Insight and pitfalls in meta-transcriptomics

Lucy Stark and Röbbe Wünschiers

HOCHSCHULE MITTWEIDA
UNIVERSITY OF APPLIED SCIENCES

Faculty of Mathematics, Natural Sciences and Computer Science

Dipl.-Ing. (FH)
Lucy Stark

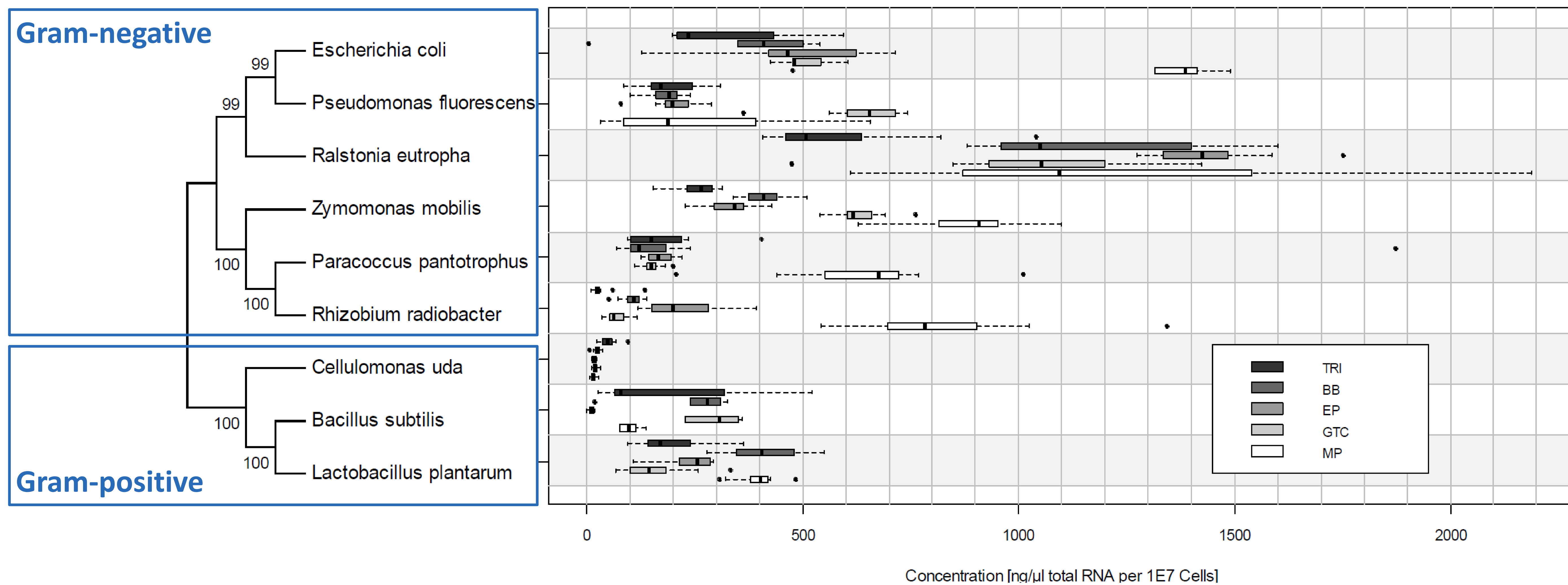
Technikumplatz 17
09648 Mittweida/Germany

Phone +49 (0) 3727 58-1181
stark@hs-mittweida.de

The multi-staged microbial conversion of biomass to biogas is prone to failures due to the difficult thermodynamics of the involved microorganisms. Monitoring via chemical parameters can be problematic because changes are mainly a consequence of disturbances between the microbes. With advances of next-generation sequencing technologies, analyzing the metatranscriptome and therefore investigation of the regulatory dynamics becomes feasible.

