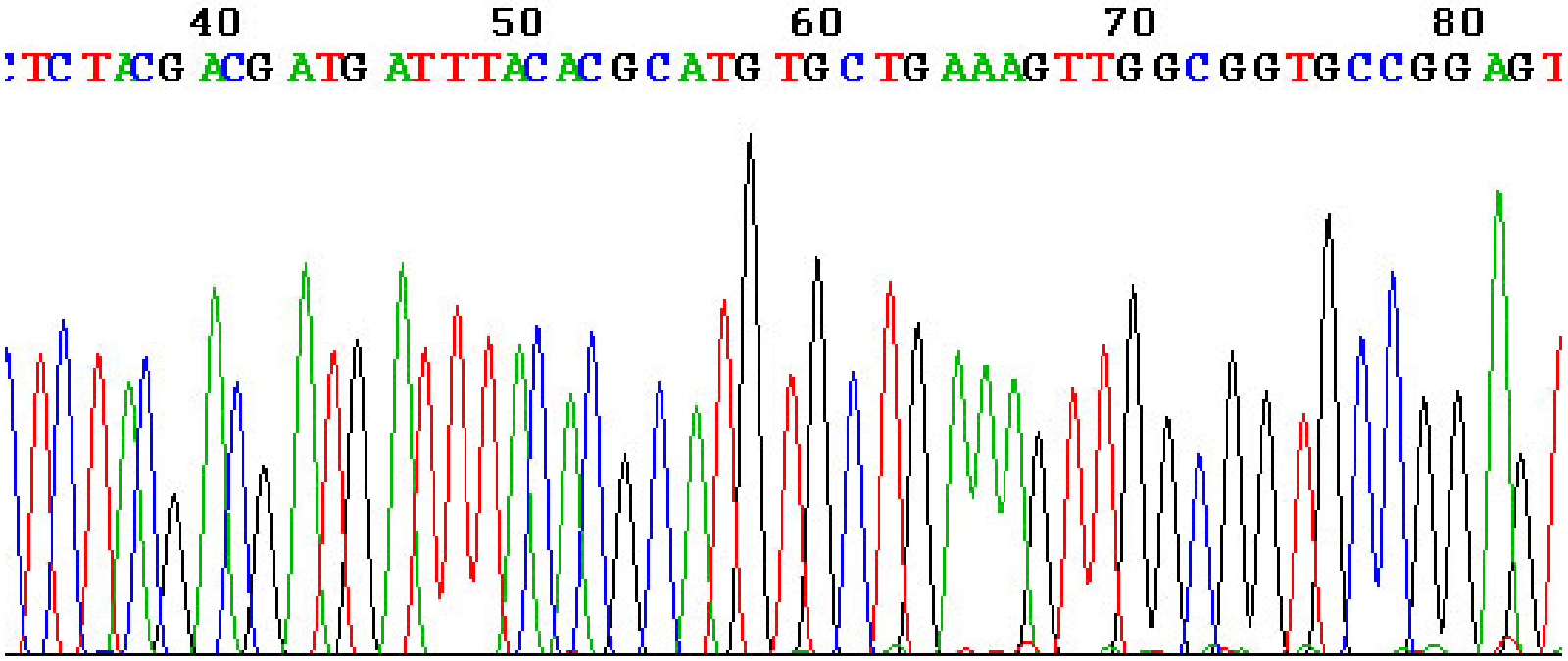
Naturhistoriska riksmuseet

Suitable primers for general species identification

- The amplified sequence should show little variation **within** and high divergence **between** species
- Specific, robust PCR amplification with ’universal’ primers
- Phylogenetically informative for species delineation
- Not ’too long’ fragment
- Animals: mtDNA CO1 (658 bp)
- Plants: chl rbcL (500 bp); matK (800 bp)

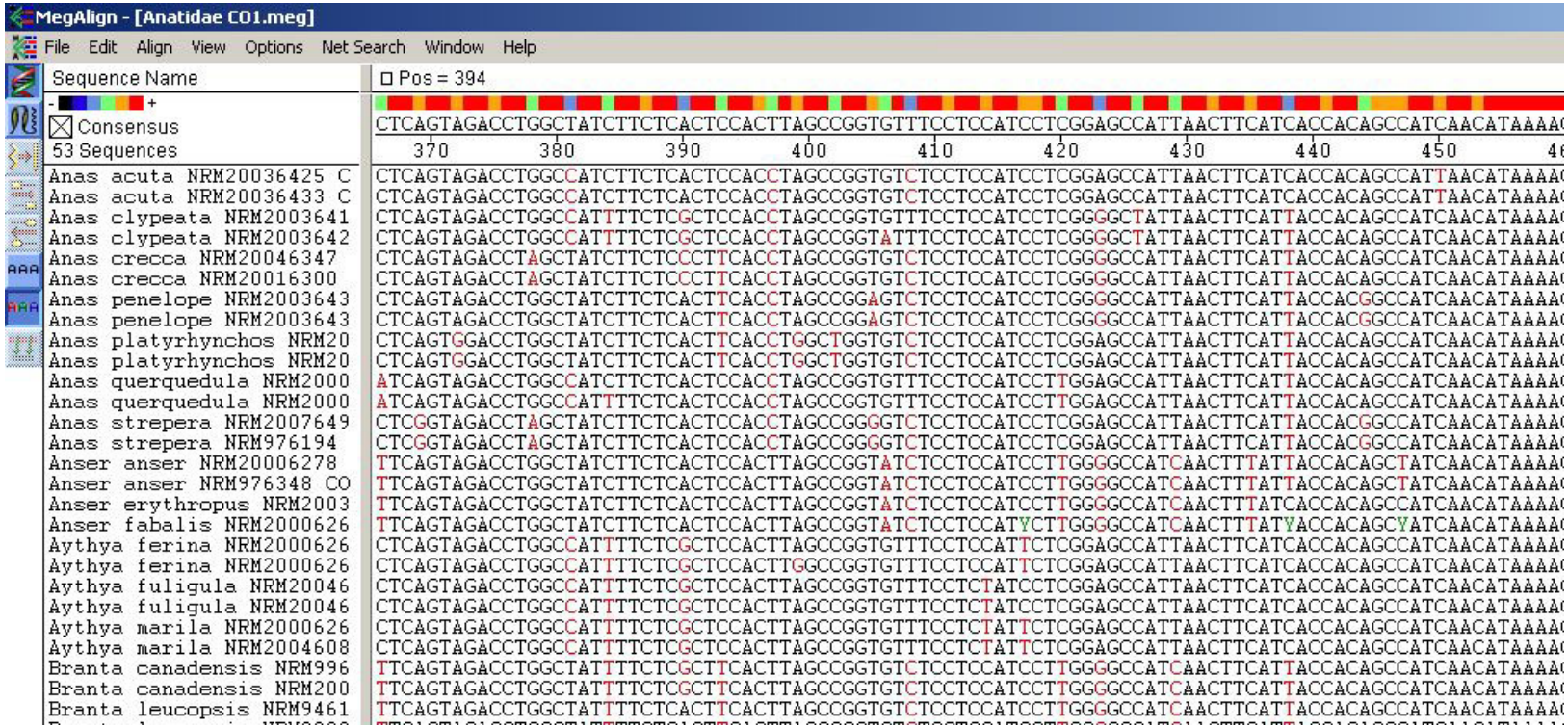
Naturhistoriska riksmuseet

Sequencing



Naturhistoriska riksmuseet

Database comparison



Improvements to environmental monitoring

- More reliable identification of taxa –to species level
- Very small amounts of source DNA needed (hair, exkretions, pawprints, insect parts, cell….)
- Expansion to new taxa otherwise difficult to identify morphologically; cryptic species; biodiversity assessments
- Any lifestages can be used (any season)
- Less dependent on taxonomic expertise
- Time and cost savings
- Early warning (invasive species, pests, etc.)
- Non-invasive tracking of rare and endangered species
- Data accessible and searchable

