(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 January 2025 (16.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/015288 A1

(51) International Patent Classification:
G01N 33/52 (2006.01) *G01N 21/64* (2006.01)
C12N 9/14 (2006.01)

(21) International Application Number:
PCT/US2024/037855

(22) International Filing Date:
12 July 2024 (12.07.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

63/526,278	12 July 2023 (12.07.2023)	US
63/641,318	01 May 2024 (01.05.2024)	US
63/643,053	06 May 2024 (06.05.2024)	US

(71) Applicant: **TRIAD NATIONAL SECURITY, LLC** [US/US]; c/o Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, New Mexico 87545 (US).

(72) Inventors: **NGUYEN, Hau Thi Bich**; c/o Los Alamos National Laboratory, P.O. Box 1663 MS A187, Los Alamos, New Mexico 87545 (US). **DALE, Taraka T.**; c/o Los Alamos National Laboratory, P.O. Box 1663 MS A187, Los Alamos, New Mexico 87545 (US). **GROSECLOSE, Thomas M.**; c/o Los Alamos National Laboratory, P.O. Box 1663 MS A187, Los Alamos, New Mexico 87545 (US).

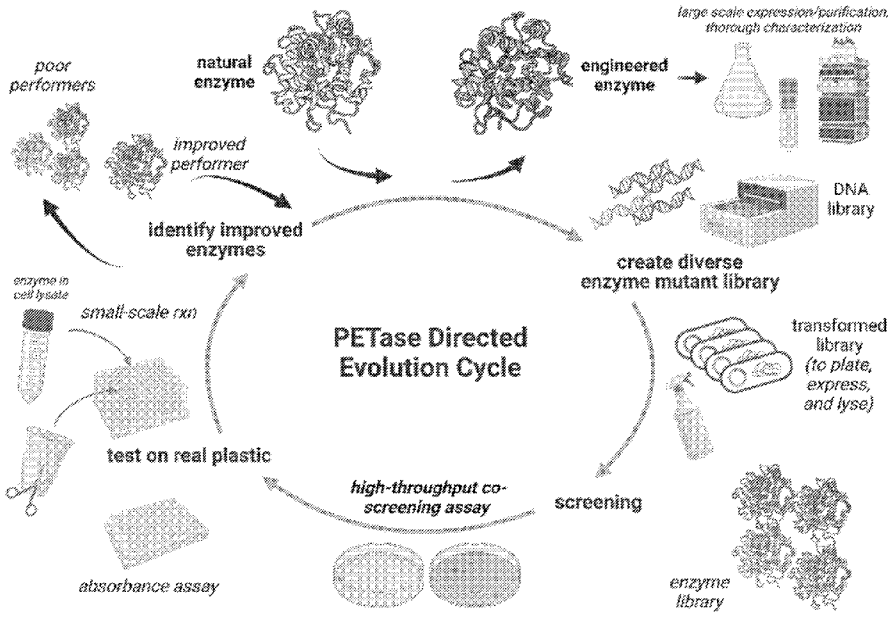
(74) Agent: **GRAF, Susan W.** et al.; Klarquist Sparkman, LLP, One World Trade Center, 121 SW Salmon St., Suite 1600, Portland, Oregon 97204 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,

WO 2025/015288 A1

(54) Title: METHODS TO IMPROVE ENZYMES THAT DEGRADE POLYETHYLENE TEREPHTHALATE

FIG. 1



(57) Abstract: Methods to identify, select, and/or produce enzymes with improved degradation activity to polyethylene terephthalate are provided. In some examples, the methods include expressing a fusion protein comprising a protein of interest linked to a reporter protein in a cell; contacting the fusion protein with a substrate for the protein of interest; detecting a signal from the reporter protein; and detecting degradation of the substrate for the protein of interest, wherein signal from the reporter protein indicates expression of the fusion protein and increased degradation of the substrate compared to a control indicates that the protein of interest has increased PET degradation activity.

[Continued on next page]

RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

METHODS TO IMPROVE ENZYMES THAT DEGRADE POLYETHYLENE TEREPHTHALATE

METHODS TO IMPROVE ENZYMES THAT DEGRADE POLYETHYLENE TEREPHTHALATE

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application No. 63/526,278, filed July 12, 2023; U.S. Provisional Application No. 63/641,318, filed May 1, 2024; and U.S. Provisional Application No. 63/643,053, filed May 6, 2024; each of which is incorporated by reference in its entirety.

FIELD

This disclosure relates to high-throughput methods for identifying enzymes with improved activity to degrade polyethylene terephthalate (PET).

ACKNOWLEDGMENT OF GOVERNMENT SUPPORT

This invention was made with government support under 89233218CNA000001 awarded by the National Nuclear Security Administration. The government has certain rights in the invention.

SEQUENCE LISTING INCORPORATION BY REFERENCE

The Sequence Listing is submitted as an XML file in the form of the file named "8472-110491-04_Sequence_Listing.xml" (25,889 bytes), which was created on July 11, 2024, which is incorporated by reference herein.

BACKGROUND

In recent years, enzymes that are able to break down human produced polymers (plastics) have gained increasing attention. These enzymes have been reported with activities that allow them to convert these otherwise recalcitrant polymers to their simplest chemical building blocks (precursors), which would facilitate complete, chemical recycling processes, "closing the loop" on the plastics economy. These enzymes have been typically sourced from natural diversity, in specific microorganisms, or from metagenomes. However, these natural enzymes lack the stability, expression, and activity required for large-scale recycling processes to be efficient and cost-effective. Thus, there is a need to generate enzymes with improved properties.

SUMMARY

The disclosed methods are high-throughput screening methods for the evolution of PET degrading enzymes. Most of enzyme evolution is currently guided by structure and computational analysis of structure, and homologous structures, which often is trial-and-error, slow, imprecise, and cannot adequately predict beneficial mutations. This is especially true for prediction of groupings of mutations, which can actually "act against" one another to lower activity/function unpredictably, but many times, multiple mutations are required for achieving the highest activity. The disclosed methods mimic natural evolution, where mutations are made, and resulting mutants are then screened in large numbers to find any rare, beneficial "hits."

In some aspects, the disclosed methods include expressing a fusion protein including a protein of interest linked to a reporter protein in a cell; contacting the fusion protein with a substrate for the protein of interest; detecting a signal from the reporter protein; and detecting degradation of the substrate for the protein of interest, wherein signal from the reporter protein indicates expression of the fusion protein and increased degradation of the substrate compared to a control indicates that the protein of interest has increased PET degradation activity.

In some aspects, the reporter protein is a fluorescent protein or portion thereof. In one example, the fluorescent protein is a green fluorescent protein (GFP) or a portion thereof. In a particular example, the reporter protein is a GFP11 split fluorescent protein tag. In examples, where the reporter protein is a GFP11 split fluorescent protein tag, detecting a signal from the reporter protein includes contacting the fusion protein with a GFP1-10 split fluorescence protein detector and detecting fluorescence signal.

In some aspects, expressing the fusion protein including the protein of interest linked to a reporter protein in a cell is performed on a solid medium. In other aspects, contacting the fusion protein with the substrate for the protein of interest is performed on a solid medium.

In certain aspects, detecting degradation of the substrate for the protein of interest is performed using a colorimetric assay. In some examples, the substrate for the protein of interest is bis(2-hydroxyethyl) terephthalate (BHET).

In additional aspects, the disclosed methods further include selecting a protein of interest identified as having increased PET degradation activity and determining activity of the selected protein of interest to degrade PET. In some examples, determining activity of the selected protein of interest includes contacting the selected protein of interest (or fusion protein) with PET and measuring presence or amount of one or more PET degradation products. In some examples, the one or more PET degradation products are measured by absorbance. In certain examples, the one or more PET degradation products include one or more of bis(2-hydroxyethyl) terephthalate

(BHET), mono-(2-hydroxyethyl) terephthalate (MHET), terephthalic acid (TPA) and ethylene glycol (EG).

