Discovering new enzymes just got easier

Industry use case: Biocatalysts Ltd.

What if you could make plants produce biodegradable plastic? Create biofuel out of the bamboo in your back garden? Make textiles without polluting the environment? Produce innovative new food ingredients?

Biotechnology researchers are always looking for ways to make these ideas a reality. To do that, they need to find enzymes – biomolecular machines, or catalysts – that can transform things or break them down.

But finding an enzyme that performs the correct transformation under the right conditions is not that easy. You can't just Google all life on Earth to find an enzyme that does the right job. But you could try using metagenomics-based analysis to identify enzyme of commercial interest. EMBL-EBI's MGnify is a good place to start. This open data resource from the European Bioinformatics Institute (EMBL-EBI) uses metagenomics-based analysis to identify enzyme of commercial interest.

9 ways biotech companies use enzymes

1. Food: better processing of proteins and carbohydrates (e.g. higher nutritional value, better preservation) and healthier flavour enhancers

2. Drink: more efficient processing of raw materials for alcohol fermentation (less waste)

3. Pharmaceutical: diagnosing certain diseases, killing disease-causing microorganisms

4. Fuel: bioethanol from starchy plants (e.g. bamboo or agricultural waste)

5. Detergents: washing powders that work well at cold temperatures (requires less energy)

6. Textiles: fabrics prepared beautifully, without the pollution

7. Industrial processes: synthetic products that require less energy and create less waste

8. Plastics: 'bioplastic' granules produced by bacterial enzymes

9. Paper: break down wood into pulp

Partnering for innovation

Innovate UK supported a collaboration between Biocatalysts Ltd and EMBL-EBI's MGnify team to tackle some of the data challenges in commercial metagenomics research.

Within the project, Biocatalysts developed MetXtra™: a proprietary software platform that integrates databases of metagenomic peptide sequences (both Biocatalysts' and public-domain databases) and links sequences to the sample in which they were found. In parallel, the team at EMBL-EBI developed MGnify, a database containing

THE CHALLENGE

The sheer size of metagenomics datasets makes it nearly impossible to manually "search" for the right enzyme". Since metagenomics is so new, people are still making the tools to analyse this jumble of big data.

This is actually a good problem to have. Before, researchers hunting for a suitable enzyme would take the 'best' sequence they could find, introduce mutations randomly and hope to hit upon the desired enzyme properties. This 'artificial' process involved multiple rounds of functional screening. It was costly, very time consuming and led only to gradual improvements.

With metagenomics, samples from nature generate a wider collection of sequences, and give insights into biological adaptations. The downside is having almost too many sequences to manage. The payoff is that with higher-quality data – and metadata – scientists can make more informed decisions when selecting new enzymes.

millions of protein sequences from environmental samples, where novel enzymes are waiting to be found.
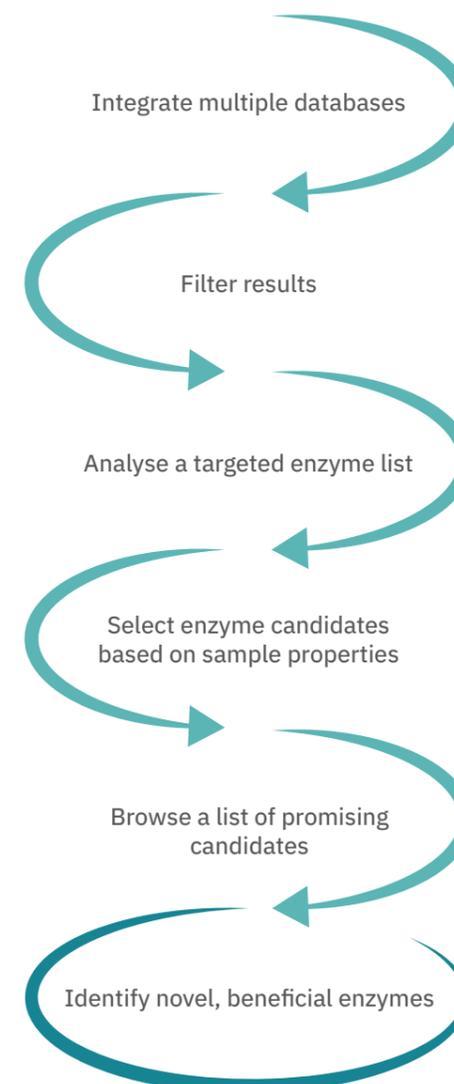
MetXtra™ and MGnify provide complimentary platforms for enzyme discovery from metagenomic libraries, with the former being specifically-designed for the identification and selection of those enzymes having a high probability of commercial success.

MetXtra™: How it works

First, Biocatalysts coupled their own metagenomic libraries with the open datasets from MGnify, creating a collection of over 340 million unique protein sequences that function in different environmental niches (e.g. in very cold, salty or acidic conditions). Without the right tools, it would have taken days to analyse this dataset and screen for enzymes that perform desirable functions in specified conditions.

Biocatalysts then developed MetXtra™, a proprietary interface, which allows users to query multiple databases and select enzyme candidates based on sample properties and sequence features (an important consideration for industrial production).

To refine a list of candidate enzymes, MetXtra™ filters results through Biocatalysts' proprietary algorithms that filter candidate enzymes for Design for Manufacture criteria (DFM) that maximises the chance of success for actually producing the active enzyme in the laboratory and, more importantly, being able to scale-up for production and manufacture. This produces a manageable, targeted list of enzymes (e.g. ones that break down oils in cold temperatures).

Integrate multiple databases

Filter results

Analyse a targeted enzyme list

Select enzyme candidates based on sample properties

Browse a list of promising candidates

Identify novel, beneficial enzymes

Successful outcomes

According to Dr Andrew Ellis, Technical Director at Biocatalysts, "MetXtra™ has already been used to successfully identify novel, beneficial speciality enzymes and is continuing to attract interest from many companies in different market sectors."

"It's been amazing to see the pace at which Biocatalysts has produced real-world solutions from their integrated metagenomics dataset," said Dr Rob Finn, Team Leader at EMBL-EBI. "We hope others will embrace the potential of 'nature's protein engineering', and use MGnify to innovate."

In the first year since the official launch of MetXtra, Biocatalysts has had 25 customer enquiries, 9 successful customer projects, and a shortlist for a Queens Award for Innovation.

Innovating on EMBL-EBI resources

EMBL-EBI is all about collaboration, from small research projects to sustainable international infrastructures. Collaborating with companies through EMBL's technology-transfer partner, EMBLEM, helps the institute's public data resources keep pace with the demands of commercial R&D.

The interaction between EMBL-EBI and Biocatalysts, an applied industrial biotechnology company, is a perfect example. Biocatalysts was able to develop new technology to filter metagenomics datasets automatically, and EMBL-EBI was able to better understand the company's data needs.

Any researcher who wants to discover new proteins for specific applications – in food technology, waste management, deep-sea mining or other bioindustries – can access these rich datasets through the MGnify website, ftp site and search tool.

Discover more

Browse MGnify on:
www.ebi.ac.uk/metagenomics

Discover other EMBL-EBI data services on:
www.ebi.ac.uk

Find out more about Biocatalysts Ltd:
www.biocatalysts.com

BIOCATALYSTS
exceeding enzyme expectations