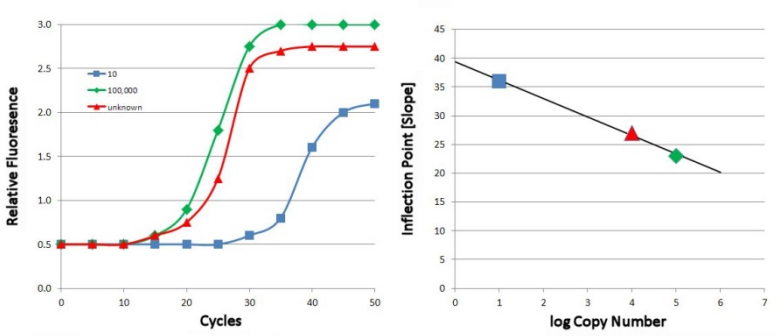
Monitoring questions

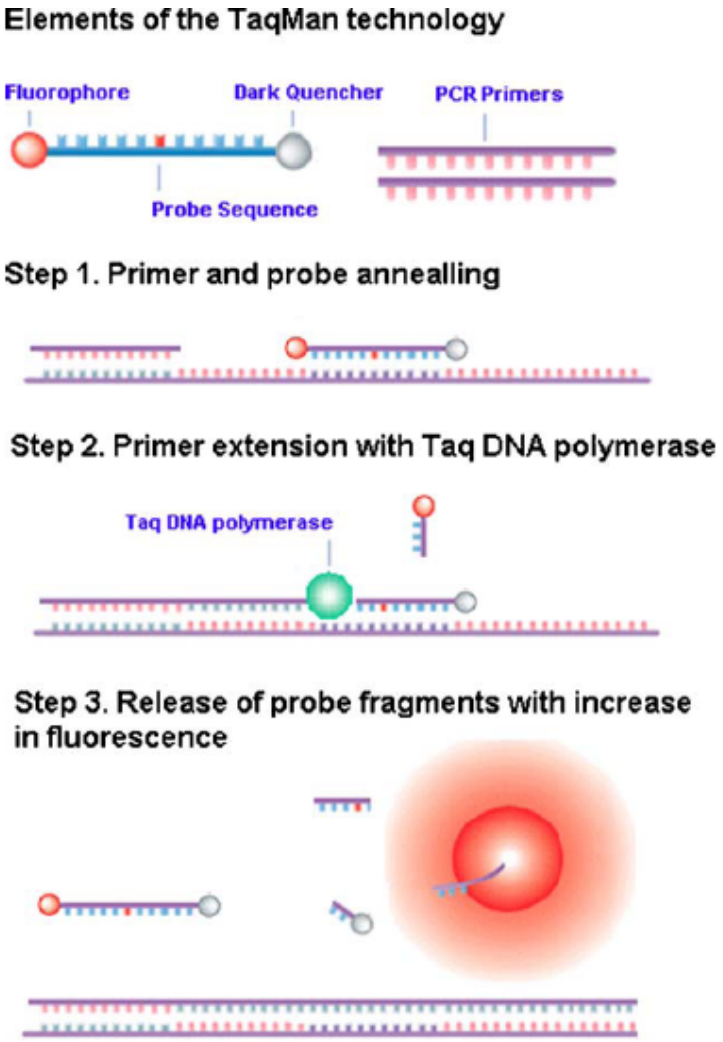
- *Which species?*
  - ’Barcoding’ – compare barcode sequence to database
- *Is species A present?*
  - Detect species specific sequence in sample
- *Species composition in community?*
  - ’Metabarcoding – identify many species in bulk sample
- *Is there a lot or little of species A?*
  - Species specific detection with quantitative PCR
- *What are the relative abundances of A, B, C..?*
  - Multiplex quantitative PCR; ’Next Generation Sequencing’

Quantitative PCR, qPCR

Från Gibson (2006),
Clinica Chemica Acta **363**

Real time





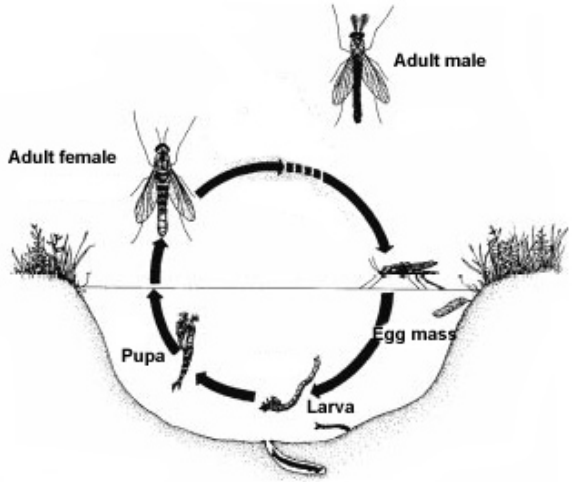
Naturhistoriska riksmuseet

Improved envirionmental monitoring in the Baltic Sea by DNA barcoding of Chironomids

- Naturvårdsverket, Havs-och Vattenmyndigheten, Gunilla Ejdung
- NRM: Yngve Brodin, Bodil Cronholm, Erik Ersmark, Veronika Nyström Edmark, Jonas Strandberg

Chironomidae (non-biting mites) (Diptera)

C:a 650 species in Sweden, about 230 in the Baltic Sea
(Brodin 2011)



Chironomid larvae

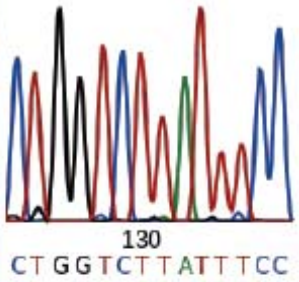
- Ecologically important; up to 30% of macro zoobenthos biomass
- Important environmental indicators, different species vary in sensitivity (oxygen depletion, eutrophication)
- Difficult to identify morphologically – experts identify <20%
- Lumped into ’Chironomids’ in traditional environmental monitoring
- Included in Benthic Quality Index (BQI)

Naturhistoriska riksmuseet

DNA barcoding of chironomids

- **Aim**: Reliable and cost-effective species identification
- **Purpose**: Improved assessment of biodiversity and of the environmental quality index BQI
- **Questions:** Which species occur in the benthic samples from different stations? In what proportions? How does this relate to environmental variables?

samlingar



DNA analys

By Everaldo Coelho and Yellowicon

Matchning mot referensdatabas

Provtagning

Artlista

Chironomus aprilinus
Chironomus pallidivittatus
Chironomus plumosus
Chironomus pseudothummi
Chironomus salinarius
Cladopelma virescens
Cladotanytarsus difficilis
Cladotanytarsus manous
.....

