Collection associated tools: taxonomy, research, information management

Erko Stackebrandt

DSMZ, Germany

Main reference:

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

Why do we need Culture Collections?

To understand the role and potential of biodiversity, the ability to maintain and identify biological resources is crucial

Though biodiversity offers more than biotechnology, e.g., understanding the evolution of the tree of life and to educate biosystematists, biotechnology is **key to meet the needs of the 21st century**. The global taxonomic impediment has been recognised and initiatives are underway to help lead to its resolution.

The Convention on Biological Diversity has specifically included microorganisms and the importance of ex-situ collections, as well as benefit sharing.

(CBD, http://www.biodiv.org/convention/default.shtml)

Article 9 of the CBD: Ex-situ Conservation

Each Contracting Party shall, as far as possible and as appropriate, and predominantly for the purpose of complementing in-situ measures:

(a) Adopt measures for the ex-situ conservation of components of biological diversity, preferably in the country of origin of such components;

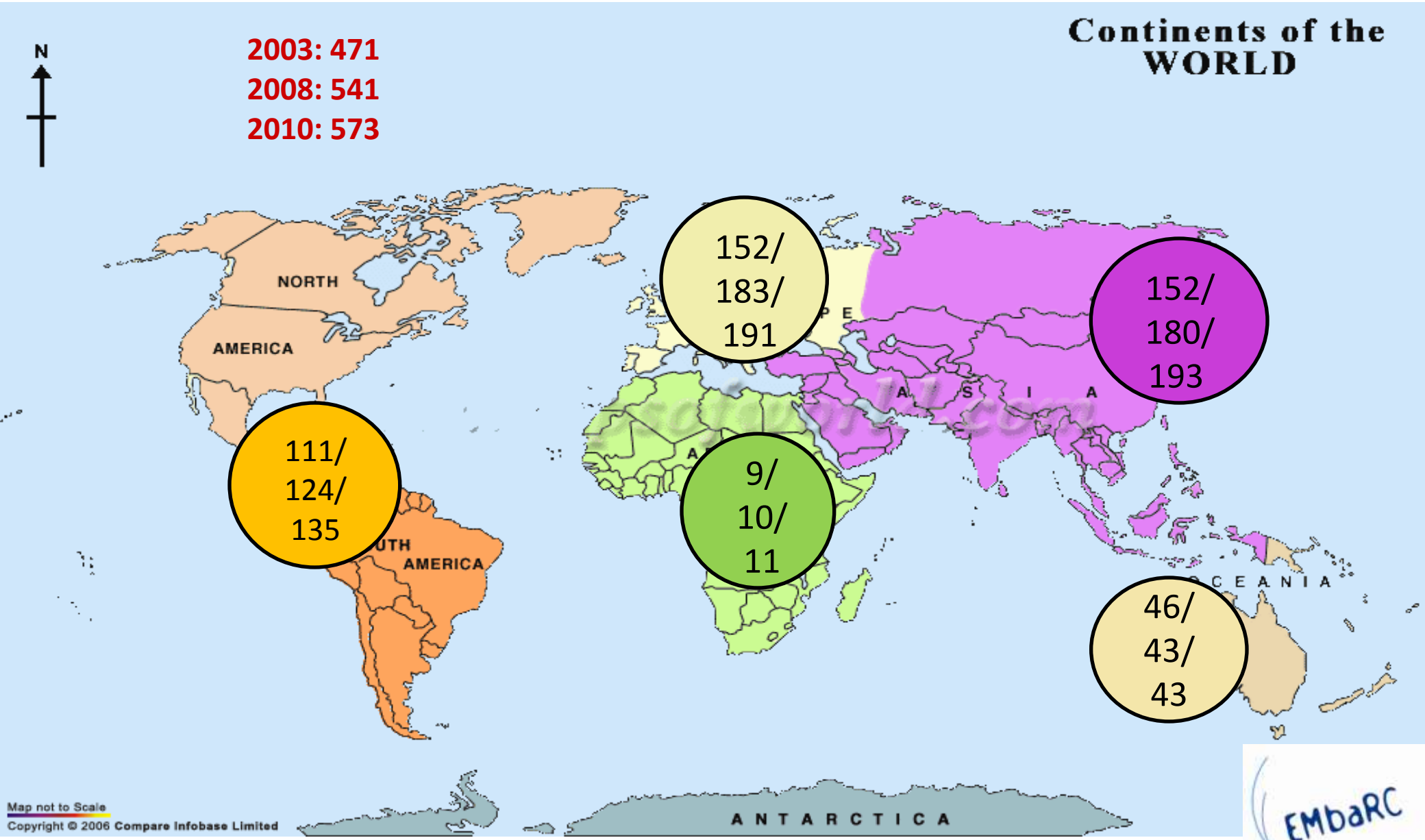
(b) Establish and maintain facilities for ex-situ conservation of and research on plants, animals and **microorganisms**, preferably in the country of origin of genetic resources;

