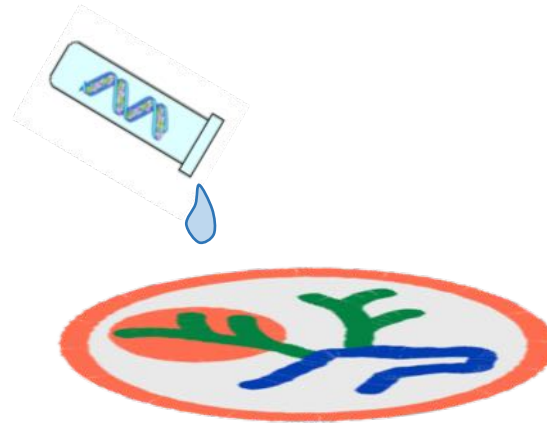




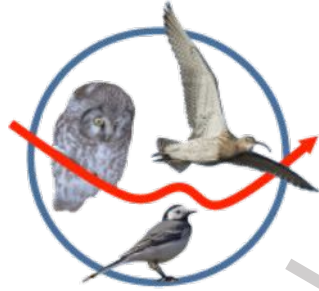
Integration of molecular data into Biodiversity Atlas Sweden (BAS)



Maria Prager (SU), Anders Andersson (KTH), Diego Brambilla (LNU),
Daniel Lundin (SU/LNU), Anna Rosling (UU)



Systematic monitoring data



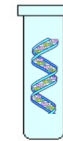
Biotelemetry (tracking) data



Natural history collection data



Molecular biodiversity data

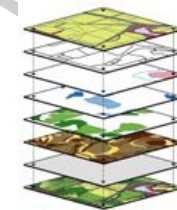


Metabarcoding
Metagenomics

python™

R
ALA4R...

R OpenSci
Mirroreum: R online



Spatial
portal



Metabarcoding (marker-gene amplicon sequencing)



DNA extraction



PCR



DNA sequencing

```

>seq1
AATGTAGTA...
>seq2
AACGTAGGT...
>seq3
...

```

Sample (water, soil, faeces, biopsy ...)

DNA

PCR-amplified genes

Barcode sequences

OTU	Species	Sample1	Sample2
1	E.coli	17	0
2	S.aurus	231	11800
3	unknown	30	0
...

OTU counts per sample

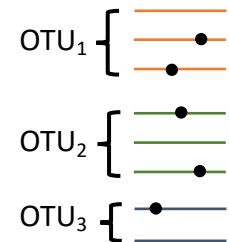


BAS

Matching of OTU representatives to reference database
 ←
 with taxonomically classified sequences (e.g. RDP, Silva)

Clustering ↓ (e.g. ≥ 99% Identity)

Operational Taxonomic Units (OTUs)





Metabarcoding (marker-gene amplicon sequencing)



DNA extraction



PCR



DNA sequencing

```

>seq1
AATGTAGTA...
>seq2
AACGTAGGT...
>seq3
...

```

Sample (water, soil, faeces, biopsy ...)

DNA

PCR-amplified genes

Barcode sequences

ASV	Species	Sample1	Sample2
1	E.coli	17	0
2	S.aurus	231	11800
3	unknown	30	0
...

ASV counts per sample

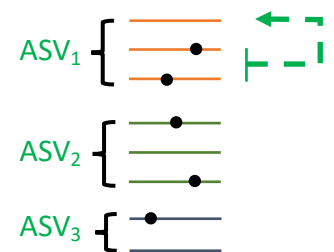


Matching of ASVs to (BAS) reference list

and deriving taxon string from reference database (e.g. RDP, Silva)

Error correction (denoising)