Miljötillståndsindex

Naturhistoriska riksmuseet

Present reference database

- > 800 sequenced individuals from the Swedish Baltic coast
- Standard barcode (mtDNA COI 650 bp)
- About 150 identified species (of about 200)
- near 100% accurate species identification

’NJ’ tree of chironomids

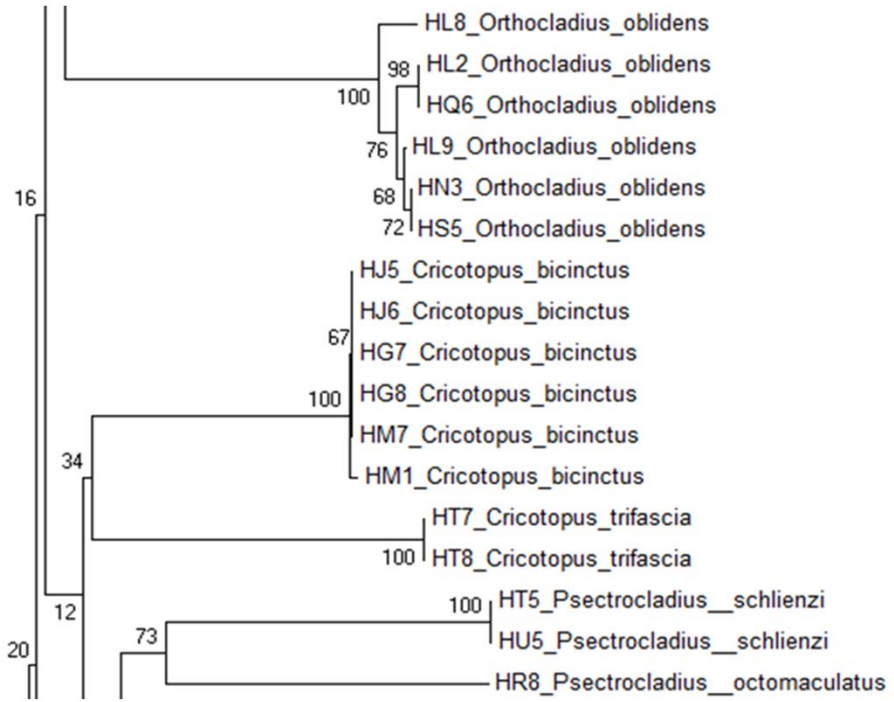
Monofyletic species clades with short branches

Long interspecific branches

Naturhistoriska riksmuseet

Different individuals of the same species in the same clade

Reliable species identification



Invasive alien species

Naturvårdsverket

Gammarids (pilot study)
Early warning – Where has it spread?

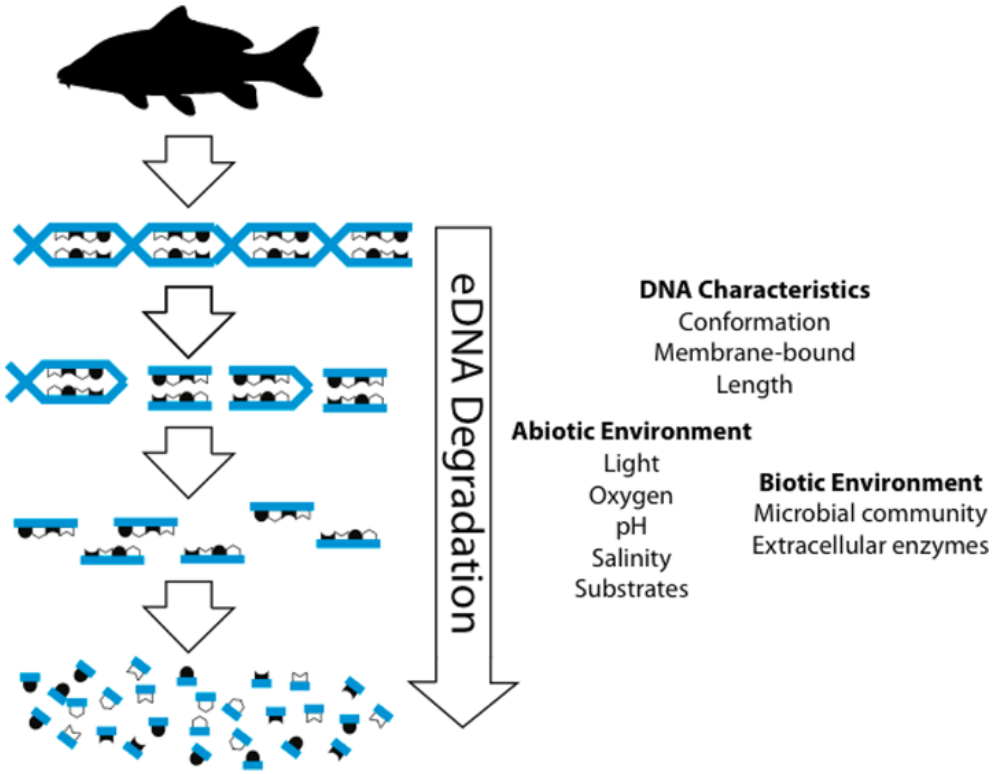
Gammarus tigrinus (invasive)

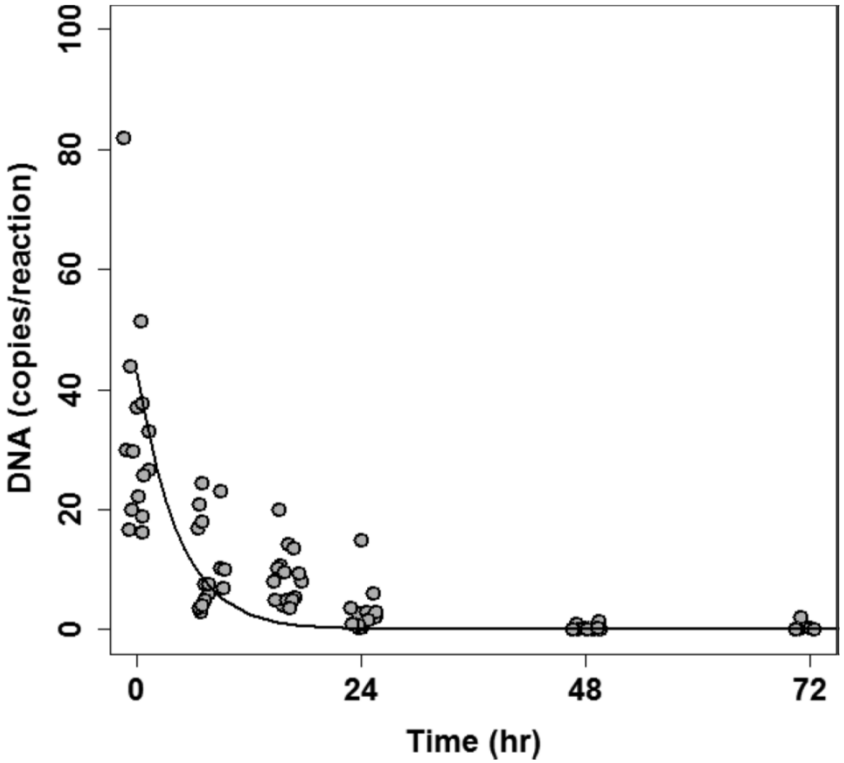
Gammarus zaddachi (native)