(c) Adopt measures for the recovery and rehabilitation of threatened species and for their reintroduction into their natural habitats under appropriate conditions;

(d) Regulate and manage collection of biological resources from natural habitats for ex-situ conservation purposes so as not to threaten ecosystems and in-situ populations of species, except where special temporary ex-situ measures are required under subparagraph (c) above; and

(e) Cooperate in providing financial and other support for ex-situ conservation outlined in subparagraphs (a) to (d) above and in the establishment and maintenance of ex- situ conservation facilities in developing countries.

Numbers of collections and percentage of holdings



Source: World Data Centre for Microorganisms (WDCM)

Of the 573 culture collections in 68 countries*

227 of them are supported by government.

54 of them are semi-governmental.

211 of them are supported by university.

14 of them are supported by industry.

21 of them are private.

Of these

245 provide storage services.

265 provide distribution services.

284 provide identification services.

237 provide training services.

247 provide consultation services

*, http://wdcm.nig.ac.jp/statistics.html; date: Feb 24, 2010

573 culture collections*

In Europe	191	633,444
Armenia	1	7,575
Belarus	1	1,175
Bulgaria	3	13,234
Czech	1	38,497
Estonia	3	6,300
Hungary	6	8,717
Kazakhstan	2	199
Latvia	1	692
Poland	9	8,464
Romania	2	760
Russian Federation	14	40,874
Slovak	3	4,916
Slovenia	2	4,160
Ukraine	3	3,286
Uzbekistan	3	1,456
Yugoslavia (?)	2	897
East Europe	56 (29%)	141,203 (22.3%)

*, http://wdcm.nig.ac.jp/statistics.html; date: Feb 24, 2010

1. Taxonomy: for authentification (in, out, viability check), service, reputation

- Phenotype: morphology, physiology, cultural properties, key properties..
- Genotype: G+C content, MLSA sequences, DNA pattern, ribopattern..
- Chemotaxonomy: peptidoglycan, polar lipid, fatty acid, isoprenoid quinone protein pattern, MALDI-TOF..

1. Taxonomy: for authentification (in, out, viability check), service, reputation

- Phenotype: morphology, physiology, cultural properties, key properties..
- Genotype: G+C content, MLSA sequences, DNA pattern, ribopattern..
- Chemotaxonomy: peptidoglycan, polar lipid, fatty acid, isoprenoid quinone protein pattern, MALDI-TOF..

2. Research: publications = reputation, capacity building, collaboration

- New descriptions
- Phylogeny: MLSA..
- New methods: genome sequences, ANI..
- Ecology, clone libraries, metagenomics, proteomics..
- Culture conditions..
- Long-term preservation..

1. Taxonomy: for authentification (in, out, viability check), service, reputation

- Phenotype: morphology, physiology, cultural properties, key properties..
- Genotype: G+C content, MLSA sequences, DNA pattern, ribopattern..
- Chemotaxonomy: peptidoglycan, polar lipid, fatty acid, isoprenoid quinone protein pattern, MALDI-TOF..

2. Research: publications = reputation, capacity building, collaboration

- New descriptions
- Phylogeny: MLSA..
- New methods: genome sequences, ANI..
- Ecology, clone libraries, metagenomics, proteomics..
- Culture conditions..
- Long-term preservation..

