

THE ABS
CAPACITY
DEVELOPMENT
INITIATIVE



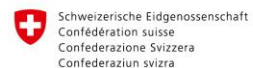
L'INITIATIVE DE
RENFORCEMENT
DES CAPACITES
POUR L'APA

Assessing (biotechnology) capacity development needs for the use of DSI – Overview about DSI users and use

02 June 2021, Technical Webinar on DSI

*Dr. Hartmut Meyer, Team Leader
ABS Capacity Development Initiative*

The ABS Initiative is funded by



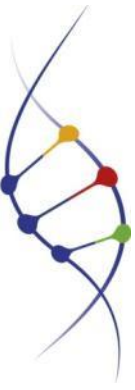
Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER
State Secretariat for Economic Affairs SECO

and implemented by



Content



DSI users and use

- **Origin of DSI in and DSI users of the International Nucleotide Sequence Database Collaboration**
(slides adapted from Amber Scholz, DSMZ, Germany and Guy Cochrane, EMBL-EBI, UK, presented at 1st Global Dialogue on DSI, Nov 2019 and the Webinar “Understanding DSI: a technical overview of its production, distribution and use”, Dec 2020, resp.)
- **Sequencing of African cattle genomes and providing DSI to databanks**
- **African bioremediation research and use of DSI from Databanks**
- **Synthetic biology approaches to bioremediation and biosafety implications**

International Nucleotide Sequence Database Collaboration (INSDC)

Scale

- **Rate:** 1 new dataset every 6 minutes
- **Data:** 2×10^9 sequences and 1×10^{16} base pairs of read data across 2×10^6 taxa
- **Usage:** 2,000 submitters; 10x thousands monthly consumers; 10x millions of monthly hits, many times this globally
- **Support:** 46 tickets per day and in-person training delivered to more than 350 users per annum
- **Infrastructure:** includes 100s of petabytes of storage, ~\$50 USD mil./year
- **Biodiversity data**
 - 7.2 million records point to natural history museums, culture collections and biobanks
 - 7.9 million sequences with geographical coordinates
 - 2.1 million sequences have place name annotation
 - 1 million raw sequencing data sets have coordinates