In additional aspects, expressing the fusion protein includes expressing a library of fusion proteins including one or more amino acid substitutions in the protein of interest compared to a wild type or control protein of interest. In some examples, the control protein of interest is a modified version of the protein of interest compared to the wild type protein of interest.

In other aspects, the disclosed methods further include measuring thermostability of the protein of interest with increased PET degradation activity (such as the fusion protein including the protein of interest with increased PET degradation activity). In some examples, measuring thermostability of the protein of interest with increased PET degradation activity includes detecting signal from the reporter protein before and after heat treatment. In some examples, the heat treatment includes incubating the fusion protein at about 65-80°C, for example, prior to or simultaneously with contacting the fusion protein with the substrate.

The foregoing and other features of the disclosure will become more apparent from the following detailed description, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic diagram of an exemplary PETase directed evolution cycle, using the disclosed method. A natural enzyme (or modified version of a natural enzyme) is chosen as the starting point for engineering. Directed evolution involves cycles of: a diverse library of mutants being created and transformed, the library being plated, expressed, and lysed on screening plates, the high-throughput screening method being used to quickly and effectively evaluate mutants for expression, activity, and stability, and verification of mutant activity on real PET plastic. This cycle is used iteratively to continue to improve the enzyme, until one or more finalized engineered enzymes are chosen, which are characterized in more detail.

FIGS. 2A-2F shows an exemplary workflow for high-throughput screening assay. FIG. 2A: A bacterial library containing mutant PETase plasmids is plated on a membrane on an LB plate, grown overnight, then induced by moving the membrane onto an LB+IPTG plate. FIG. 2B: The membrane is then moved onto a BHET screening plate, and colonies are partially lysed. Enzymes diffuse into the plate. FIG. 2C: The membrane is removed and the BHET plate is incubated at the reaction temperature (optionally, first incubated at an elevated temperature for heat treatment). Clearing zones appear where enzymes have high activity on BHET. FIG. 2D: A solution of GFP1-10 is added to the plate to complement the GFP11 tags on enzymes and incubated at room

temperature. FIG. 2E: Fluorescence of clearing zones is measured. FIG. 2F: The membrane is oriented back onto the plate and colonies are picked based upon their activity and expression.

FIGS. 3A-3B illustrate complementation of GFP11 tag with GFP1-10. FIG. 3A: If a mutant enzyme is soluble and folded (potentially functional), its GFP11 tag is accessible to GFP1-10 when complemented. The association of GFP11 and GFP1-10 reconstitute full-length GFP, giving a fluorescent signal correlating with the concentration of enzyme of interest present. However, if a mutant is aggregated or misfolded, complementation cannot occur. FIG. 3B: A standard curve allows the quantification and normalization of an enzyme (or library of enzymes) in cell lysate based upon its fluorescence in solution.

FIGS. 4A-4B show screening plates for high-throughput assay. FIG. 4A: BHET plates are initially opaque, but after colonies containing the library are lysed on the plate, and the plate is incubated at the reaction temperature, clearing zones begin to appear, indicating activity on BHET. GFP1-10 can then be used on the plate to complement (quantify) enzyme:GFP11 proteins. When screening, colony activity/expression are evaluated. In this example, a high activity/low expression colony (arrow) and a high activity/high expression colony (arrowhead) would be chosen, while low activity/high expression and low activity/high expression colonies would not. FIG. 4B: Colonies containing plasmids with mCherry were plated alongside a PETase library and lysed, with red fluorescence imaged to demonstrate approximately equal lysis of library colonies across a plate.

FIG. 5 shows degradation of PET for engineered PHL7 mutants, wild-type PHL7, and LCC-ICCG. Reactions were performed at 70°C, pH8, in 1 M potassium phosphate buffer (0.1 M for LCC ICCG), with 2.9% w/v loading by mass PET amorphous film (8.5% crystallinity), 0.69 μM (left) or 0.345 μM (right) enzyme (0.7 mg enzyme/g PET for PHL7). Measurements are of aggregate aromatic products, measured by absorbance at 240 nm. Left, initial rate at higher enzyme concentration; right, activity over time at lower enzyme concentration.

FIG. 6 shows SDS-PAGE gels showing increase in expression of PHL7 mutants in *E. coli* cells. Left, wild-type PHL7 following concentration. Right, 5 PHL7 optima. Band corresponding to the molecular weight of PHL7 and mutants is denoted with a red arrow. Complementation of PHL7 mutants quantifies this difference as a 20-30-fold increase, depending on the mutant.

SEQUENCE LISTING

The nucleic acid and amino acid sequences listed herein and in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases and amino acids. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

SEQ ID NO: 1 is the nucleic acid of an exemplary wild-type PHL7 enzyme:

ATGGCGAATCCGTATGAGCGTGGCCCGGACCCGACCGAGAGCAGCATTGAAGCGGTTC
GTGGCCCGTTTGCGGTTGCGCAGACCACCGTTAGCCGTCTGCAGGCGGATGGCTTCGG
TGGTGGCACCATCTACTATCCGACCGACACCAGCCAAGGCACCTTCGGTGCGGTGGCG
ATTAGCCCGGGCTTTACCGCGGGTCAGGAGAGCATTGCGTGGCTGGGTCCGCGTATTG
CGAGCCAAGGTTTTGTGGTTATCACCATTGATACCATCACCCGTCTGGATCAGCCGGAC
AGCCGTGGCCGTCAGCTGCAAGCGGCGCTGGATCACCTGCGTACCAACAGCGTGGTTC
GTAACCGTATTGACCCGAACCGTATGGCGGTTATGGGTCACAGCATGGGCGGTGGtGG
TGCGCTGAGCGCTGCGGCGAACAACACCAGCCTGGAAGCGGCGATCCCGCTGCAGGG
TTGGCACACCCGTAAGAACTGGAGCAGCGTGCGTACCCCGACCCTGGTGGTTGGTGCG
CAACTGGATACCATTGCGCCGGTTAGCAGCCACAGCGAGGCGTTCTACAACAGCCTGC
CGAGCGATCTGGACAAAGCGTATATGGAACTGCGTGGTGCGAGCCACCTGGTGAGCA
ACACCCCGGACACCACCACcGCGAAGTACAGCATCGCGTGGCTGAAACGTTTTGTTGA
CGATGACCTGCGTTATGAGCAATTTCTGTGCCCGGCGCCGGACGACTTTGCGATTAGC
GAATACCGTAGCACCTGCCCGTTCCTCGAG

SEQ ID NO: 2 is the amino acid sequence of an exemplary wild-type PHL7 enzyme:

MANPYERGPDPTESSIEAVRGPFAVAQTTVSRLQADGFGGGTIYYPTDTSQGTFGAVAISP
GFTAGQESIAWLGPRIASQGFVVITIDTITRLDQPDSRGRQLQAALDHLRTNSVVRNRIDPN
RMAVMGHSMGGGGALSAAANNTSLEAAIPLQGWHTRKNWSSVRTPTLVVGAQLDTIAP
VSSHSEAFYNSLPSDLDKAYMELRGASHLVSNTPDTTTAKYSIAWLKRFVDDDLRYEQFL
CPAPDDFAISEYRSTCPFLE

SEQ ID NO: 3 is the nucleic acid sequence of PHL7 variant #1:

ATGGCGAATCCGTATGAGCGTGGCCCGGACCCGACCGAGAGCAGCATTGAAGCGGTTC
GTGGCCCGTTTGCGGTTGCGCAGACCACCGTTAGCCGTCTGCAGGTGGATGGCTTCGG
TGGTGGCACCATCTACTATCCGACCGACACCAGCCAAGGCACCTTCGGTGCGGTGGCG
ATTAGCCCGGGCTTTACCGCGGGTCAGGAGAGCATTGCGTGGCTGGGTCCGCGTATTG
CGAGCCAAGGTTTTGTGGTTATCACCATTGATACCATCACCCGTCTGGATTATCCGGAC
AGCCGTGGCCGTCAGCTGCAAGCGGCGCTGGATCACCTGCGTATCAACAGCGTGGTTC
GTAACCGTATTGACCCGAACCGTATGGCGGTTATGGGTCACAGCATGGGCGGTGGTGG
TGCGCTGAGCGCTGCGGCGAACAACACCAGCCTGGAAGCGGCGATCCCGCTGCAGGG
TTGGCACACCCGTAAGAACTGGAGCAGCGTGCGTACCCCGACCCTGGTGGTTGGTGCG
GAACTGGATACCATTGCGCCGGTTAGCAGCAACAGCGAGGCGTTCTACAACAGCCTGC
CGAGCGATCTGGACAAAGCGTATATGGAACTGCGTGGTGCGAGCCACCTGGTGAGCA
ACACCCCGGACACCACCACCGCGAAGTACAGCATCGCGTGGCTGAAACGTTTTGTTGA