Naturhistoriska riksmuseet

eDNA (’environmental DNA’); metabarcoding

- DNA in environmental samples (water, air, soil, etc)
- Diet samples (faeces, gut, excretions, blood, etc)
- No manual separation of individuals
- Complex mixture of organisms/DNA
- Small, degraded pieces of DNA
- Cellular, extracellular, free DNA or particle bound
- New ’minibarcodes’ needed





Small, degraded pieces of DNA

DNA detected in water 1 day-1 week

Barnes et al. 2014, Environ. Sci. Technol. 48

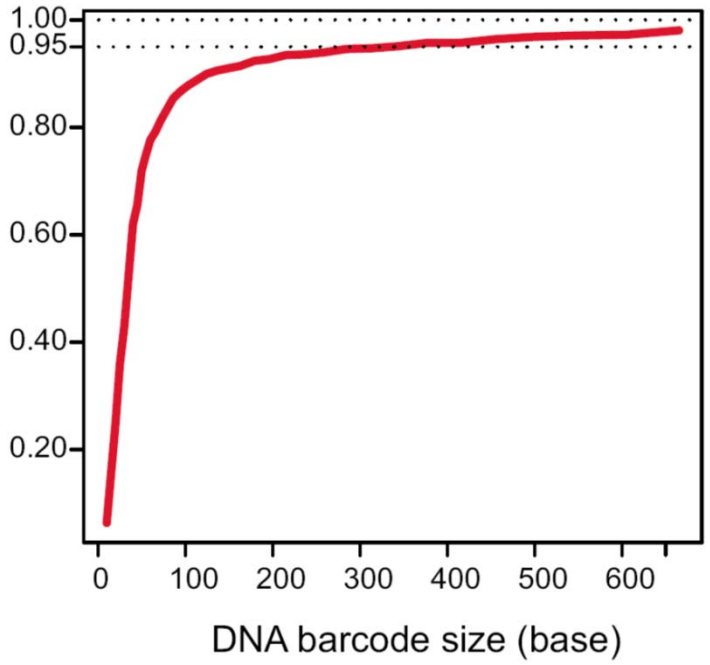
Universal DNA 'mini-barcode' (Meusnier et al. 2008)

Accuracy mtCOI:

650 bp: 97%

250 bp: 95%

100 bp: 90%



Problematic to design universal internal primers covering variable region?

Develop new barcodes.

Naturhistoriska riksmuseet

Bullfrog, 72 bp mtDNA cytb (Ficetola et al. 2008)

Table 1. Rate of bullfrog detection in water samples.

pond	bullfrog presence and relative abundance	water samples positives at least once	positive PCRs
1	yes-low	2/3	2/9
2	yes-low	3/3	6/9
3	yes-low	2/3	2/9
4	yes-high	3/3	8/9
5	yes-high	3/3	6/9
6	yes-high	3/3	8/10
7	no	0/3	0/9
8	no	0/3	0/9
9	no	0/3	0/15

Naturhistoriska riksmuseet

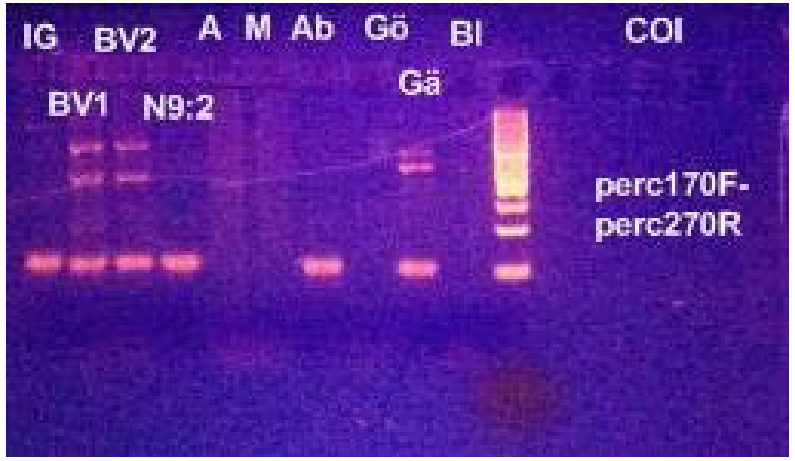
Example: Perch in the water?

PCR primers for perch

Yes! Perch detected (Ab)

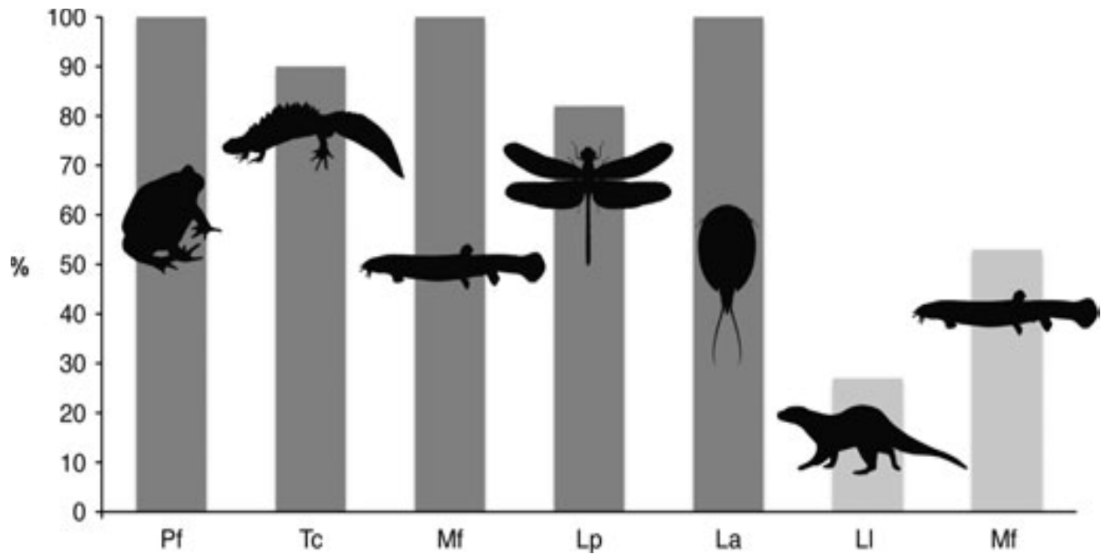
But also ruffe!(Gä)

More specific primers needed (or optimized PCR)

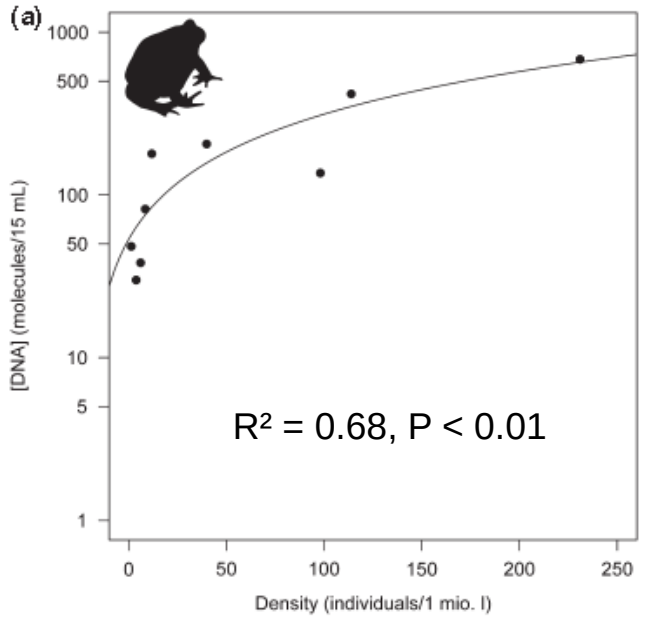


Naturhistoriska riksmuseet

Detection of limnic fauna using qPCR (finding selected species)