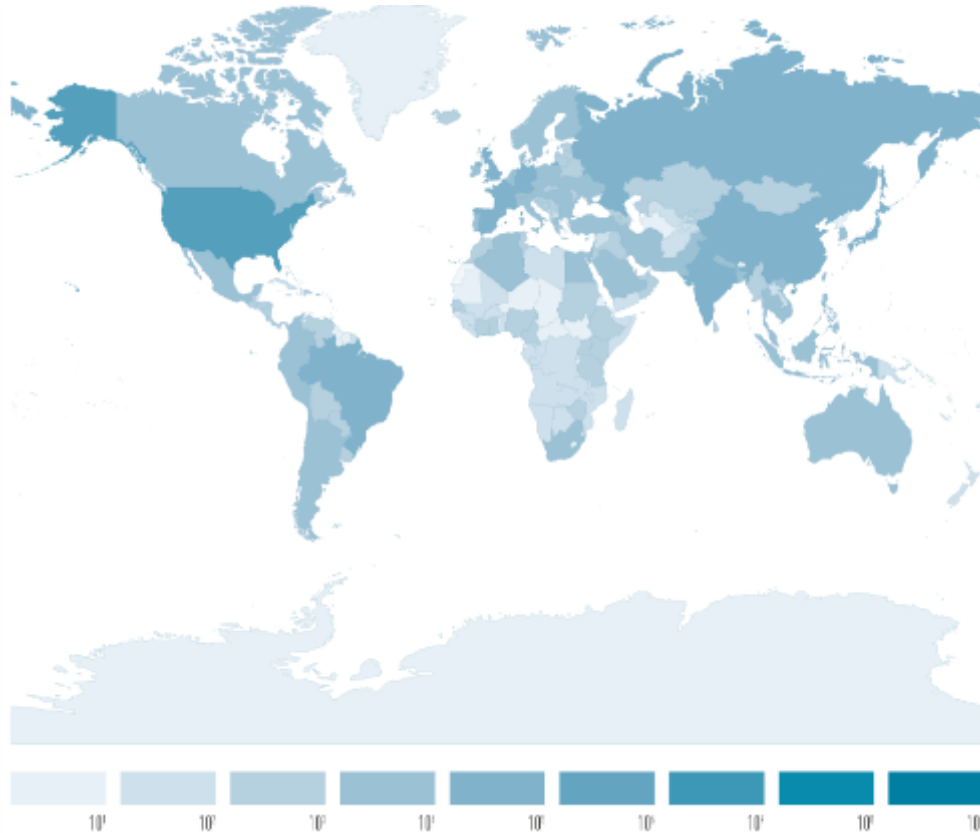


<http://www.insdc.org/>

**There are 10-15 million total users of INSDC.
They live in every country in the world.**

1. United States	22.69 %
2. China	15.42 %
3. India	6.16 %
4. Japan	3.97 %
5. Germany	3.67 %
6. United Kingdom	3.45 %
7. France	2.84 %
8. Brazil	2.83 %
9. Spain	2.31 %
10. Russian Federation	2.25 %

Costs: \$3-5 per user
50% of users live in countries that do not contribute to NSD infrastructure costs

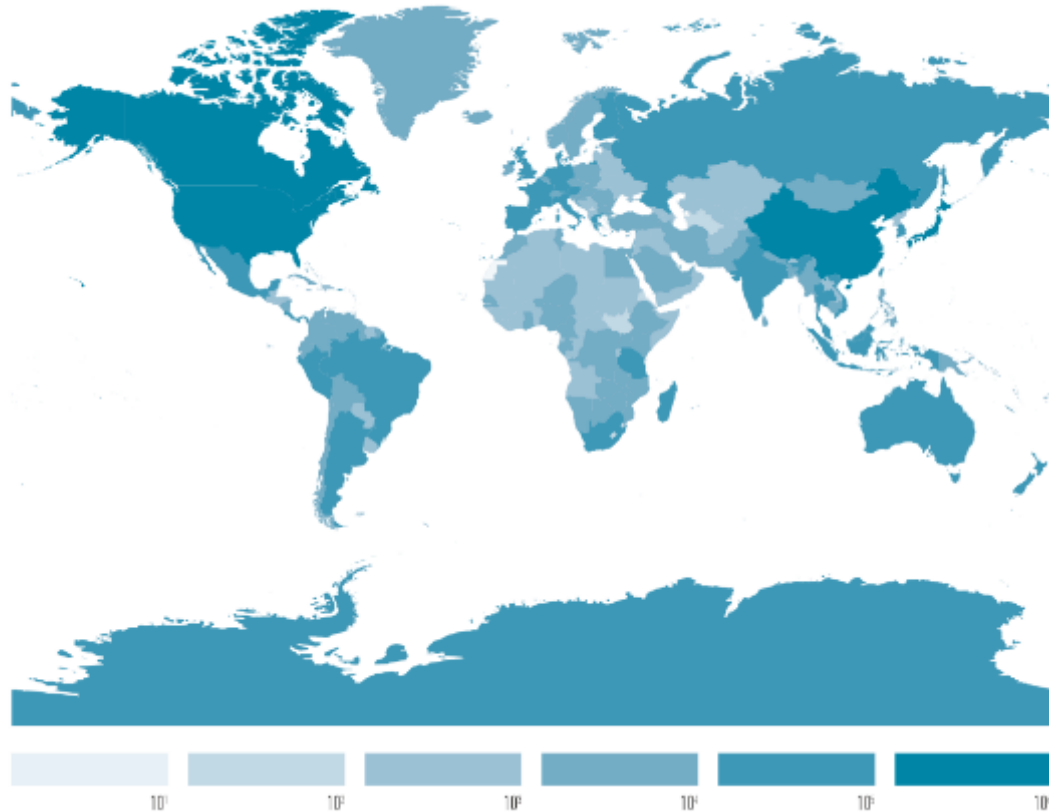


Where does the original GR for these NSD come from?

What is the country of origin for non-human NSD?

