Deliverable 2.4

Workpackage 2

Responsible Partner: 2-AGES, 36-INSA

Contributing partners: 25-NUIG

GENERAL INFORMATION

European Joint Programme full title	**Promoting One Health in Europe through joint actions on foodborne zoonoses, antimicrobial resistance and emerging microbiological hazards**
European Joint Programme acronym	**One Health EJP**
Funding	**This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 773830.**
Grant Agreement	**Grant agreement n° 773830**
Start Date	**01/01/2018**
Duration	**60 Months**

DOCUMENT MANAGEMENT

Title OHEJP deliverable	**D-JRP15-FED-AMR-WP2.4 (Determination of naturally transformable bacteria in tested environmental compartments (T2.5))**
WP and task	WP2, JRP15-R2-WP2-T5
Leader	2-AGES, 36-INSA
Other contributors	25-NUIG
Due month of the deliverable	M46
Actual submission month	M60
Type *R: Document, report* *DEC: Websites, patent filings, videos, etc.; OTHER*	R **Save date:** 14.12.2022
Dissemination level *PU: Public (default)* *CO: confidential, only for members of the consortium (including the Commission Services).*	PU *See updated Grant Agreement*
Dissemination *Author's suggestion to inform the following possible interested parties.*	OHEJP WP 1 ☐ OHEJP WP 2 ☐ OHEJP WP 3 ☒ OHEJP WP 4 ☐ OHEJP WP 5 ☐ OHEJP WP 6 ☐ OHEJP WP 7 ☐ Project Management Team ☒ Communication Team ☒ Scientific Steering Board ☐ National Stakeholders/Program Owners Committee ☐ EFSA ☐ ECDC ☐ EEA ☐ EMA ☐ FAO ☐ WHO ☐ OIE ☐ Other international stakeholder(s): …… Social Media: ……

	Other recipient(s): .....

D-JRP15-FED-AMR-WP2.4

I. Introduction

1. Description of the task JRP15-R2-WP2-T2.5

This deliverable is associated to task JRP15-R2-WP2-T2.5 (*Identify naturally transformable bacterial species in the tested compartments (NGS)*) and in the original project proposal was listed as D-JRP15-FED-AMR-WP2.5. The task aimed at analysing Illumina 16S sequencing data obtained in task JRP15-R2-WP2-T2.3 to detect naturally transformable bacterial species.

In this task, to identify naturally transformable bacteria in the compartments tested, a query database containing bacterial species proven to be able to take up extracellular DNA under naturally occurring environmental conditions was generated. For this purpose, a literature search for naturally competent bacteria covering the years from 1994 to 2021 was performed using defined search strings and PubMed and SCOPUS as reference databases. Only hits describing the development of competence in bacteria under physiological conditions were eligible for entering the query database. The bacterial species names retrieved from literature were annotated and the appropriate phylogenetic levels were allocated according to the taxonomic classification system as laid down in SILVA 138.1. For the annotation and classification of 16S NGS data generated in the frame of the project the same version of SILVA was employed. The resulting collection of observed bacterial species in FED-AMR samples was used as the searched database. This search database was checked for the presence of bacteria sampled in the query database using the level “genus” as qualifier. The research questions were: 1. Are naturally transformable bacteria present in FED-AMR isolates? 2. If yes, in which compartments? 3. Which competent bacteria are to be found in these compartments?

2. Deliverable description

The query database contained 171 entries (= bacterial species) with empirically obtained evidence for their natural transformability. The searched database contained the 16S sequence information of 488 samples and 5763 different bacterial species. Naturally transformable bacteria were identified in almost all samples (except for a sample from soil and from a farmer – both of exDNA as sample type), in all sample types and environmental compartments. The overall most prevalent taxons encountered in FED-AMR samples were Pseudomonads and Sphingomonads. Lactobacilli were found in a significantly lower number of samples but showed highly abundant genus-specific sequence reads. Samples from wildlife (exDNA), farmers (ex/total DNA) and crops (ex/total DNA) carried the lowest abundance of competent bacteria. The highest loads with competent bacteria could be observed in feed (exDNA), groundwater (total DNA) and field drainage (total DNA).

We can conclude that bacteria known to be able to take up extracellular DNA are present in all environmental compartments and may serve as receptor for ARG encoded on free exDNA.

To further investigate the bacterial transformation in the targeted compartments, an in-silico analysis of genes related to bacterial competence was initiated. To do this, the *comEC* gene was chosen as the ComEC protein mediated DNA import during transformation. The ComEC protein sequence has been characterised for known naturally-competent bacteria, providing a suitable genetic marker for bioinformatic searches in metagenomic datasets. This analysis allows us to ask the following questions: 1) are we able to determine the presence of genes that encode for bacterial transformation in the metagenomes? 2) How similar are the FED-AMR *comEC* genes compared to those of known bacteria?

The ComEC protein sequences from naturally competent bacteria were downloaded from the NCBI RefSeq database using gene accessions available in Pimentel *et al*. (2018) (doi: 10.3389/fmicb.2018.02980). The 30 ComEC protein sequences available were used to build a custom DIAMOND-Blast database for the subsequent alignment of translated nucleotide sequences (Diamond BLASTX). Before DIAMOND BLASTX, FED-AMR fastq files obtained from Irish deer faeces and wastewater samples were filtered to remove any non-biological sequences and poor-quality reads.

DIAMOND-BLASTX showed the presence of 185 sequences with sequence similarity to known ComEC proteins. The numbers in wastewater samples varied from 1 to 3 sequences, while deer faeces had 0-64 sequences with similarity to ComEC proteins. The taxonomic profile of the sequences matching ComEC proteins were determined in the Metagenome Analyser Software (MEGAN) using the Lowest Common Ancestor criteria (LCA), and the most abundant taxonomies of the detected comEC sequences were *Streptococcus oralis* (86 sequences), *Deinococcus radiodurans* (10), unclassified *Streptomyces* (9), *Pseudomonas fluorescens* (8), *Escherichia coli* (8), *Bacillus amyloliquefaciens* (6) and Ralstonia (6). These taxonomies require confirmation with phylogenetic analyses.