CGATGACCTGCGTTATGAGCAATTTCTGTGCCCGGCGCCGGACGACTTTGCGATTAGC
GAATACCGTAGCACCTGCCCGTTCCTCGAG

SEQ ID NO: 4 is the amino acid sequence of PHL7 variant #1:

MANPYERGPDPTESSIEAVRGPFAVAQTTVSRLQVDGFGGGTIYYPTDTSQGTFGAVAISP
GFTAGQESIAWLGPRIASQGFVVITIDTITRLDYPDSRGRQLQAALDHLRINSVVRNRIDPNR
MAVMGHSMGGGGALSAAANNTSLEAAIPLQGWHTRKNWSSVRTPTLVVGAELDTIAPVS
SNSEAFYNSLPSDLDKAYMELRGASHLVSNTPDTTTAKYSIAWLKRFVDDDLRYEQFLCP
APDDFAISEYRSTCPFLE

SEQ ID NO: 5 is the nucleic acid sequence of PHL7 variant #2:

ATGGCGAATCCGTATGAGCGTGGCCCGGACCCGACCGAGAGCAGCATTGAAGCGGTTC
GTGGCCCGTTTGCGGTTGCGCAGACCACCGTTAGCCGTCTGCAGGCGGATGGCTTCGG
TGGTGGCACCATCTACTATCCGACCGACACCAGCCAAGGCACCTTCGGTGCGGTGGCG
ATTAGCCCGGGCTTTACCGCGGGTCAGGAGAGCATTGCGTGGCTGGGTCCGCGTATTG
CGAGCCATGGTTTTGTGGTTATCACCATTGATACCATCACCCGTCTGGATTATCCGGAC
AGCCGTGGCCGTCAGCTGCAAGCGGCGCTGGATCACCTGCGTACCAACAGCGTGGTTC
GTAACCGTATTGACCCGAACCGTATGGCGGTTATGGGTCACAGCATGGGCGGTGGTGG
TGCGCTGAGCGCTGCGGCGAATAACACCAGCCTGGAAGCGGCGATCCCGCTGCAGGGT
TGGCACACCCGTAAGAACTGGAGCAGCGTGCGTACCCCGACTCTGGTGGTTGGTGCGG
AACTGGATACCATTGCGCCGGTTAGCAGCAACAGCGAGGCGTTCTACAACAGCCTGCC
GAGCGATCTGGACAAAGCGTATATGGAACTGAAAGGTGCGAGCCACCTGGTGAGCAA
CACCCCGGACACCACCACCGCGAAGTACAGCATCGCGTGGCTGAAACGTTTTGTTGAC
GATGACCTGCGTTATGAGCAATTTCTGTGCCCGGCGCCGGACGACTTTGCGATTAGCG
AATACCGTAGCACCTGCCCGTTCCTCGAG

SEQ ID NO: 6 is the amino acid sequence of PHL7 variant #2:

MANPYERGPDPTESSIEAVRGPFAVAQTTVSRLQADGFGGGTIYYPTDTSQGTFGA
VAISPGFTAGQESIAWLGPRIASHGFVVITIDTITRLDYPDSRGRQLQAALDHLRTNSVVRN
RIDPNRMAVMGHSMGGGGALSAAANNTSLEAAIPLQGWHTRKNWSSVRTPTLVVGAELD
TIAPVSSNSEAFYNSLPSDLDKAYMELKGASHLVSNTPDTTTAKYSIAWLKRFVDDDLRYE
QFLCPAPDDFAISEYRSTCPFLE

SEQ ID NO: 7 is the nucleic acid sequence of PHL7 variant #3:

ATGGCGAATCCGTATGAGCGTGGCCCGGACCCGACCGAGAGCAGCATTGAAGCGGTTC
GTGGCCCGTTTGCGGTTGCGCAGACCACCGTTAGCCGTCTGCAGGCGGATGGCTTCGG
TGGTGGCACCATCTACTATCCGACCGACACCAGCCAAGGCACCTTCGGTGCGGTGGCG
ATTAGCCCGGGCTTTACCGCGGGTCAGGAGAGCATTGCGTGGCTGGGTCCGCGTATTG

CGAGCCAAGGTTTTGTGGTTATCACCATTGATACCATCACCCGTCTGGATCAGCCGGAC
AGCCGTGGCCGTCAGCTGCAAGCGGCGCTGGATCACCTGCGTGCCAACAGCGTGGTTC
GTAACCGTATTGACCCGAACCGTATGGCGGTTATGGGTCACAGCATGGGCGGTGGCGG
TGCGCTGAGCGCTGCGGCGAACAACACCAGCCTGGAAGCGGCGATCCCGCTGCAGGG
TTGGCACACCCGTAAGAACTGGAGCAGCGTGCGTACCCCGACCCTGGTGGTTGGTGCG
GAACTGGATACCATTGCGCCGGTTAGCAGCAACAGCGAGGCGTTCTACAACAGCCTGC
CGAGCGATCTGGACAAAGCGTATATGGAACTGCGTGGTGCGAGCCACCTGGTGAGCA
ACACCCCGGACACCACCACCGCGAAGTACAGCATCGCGTGGCTGAAACGTTTTGTTGA
CGATGACCTGCGTTATGAGCAATTTCTGTGCCCGGCGCCGGACGACTTTGCGATTAGC
GAATACCGTAGCACCTGCCCGTTCCTCGAG

SEQ ID NO: 8 is the amino acid sequence of PHL7 variant #3:

MANPYERGPDPTESSIEAVRGPFAVAQTTVSRLQADGFGGGTIYYPTDTSQGTFGAVAISP
GFTAGQESIAWLGPRIASQGFVVITIDTITRLDQPDSRGRQLQAALDHLRANSVVRNRIDPN
RMAVMGHSMGGGGALSAAANNTSLEAAIPLQGWHTRKNWSSVRTPTLVVGAELDTIAPV
SSNSEAFYNSLPSDLDKAYMELRGASHLVSNTPDTTTAKYSIAWLKRFVDDDLRYEQFLCP
APDDFAISEYRSTCPFLE

SEQ ID NO: 9 is the nucleic acid sequence of PHL7 variant #4:

ATGGCGAATCCGTATGAGCGTGGCCCGGACCCGACCGAGAGCAGCATTGAAGCGGTTC
GTGGCCCGTTTGCGGTTGCGCAGACCACCGTTAGCCCTCTGCAGGCGGATGGCTTCGG
TGGTGGCACCATCTACTATCCGACCGACACCAGCCAAGGCACCTTCGGTGCGGTGGCG
ATTAGCCCGGGCTTTAGCGCGGGTCAGGAGAGCATTGCGTGGCTGGGTCCGCGTATTG
CGAGCCAAGGTTTTGTGGTTATCACCATTGATACCATCACCCGTCTGGATCAGCCGGAC
AGCCGTGGCCGTCAGCTGCAAGCGGCGCTGGATCACCTGCATACCAACAGCGTGGTTC
GCAACCGGATTGACCCGAACCGTATGGCGGTTATGGGTCACAGCATGGGCGGTGGCGG
TGCGCTGAGCGCTGCGGCGAACAACACCAGCCTGGAAGCGGCGATCCCGCTGCAGGG
TTGGCACACCCGTAAGAACTGGAGCAGCGTGCGTACCCCGACCCTGGTGGTTGGTGCG
GAACTGGATACCATTGCGCCGGTTAGCAGCAACAGCGAGGCGTTCTACAACAGCCTGC
CGAGCGATCTGGACAAAGCGTATATGGAACTGCGTGGTGCGAGCCACCTGGTGAGCA
ACACCCCGGACACCACCACCGCGAAGTACAGCATCGCGTGGCTGAAACGTTTTGTTGA
CGATGACCTGCGTTATGAGCAATTTCTGTGCCCGGCGCCGGACGACTTTGCGATTAGC
GAATACCGTAGCACCTGCCCGTTCCTCGAG