3. Information management: outreach, visibility, improved efficiency

- Online catalogues: what, where, when, how, by whom, which methods, references..
- In-house network: linking access, dispatch and research with online catalogue
- National and international networks: linking online catalogues
- Added value entries: separate databases on basically anything
- Improved capacity building for bioinformation scientists

EMbaRC

1. Taxonomy: for authentification (in, out, viability check), service, reputation

- Phenotype: morphology, physiology, cultural properties, key properties..
- Genotype: G+C content, MLSA sequences, DNA pattern, ribopattern..
- Chemotaxonomy: peptidoglycan, polar lipid, fatty acid, isoprenoid quinone protein pattern, MALDI-TOF..

2. Research: publications = reputation, capacity building, collaboration

- New descriptions
- Phylogeny: MLSA..
- New methods: genome sequences, ANI..
- Ecology, clone libraries, metagenomics, proteomics..
- Culture conditions..
- Long-term preservation..

3. Information management: outreach, visibility, improved efficiency

- Online catalogues: what, where, when, how, by whom, which methods, references..
- In-house network: linking access, dispatch and research with online catalogue
- National and international networks: linking online catalogues
- Added value entries: separate databases on basically anything
- Improved capacity building for bioinformation scientists

EMbaRC

1. Taxonomy: for authentification (in, out, viability check), service, reputation

- Phenotype: morphology, physiology, cultural properties, key properties..
- Genotype: G+C content, MLSA sequences, DNA pattern, ribopattern..
- Chemotaxonomy: peptidoglycan, polar lipid, fatty acid, isoprenoid quinone protein pattern, MALDI-TOF..

2. Research: publications = reputation, capacity building, collaboration

- New descriptions
- Phylogeny: MLSA..
- New methods: genome sequences, ANI..
- Ecology, clone libraries, metagenomics, proteomics..
- Culture conditions..
- Long-term preservation..

3. Information management: outreach, visibility, improved efficiency

- Online catalogues: what, where, when, how, by whom, which methods, references..
- In-house network: linking access, dispatch and research with online catalogue
- National and international networks: linking online catalogues
- Added value entries: separate databases on basically anything
- Improved capacity building for bioinformation scientists

EMbaRC

ad 1. Identification and authenticity check

Automated identification systems, including the

- API system (Biomerieux, S.A. France) and
- Biolog plates (Biolog Inc, USA).
- MIDI FAME analysis

