

Introducing a SEQuence DataBase management system at the Swedish Museum of Natural History [and how to export to external systems]

Ingimar Erlingsson

Systemdeveloper Bioinformatics and Genetics at the Swedish Museum of Natural History

Ingimar.Erlingsson@nrm.se



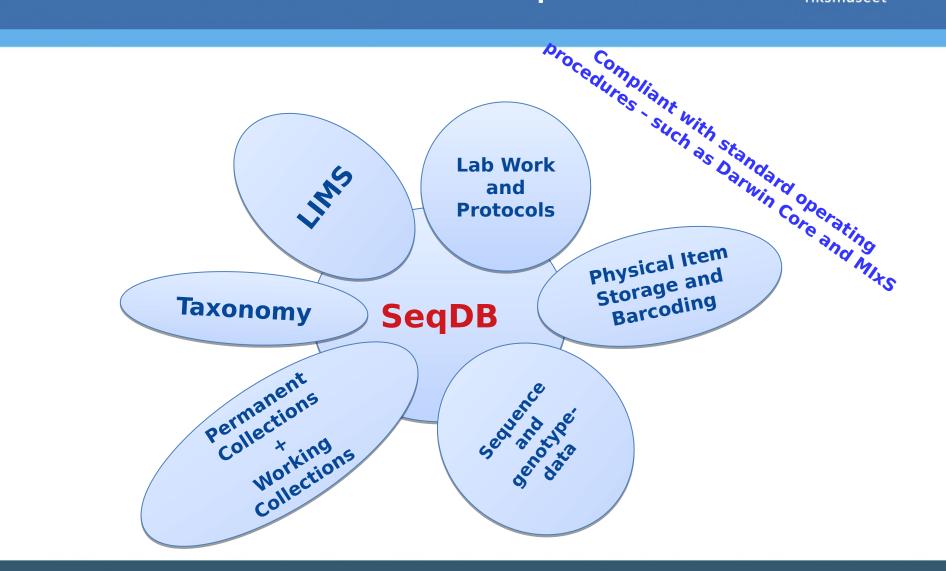
SeqDB is a modern information system

- A web accessible relational database
- Robust data validation
- Standardized and structured
- Powerful searching and filtering
- Bulk Editing (import / export of spreadsheets)
- Export to external Systems (BOLD)

What is 'SeqDB'?

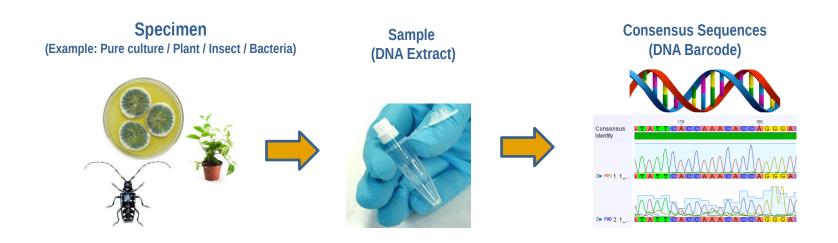
Provides standards compliant information management for Collection Record to DNA Sequence Workflows





The **SeqDB** Sanger barcoding workflow





The **SeqDB** Genotyping workflow





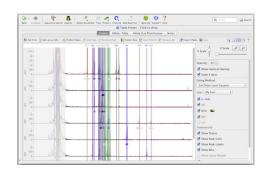


Sample (DNA Extract)





Genotyping Analysis



The **SeqDB** NGS workflow







Sample (DNA Extract)





Sequencing Data (FASTQ)