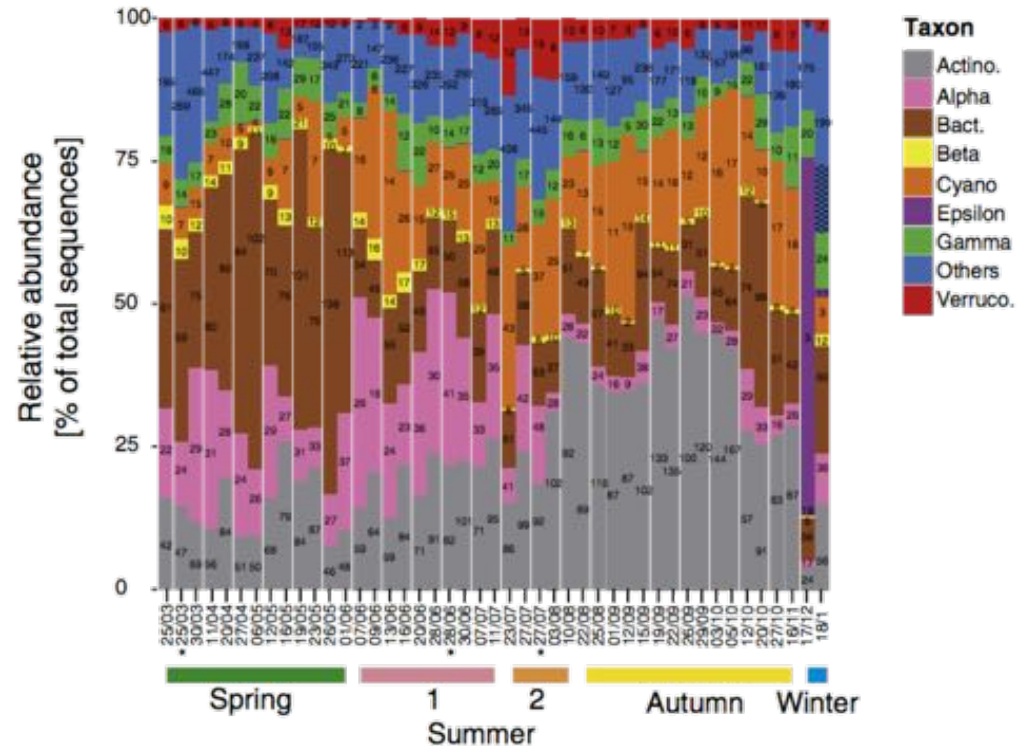
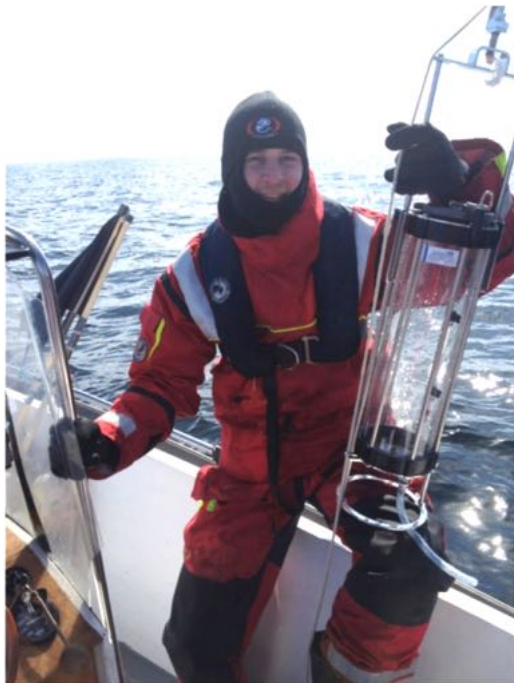
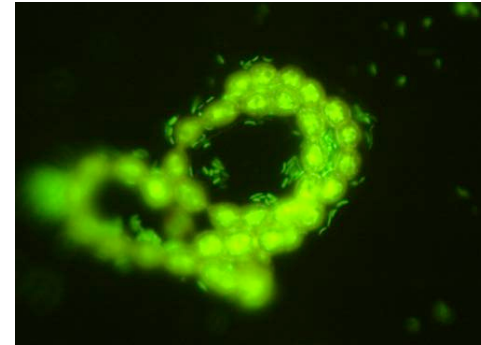
Amplicon Sequence Variants (ASVs)





Metabarcoding datasets – example

Linnaeus
Microbial
Observatory
(LMO)
time series





Services for: Data mobilization & interoperability

Sequencing provider



Lat:	>seq1
Lon:	AATGTAGTA...
Temp:	>seq2
Salinity:	AACGTAGGT...
...	...

Metadata Barcode sequences

Simplified
submission
of reads &
metadata



Sequence & metadata repositories





Services for: Data mobilization & interoperability

Sequencing provider



Sequence & metadata repositories



Lat:	>seq1
Lon:	AATGTAGTA...
Temp:	>seq2
Salinity:	AACGTAGGT...
...	...

Simplified submission of reads & metadata



Metadata Barcode sequences



Pipeline for denoising & taxon. annotation of OTUs (16S rRNA, Illumina MiSeq or PacBio)



OTU	Species	Sample1	S
1	E.coli	17	
2	S.aurus	231	
3	unknown	30	
...



Services for: Data mobilization & interoperability

Sequencing provider



Sequence & metadata repositories



Lat:	>seq1
Lon:	AATGTAGTA...
Temp:	>seq2
Salinity:	AACGTAGGT...
...	...

Simplified submission of reads & metadata

Reference databases

BAS ASVs



international BARCODE OF LIFE

UNIEUK

silva
high quality ribosomal RNA databases

unite
community



Taxonomy backbone

Pipeline/guide for submission of taxon. annotated OTU counts & metadata





Services for: Analysis & visualization

Visualisation of relative abundances



Aggregation at higher taxonomic levels

- Cyanobacteria
- Verrucomicrobia
- Alphaproteobacteria
- Betaproteobacteria
- Gammaproteobacteria
- Deltaproteobacteria
- Actinobacteria
- Bacteroidetes
- Planctomycetes
- Chloroflexi





How can we help you publish, access and use metabarcoding data?

- Please complete and/or spread the word about our online survey: <https://bit.ly/2MGxy4d>

Metabarcoding data in Biodiversity Atlas Sweden (BAS) - Questions for prospective users and contributors

Need for services

Which services would you have need for in your research?

- Easier submission of raw sequence data & metadata to repositories such as ENA
- Pipeline for denoising / OTU clustering, and taxonomic annotation of sequence data
- Online access to molecular biodiversity data, including tools for search, display and download

- or email me: maria.prager@scilifelab.se