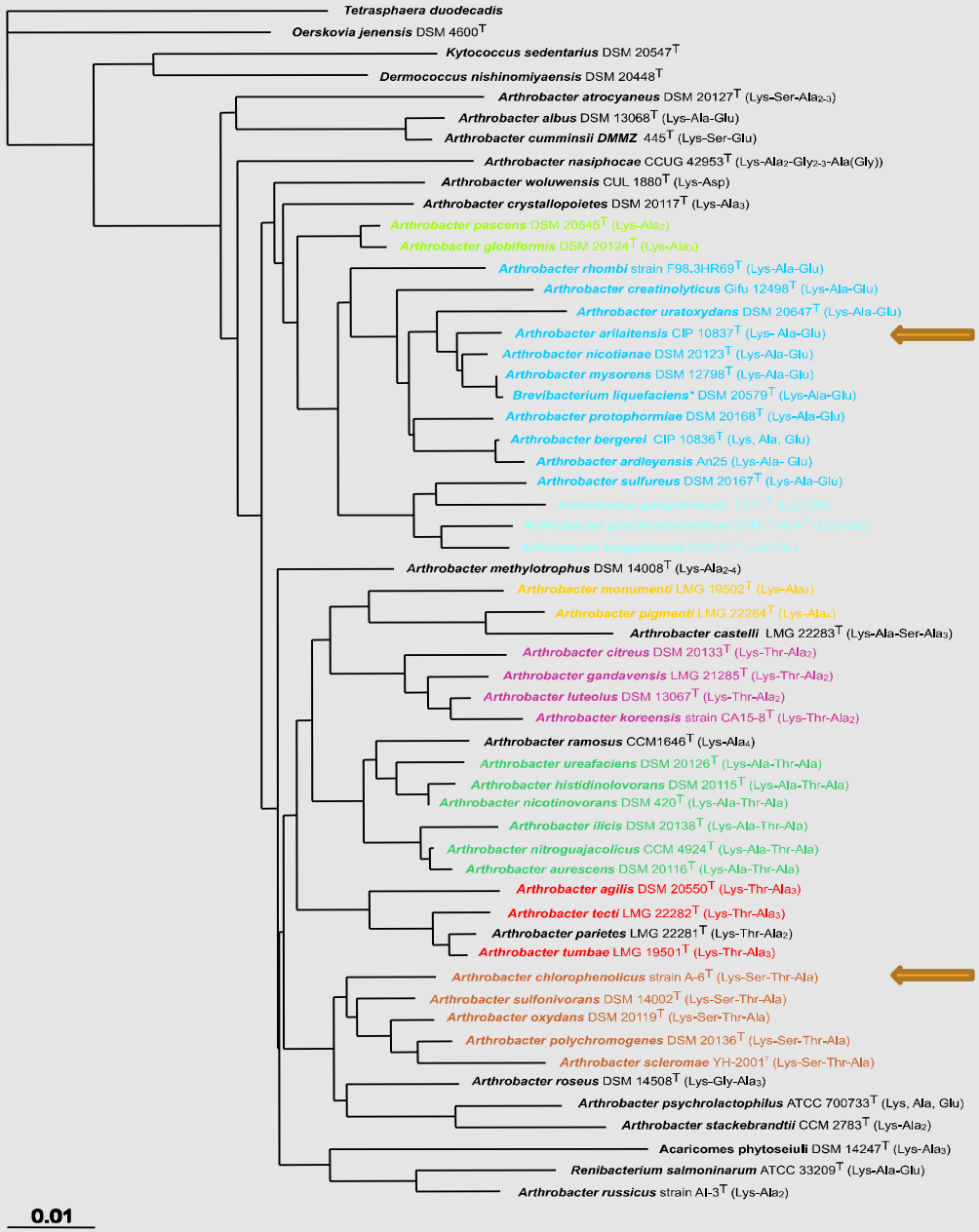
MALDI-TOF

Analysis and characterisation of nucleic acids.

Sequencing of DNA genes, operons and spacers

Fingerprinting techniques

- Riboprinting
- RFLPs (Restriction Fragment Length Polymorphism),
- RAPDs (Random Amplified Polymorphic DNA),
- AFLP (Arbitrary Fragment Length Polymorphism),
- SSCP (Single Strand Conformation Polymorphism).
- Variable number tandem repeat (VNTR) PCR,
- Repetitive extragenic palindromic (REP) elements