Proportion of positive detections at known occurrences



DNA concentration at varying frog densities

Thomsen et al, 2012, Mol Ecol 21

Naturhistoriska riksmuseet

Some sampling issues

- Replicates - increase detection probability, account for spatial variation, temporal variation, etc
- Blanks – negative controls to check for contamination (sampling, extraction, PCR)
- Positive controls – PCR reactions working, primer design, etc.
- Plan sampling together with the lab

Naturhistoriska riksmuseet

Alternative ’minibarcode’ gene regions

A collection of metabarcoding primers

Taxonomic group	Gene	Length	Accuracy (Bs)
Angiosperms/Gymnosperms	cpDNA *trn*L intron	10-100 bp	Genus/Species
Poaceae	ITS1	54-88 bp	Species
Fungi	ITS1	~ 200 bp	Species ?
Vertebrates	mtDNA 12S V05	76-110 bp	Genus/Species
Teleost fishes	mtDNA 12S	60-70 bp	Species
Batrachia	mtDNA 12S	~ 42-57 bp	Species
Earthworms	mtDNA 16S (ewB/ewC)	~ 30 bp	Species
Earthworms	mtDNA 16S (ewD/ewE)	~ 70 bp	Species
Oligochaetes	mtDNA 16S (ewB/ewE)	~ 120 bp	Species
Arthropods/Mollusks	mtDNA 16S	35-40 bp	Family/Genus
Termites	mtDNA 12S	~ 30 bp	Species ?
Termites	mtDNA 12S	~ 70 bp	Species ?
Collembola	mtDNA 12S	39-44 bp	Species ?
Collembola	mtDNA 12S	125-138 bp	Species ?

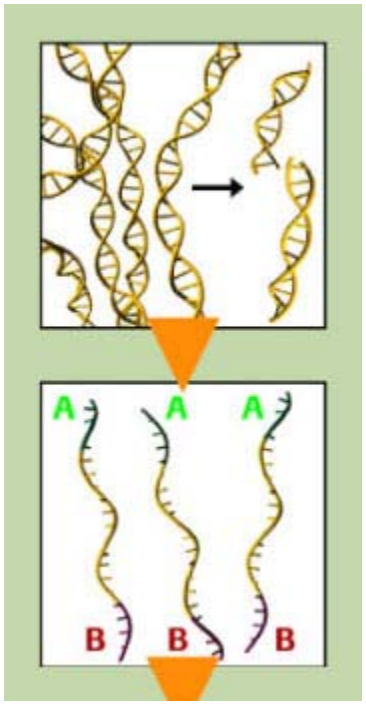
More information soon on www.metabarcoding.org

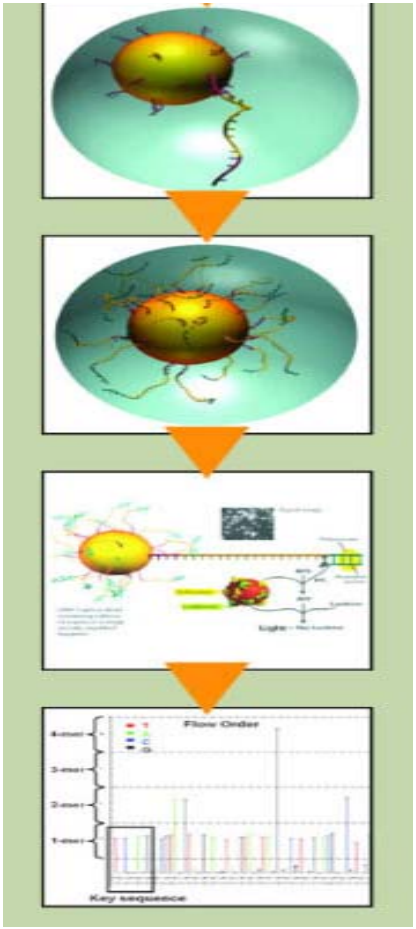
’Ecoprimers’
www.grenoble.prabi.fr/trac/ecoPrimers

Taxonomic stringency?

Next Generation Sequencing (454, Illumina, etc.)

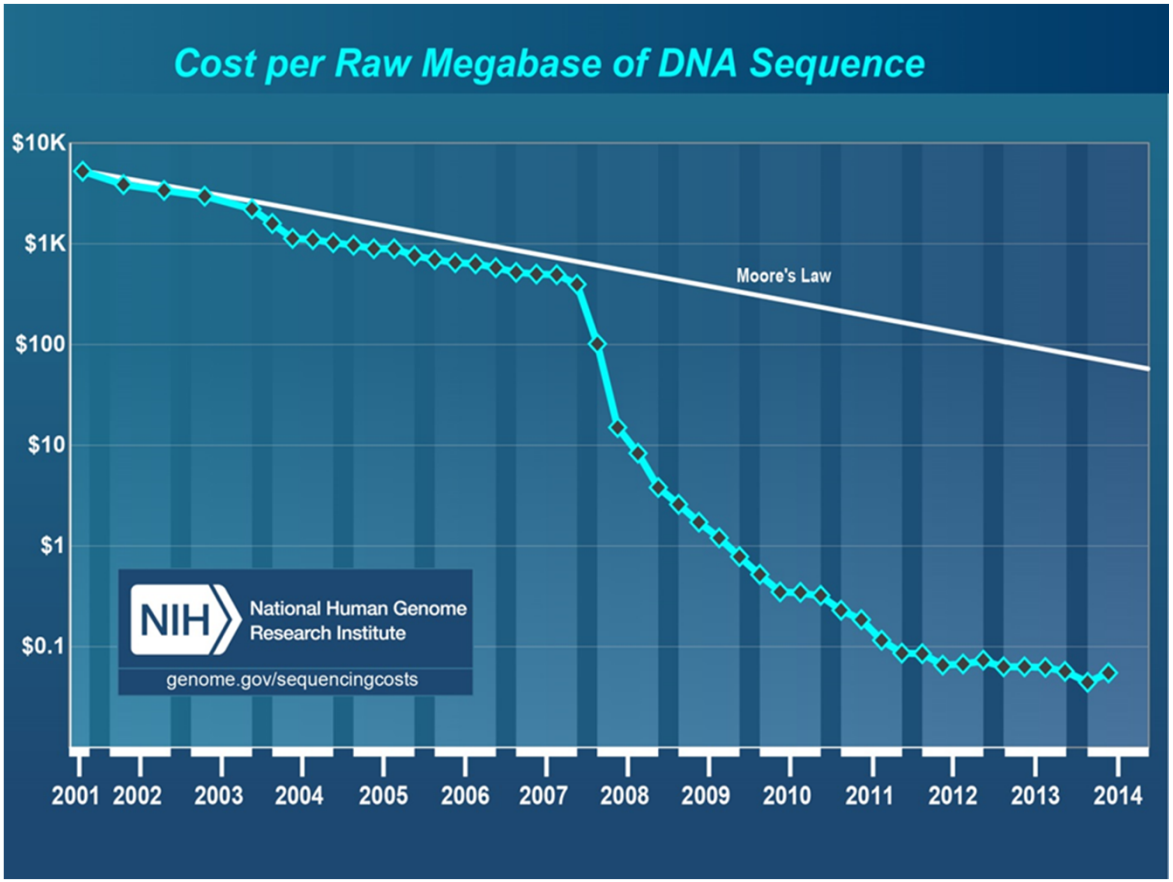
- 200-1000 baspairs
- 100 000 – billions of sequences
- Hours to days
- Clonal sequencing – each starting from single molecule amplified millionfold
- ’Semi-quantitative’
- Gbp per day