1. China	18.23 %
2. United States	17.39 %
3. Canada	9.10 %
4. Japan	7.24 %
5. India	3.46 %
6. Australia	2.66 %
7. Mexico	2.54 %
8. Brazil	2.30 %
9. Germany	1.83 %
10. Spain	1.58 %

52% of NSD comes from 4 countries



Creating DSI from African cattle



nature
genetics

ARTICLES

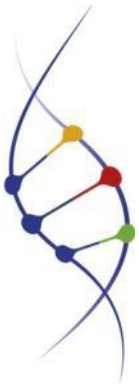
<https://doi.org/10.1038/s41588-020-0694-2>



The mosaic genome of indigenous African cattle as a unique genetic resource for African pastoralism

Kwondo Kim ^{1,2}, Taehyung Kwon ¹, Tadelle Dessie³, DongAhn Yoo⁴, Okeyo Ally Mwai⁵, Jisung Jang⁴, Samsun Sung², SaetByeol Lee², Bashir Salim ⁶, Jaehoon Jung¹, Heesu Jeong⁴, Getinet Mekuriaw Tarekegn^{7,8}, Abdulfatai Tijjani ^{3,9}, Dajeong Lim¹⁰, Seoae Cho², Sung Jong Oh¹¹, Hak-Kyo Lee¹², Jaemin Kim¹³, Choongwon Jeong ¹⁴, Stephen Kemp^{5,9}, Olivier Hanotte ^{3,9,15}  and Heebal Kim ^{1,2,4} 

Creating DSI from African cattle



DSI used for characterising cattle breeds

- Sequencing genomes of 114 individuals of 12 indigenous African cattle breeds from Ethiopia, Kenya, Sudan and Uganda
- ~3.50 terabytes of sequences, uploaded to GenBank
- Comparison with 217 publicly available genomes
- Statistical analysis revealing mosaic genome structure and evolutionary developments

Bos indicus (zebu cattle)

Accession: PRJNA574857

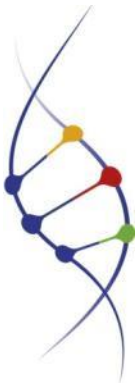
African cattle raw sequence reads

Raw sequence reads for 12 indigenous African cattle breeds (114 individuals) and African buffalo (2 individuals)

Accession	PRJNA574857
Data Type	Raw sequence reads, Genome sequencing
Scope	Multiisolate
Organism	Bos indicus [Taxonomy ID: 9916] Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos; Bos indicus
Submission	Registration date: 29-Sep-2019 Seoul National University
Relevance	Agricultural

- Goal: publicly available genome data to “open the door to sustainable crossbreeding programs combining local environmental adaptation and increased exotic productivity”
- Sources: Kim et al. (2020). Nature Genetics 52: 1099; GenBank Acc. No. PRJNA574857

DSI for African bioremediation research



SCIENTIFIC REPORTS

OPEN

Metagenomic profiling for assessing microbial diversity and microbial adaptation to degradation of hydrocarbons in two South African petroleum-contaminated water aquifers

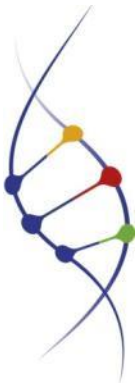
Leonard Kachienga¹, Keshri Jitendra ² & Maggy Momba¹

Received: 14 December 2017

Accepted: 2 May 2018

Published online: 15 May 2018

DSI for African bioremediation research



Methods applied / DSI used for characterising microbes

- **Extraction of metagenomic DNA**
- **Amplification of 18 S rRNA by PCR**
18 S rRNA sequences are used in taxonomic research for identification of taxonomic groups as genus or family in, e.g., a sample of environmental DNA
- **Sequencing of genomic DNA**
performed by a commercial DNA sequencing company
- **Analysis of sequence data**
BLAST-Analysis (using DSI stored in databanks) and statistical analysis based on standard software