16S rRNA gene phylogeny is generally consistent with peptidoglycan types

Lys-Ala-Glu

Lys-Thr-Ala

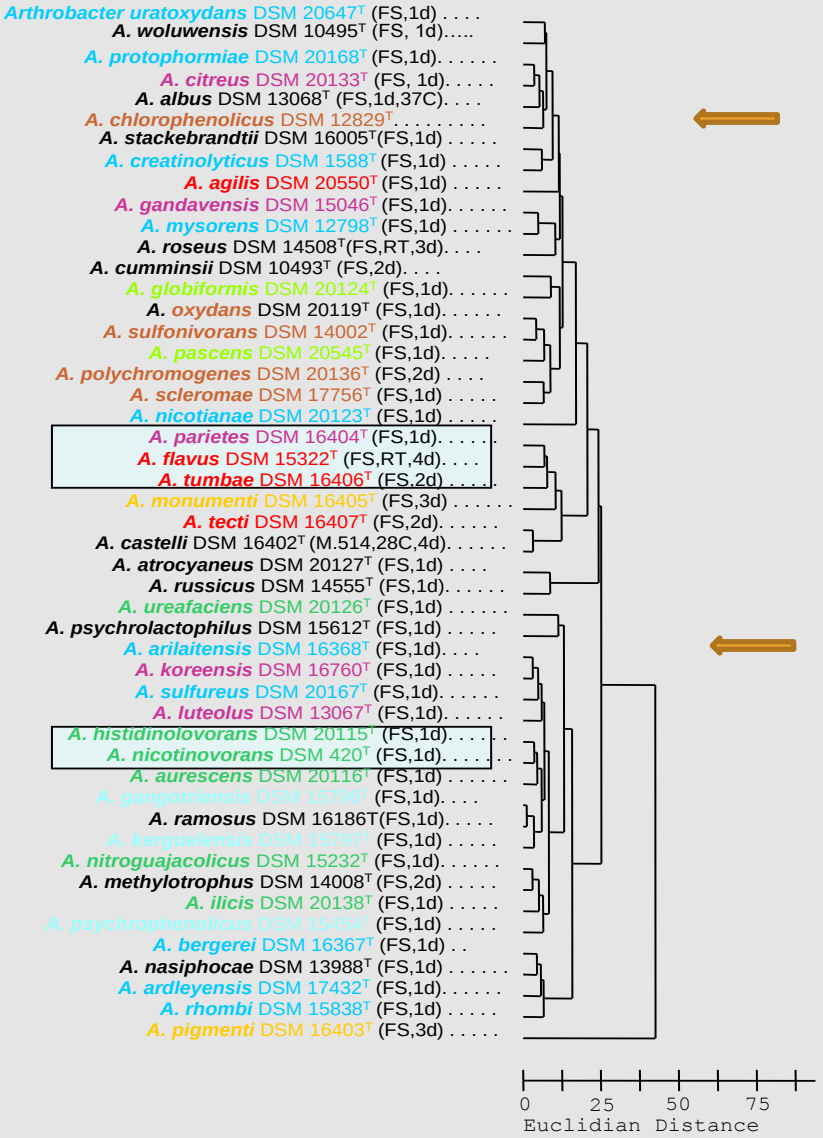
Lys-Ala-Thr-Ala

Lys-Ser-Thr-Ala

Peter Schumann, DSMZ

EMbaRC

DSMZ
Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH



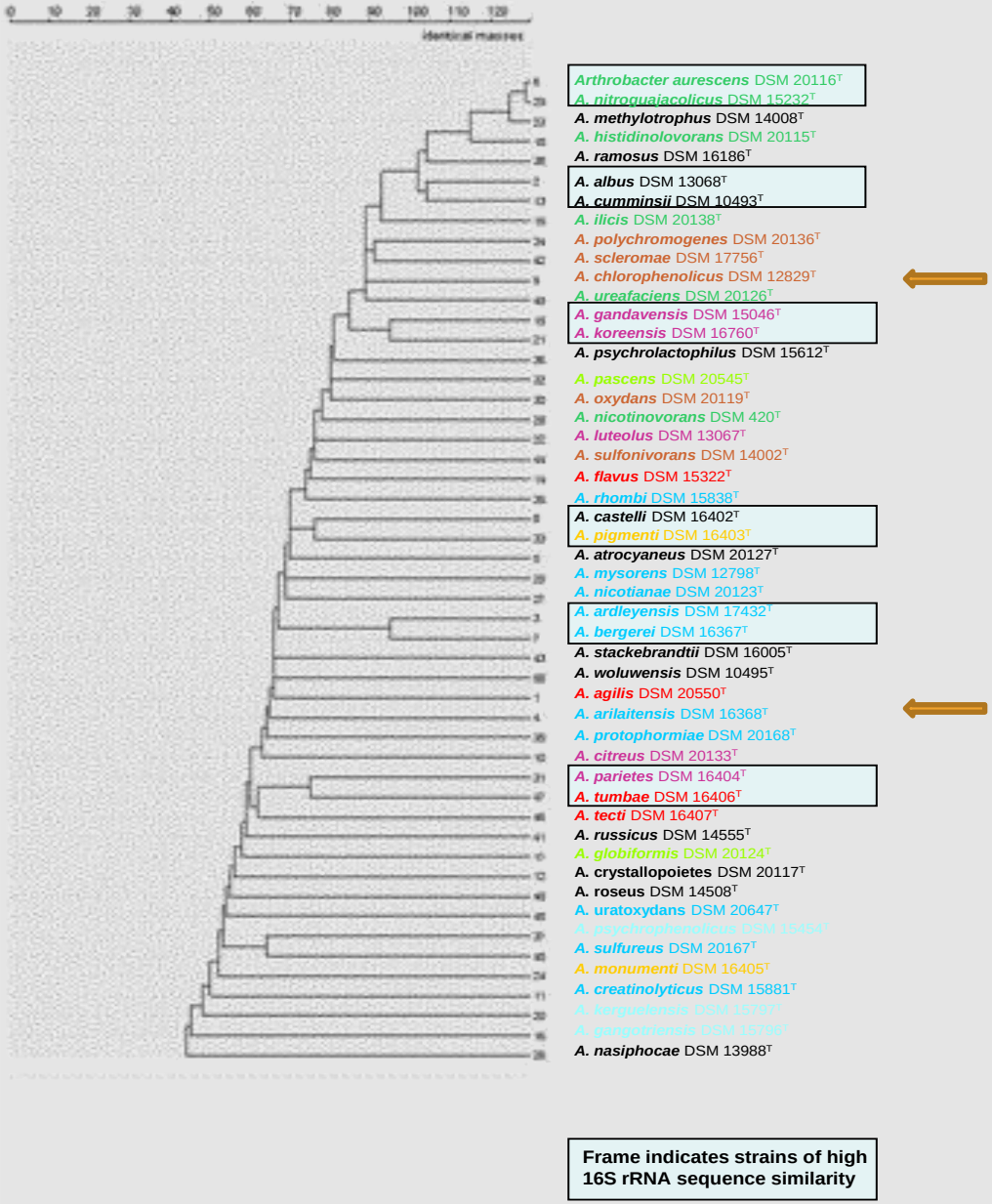
Frame indicates strains of high 16S rRNA sequence similarity

