Horizon 2020 Work programme

Food Security, Sustainable Agriculture and Forestry, Marine, Maritime and Inland Water Research and the Bioeconomy

Call

H2020-FNR-2020: Food and Natural Resources

Topic name

FNR-16-2020: ENZYMES FOR MORE ENVIRONMENT-FRIENDLY CONSUMER PRODUCTS

FuturEnzyme:

Technologies of the Future for Low-Cost Enzymes for Environment-Friendly Products

21/07/2022

SET OF 180 ENZYMES FOR EXPERIMENTAL FOCUS

D2.4

VÍCTOR GUALLAR
BSC
Jordi Girona 31, BARCELONA 08034,Spain

Document information sheet

Work package:	WP2, Machine learning enzyme bioprospecting integrated into an industrial context
Authors:	BSC (Víctor Guallar), CSIC (Manuel Ferrer, Patricia Molina)
Document version:	1
Date:	21/07/2022
Starting date:	01/06/2021
Duration:	48 months
Lead beneficiary:	BSC
Participant(s):	BSC, CSIC, BANGOR, UHAM, UDUS, CNR, IST-ID, SCHOELLER , HENKEL, EVO
Dissemination Level:	Confidential, only for consortium's members (including the Commission Services)
Type	Other
Due date (months)	14
Contact details:	Víctor Guallar (victor.guallar@bsc.es), Manuel Ferrer (mferrer@icp.csic.es), Patricia Molina (patricia.molina@icp.csic.es)

Summary

SET OF 180 ENZYMES FOR EXPERIMENTAL FOCUS

1. Scope of Deliverable

This deliverable consists in a fasta file containing at least 180 full-length top priority candidate sequences encoding enzymes with high probability to fulfil manufacturer specifications. These sequences have been selected in Task 2.4 by iterative and decision-making hierarchical procedure applied to the 1,000 full-length candidate sequences delivered in D2.3. The fasta file will be deposited in the FuturEnzyme internal repository with a report detailing the hierarchical procedure applied for the selection, as well as the annotation and details of the sequence origin.

2. Methodology & results

The starting point of this deliverable is multiple. On the one side, the sequences encoding enzymes relevant to the project pre-selected using homology, machine learning and Hidden Markov Models pipelines, detailed in the deliverable D2.3. On the other side, the enzymes and microorganisms (from previous projects or newly isolated) that according to the activity tests detailed in deliverable D.2.3 and to the activity and stability criteria set out in deliverables D2.2 and D2.3, have shown characteristics of interest. In total, 205 enzymes and 54 microbes containing enzymes have been pointed out to be thoroughly analysed in search of the best candidates to accomplish the objectives of the manufacturers, broadly covering the goal of this deliverable. Out of the 304 candidates, the sequences of 258 enzymes are available, and the genomes of 13 microbial have been sequenced to date (see details in Annex 1 and 2-D2.4, available at the FuturEnzyme web intranet's section shared data).

In **Figure 1** the number of candidates (either single enzymes or isolates) selected by activity is summarized.

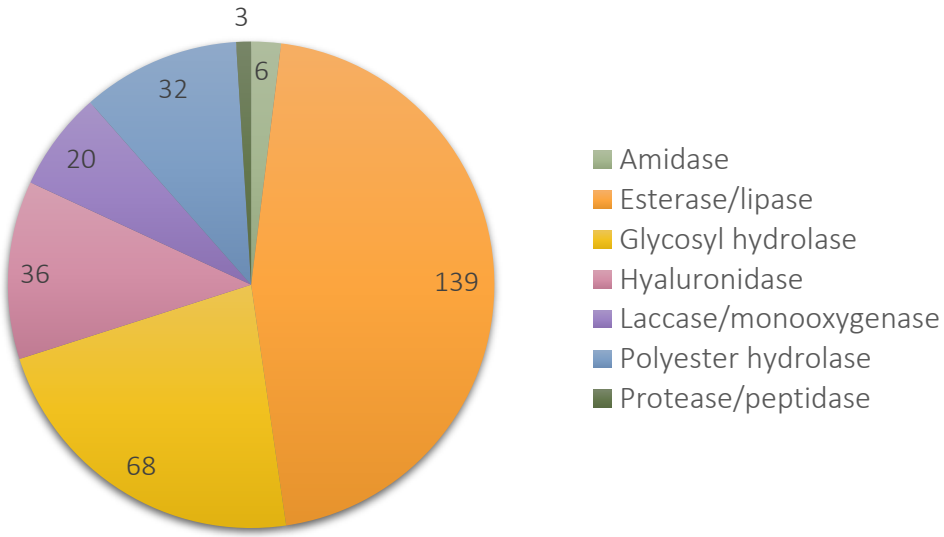

Fig. 1. Number of each type of enzyme selected by activity.