SEQ ID NO: 10 is the amino acid sequence of PHL7 variant #4:

MANPYERGPDPTESSIEAVRGPFAVAQTTVSPLQADGFGGGTIYYPTDTSQGTFGAVAISPG
FSAGQESIAWLGPRIASQGFVVITIDTITRLDQPDSRGRQLQAALDHLHTNSVVRNRIDPNR

MAVMGHSMGGGGALSAAANNTSLEAAIPLQGWHTRKNWSSVRTPTLVVGAELDTIAPVS
SNSEAFYNSLPSDLDKAYMELRGASHLVSNTPDTTTAKYSIAWLKRFVDDDLRYEQFLCP
APDDFAISEYRSTCPFLE

SEQ ID NO: 11 is the nucleic acid sequence of PHL7 variant #5:

ATGCCGAATCCGTATGAGCGTGGCCCGGACCCGACCGAGAGCAGCATTGAAGCGGTTC
GTGGCCCGTTTGCGGTTGCGCAGACCACCGTTAGCCGTCTGCAGGCGGATGGCTTCGG
TGGTGGCACCATCTACTATCCGACCGACACCAGCCAAGGCACCTTCGGTGCGGTGGCG
ATTAGCCCGGGCTTTACCGCGGGTCAGGAGAGCATTGCGTGGCTGGGTCCGCGTATTG
CGAGCCAAGGTTTTGTGGTTATCACCATTGATACCATCACCCGTCTGGATTATCCGGAC
AGCCGTGGCCGTCAGCTGCAAGCGGCGCTGGATCACCTGCGTACCAACAGCGTGGTTC
GTAACCGTATTGACCCGAACCGTATGGCGGTTATGGGTCACAGCATGGGCGGTGGCGG
TGCGCTGAGCGCTGCGGCGAACAACACCAGCCTGGAAGCGGCGATCCCGCTGCAGGG
TTGGCACACCCGTAAGAACTGGAGCAGCGTGCGTACCCCGACCCTGGTGGTTGGTGCG
GAACTGGATACCATTGCGCCGGTTAGCAGCAACAGCGAGGCGTTCTACAACAGCCTGC
CGAGCGATCTGGACAAAGCGTATATGGAACTGAAAGGTGCGAGCCACCTGGTGAGCA
ACACCCCGGACACCACCACCGCGAAGTACAGCATCGCGTGGCTGAAACGTTTTGTTGA
CGATGACCTGCGTTATGAGCAATTTCTGTGCCCGGCGCCGGACGACTTTGCGATTAGC
GAATACCGTAGCACCTGCCCGTTCCTCGAG

SEQ ID NO: 12 is the amino acid sequence of PHL7 variant #5:

MPNPYERGPDPTESSIEAVRGPFAVAQTTVSRLQADGFGGGTIYYPTDTSQGTFGAVAISPG
FTAGQESIAWLGPRIASQGFVVITIDTITRLDYPDSRGRQLQAALDHLRTNSVVRNRIDPNR
MAVMGHSMGGGGALSAAANNTSLEAAIPLQGWHTRKNWSSVRTPTLVVGAELDTIAPVS
SNSEAFYNSLPSDLDKAYMELKGASHLVSNTPDTTTAKYSIAWLKRFVDDDLRYEQFLCP
APDDFAISEYRSTCPFLE

SEQ ID NO: 13 is the nucleic acid sequence of an exemplary LCC-ICCG enzyme:

ATGTCTAACCCGTACCAGCGCGGACCGAACCCGACCCGTTCTGCGTTAACCGCTGATG
GTCCGTTTTCCGTGGCTACCTACACCGTTTCTCGTCTGTCCGTTTCCGGTTTTGGTGGTG
GTGTTATCTACTATCCGACTGGTACCTCTCTGACCTTCGGCGGTATCGCGATGTCCCCG
GGTTACACCGCTGATGCTTCCTCTCTGGCGTGGCTGGGTCGTCGCCTGGCGAGCCACG
GTTTTGTTGTTCTGGTTATCAACACGAACTCTCGTTTCGACGGCCCGGACTCCCGTGCC
TCGCAACTGTCTGCTGCGCTGAACTACCTGCGTACGTCGTCACCTTCAGCGGTCCGTGC
ACGCCTGGATGCCAATCGTCTGGCTGTGGCGGGTCACAGCATGGGCGGTGGCGGTACC
CTGCGTATTGCTGAACAGAACCCGTCCCTGAAAGCTGCAGTGCCACTGACTCCGTGGC
ATACCGACAAAACGTTCAACACCAGTGTTCCGGTACTGATCGTAGGCGCAGAAGCGG

ACACCGTAGCACCGGTTTCCCAGCACGCAATCCCGTTCTACCAGAACCTGCCGAGCAC
CACTCCAAAAGTATACGTTGAACTGTGCAACGCCTCGCACATTGCTCCGAACTCGAAC
AACGCTGCGATTAGCGTGTACACCATCTCCTGGATGAAACTGTGGGTTGATAACGATA
CCCGTTATCGCCAATTCCTGTGTAACGTGAACGATCCGGCTCTCTGCGATTTTCGTACC
AACAACCGTCATTGCCAA

SEQ ID NO: 14 is the amino acid sequence of an exemplary LCC-ICCG enzyme:

MSNPYQRGPNPTRSALTADGPFSVATYTVSRLSVSGFGGGVIYYPTGTSLTFGGIAMSPGY
TADASSLAWLGRRLASHGFVVLVINTNSRFDGPDSRASQLSAALNYLRTSSPSAVRARLD
ANRLAVAGHSMGGGGTLRIAEQNPSLKAAVPLTPWHTDKTFNTSVPVLIVGAEADTVAP
VSQHAIPFYQNLPSTTPKVYVELCNASHIAPNSNNAAISVYTISWMKLWVDNDTRYRQFL
CNVNDPALCDFRTNNRHCQ

SEQ ID NO: 15 is the nucleic acid sequence of LCC variant #2B8:

ATGTCTAACCTGTACCAGCGCGGACCGAACCCGACCCGTTCTGCGTTAACCGCTGATG
GTCCGTTTTCCGTGGCTACCTGCACCGTTTCTCGTCTGTCCGTTTCCGGTTTTGGTGGTG
GTGTTATCTACTATCCGACTGGTACCTCTCTGACCTTCGGCGGTATCGCGATATCCCCG
GGTTACACCGCTGATGCTTCCTCCCTGGCGTGGCTGGGTCGTCGCCTGGCGAGCCACG
GTTTTGTTGTTCCGGTTATCAACACGAACTCTCGTTTCGACGGCCCGGACTCCCGTGCC
TCGCAACTGTCTGCTGCGCTGAACTACCTGCGTACGTCGTCACCTTCAGTGGTCCGTGC
ACGCCTGGATGCCAATCGTCTGGCTGTGGCGGGTCACAGCATGGGCGGTGGCGGTACC
CTGCGTATTGCAGAACAGAACCCGTCCCTGAAAGCTGCAGTGCCACTGACTCCGTGGC
ATACCGACAAAACGTTCAACACCAGTGTTCCGGTACTGATCGTAGGCGCAGAAGCGGA
CACCGTAGCACCGGTTTCCCAGTACGCAATCCCGTTCTACCATAACCTGCCGAGCACC
ACTCCAAAAGTATACGTTGAACTGTGCAACGCCTCGCACATTGCTCCGAACTTGAACA
ACGCTGCGATTAGCGTGTACATCATCTCCTGGATGAAACTGTGGGTTGATAACGATAC
CCGTTATCGCCAATTCCTGTGTAACGTGAACGATCCGGCTCTCTGCGATTTTCGTACCA
ACAACCGTCATTGCCAA