Fatty acid profiles of *Arthrobacter* type strains consist of 4-5 iso/ anteiso branched major components, **are highly similar and do not allow unambiguous differentiation**. Strain DSM 16403^T contains additionally $C_{20:0}$ and 2 unidentified late eluting components.

Peter Schumann, DSMZ

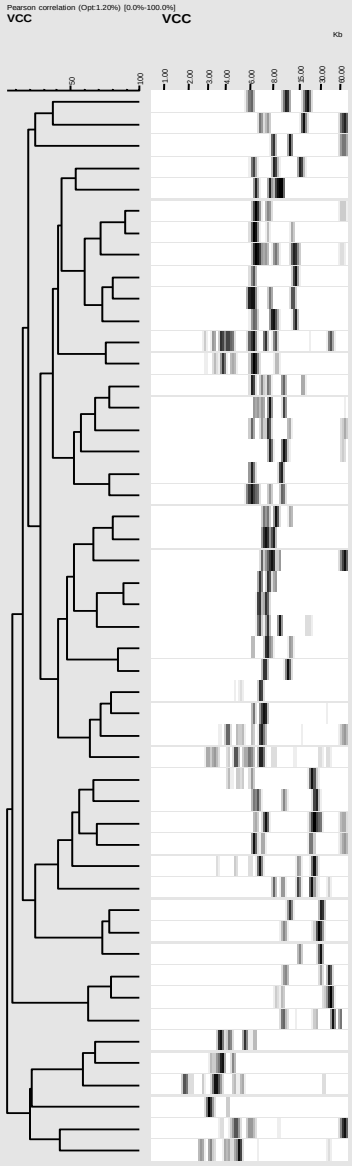
EMbaRC

DSMZ
Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH



MALDI-TOF mass spectra differentiate *Arthrobacter* type strains (except of DSM 20116^T and DSM 15232^T) and **allow their identification**. Six pairs of type strains with similar 16S rRNA gene sequences show also similar mass spectra, demonstrating that MALDI-TOF reflects the phylogenetic relationship of highly related *Arthrobacter* species.