The origin of the selected candidates is highly diverse with the purpose of covering as much environmental, taxonomic and functional diversity as possible. In summary, the selected hits were retrieved from at least 115 different sources (see details in Annex 1-D2.4), which can be classified in 4 main categories (**Figure 2**):

- Non-cultivable microorganisms - metagenome (public database);
- Non-cultivable microorganisms - Metagenome (environmental sample);
- Non-cultivable microorganisms - Metagenome (anaerobic digestor); and
- Cultivable microorganisms and their genomes

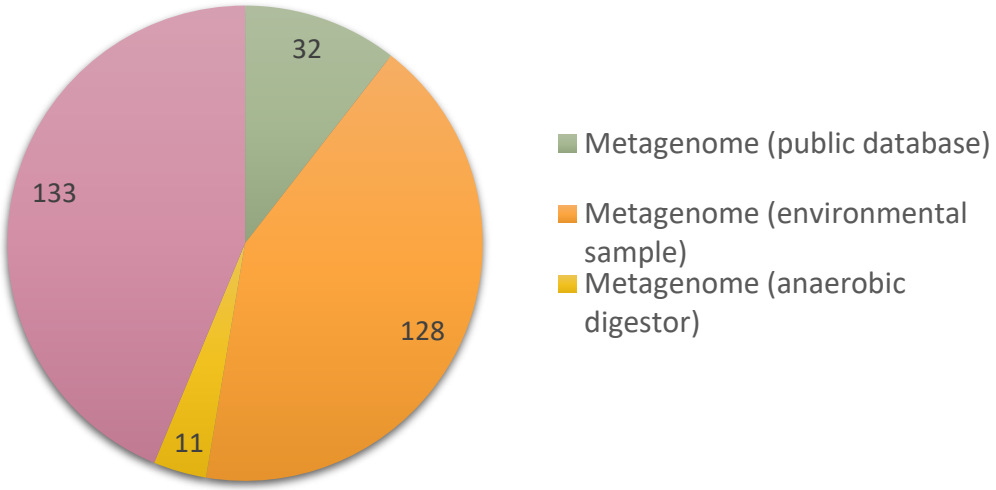

Fig. 2. Number of enzymes by origin.

The different screening methodologies that have been employed to select the candidates of interest in here described are summarized in **Figure 3**, which include:

- Functional screen
- Homology screen
- Homology and computational screen
- Patent screen (for comparative purposes)

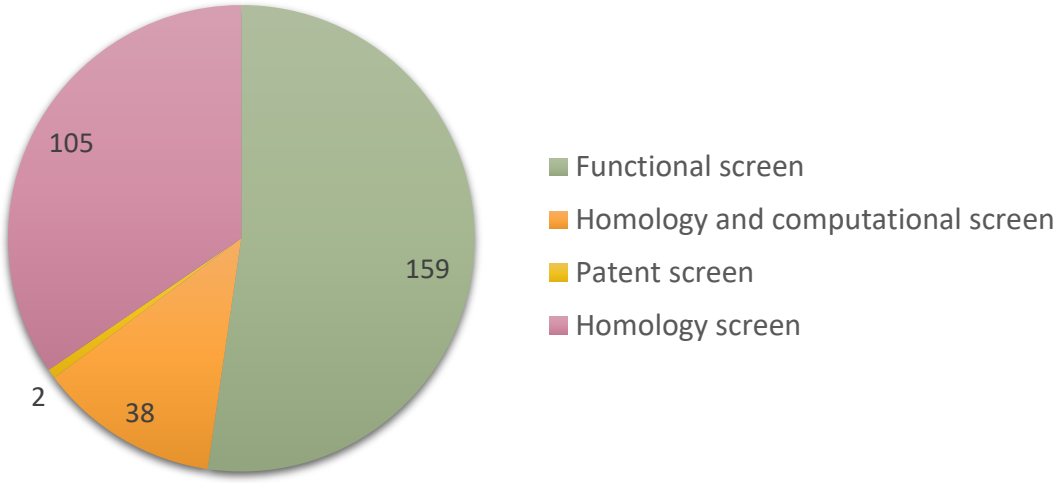

Fig. 3. Number of enzymes by screen method employed for their selection.

The 304 candidates (enzymes and isolates) were selected according to the activity tests performed as detailed in deliverable D.2.3, and to the activity and stability criteria set out in deliverables D2.2 and D2.3. **Annex 1-D2.4** summarises the quantitative or qualitative measurements, including, denaturing temperature (Td), optimal temperature and pH, substrate profile, and stability, to mention most significant characteristics, on the basis of which they were selected. Among most significant parameters, they include, activity towards long-chain lipids relevant to detergent and textile applications, activity either at low or high temperatures as requested for detergent and textile applications, activity against polyesters relevant to textiles, and activity towards hyaluronic acid as requested for cosmetic applications.

These 304 enzymes will be further subjected to in deep characterization, including experimental, and computational and supramolecular engineering, to find both the best candidates to focus on for the next steps, and design improved variants.

3. Annex 1_D2.4

Annex 1-D2.4. Detailed information of the 304 selected hits for experimental focus. Shown are: 1) Enzymatic activity; 2) Name of the candidate; 3) Screen method; 4) Amino acid sequence or genome sequencing status; 5) Origin; 6) Details of activity features, including denaturing temperature (Td), optimal temperature and pH, substrate profile, stability, etc.

Document available at the FuturEnzyme web intranet's section shared data.

4. Annex 2_D2.4

Annex 2_D2.4. Fasta file of the sequences for this deliverable.

Document available at the FuturEnzyme web intranet's section shared data.