SEQ ID NO: 16 is the amino acid sequence of LCC variant #2B8:

MSNLYQRGPNPTRSALTADGPFSVATCTVSRLSVSGFGGGVIYYPTGTSLTFGGIAISPGYT
ADASSLAWLGRRLASHGFVVPVINTNSRFDGPDSRASQLSAALNYLRTSSPSVVRARLDAN
RLAVAGHSMGGGGTLRIAEQNPSLKAAVPLTPWHTDKTFNTSVPVLIVGAEADTVAPVSQ
YAIPFYHNLPSTTPKVYVELCNASHIAPNLNNAAISVYIISWMKLWVDNDTRYRQFLCNVN
DPALCDFRTNNRHCQ

SEQ ID NO: 17 is the nucleic acid sequence of LCC variant #EC9:

ATGTCTAACCTGTACCAGCGCGGACCGAACCCGACCCGTTCTGCGTTAACCGCTGATG
GTCCGTTTTCCGTGGCTACCTGCACCGTTTCTCGTCTGTCCGTTTCCGGTTTTGGTGGTG
GTGTTATCTACTATCCGACTGGTACCTCTCTGACCTTCGGCGGTATCGCGATATCCCCG
GGTTACACCGCTGATGCTTCCTCCCTGGCGTGGCTGGGTCGTCGCCTGGCGAGCCACG
GTTTTGTTGTTCCGGTTATCAACACGAACTCTCGTTTCGACGGCCCGGACTCCCGTGCC
TCGCAACTGTCTGCTGCGCTGAACTACCTGCGTACGTCGTCACCTTCAGTGGTCCGTGC
ACGCCTGGATGCCAATCGTCTGGCTGTGGCGGGTCACAGCATGGGCGGTGGCGGTACC
CTGCGTATTGCAGAACAGAACCCGTCCCTGAAAGCTGCAGTGCCACTGACTCCGTGGC
ATACCGACAAAACGTTCAACACCAGTGTTCCGGTACTGATCGTAGGCGCAGAAGCGGA
CACCGTAGCACCGGTTTCCCAGCACGCAATCCCGTTCTACCAGAACCTGCCGAGCACC
ACTCCAAAAGTATACGTTGAACTGTGCAACGCCTCGCACATTGCTCCGAACTTGAACA
ACGCTGCGATTAGCGTGTACACCATCTCCTGGATGAAACTGTGGGTTGATAACGATAC
CCGTTATCGCCAATTCCTGTGTAACGTGAACGATCCGGCTCTCTGCGATTTTCGTACCA
ACAACCGTCATTGCCAA

SEQ ID NO: 18 is the amino acid sequence of LCC variant #EC9:
MSNLYQRGPNPTRSALTADGPFSVATCTVSRLSVSGFGGGVIYYPTGTSLTFGGIAISPGYT
ADASSLAWLGRRLASHGFVVPVINTNSRFDGPDSRASQLSAALNYLRTSSPSVVRARLDAN
RLAVAGHSMGGGGTLRIAEQNPSLKAAVPLTPWHTDKTFNTSVPVLIVGAEADTVAPVSQ
HAIPFYQNLPSTTPKVYVELCNASHIAPNLNNAAISVYTISWMKLWVDNDTRYRQFLCNV
NDPALCDFRTNNRHCQ

DETAILED DESCRIPTION

Disclosed herein is a high-throughput screening platform for the optimization and/or engineering of enzymes that can degrade polyethylene terephthalate (PET), commonly used in food packaging and bottles and textiles. This method allows for the screening of large libraries of enzyme variants (10^5-10^6). The powerful technique of directed evolution has been traditionally limited for these enzymes by the sizes of libraries that are able to be effectively screened. The disclosed method, however, is able to screen libraries of unprecedented sizes quickly and effectively using simple plate-based assays. The methods allow simultaneously evaluating protein activity and expression using a model substrate based plate assay and a split-green fluorescent protein assay. In addition, the method can be optionally used to evaluate thermostability through the heat treatment of enzyme libraries and subsequent evaluation of their protein concentration and activity. This method can facilitate significantly accelerated discovery of the enzymes that can degrade PET and truly allow for large-scale laboratory evolution (often necessary to discover rare,

beneficial mutations), as current efforts for engineering these enzymes must be done by computational rational design and trial-and-error.

In some aspects, the disclosed methods include expressing a fusion protein including a protein of interest linked to a reporter protein in a cell; contacting the fusion protein with a substrate for the protein of interest; detecting a signal from the reporter protein; and detecting degradation of the substrate for the protein of interest, wherein signal from the reporter protein indicates expression of the fusion protein and increased degradation of the substrate compared to a control indicates that the protein of interest has increased PET degradation activity. The steps of the methods can be performed in any order, for example, detecting a signal from the reporter protein my be carried out prior to contacting the fusion protein with a substrate for the protein of interest and/or detecting degradation of the substrate for the protein of interest. Alternatively, contacting the fusion protein with a substrate for the protein of interest and/or detecting degradation of the substrate for the protein of interest can be carried out prior to detecting a signal from the reporter protein. In additional examples, detecting signal from the reporter protein and contacting the fusion protein with a substrate for the protein of interest and detecting the reporter protein can be carried out simultaneously or substantially simultaneously.

In some aspects, the methos include expressing a fusion protein including a protein of interest linked to a reporter protein in a cell (such as a bacterial cell, for example, an *E. coli* cell). In other aspects, the fusion protein is an isolated or purified fusion protein. In particular examples, the fusion protein includes the protein of interest or a variant thereof. In some examples, the reporter protein is linked to the C-terminus of the protein of interest. In other examples, the reporter protein is linked to the N-terminus of the protein of interest.

In some examples, the reporter is a fluorescent protein or a portion thereof. Exemplary fluorescent proteins include green fluorescent protein (GFP), a superfolder GFP (sfGFP), eGFP, red fluorescent protein (RFP), superfolder RFP (sfRFP), mCherry, or sfCherry, mStrawberry, mOrange, or dTomato). In other examples, the reporter protein is a split or fragmented protein that can be used in a protein fragment complementation assay (PCA). In further examples, the reporter protein is an enzyme, for example, β-lactamase, horseradish peroxidase, β-galactosidase, luciferase, or dihydrofolate reductase.

In some examples, the fusion protein includes the protein of interest (or variant thereof) linked to a GFP11 fragment (*e.g.*, a split fluorescence protein tag). In a particular example, the GFP protein or portion thereof is a GFP11 tag. In some examples, GFP1-10 (*e.g.*, a split fluorescence protein detector) is added to the enzyme (*e.g.*, on a solid medium, such as a petri dish plate, or in solution). GFP1-10 complements GFP11 to reassemble functional GFP, and gives a

fluorescence signal based on the concentration of the enzyme (protein of interest) present. Therefore, in examples of the disclosed method where the reporter protein is GFP11, the method further includes contacting the fusion protein with GFP1-10 split fluorescence protein detector and detecting fluorescence signal. Exemplary, non-limiting split GFP systems are described in U.S. Pat. No. 9,081,014 and U.S. Pat. Publ. No. 2015/0099271. Detecting signal from the reporter protein (such as a fluorescent signal) indicates expression of the protein of interest. A schematic diagram showing an exemplary method of detecting signal from the reporter protein is shown in FIGS. 2D-2F.

In some examples, detecting the signal from the reporter protein is qualitative, quantitative, or semi-quantitative In some examples, the disclosed methods allows both measuring concentration of the protein of interest in screening, and to quickly quantify and normalize enzymes for controlled assays (for example, without the need for traditional, large-scale protein purification and quantification, which cannot be done in high-throughput).