CD TYACGASATEGISIANA ESIG TYCACERA SIA POR COSASIA DISCRIPTO DATISCIS -BPG-BBS121+4+100+1783+550+971	Lacuston
ERIC -BEILLI (4.100 : TRI : 55040/1 COTTACOUNTCH BARANCOUNTCACANAMATOCOANACHANT TOTAT COUNTROCHT -BEILLI (4.100 : TRI : 55040/1 DARA "D_DA" DA "YAR AS" "21" E TRO EVO EVO DERSONNE EXONECHTES	DETEROLATACASSESS (TODGSA)
RENG-BELL 21 (4100 t 193) 55500/12 CHYMOSOLAPPORIDAMIAN CONTROLOGIC BELANDS CONSERVING THE PRANTICION CONTROLOGIC BELAN- HER -BELL 21 (30) 10 (193) 55500/12 SERVING BELL 21 (30) 10 (193) 55500/12 SERVING BELL 21 (30) 10 (193) 5500/12 SERVING BELL 21 (30) 10 (193) 500/12 SERVING BELL 21 (30) 500/12 SERVING BELL 21	PACASESS TOSSAA STITECTS
CONTROCADATION AND CONTROL OF TRANSPORTATION AND CONTROL TO TATALOGUE TO CONTROL AND	
pagna" b_pa" pa" yat jat" atm" t jita jita(jitaganteensonneetempertkripteooneilykeo Ring-mailila-100:1743:161186/1	
OGETOGORATTACEATOGEAGEATOGOTTOCOCCEGOREAGOREAGOREAGOCTOCOCTTAGOCTTOGOCTTOGOREA -880-88812144100:1789/16118/N1 -870-8812144100:1789/16118/N1	
eri -rreizi (ul 100 : 1783) 122#0/1 Odfiladettypicaleropyjanotyracostajatytymriotygotestajtotracostatatytyt Oriklari (1 - 4 - 10) - 1781 - 1782 - 1783	
New "seasablebash blobbbb bbbb bbbbbb bbbbbb bbbw", a 'a') 'aT[a_51]] 'a'[a_sbbw] BBC-386721410017791779477	_X\T_F00 S\NSHID
GEORGEPTENTEGETEGGEGEGETECTCCCCCCCCCCCCCCCCCCCC	CANADACA MAZATZ
""(an'b" [sabb) 'n mbb's"bobbebsabasab FEn 'beb X'(a'FF] [' X'T FO	OS INCHES dealers_n
HERM-ANDELE EN EIN DE EN BETON 1997 DE ENTEISBERTEISBARRE EIN TYPEN BEREISBANGE EINBERCHTE TRETECKERTE PER FOR TERARAMARAKAN HERT-BASTIE 14 1 100 1 TYPEN 201907 1	ANCAZATZ
-end_ass_filed_bos[a_0 a^asa^as[a]"/[x_a"f518_\8_[1]\2f50126502x][FCF\CS\WSHI Spr_ass_filed[b] = 00.1984-4548//	Disabas_s
OCCUPANTICION CONCANDETRA PORCUTE CRICA ARRAMANTA CAPITATO CONTO CARCO ATPUNICA ARTINICA PAT HTML: RMIT 12 4 100 1782 45540 / 1	7
alth balbalmalitechietelebistelebist 'n' abboutelembistelebasteleb 'bh' ab S bair S bindan	

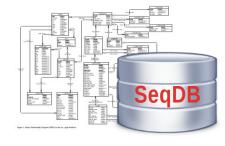
The **SeqDB** workflow



Data (Sequencing, genotyping)

Analysis Results & Reports

Analysis Tools and Pipelines



The SeqDB and The Swedish Natural History Museum

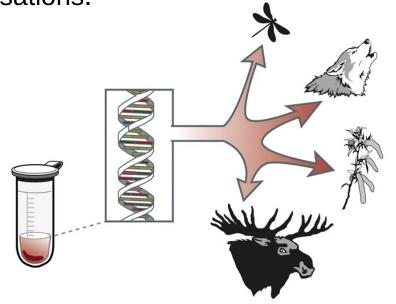


User CGI: The Center for Genetic Identification

"The Center for Genetic Identification is assignment based and offers

DNA-analyses to institutions and organisations."

1 Use case : 'the Bear-project'With individual genotype data



The SeqDB and The Swedish Natural History Museum



User: DNA-laboratorium

Mission: "We help researchers to develop and carry out DNA analyses of plants and animals, and to analyse genetic data. This can include the reconstruction of evolutionary relationships between species, inferring short-term processes within a species, and the use of DNA-based methods for investigating biodiversity."

2 Use cases:

- 1. 'Voucher specimens/DNA archive'
- 2. 'Working DNA samples for temporary research projects'

Implementing SeqDB at the Museum



Gathering requirements from CGI and the DNA-Lab

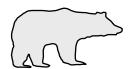
- Does SeqDB support our 3 use cases ?
- Workshops with the 'domain'-experts:
 - There has been a need to customize the software
 - Driver: "The 'Bear'-project"
 - Work: requirements, development, locally deployed, Testing
 - Infratructure; Installation at the Museum (backup)
 - Support of: import of Genotypes, to view Genotypes (Alleles)
 - Need for a public search-portal, for the Swedish EPA

Our Customization:

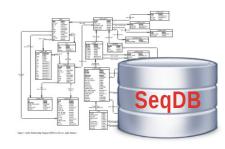
- Integrated and developed support for a
 - Micronic scanner, Rack with 96 wells **preprinted** viles
- DevOps: We have packaged the SeqDB with Docker.
 - Driver : A harmonized production environment

CGI – project : The Bear-project





- Monitoring the scandinavian brown bear population
- Citizen science
- Genotype data
- Provide basis for management decision
- ✓ Metadata from other data base
- ✓ Genotype data
- ✓ Physical storage information (DNA)





- Seqdb demo ... 2:55
 - https://archive.org/details/simplescreenrecorder_seq-2018-09-07_17.30.29
 - https://archive.org/details/simplescreenrecorder_seq-2018-09-07_17.57.14



Credits to the AAFC

Credits to the Agriculture and Agri-Food Canada (AAFC), SeqDB-Developement