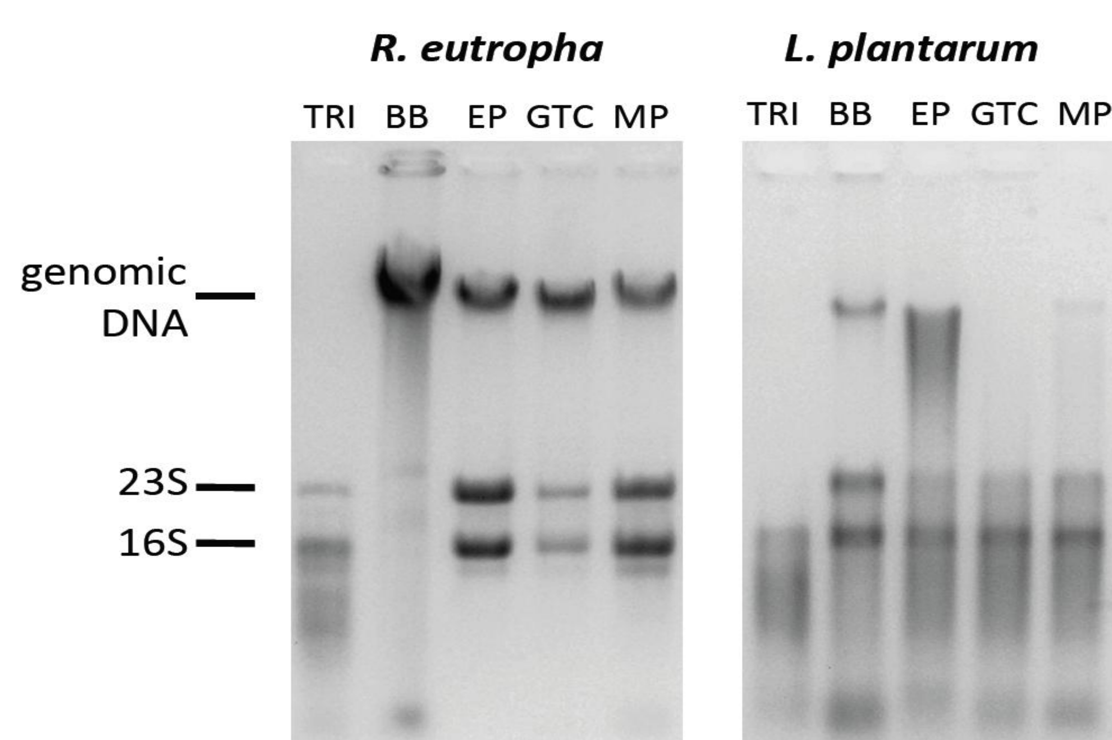
High quality RNA is essential for reliable RNA-Seq data. In order to evaluate unbiased isolation of transcripts from mixed-species environmental samples, we applied five different RNA-extraction protocols to nine taxonomic diverse bacterial. We found that the extraction efficiency of different methods depends strongly on the target organism. Transferring our results to mixed-species investigations leads to the conclusion that particular microorganisms might be over- or underrepresented depending on the method applied.

With keeping this fact in mind, the active microbial transcriptome from a biogas plant fed with different renewable substrates was analyzed with the automated analysis platform MG-RAST. From nearly 7.5 Mio hits a little more than 1.7 Mio sequence reads could be annotated to proteins, which are related to approximately 200.000 functional categories.

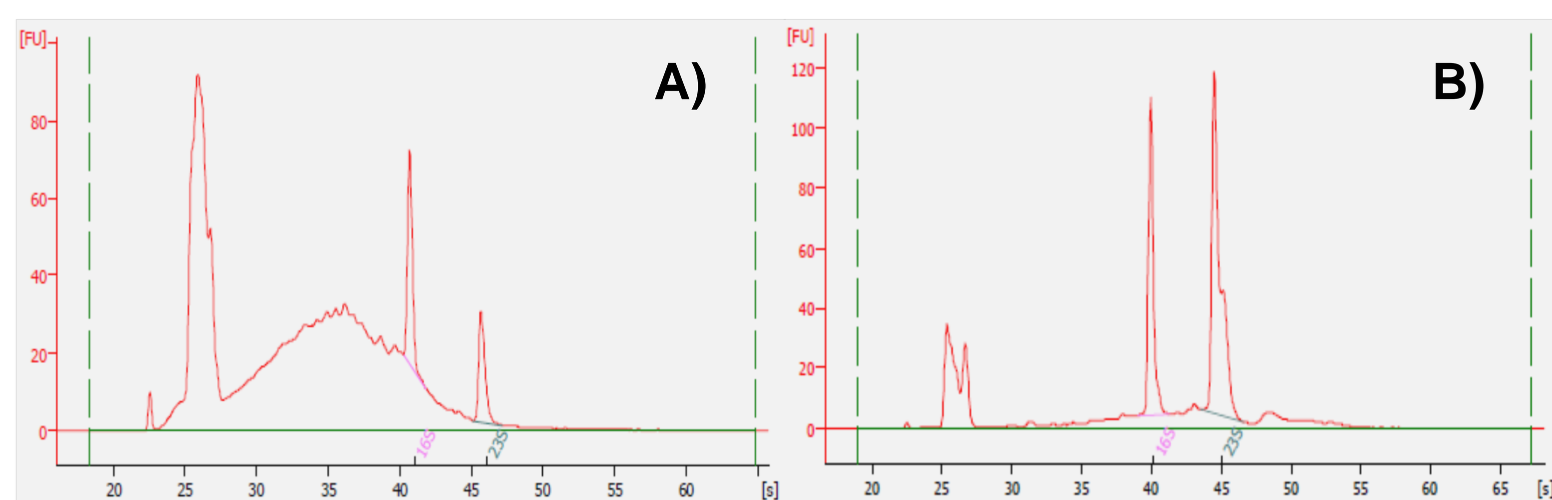


1) RNA yield is plotted for each species and each method, respectively. Bacterial species are ordered according to their 16S rRNA relationship as analyzed by the neighbor-joining method. Branching points are labeled with percentage concordance for 1,000 bootstrap repetitions. Boxes contain the 2nd and 3rd quartile (Q) of each data set, separated by median. Whiskers extend to the minimum and maximum of the distribution, respectively. If the maxima extend to more than 1.5-times the interquartile range, they are plotted as individual dots. One outlier for *R. eutropha* (method MP) with a value of 3,400 has been omitted.

