

COST Action “Endophytes for Biotechnology and Agriculture” (FA1103) Endophyte Database Workshop

8-9 January 2013
Tulln, Austria

PARTICIPANTS

Gerhard Rambold, Pablo Hardoim, Nele Weyens, Brian Murphy, Nicole Brinkmann, Hatice Özaktan, Silvia Fluch, Birgit Mitter, Stephane Compant, Friederike Trognitz, Angela Sessitsch

PROTOCOL

S. Fluch gives an overview of the Evoltree database and e-lab and the endophyte database established already at AIT; B. Mitter explains how the parameters were defined for the AIT endophyte database; Furthermore, S. Fluch gives an introduction to legal aspects. Discussion on presentation

Benefits of an endophyte database

- Overview of existing material > scientific and commercial value
- Endophyte information currently too scattered > more focused in a database
- Linking traits / functions with sequence information (whereas in e.g. NCBI only sequences are deposited)
- Better visibility of COST Action / of endophyte research
- Scientific publications
- Applications / interest by companies
- Access to information on references and key experimentators
- Organisation of own strains

Who will be the stakeholders / use the database?

- Important to define
- Only database, not strain collection (strains still with owners)

Which parameters have to be in the database to be interesting for users (i.e. those depositing data and those accessing the database)?

- Questionnaire for COST FA1103 members and other scientists to obtain information on
 - Information on app. how many strains can be provided?
 - Which data have been recorded so far?
 - Which type of data is considered as important?
 - Which kind of queries would be run?
 - See also below

Other points (miscellaneous)

- Catchy name! “EndoBase”?
- Core / steering group to further develop procedures, inform COST FA1103 and push database idea – workshop participants + Carolin Schneider
- Data should be downloadable (xls or other format)

- Check Microbial Common Language Project > use of correct microbiology terms
- Unknown names (taxa) – give recommendation how to do this
- Check “Diversity Workbench” (www.diversityworkbench.net)
- Minimal recommended information: geography, time, plant species, plant cv, plant tissue, vegetation stage, greenhouse/field, barcode sequence, how was strain isolated, reference how; reject only strains without the information on plant, geography, time, barcode sequence (16S rRNA / ITS)
- High-image resolution images of preliminary herbal imprints for information on plant species etc would be important for determining plant diagnostic characteristics
- Info on soil type, deposit of plant material / DNA, accompanying microbial community analysis could be additional interesting aspects
- Information on protocols used should be integrated in database > could be also useful to in future recommend procedures

Funding situation

- No funds available within COST except for STSM and meetings
- Database curator needed for good database management
- EndoBase v.1: could be established in xml-based data exchange format or based on AIT database; first version to see how it is adopted; further funding has to be obtained e.g. from Horizon2020 (infrastructure programme) to make it more professional (EndoBase v.2)
- “Membership” system to (co)-fund running costs, different types of memberships – academic / commercial, type of information required, data provided vs no data provided etc.; open access: plant, taxon, PGP yes/no, biocontrol yes/no etc., for more information membership needed
- Workshop could be eventually held with 1 member/student of key data providers to get data together in a format, which can be easily uploaded in the database > this maybe could be funded by COST

Questionnaire

- Questionnaire preparation: group of elements (e.g. identification), element (mitospore), attribute (mitospore length), value; we agree first on group of elements, elements etc. in a hierarchical manner
- Categorical type of data (diff categories), ordered data (negative, low, high, very high etc.), not ordered data, binary data (yes/or), value information types, text dat type of input
- How many strains do you have to deposit in the DB (1-10, 11-100, etc.)
- Type of queries you would like to run?
- Traits analysed so far (incl information on methods / category of method) / data, which are considered as important
 - Host data: plant species, cultivar, clone, history of plant material (e.g. important for plant tissue cultures)
 - Sampling conditions recorded: geographic information, vegetation stage, time, mode of isolation (method of surface sterilization or only mechanical removal of surface microbiata/fungi), media, cultivation time, soil type, soil pH, cropping history, pathogen stress, availability of external, metereological data, plant tissue,
 - Storage of material: methof storage (lyphilization, glycerol stocks (-20 or -80°C), how many times has a microbe been sub-cultured, storage of plant material at -20 or -80°C, DNA isolation
 - Identification: molecular (16S, ITS) is obligatory for the DB; fatty acids, morphology, intracellular structures (fungi), substrate utilization, MALDI-TOF, API, BIOLOG, phylogenetic analysis, which markers besides 16S and ITS, Gram-staining, strain-specific identification, microsatellites, whole genome sequences, plasmid information, DNA-DNA hybridization, deposited already in official culture collections, risk group information
 - Strain / DNA / RNA public availability? GM variants? Patents?

- Plant colonization (conditions in vitro, greenhouse, field; microscopy, re-isolation, cv-independent, qualitative vs. quantitative, RT-PCR, flow cytometry, immune methods, same or different plants in comparison to original host plant, spatial distribution, inter- and/or intracellular colonization)
- Plant growth promotion: IAA, ACCD, siderophore, N-fixation (sequence, acetylene reduction, ¹⁵N), P-solubiliz, surfactants, proline production, cytokinins, VOC, vitamins, S solubilisation, exopolysaccharides, organic acids, antioxidative potential (respiration activity in presence of oxidative substances), trehalose, QS; include information on qual., quant.
- Effects on system plant-microbe: if tested in planta (conditions), biomass production/crop yield, effects on photosynthesis (fluorescence intensity, CO₂ rates), IAA, chlorophyll, agronomic data (biomass, flower onset, length of radicle and hypocotyledon, seed emergence, stomatal resistance, antioxidative enzymes, QS, N-fixation, element composition, lignin content; should be maybe a separate DB linked to this as all the plant metadata are needed)
- Biocontrol: direct mechanisms towards bacterial, fungal or oomycotal phytopathogens or pests (antibiotics, siderophores, lytic enzymes secretion, parasitism, production of toxins, reduction of pathogen virulence), indirect mechanism (ISR or SAR) towards pathogens and abiotic stresses, outcompetition
- Trace element-related traits (Nele will define parameters)
- Organic compound degradation (Nele will define parameters)
- Enzyme production
 - People should type in which enzymes they found > how was analysis done (substrate enzyme assay, expression, sequence found...)
- Metabolites
 - People should type in which metabolites they found > how was analysis done (chemical identification, expression, sequence found...)
- Secondary metabolites
 - People should type in which secondary metabolites they found > how was analysis done (chemical identification, expression, sequence found...)

Next steps

- Set up of questionnaire (workshop participants)
- Info on FA1103 webpage
- Brief presentation on the database idea and the questionnaire in Berlin (poster + few slides in frame of a presentation) to inform and raise interest
- Sending out questionnaire to COST FA1103 members and to other endophyte scientists (outside Europe)