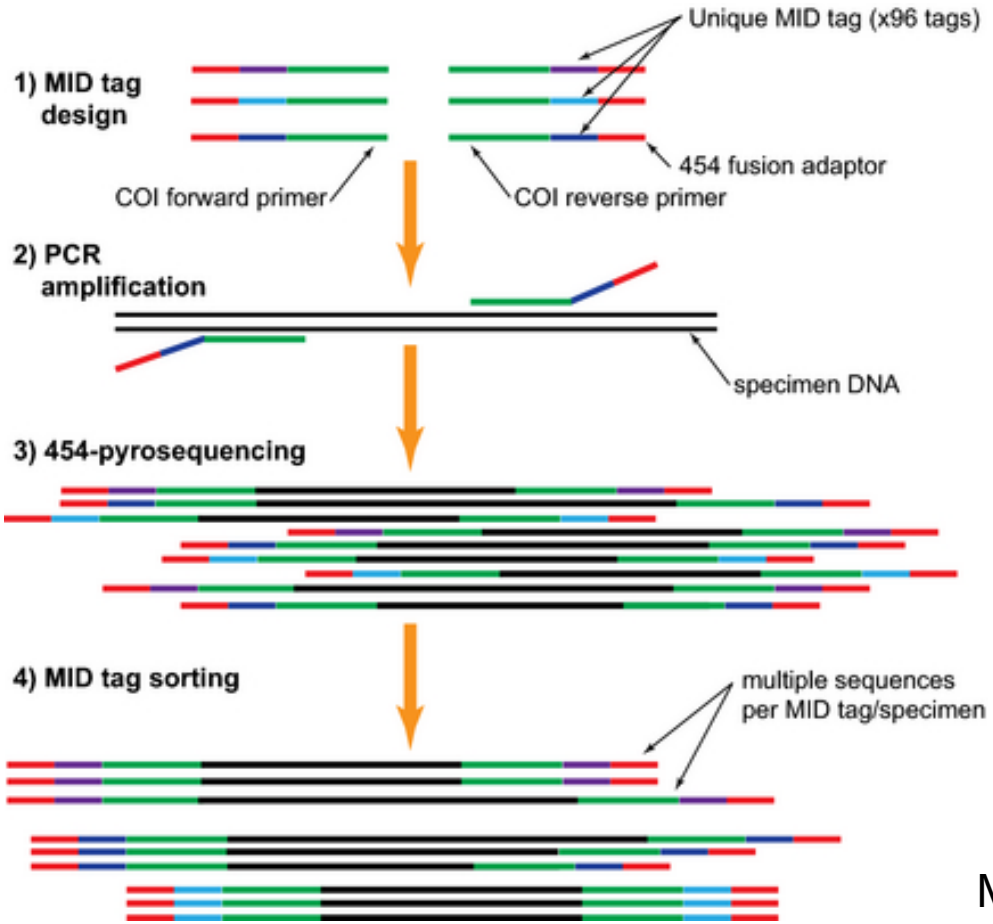
Naturhistoriska riksmuseet

Fast technological development, price drops



Naturhistoriska riksmuseet

Tagging (other ’barcodes’) of samples to be able to sequence together and separate after



Mol. Ecol. Res. **14**, 892-901.

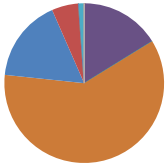
Naturhistoriska riksmuseet

Metabarcoding of chironomids in benthic samples

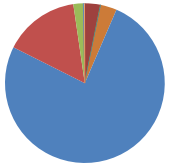
- 20 benthic samples per year from Bothnian Bay (UMF)
- Bulk samples from 8-10 stations sequenced together in NGS
- 150 bp of CO1 barcode sequence (minibarcode)
- Samples tagged to separate the stations
- BLAST matching against local database
- Proportion of different species at the different stations

