

Bioinformatic tools for bacterial identification and characterization

Giovanna Felis

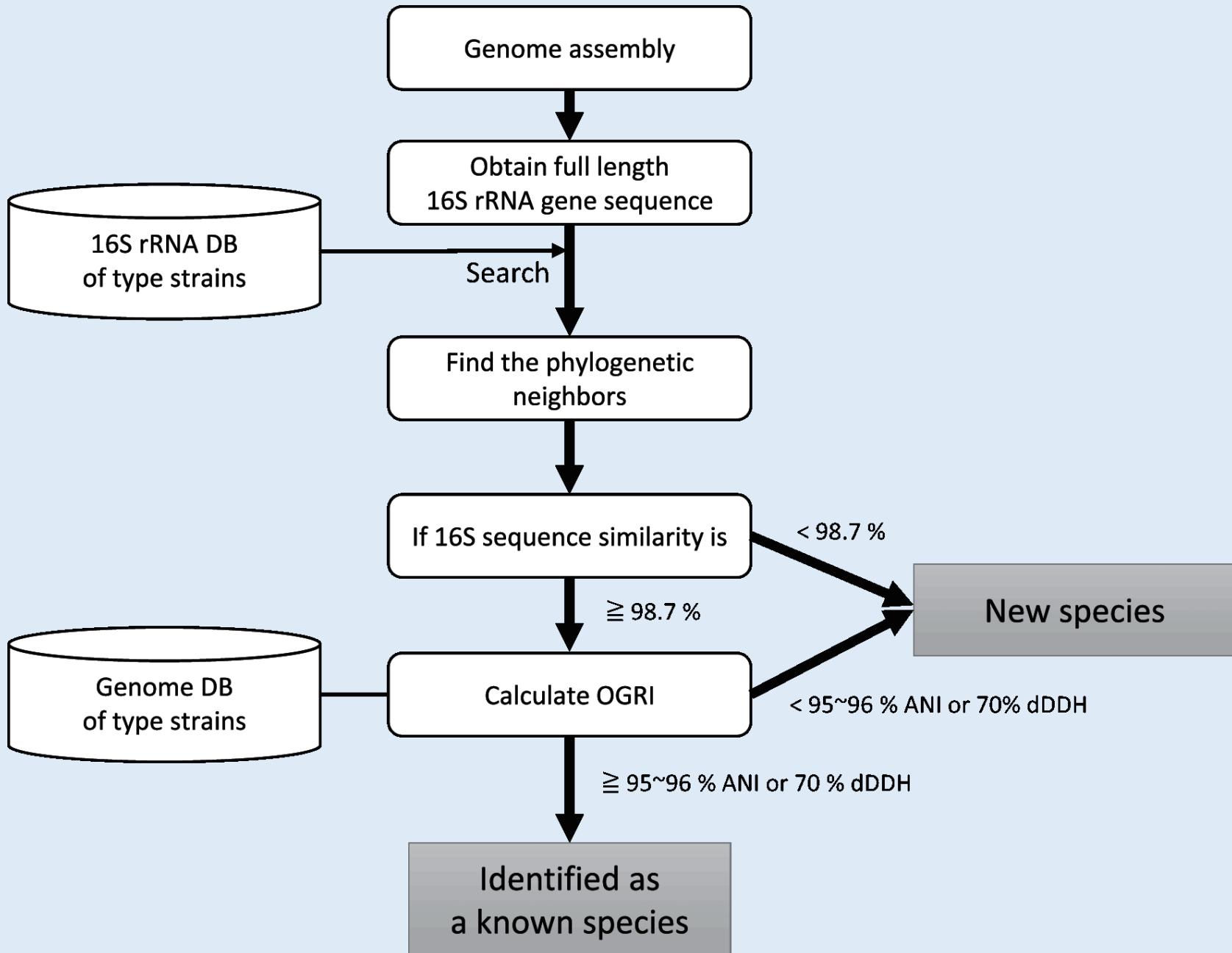
University of Verona – Dept. Biotechnology

giovanna.felis@univr.it

@FelisGiovanna

Today – day2

- **Identification** – species level
 - SILVA
 - EZ biocloud
 - ANI calculation
- **Nomenclature**
 - List of Prokaryotic Names with standing in Nomenclature (LPNS)
- **Characterization** – strain level
 - Antibiotic-resistance
 - Virulence genes
 - Carbohydrate active enzymes
 - CRISPR/phages
 - Reg precise (transcriptional regulators)
 - KEGG – Kyoto Encyclopedia of genes and genomes
 - <https://microbe.directory/>
- **Phylogenetic** analysis



16S rRNA gene analysis

- Download partial/complete sequence (e.g., from DGGE band)
 - Acc number: AJ965482
- Tools for identification
 - NCBI, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
 - SILVA, <https://www.arb-silva.de/> (most recent, Glöckner et al. 2017 *J. Biotechnol.* 261: 169-176)
 - EZ Biocloud, <https://www.ezbiocloud.net/> (Yoon et al. 2017 *IJSEM* 67:1613-17)
 - (<http://jspecies.ribohost.com/jspeciesws/#home>)
- Discussion of results