Peter Schumann, DSMZ



Arthrobacter	*albus*	DSM13068T
Arthrobacter	*polychromogenes*	DSM20136T
Arthrobacter	*russicus*	DSM14555T
Arthrobacter	*cumminsii*	DSM10493T
Arthrobacter	*creatinolyticus*	DSM15881T
Arthrobacter	*sulfonivorans*	DSM14002T
Arthrobacter	*nitroguajacolicus*	DSM15232T
Arthrobacter	*ureafaciens*	DSM20126T
Arthrobacter	*castelli*	DSM16402T
Arthrobacter	*aurescens*	DSM20116T
Arthrobacter	*ilicis*	DSM20138T
Arthrobacter	*bergerei*	DSM16367T
Arthrobacter	*ardleyensis*	DSM17432T
Arthrobacter	*histidinolovorans*	DSM20115T
Arthrobacter	*citreus*	DSM20133T
Arthrobacter	*chlorophenolicus*	DSM12829T
Arthrobacter	*rhombi*	DSM15838T
Arthrobacter	*pigmenti*	DSM16403T
Arthrobacter	*nicotinovorans*	DSM420T
Arthrobacter	*luteolus*	DSM13067T
Arthrobacter	*flavus*	DSM15322T
Arthrobacter	*gandavensis*	DSM15046T
Arthrobacter	*koreensis*	DSM16760T
Arthrobacter	*scleromae*	DSM17756T
Arthrobacter	*parietes*	DSM16404T
Arthrobacter	*stackebrandtii*	DSM16005T
Arthrobacter	*agilis*	DSM20550T
Arthrobacter	*kerguelensis*	DSM15797T
Arthrobacter	*oxydans*	DSM20119T
Arthrobacter	*gangotriensis*	DSM15796T
Arthrobacter	*sulfureus*	DSM20167T
Arthrobacter	*nasiphocae*	DSM13988T
Arthrobacter	*globiformis*	DSM20124T
Arthrobacter	*ramosus*	DSM16186T
Arthrobacter	*pascens*	DSM20545T
Arthrobacter	*crystallopoietes*	DSM20117T
Arthrobacter	*atrocyaneus*	DSM20127T
Arthrobacter	*monumenti*	DSM16405T
Arthrobacter	*tecti*	DSM16407T
Arthrobacter	*tumbae*	DSM16406T
Arthrobacter	*methylotrophus*	DSM14008T
Arthrobacter	*psychrolactophilus*	DSM15612T
Arthrobacter	*mysorens*	DSM12798T
Arthrobacter	*psychrophenolicus*	DSM15754T
Arthrobacter	*protophormiae*	DSM20168T
Arthrobacter	*arilaitensis*	DSM16368T
Arthrobacter	*nicotianae*	DSM20123T
Arthrobacter	*woluwensis*	DSM10495T
Arthrobacter	*uratoxydans*	DSM20647T

Frame indicates strains of high 16S rRNA sequence similarity

Riboprints:

*Pvu*II-RiboPrint patterns are useful for differentiation of *Arthrobacter* type strains but do not correlate with their phylogenetic relationship

Peter Schumann, DSMZ

EMbaRC

DSMZ
Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH

ad 1. Despite recognized needs we see a reduction in numbers of taxonomists

Today, **traditional microbiology found its alliance in molecular biology** and modern non-culture tools are now available to recognize the vast diversity of microorganisms, a small fraction of which has been grown in culture. The huge numbers of microbial species yet to be discovered requires skilled taxonomists, innovative isolation strategies, automated identification and a high quality global network of bioinformation of properties of organisms already in culture.

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

Conclusion ad 1. Improve taxonomy

Taxonomic expertise is absolutely essential to ensure microorganisms are correctly identified and culture collections need access to such skills to ensure that the identity can be monitored during storage, handling and distribution.

Maintain the available skills: do not fully replace trained taxonomists by molecular biologist

Expand the molecular skills: apply sequence analysis and rapid and reliable DNA pattern analysis

Be involved in genome sequencing projects

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

ad 2. Storage and maintenance

The primary objective of preserving and storing an organism maintaining it in a viable state **without morphological, physiological, or genetic change** until it is required for future use.

Complete viability and stability should be achieved, especially for **important** research and industrial isolates.

Preservation techniques range from

continuous growth methods

to methods that reduce rates of metabolism

to the ideal situation where metabolism is suspended.