- **Goals: Isolation, selection and use of microbes**

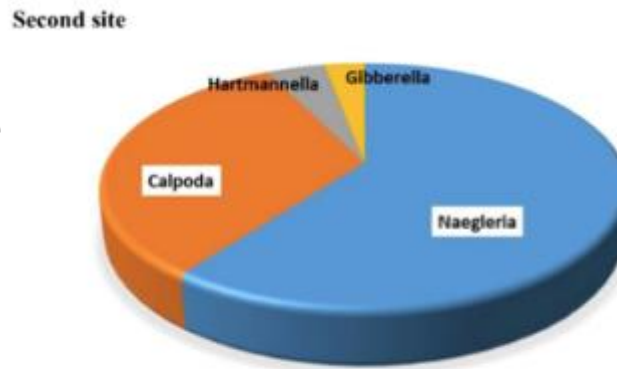
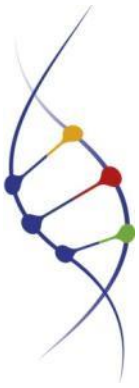


Figure 3. Different genera of the microbes observed in the first and second sites in petroleum-polluted water.

- Source: Kachienga et al. (2018). Scientific Reports 8: 7564

Synbio approaches to bioremediation

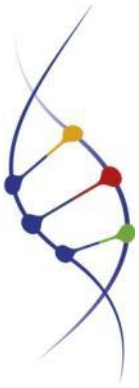


Alternative Strategies for Microbial Remediation of Pollutants via Synthetic Biology

*Shweta Jaiswal and Pratyosh Shukla**

Enzyme Technology and Protein Bioinformatics Laboratory, Department of Microbiology, Maharshi Dayanand University, Rohtak, India

Synbio approaches to bioremediation



Methods suggested / DSI also used for engineering microbes

- Synthetic microbial biosensor systems
- Metabolic engineering
- Genome editing
- Data mining
- Developing synthetic microbial communities

- Goals: Production of GM microbes
- Biosafety risk assessments, depending on national laws

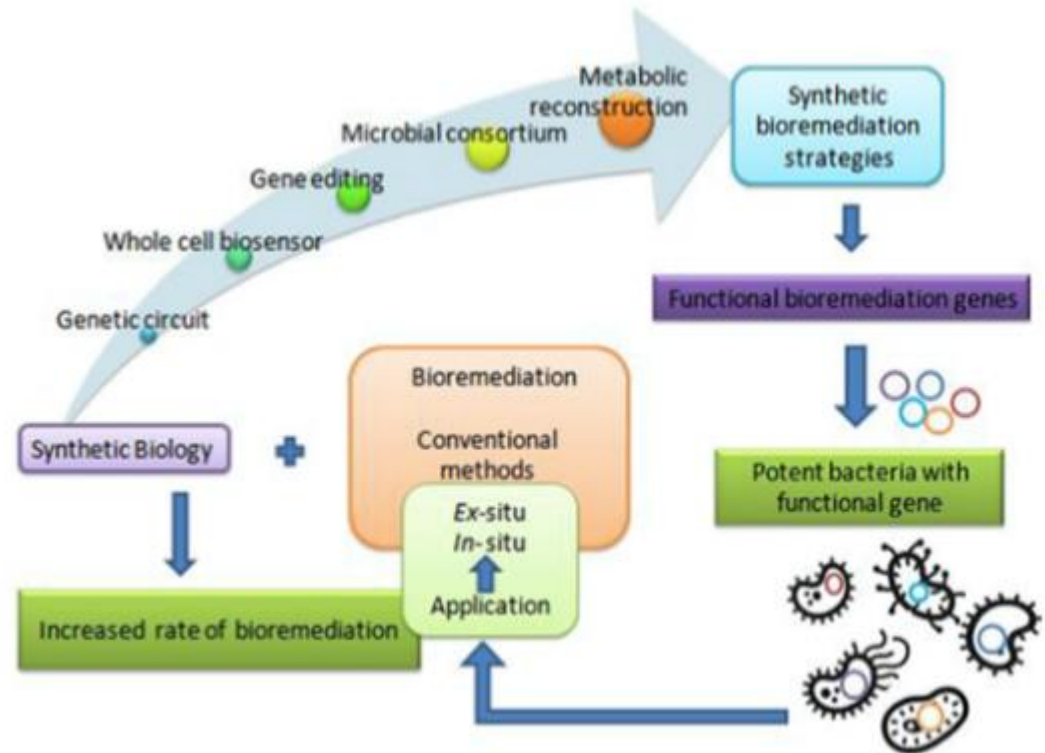


FIGURE 1 | The strategies of synthetic biology applicable for bioremediation.