In particular aspects, the methods include expressing one or more variants of the protein of interest (such as a library of variants) linked to the reporter protein. The variants may be produced by rational design directed evolution, or both. In some examples, the library may be produced using site-directed mutagenesis and/or site-saturation mutagenesis. In some aspects, expressing the fusion protein includes expressing a library of fusion proteins including one or more amino acid substitutions in the protein of interest compared to a wild type or control protein of interest. In some examples, the control protein of interest is a modified version of the protein of interest compared to the wild type protein of interest.

In additional aspects of the disclosed methods, the fusion protein is contacted with a substrate for the protein of interest and degradation of the substrate is detected. In some examples, the protein of interest is a PETase and the fusion protein is contacted with PET or a small molecule version of PET that is a substrate for the PETase. In one example, the substrate is PET. In another example, the substrate is Bis(2-hydroxyethyl) terephthalate (BHET). In a further example, the substrate is impranil. In some aspects, the methods include contacting the fusion protein with the substrate on a solid medium. In some examples, a solid medium (such as an agar or agarose plate) is implanted with the substrate. The fusion protein is contacted with the substrate, and activity of the protein of interest is detected. In some examples, the fusion protein is expressed in cells (such as bacterial cells, for example, *E. coli* cells) on the plate or on a semi-permeable membrane on the plate. The fusion protein is released from the cells (for example, by lysis or partial lysis of the cells) thereby contacting the fusion protein with the substrate. Detecting degradation of the substrate may be by a colorimetric assay. For example, if the substrate is BHET, degradation of

BHET by the protein of interest portion of the fusion protein is detected by the presence of clear areas on the plate, as the opaque BHET plate becomes more transparent when BHET is degraded to more water-soluble precursors. An exemplary method is illustrated schematically in FIGS. 2A-2C. The colonies the exhibit degradation of the substrate can be traced to the original colony on the membrane and identified for further screening and/or sequencing.

In some aspects, the methods further include measuring thermostability of the protein of interest. In some examples, prior to contacting the fusion protein with the substrate, a heat treatment is performed (for example, on the fusion protein present on the solid medium). In other examples, the heat treatment is performed simultaneously with contacting the fusion protein with the substrate. In some examples, only thermostable enzymes will "survive" to react with the substrate. In other examples, signal from the reporter protein is detected before and after the heat treatment. In some examples, the heat treatment includes incubating the fusion protein or protein of interest at a temperature of about 65-80°C (such as about 65-70°C, about 68-75°C, or about 72-80°C, for example, about 65°C, about 66°C, about 67°C, about 68°C, about 69°C, about 70°C, about 71°C, about 72°C, about 73°C, about 74°C, about 75°C, about 76°C, about 77°C, about 78°C, about 79°C, or about 80°C).

Once proteins of interest or variants thereof with increased PET degradation activity are selected, the selected proteins of interest or variants thereof are screened for activity to degrade PET on a larger scale. In some examples, the selected protein of interest or fusion protein including the selected protein of interest is contacting with PET or a PET substrate and presence or amount of one or more degradation products is measured. In some examples, the selected protein of interest or fusion protein including the protein of interest is contacted with PET powder or film, or PET of various substrate crystallinities. Product concentration, such as amount of bis(2-hydroxyethyl) terephthalate (BHET), mono-(2-hydroxyethyl) terephthalate (MHET), terephthalic acid (TPA), ethylene glycol (EG), or a combination of two or more thereof is measured. In some examples, product concentration is measured by absorbance (such as UV absorbance) or using HPLC assays.

In some aspects, one or more additional rounds of screening using the disclosed methods are carried out on selected protein variants identified in an initial (or subsequent) screen. Variants with increased PET degradation activity, thermostability, and/or expression may be combined in subsequent rounds of screening. After selection of top or "final" variants, the protein of interest may be expressed and purified in large scale for in-depth characterization, using PET substrates at various, industrially-relevant conditions/testing.

In particular aspects, the protein of interest is a protein capable of degrading PET, such as a PET hydrolase. In some example, a PET hydrolase includes cutinases (EC 3.1.1.74), lipases (EC 3.1.1.3), carboxylesterases (EC 3.1.1.1), PETases (EC 3.1.1.101), MHETases (EC 3.1.1.102), and esterases, for example, from bacterial sources. Exemplary PET hydrolases include leaf-branch compost cutinase (LCC), polyester hydrolase Leipzig 7 (PHL7), *Ideonella sakaiensis* PETase (*Is*PETase), TfCut2, Cut190, FsC, and HiC. Additional PET hydrolases can be identified by one of skill in the art (see, *e.g.*, Anuar *et al.*, *In. J. Mol. Sci.* 23:12644, 2022; Erickson *et al.*, *Nat. Commun.* 13:7850, 2022).

In some aspects, the disclosed methods can be used to identify, select, and/or characterize one or more proteins of interest (such as variant of a protein of interest) with one or more improved characteristics compared to a control protein of interest (such as the wild type protein or a previously identified variant of the protein). In some examples, the methods are utilized to identify a protein of interest or variant that has improved or increased PET degradation activity compared to a control (such as an increase of in PET degradation activity of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 2-fold, at least 2.5-fold, at least 3-fold, at least 4-fold, at least 5-fold, or more compared to a control. In other examples, the methods are utilized to identify a protein of interest or variant that has increased thermostability compared to a control, such as the wild type protein or a previously identified variant of the protein. In further examples, the methods are utilized to identify a protein of interest or variant thereof with increased expression compared to a control (such as the wild type protein or a previously identified variant of the protein), for example, an increased expression of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 2-fold, at least 2.5-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold or more compared to a control.

EXAMPLES

The following examples are provided to illustrate particular features of certain aspects of the disclosure, but the scope of the claims should not be limited to those features exemplified.

Example 1

High throughput screening assay for identifying PETases with improved activity

A new platform for engineering PET hydrolases, which is capable of simultaneously screening large, random mutagenesis enzyme libraries for improved (i) protein expression and (ii) activity has been developed. The method is summarized briefly in FIG. 1. To evaluate expression,

enzymes in the libraries (such as LCC or PHL7) were expressed with C-terminal GFP11 tags, which allows their quantification in crude cell lysates, in solutions or on plates, via a fluorescent readout when complemented with GFP1-10 (Cabantous *et al.*, *Nature Biotechnology* 23:102-107, 2005). Activity was evaluated step-wise, on model, then on PET substrates (FIG. 1). A colorimetric, agar plate screen using bis(2-hydroxyethyl) terephthalate (BHET) as a model substrate was used to screen bacterial colony libraries (with $>10^5$ constituents) simultaneously, monitoring cell lysates causing the appearance of clearing (transparent) zones. Coupling the fluorescence and colorimetric assays, it was possible to quickly and precisely select improved enzymes based on both activity and protein expression. Selected enzymes were then expressed in small scale (2-25 mL) and screened in medium-throughput assays on PET substrates, in solution, with aggregate product concentration measured in microwell plates (~10^2), with plate-based absorbance protocols. Any superior enzyme compared to the parent enzymes from the previous round were sequenced and selected as parents for the next round of evolution.

Materials and data analysis: Unless noted, materials were obtained from the following sources. Oligonucleotides were purchased from Integrated DNA Technologies. Genes were synthesized by Twist Biosciences. Sanger sequencing was performed by Genewiz. Enzymes were purchased from New England Biolabs. Amorphous PET films (product ES301445; 8.5% crystallinity) and high crystallinity PET powder (product ES306031; 39.7% crystallinity) were purchased from Goodfellow Cambridge Ltd. A micronized amorphous powder was produced from the PET film by cryo-milling, first in a SM300 cutting mill (Retsch), then in a ZM200 centrifugal mill (Retsch), as described previously (Erickson *et al., Nature Commun.* 13:7850, 2022), but using a ring sieve with a larger pore size (0.5 mm) in the second step. The powder was thoroughly dried at 45°C for over 24 hours before use as a substrate. Chemicals were purchased from Fisher Scientific or Millipore Sigma. Kits were purchased from Qiagen. Data analysis and curation was performed in Microsoft Excel, GraphPad Prism, and Agilent OpenLab CDS. Sequencing and gene design were performed using ApE (M. Wayne Davis) and SnapGene (Dotmatics).