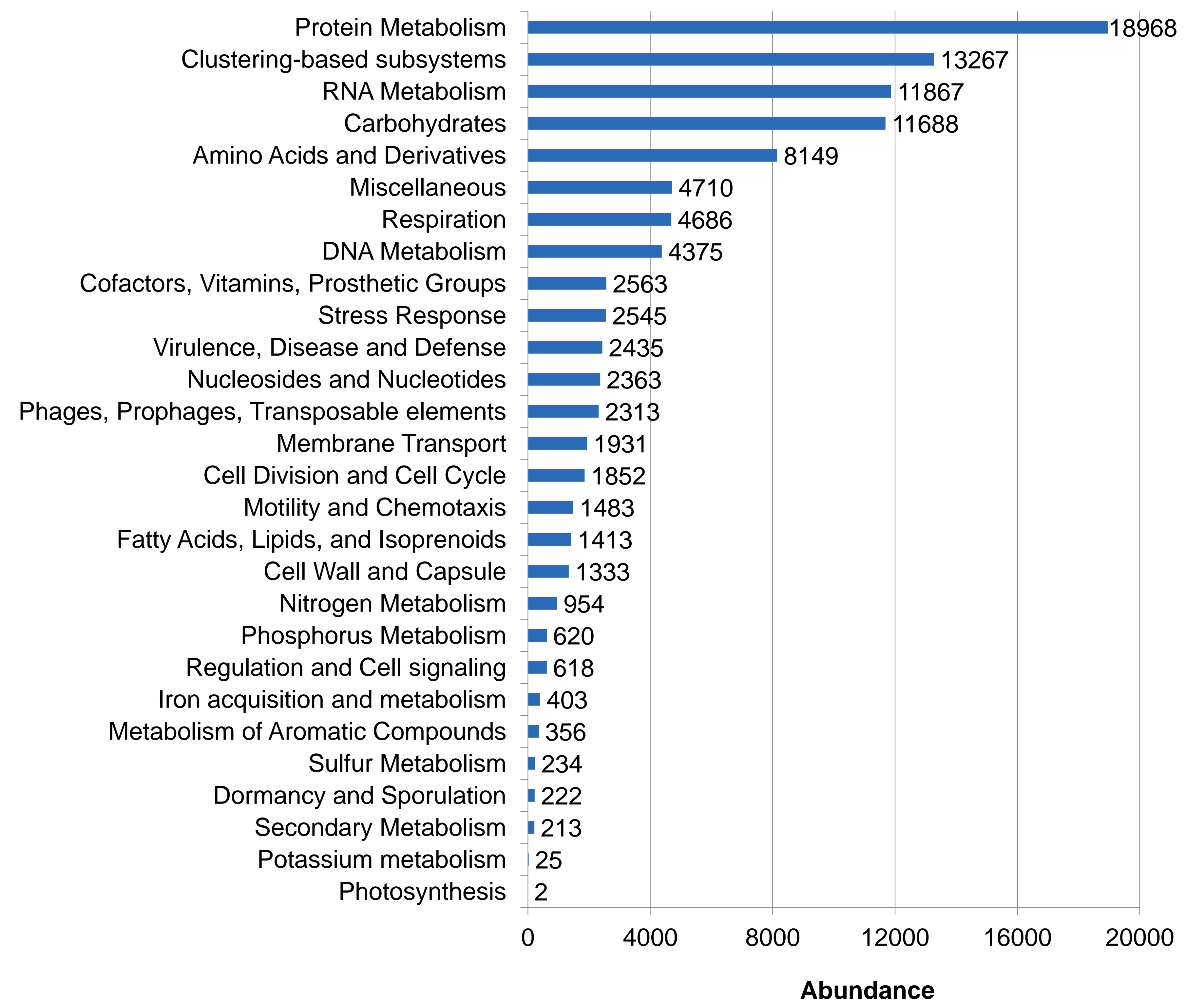
Method shortcuts: **TRI** = TRI Reagent LS and ultrasonic disintegration; **BB** = Bead Beating; **EP** = Extraction buffer and Proteinase K; **GTC** = guanidinium thiocyanate; **MP** = MasterPure



2) Agarose gel images of *L. plantarum* and *R. eutropha* nucleic acid extraction prior to DNase treatment. Samples run on two different gels. Each lane was loaded with 3.5 µg total nucleic acid.



3) Quality Control was done with Agilent's Bioanalyzer 2100. An *E.coli* overnight culture was lysed with **A)** Ultrasonic and TRI Reagent LS (c = 1.036 µg/µl, RIN =3.8) and **B)** MasterPure RNA Purification Kit (c = 1.984 µg/µl, RIN = 9.2)



4) Functional hierarchical classification. Annotated proteins sorted to COGs in subsystem protein database with at least 80% identity, min. alignment length of 15 amino acids and an Max. e-Value of 1e-5.

Stark L, et al., Efficiency of RNA extraction from selected bacteria in the context of biogas production and metatranscriptomics, Anaerobe (2013), http://dx.doi.org/10.1016/j.anaerobe.2013.09.007

Europa fördert Sachsen.
ESF
Europäischer Sozialfonds