Proportions of chironomid spp at Bothnian Bay benthic stations

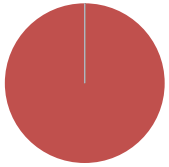
G14



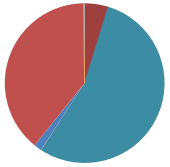
G5



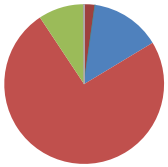
GR4



R1-6-42F



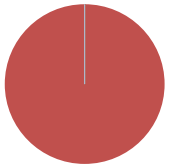
R1-6-LCO



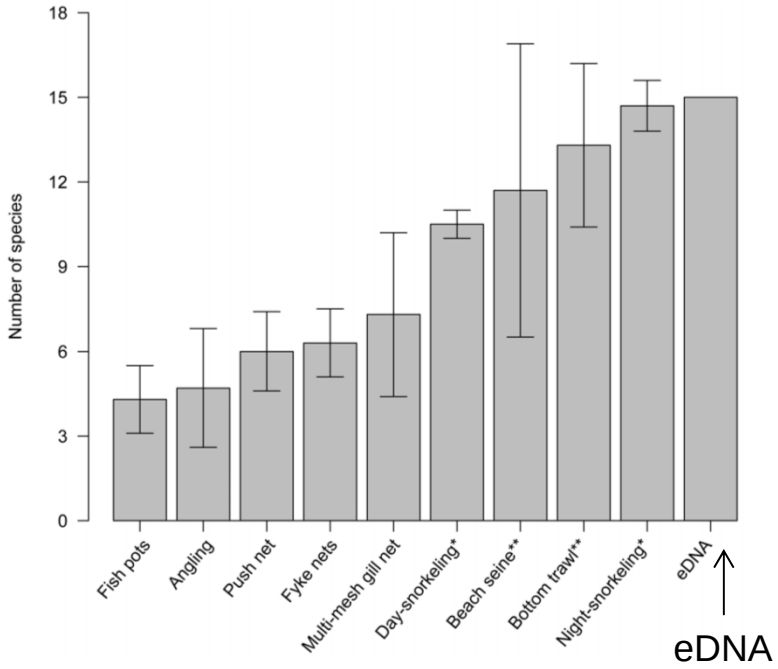
R2-19-42F

R2-19-LCO

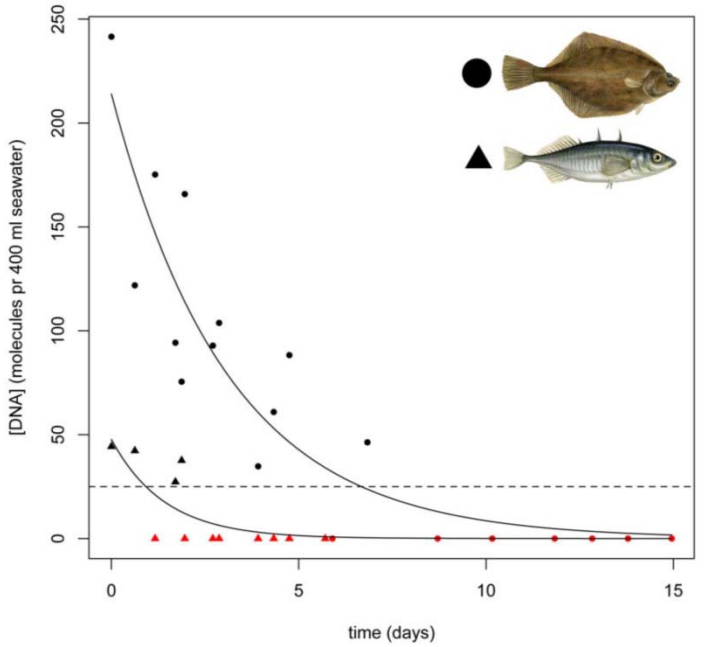
R2-19X



Survey of fish fauna in Danish coastal waters



Fishing methods



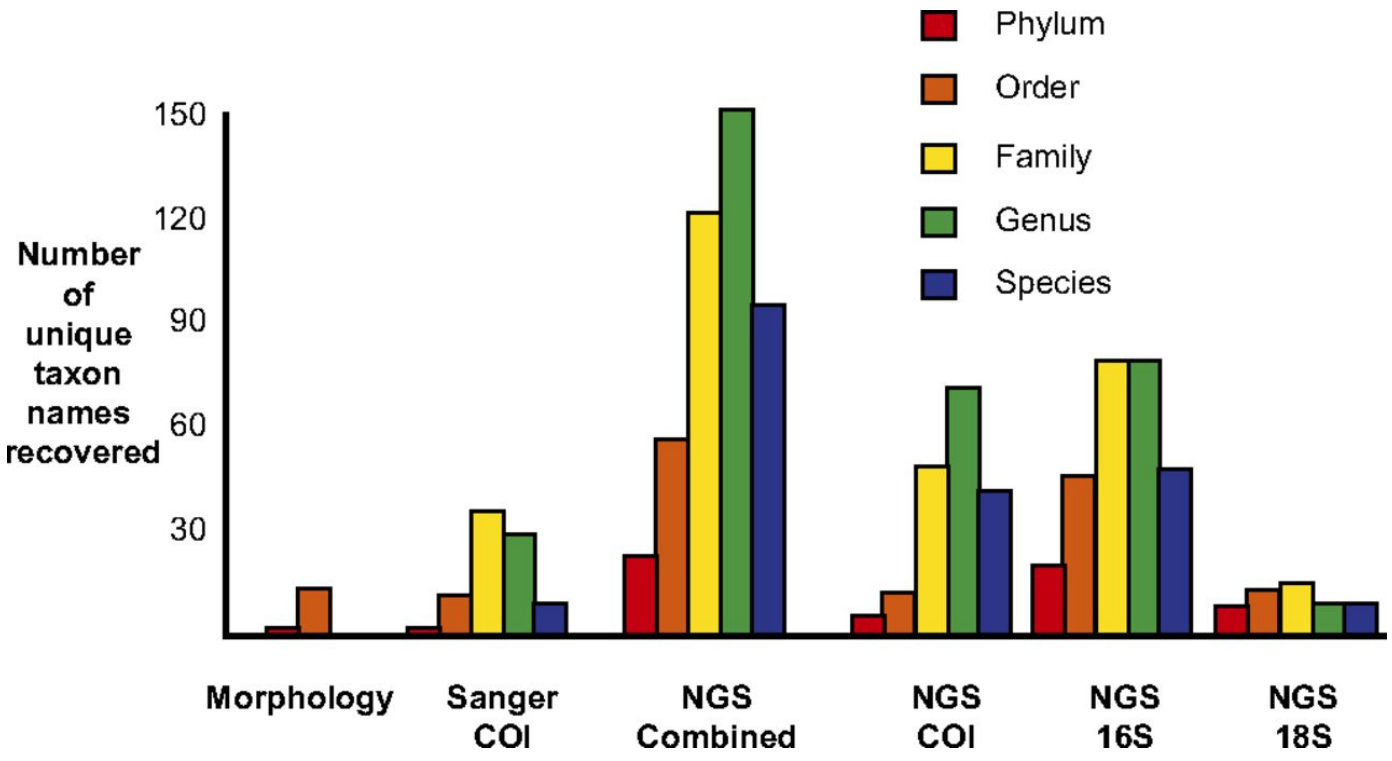
DNA degradation in aquarium (< 1 week)

Number of species detected

Thomsen et al. 2012, PLoSOne 7

Naturhistoriska riksmuseet

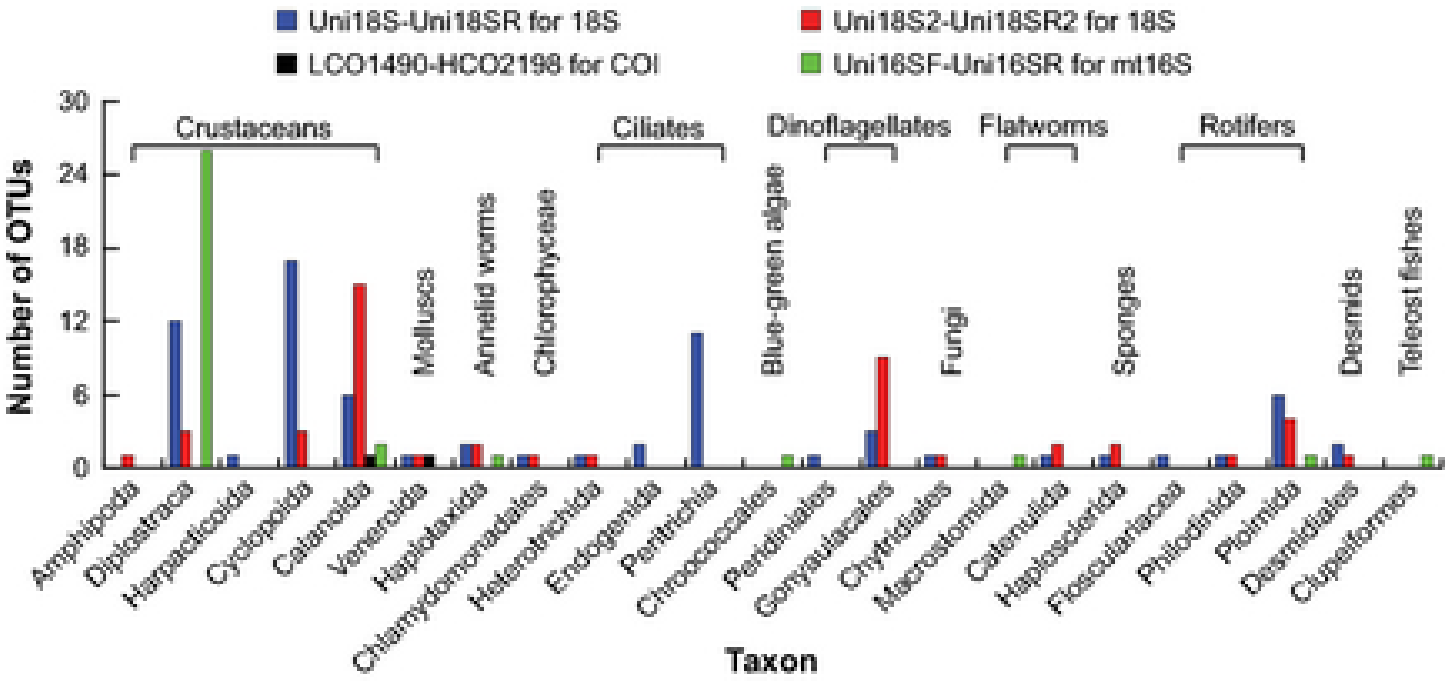
Number of unique taxon names recovered by three different methods: morphological identification, Sanger sequencing of individuals, and NGS of COI, 16S, and 18S gene regions.