Cloning, mutagenesis, and library creation: A codon-optimized, synthesized gene encoding PHL7 or LCC was cloned into the pET21b(+)-GFP11 screening vector, between the NdeI and BamHI sites. Selected engineered mutants were amplified from the pET21b(+)-GFP11 vector with NdeI and XhoI sites and subcloned into pET21b(+), as necessary, for expression and purification using the His_6 system. Plasmids were transformed into *E. coli* BL21 (DE3) Gold cells (B F^- *ompT hsdS*(rB^- mB^-) dcm^+ Tet^r gal λ(DE3) *endA* Hte). Chemical transformation was used for routine cloning, while library transformations used in-house electrocompetent cells. Cells were

cultured either using LB Miller agar or LB Miller liquid media, with relevant antibiotics (carbenicillin, 100 μg/mL).

Site-directed mutagenesis (SDM) was performed by inverse PCRs using 5'-phosphorlyated oligo primers, followed by treatment with DpnI and T4 DNA Ligase at 30°C overnight. Site-saturation mutagenesis (SSM) was performed in the same way, except with the use of degenerate (NNK) oligos at the position of interest. For SDM, single colonies were picked, cultured, subjected to plasmid isolation, and sequence verified. For SSM, 96 colonies were picked for screening into media in a 96-well plate.

Libraries were constructed using a DNA shuffling protocol. Briefly, gene templates were amplified by Q5 DNA Polymerase (NEB), then fragmented with DNAseI (Invitrogen). Fragmented DNA was re-assembled and amplified using Exo(-) Pfu DNA Polymerase (Agilent). Full-length library gene fragments were cloned into pET21b(+)-GFP11 between the NdeI and BamHI sites, after digestion with restriction enzymes and ligation with T4 DNA Ligase. The ligated library was transformed into *E. coli* cells, which was selected for on LB Miller plates with carbenicillin. Colonies on plates were streaked into LB liquid media, were prepared as 1.0 OD_{600} glycerol stocks, and were stored at -80°C until use.

High-throughput co-screening assay: Briefly, transformed bacterial libraries were plated on Durapore PVDF 0.45 μm 47 mm membrane filters (product HVLP14250) on LB agar plates. To yield a well-spread, yet pickable density of cells on the plate, libraries were plated at approximately a 2.5×10^5 dilution from a 1.0 OD_{600} freezer cell stock. Library plates were then grown overnight. The next day, Durapore membranes (with cells) were transferred onto LB agar plates with IPTG (1 mM) and incubated for 2 hours to induce protein expression. Membranes were then transferred to BHET screening plates. To cast BHET screening plates, first, a 0.7% (w/v) agarose in [500 mM potassium phosphate pH 8] buffer solution was made. BHET solution (at a working concentration of 500 mM BHET in 100% DMSO) was then added to the agarose solution (in 50 mL total aliquots) to the appropriate concentration (ranging from 20-120 mM BHET), then mixed well, pouring into a 50 mm petri dish, then cooled. 500 mM buffer was used in screening plates due to solubility limitations of agarose at
1 M buffers.

Library colonies were lysed on screening plates by spraying membranes with BugBuster (Millipore) 4 times from a spray bottle, rotating the plate. This method ensures an even coverage of BugBuster and lysed cells across the plate. Membranes were then removed from plates and stored at 4°C on original LB agar plates. Screening plates were then incubated at relevant heat treatment and screening temperatures. Incubations and reactions were done in VWR Hybridization

Ovens (model 5420), for 2 to 24 hours. After reactions were completed, solutions of refolded GFP1-10 in [100 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% (v/v) glycerol] (TNG buffer) were put on screening plates. GFP1-10 was refolded from inclusion bodies then incubated 4 hours to overnight. Plates were imaged using a ChemiDoc MP Imager, detecting colorimetric blot and Alexa 488 signals. Membranes (with partially-lysed colonies) were then re-aligned on screening plates and colonies were picked into LB in 96-well plates for next steps of screening.

Selected colonies from libraries were grown out in plates overnight, then replica plated (Boekel Scientific) onto Durapore membranes. The screening process was repeated as above, but with 8 µL of BugBuster pipetted onto each colony for cell lysis. Colonies chosen from this fine screening were chosen as putative improved variants.

Medium-throughput screening assays: Putative improved mutants were expressed in small-scale, 2 to 25 mL expressions. Starter cultures of colonies grown overnight in LB were inoculated 1:100 into 2 to 25 mL of 2XYT media with antibiotic in Falcon tubes (Fisher Scientific) or deep-well microwell plates (USA Scientific) and grown to 0.6 to 0.8 OD_{600} at 37°C, 250 rpm. Cultures were then placed on ice or at 4°C for 10 minutes before 1 mM IPTG was added to induce expression, which were then grown for an additional 16-20 hours at 20°C, 150 rpm. Cultures were pelleted at 3500 rpm for 20 minutes, supernatant was removed, and pellets were resuspended in 500 µL of lysis buffer [100 mM potassium phosphate pH 8, 200 mM NaCl] then lysed by sonication with a Fisherbrand Model 50 Sonic Dismembrator (Fisher Scientific). Sonication was 5×20 seconds, on ice, centrifuging at 14,000 rpm for 3 minutes at 4°C between cycles, with a final centrifuge for 30 minutes to clarify cell lysate.

Enzyme concentration in cell lysates was measured *via* plate reader (detecting GFP fluorescence intensity; excitation: 488 nm, emission: 520 nm) after complementation with GFP1-10. Briefly, 20 µL of cell lysate was added to Corning MaxiSorp 96-well plates with 180 µL of refolded GFP1-10 in TNG buffer. Plates were then incubated overnight at room temperature with shaking. Proteins were quantified *via* a standard curve from 2-fold serial dilutions of a purified sulfide reductase-GFP11 construct (from 0.11 to 14.26 µM). Background fluorescence was subtracted from all samples using the cell lysate of an expression construct lacking the GFP11 tag [PHL7 in pET21b(+)]. Fluorescence was measured using a Tecan M Plex Plate Reader. GFP1-10 complementation was performed in triplicate.

Proteins were diluted to 0.5 µM using lysis buffer and added 1:10 (to 0.1 µM, in 500 µL total) in reactions containing [1 M potassium phosphate buffer, pH 8] (PHL7) or [100 mM potassium phosphate buffer, pH 8] (LCC-ICCG) reaction buffer, and 0.92% (w/v) PET coupons as 3 mm hole-punched circles (approximately 2.5 mg each; Fiskars). Reactions were then incubated

in deep-well 96-well plates at 70°C, with aliquots drawn at each time point: 2, 4, 6, 8, 24, 48, and 72 hours. Absorbance at 240 nm was measured using a Tecan M Plex Plate Reader to detect aggregate aromatic products released (Baath *et al., Analytical Biochemistry* 607:113873, 2020), with baseline (t=0) absorbance for each enzyme subtracted from timepoints. BHET equivalent concentrations were determined from a standard curve of absorbance of serially-diluted BHET. Promising enzyme variants were grown out and plasmids were isolated and sequenced. Plasmids from any promising variants were used as parents for additional rounds of evolution.

Protein expression and purification: Proteins were expressed using the pET21b(+) expression vector, using His_6 tag purification with Co TALON Resin (Takara Bio). Colonies were streaked out on LB selection plates, picked, and grown out overnight in LB media at 37°C, 250 rpm. Cultures were then inoculated 1:100 into 500 mL 2XYT media with carbenicillin, grown to 0.6 to 0.8 OD_{600} at 37°C, 250 rpm, and induced with 1 mM IPTG after being cooled for 10 minutes on ice or at 4°C. Cultures were then grown for an additional 16-20 hours at 20°C, 150 rpm. Cells were harvested for 20 minutes at 3,500 rpm and stored at -80°C until purification.