Whole genome comparison

- Choose a species
- Genome sequence accession numbers
- tools
 - <http://ggdc.dsmz.de/ggdc.php#>
 - <http://enve-omics.ce.gatech.edu/ani>

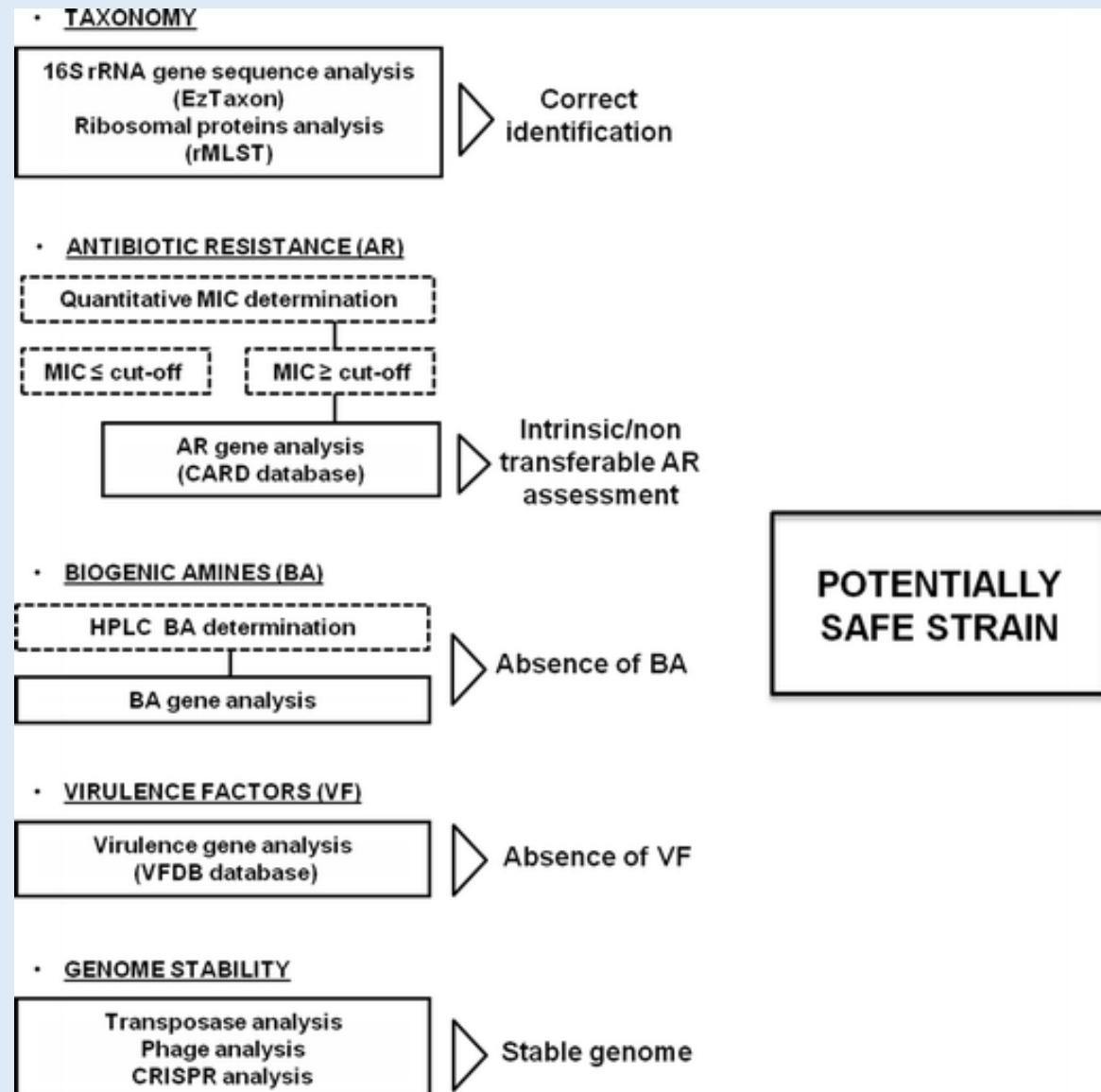
Genome-level comparison

Algorithm	Function	Type	URL/Reference
OrthoANI with usearch	Calculation of ANI	Standalone	https://www.ezbiocloud.net/tools/orthoaniu [9]
OrthoANI with usearch	Calculation of ANI	Web service	https://www.ezbiocloud.net/tools/ani [9]
Genome-to-Genome Distance Calculator	Calculation of dDDH	* Web service	http://ggdc.dsmz.de/ggdc.php [7]
ANI calculator	Calculation of ANI	* Web service	http://enve-omics.ce.gatech.edu/ani/
JSpecies	Calculation of ANI	Standalone	http://imedea.uib-csic.es/jspecies/ [5]
JSpeciesWS	Calculation of ANI	Web service	http://jspecies.ribohost.com/ [30]
CheckM	Checking contamination	Standalone	http://ecogenomics.github.io/CheckM/ [29]
ContEst16S	Checking contamination	Web service	https://www.ezbiocloud.net/tools/contest16s [28]
BBMap	Calculation of sequencing depth of coverage	Standalone	https://sourceforge.net/projects/bbmap/
Amphora2	Phylogenomic treeing	Standalone	http://wolbachia.biology.virginia.edu/WuLab/Software.html [21]
BIGSdb	Phylogenomic treeing	Standalone	https://pubmlst.org/software/database/bigsdb/ [31]
bcgTree	Phylogenomic treeing	Standalone	https://github.com/iimog/bcgTree [32]
Phylophlan	Phylogenomic treeing	Standalone	https://huttenhower.sph.harvard.edu/phylophlan [22]
UBCG	Phylogenomic treeing	Standalone	https://www.ezbiocloud.net/tools/ubcg

TABLE 1. Web-services and standalone software tools for taxonomic purposes

Strain characterization

In silico safety assessment



Antibiotic resistance - genotype

- **Comprehensive Antibiotic Resistance Database (CARD)**
<https://card.mcmaster.ca/> with BLASTP
- **Virulence Factor Database (VFDB)**
<http://www.mgc.ac.cn/VFs/main.htm>

Coverage	similarity	E	Number of genes
> 70 %	> 30 %	< 1e-2	109

Other relevant genetic traits

- Putative prophage sequences
 - **ProphageFinder**, <http://bioinformatics.uwp.edu/~phage/DOEResults.php>
 - **PHASTER**, <http://phaster.ca/>
- Clustered regularly interspaced short palindromic regions (CRISPR) with **CRISPRFinder** <http://crispr.i2bc.paris-saclay.fr/Server/>

Other relevant genetic traits