Gibson J et al. PNAS 2014;111:8007-8012

Naturhistoriska riksmuseet

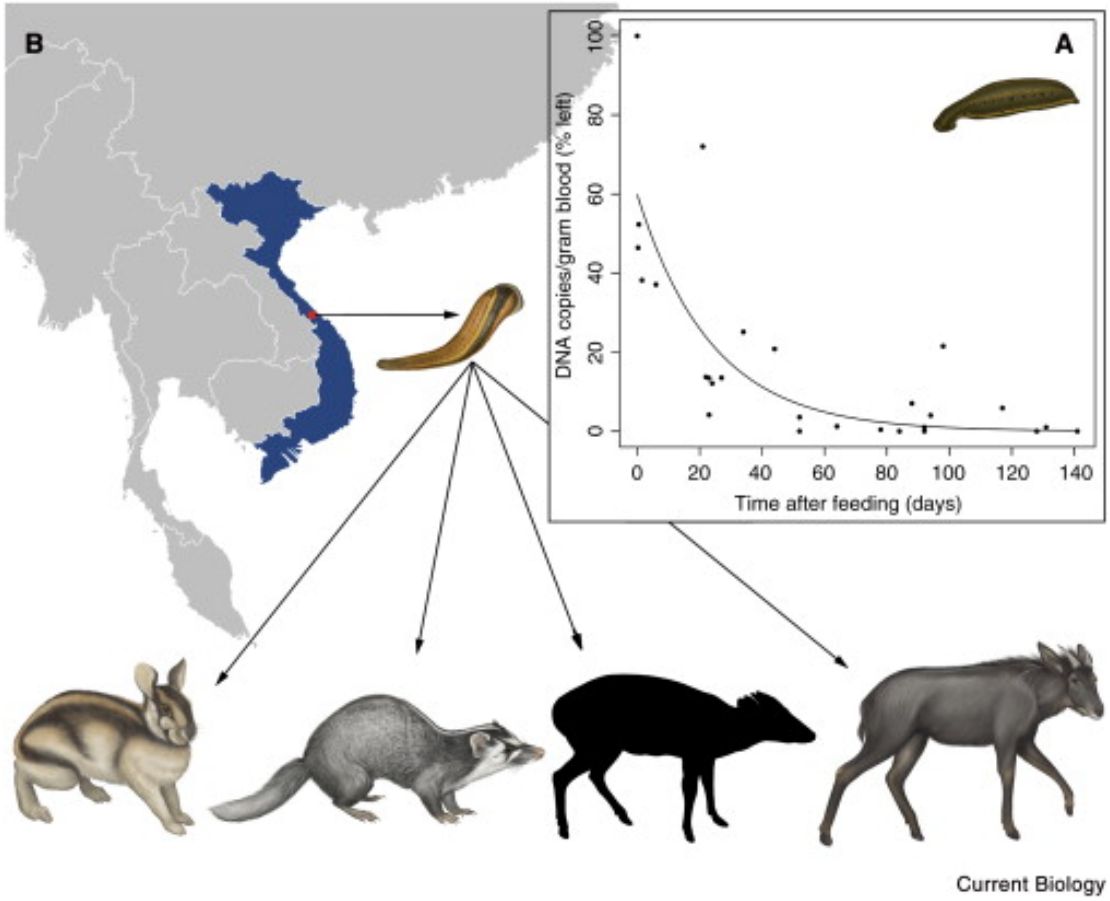
Performance comparison of genetic markers for high-throughput sequencing-based biodiversity assessment in complex communities



Plankton community, Ontario, Ca

Mol Ecol Res **14**, 1049-1059

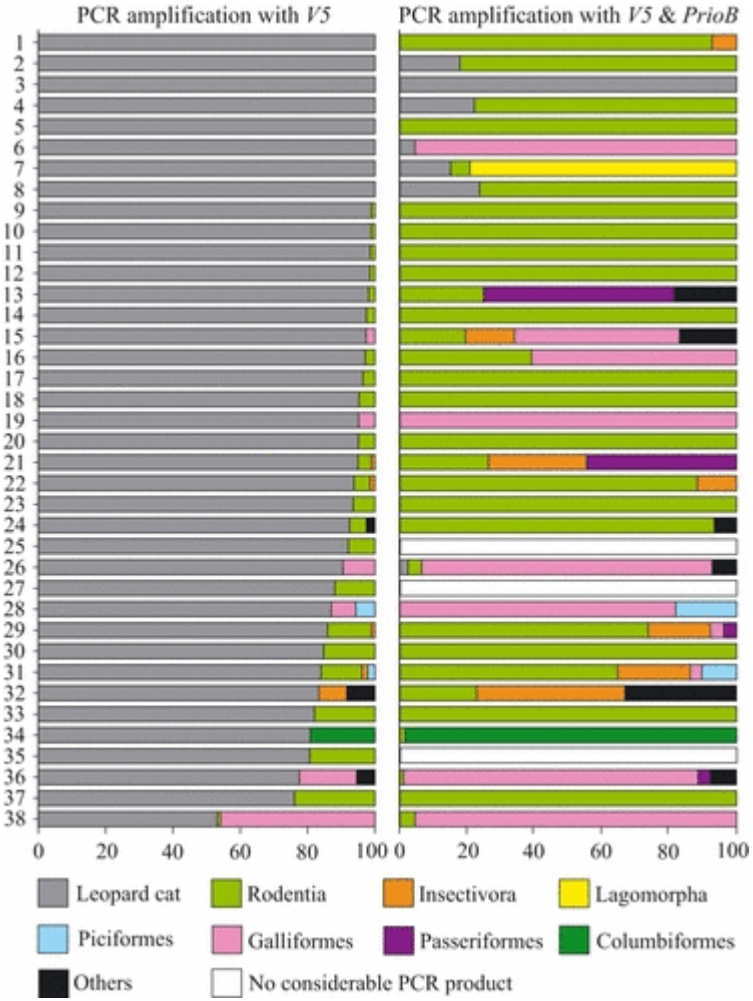
Screening mammal biodiversity using DNA from leeches



Current Biology, Volume 22, Issue 8, 2012, R262 - R263

Naturhistoriska riksmuseet

Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan



Blocking primer masking leopard cat (right)

Swedish Museum of Natural History www.nrm.se

- Dep of Bioinformatics and genetics
- Centre for genetic identification (CGI) cgi@nrm.se
- Taxonomic expertise
- Large collections, searchable databases www.naturarv.se
- DNA reference database portal: www.dnanyckeln.se
- Tissue collection

Naturhistoriska riksmuseet

Naturhistoriska riksmuseet

Conclusions

- DNA barcoding gives quality assured identifications
- Potential for large scale environmental monitoring, NGS, eDNA
- Invasive species monitoring
- Non-invasive tracking of endangered species
- Difficult taxa and life stages
- Prioritized taxa – Which taxonomic level is required?
- Reference database involving taxonomic expertise