No preservation technique has been successfully applied to all microorganisms, although storage in liquid nitrogen appears to approach the ideal.

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

ad 3. Quality management

The global system for Biological Resource Centres needs a **common standard** that can be worked to by all its members **to ensure conformity** and therefore at the very least the system chosen must be based on a common general standard.

Several collections have already adopted ISO 9000 series certification, a system that ensures quality through critical management of processes.

The system requires that

procedures and practices are documented and that

auditing procedures are put in place

to ensure that what is said is done is actually carried out.

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

ad 3. Information technology

There are enormous possibilities for generating information on microorganisms from descriptive text on

- morphology,
- information on isolation and geographic location,
- host and substrate etc.,
- to digital images,
- metabolic
- genomic data.

Scatterered information is available:
- World Data Centre for Microorganisms (http://wdcm.nig.ac.jp),
- Global Information Facility (http://www.gbif.org),
- EMBL and GENBANK

To ensure the data provided by culture collections is authentic, of high quality and relevance they must institute quality assurance measures for recording, management and exchange.

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

ad 3.: No collection is working fully independently

Actions like handling, maintenance, storage, shipping etc must be carried out safely and compliant with the various legislation and regulations that control these matters. Not only does the legislation exist but also from time to time it is changed or added to (http://wdcm.nig.ac.jp/wfcc/wfccreports.pdf).

CCs/BRCs must comply with biosafety requirements.

These responsibilities are wide ranging and incorporate:

- Health and Safety requirements
- Classification of Micoorganisms on the Basis of Hazard
- Quarantine regulations
- Ownership of Intellectual Property Rights (IPR)
- Convention on Biological Diversity
- Safety information provided to the recipient of microorganisms
- Regulations governing shipping of cultures
- Control of Distribution of Dangerous Organisms

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

Summary 1:
In order to cope with the anticipated massive expansion of biological resources, including living biological materials and data on genomics, CC/BRCs need to:

- Contribute to the **co-ordination** of efforts to **conserve biodiversity** and to **provide access** to natural and engineered biological resources.
- Assist in the development of a **co-ordinated** international system for decision making to guide appropriate acquisition, maintenance and distribution of biological resources so as to avoid unnecessary duplication of effort while preserving critical levels of biodiversity.
- **Modernise** to incorporate the latest developments in web-based electronic communication, bioinformational science and informatics technologies.
- **Co-ordinate** and unify catalogues and databases to meet the requirements of science in the developing post-genomics era.
- **Develop** new systems and technologies for the long-term maintenance and distribution of large numbers of diverse biological resources.
- **Co-ordinate** curation, as well as development and networking of informatics tools for data analysis, comparison and visualisation.
- Ensure that the scientific community has access to affordable products and services.

Summary 2.: Culture collection benefits

- Recognition that they operate to international scientifically based quality criteria
- An international mark of quality
- Raised profile
- Sharing of tasks
- Common policies and procedures
- Competitive edge
- Common access to data enabling links to be made to other international initiatives without duplication of effort
- Common approach to data access, sharing and interoperability
- Improved data usage
- Collaborative research and development

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

Summary 3.: Potential Income Streams Anticipated for CCs/BRCs (examples)

- cDNA libraries, genomic libraries, filter sets, clones, plates, PCR products
- Microarrays and reagents
- Accreditation/standardization-added value products and services
- Data storage and retrieval
- Software development/collaborations - data mining tools
- Technology development/collaborations
- Sequence database annotation/phenotypic analysis
- Linking genomics databases to proteomics

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.

At the end: where are the Culture Collection benefits from increasing its performance?

- Recognition that they operate to international scientifically based quality criteria
- An international mark of quality
- Raised profile
- Sharing of tasks
- Common policies and procedures
- Competitive edge
- Common access to data enabling links to be made to other international initiatives without duplication of effort
- Common approach to data access, sharing and interoperability
- Improved data usage
- Collaborative research and development

Smith D, Ryan MR, Stackebrandt E. Doelle, H. W. and Da Silva, E. J. Encyclopedia of Life Support systems (EOLSS)[Biotechnology]. 2008. UNESCO, Eolss Publishers, Oxford, UK.