- Genes related to **biogenic amine** production (mainly aminoacids decarboxylases) **BLASTX**, with seed sequences:
 - complete operon sequence of tyrosine decarboxylase from *B. cereus* BAG2X1-1 (JH792376.1) for the production of tyramine,
 - histidine decarboxylase from a *B. coagulans* strain (AB553281.1) for the production of histamine,
 - arginine decarboxylase from *Bacillus thuringiensis* HD1 (CP010005.1),
 - agmatine deiminase from *B. cereus* 2_A_57_CT2 (NZ_GL635753.1)
 - putrescine carbamoyl transferase from *Bacillus massilioanorexius* AP8 (NZ_CAPG01000089.1),
 - N-carbamoylputrescine amidase from *Bacillus cellulosilyticus* DSM 2522 (CP002394.1),
 - arginine deiminase from *B. cereus* Al Hakam (CP009651.1)
 - ornithine carbamoyltransferase from *Bacillus cytotoxicus* NVH 391-98 (CP000764.1) for putrescine.
- **Enterotoxin genes:**
 - haemolysin BL (HBL complex; hblC, hblD, hblA and hblB: AJ007794),
 - non-haemolytic enterotoxin NHE (NHE complex; nheA, nheB and nheC: Y19005),
 - Enterotoxin T (bceT; D17312),
 - cytotoxin K (cytK; AJ277962) (Guinebretière et al. 2002)
 - cereulide (cesA, cesH, cesP, cesT, cesB, cesC, cesD; DQ360825) (Ehling-Schulz et al. 2006)
 - Synthesis of lipopeptides, as fengycin (fenA, AF023464; fenD, CAA09819; fenE, AF023465),
 - Surfactins (srfAA, D13262; srfAB, AF233756; sfrAC, CAB12145) (Tapi et al. 2010) and
 - lichenisin (lchAA, lchAB, lchAC; AJ005061) (Yakimov et al. 1998)

Other interesting tools

- web server and DataBase for automated Carbohydrate-active enzyme Annotation
<http://csbl.bmb.uga.edu/dbCAN/>
 - Analysys on your own genome
- Pathogen-Host Interactions database, <http://www.phi-base.org>
 - Genes to be used as seed against your genome
- KEGG - Kyoto Encyclopedia of Genes and Genomes <http://www.genome.jp/kegg/>
- database of Clusters of Orthologous Groups of proteins (COG)
 - <http://www.ncbi.nlm.nih.gov/COG/>
 - <ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/static/lists/listCOGs.html>
- <https://microbe.directory/>
- Reg precise (transcriptional regulators)
 - <http://regprecise.lbl.gov/RegPrecise/>

Phylogenetic analysis

Online tools

- List of tools: <https://molbiol-tools.ca/Phylogeny.htm>
- Tree of Life: <https://itol.embl.de/>
- https://www.ebi.ac.uk/Tools/phylogeny/simple_phylogeny/
- <http://www.phylogeny.fr/>
- Standing alone
 - MEGA software <https://www.megasoftware.net/>