For purification, pellets were thawed and resuspended in 30 mL column buffer [100 mM potassium phosphate pH 8, 200 mM NaCl, 10% (v/v) glycerol], then sonicated using a Branson Digital Sonifier 450 at 80% amplitude for 10 minutes on ice at 20°C. Lysate was clarified by centrifuging 1 hour at 4 °C and 40,000 × g, then filtered with a 0.45 μm syringe filter before loading onto 2.5 mL packed, equilibrated resin. The lysate was incubated with the resin, rocking at 4°C overnight. Purification was performed manually. Flow-through was discarded and the resin was washed with 15 column volumes (CV) of column buffer [100 mM potassium phosphate pH 8, 200 mM NaCl, 10% (v/v) glycerol], 10 CVs of column buffer with 5 mM imidazole, and finally eluted with 5 CVs with column buffer with 250 mM imidazole. Proteins were verified for correct size and purity by SDS PAGE gel by running alongside Protein Kaleidoscope Protein Standards (Bio-Rad). Purified protein samples were boiled in Laemmli Buffer at 100°C for 20 minutes before loading on a gel. Purity of purified proteins was >90% (evaluated with Image Lab, Bio-Rad). Enzymes were then buffer exchanged using an Amicon 10 kDa cutoff filter (Millipore Sigma) with [100 mM potassium phosphate pH 8, 200 mM NaCl], using the manufacturer's protocol. Protein concentration was quantified by Pierce BCA Protein Assay (Fisher Scientific) using the manufacturer's protocol. Aliquots of the enzymes were stored at -80°C.

Protein thermostability assay: Enzymes in cell lysates were normalized to the same concentration, 0.5 μM, and incubated for 1 hour in a thermal cycler (MJ Research; model PTC-200) at a range of temperatures, from 60°C to 85°C, in reaction buffer [1 M potassium phosphate, pH 8] in PCR tubes. Following heat treatment, samples were removed, transferred to 1.5 mL microtubes,

and centrifuged at 14,000 × g for 3 minutes to separate aggregated protein and cell debris. Supernatant was removed and GFP complementation was used to quantify the amount of soluble enzyme remaining by diluting 1:10 in a solution of GFP1-10 in TNG buffer and incubated for 4 hours to overnight, shaking, at room temperature, in the wells of Corning MaxiSorp 96-well plates. Background fluorescence was subtracted from all samples. Fluorescence was measured using a Tecan M Plex Plate Reader (ex: 488 nm, em: 520 nm). All samples were performed in triplicate. Remaining protein was compared to initial concentrations to determine fraction of protein retained.

Small-scale PET hydrolysis reactions: Reactions were performed with 0.69 or 0.345 μM enzyme and 2.9% (w/v) loading PET (0.35 or 0.7 mg enzyme/g PET for PHL7-WT) in 500 μL evaporation-proof cryo-vials (Simport Scientific; product T309-2A). Reactions were composed of PET, enzymes (diluted with lysis buffer), and appropriate potassium phosphate buffer (of varied pH and concentration). PET was either in the form of milled powder (added prior to reaction buffer and aliquoted into reactions after re-suspension) or as film in the form of 3 mm hole-punched circular coupons (Fiskars). Time points were taken at 2, 4, 6, 8, 24, 48, and 72 hours, incubating at the reaction temperature. Samples were taken for absorbance measurement and HPLC analysis. For HPLC, samples were immediately diluted 50% (v/v) with methanol and then filtered using a 0.2 μm plate filter using MultiScreen HTS Filter Plates (Millipore Sigma; product MSGVN2250). Absorbance measurement was performed as above. Samples were stored at -20°C until analysis. As necessary, samples for absorbance and HPLC analysis were diluted with ultrapure water. All reactions were performed in triplicate.

Monomer quantification: Concentrations of monomers TPA, MHET, and BHET were quantified by HPLC using an Agilent Technologies Infinity II 1260, equipped with a G7115A diode array detector (DAD), detecting signal at 240 nm. 10 μL of sample maintained at 10°C were injected onto a Phenomenex Luna C18(2) (100 Å, 150 mm × 4.6 mm, 5 μm) 40°C. The mobile phase consisted of (A) 20 mM phosphoric acid in ultrapure water and (B) 100% methanol. The flow rate was a constant 1.2 mL/min for a total time of 10 min per sample. An A:B gradient program was used, as follows: 80:20 at t = 0 min; a gradient to 35:65 by t = 7.5 min; and held constant at 80:20 from t = 7.51 min to 10 min. A calibration curve, from 0.1 to 500 mg/L, was used for each analyte to determine concentrations.

PET hydrolysis in pH-controlled bioreactors: Enzymatic PET hydrolysis reactions at 200 mL scale were carried out in duplicate using Applikon MiniBio bioreactor systems with 250 mL glass vessels (Getinge AB, Sweden) equipped with one marine impeller. Amorphous PET film of 0.25 mm thickness (Goodfellow) was cut into approximately 10 x 10 mm squares, washed with 70% EtOH, and incubated at 40°C until completely dry. These PET film squares were added to the

reactor at a given solids loading [2.9%, 5.8%, or 20% (w/v)] in 1 M sodium phosphate buffer, pH 8. The suspension was pre-equilibrated to 65°C with stirring at 400 rpm. The reaction was initiated by the addition of enzyme to 1 mg/g PET. Depolymerization reactions proceeded for 48 hours with continuous pH control through the intermittent addition of 6 or 9.5 M NaOH using a peristaltic pump control module (Applikon my-Control). At the end of the reaction, any remaining substrate was recovered by filtration through a Whatman glass microfiber filter (Cytiva) using a Büchner funnel. The retained solid residue was washed with ultrapure water to remove any precipitated salts, and dried at 40°C overnight prior to obtaining the residual dry weight, from which the percentage mass loss was calculated.

Example 2

PHL7 and LCC polypeptides with improved activity

The disclosed methods were applied to polyester hydrolase Leipzig 7 (PHL7). The engineered PHL7 enzymes described herein have improved expression and activity on PET, including activity over time, compared to the wild-type PHL7 enzyme. Enzymes are the result of 4 rounds of directed evolution. These enzymes give 20-30 fold higher expression in *E. coli* cells than the wild-type enzyme, making them far easier and cheaper to produce in large quantities, and exhibit approximately 3- to 4-fold higher initial rates compared to wild-type and LCC ICCG (in breaking down PET within 8 hours) and improved activities over time, with up to approximately 2-fold greater extents of reaction by 72 hours. Thus, these enzymes can break down more PET faster (allowing reactions to be done in less time), as well as have demonstrated performance in reactions over time, as necessary.

The disclosed methods identified five PHL7 variant enzymes including the following mutations: PHL7 variant #1 (A35V, Q95Y, T112I, Q175E, H185N), PHL7 variant #2 (Q80H, Q95Y, Q175E, H185N, R205K), PHL7 variant #3 (T112A, Q175E, H185N), PHL7 variant #4 (R32P, T64S, R111H, Q175E, H185N), and PHL7 variant #5 (A2P, Q95Y, Q175E, H185N, R205K) (all numbering corresponding to positions in SEQ ID NO: 2).

The disclosed methods were also applied to leaf-branch compost cutinase (LCC). For LCC enzyme, LCC-ICCG (Tournier *et al., Nature* 580:216-219, 2020) was used as a starting template for directed evolution using the disclosed screening method. After 4 rounds of directed evolution, selection and characterization, two enzyme variants with 5-10 fold higher initial rates compared to LCC-ICCG were obtained. These enzymes can break down more PET faster, allowing reactions to be done in shorter amount of time.

These methods identified two LCC variant enzymes including the following mutations: LCC variant #2B8 (P38L, Y61C, M91I, L117P, A149V, H218Y, Q224H, S247L, T256I) and LCC variant #EC9 (P38L, Y61C, M91I, L117P, A149V, S247L) (all numbering corresponding to positions in SEQ ID NO: 14).

It will be apparent that the precise details of the methods or compositions described may be varied or modified without departing from the spirit of the described aspects of the disclosure. We claim all such modifications and variations that fall within the scope and spirit of the claims below.

We claim:

1. A method of screening for increased polyethylene terephthalate (PET) degradation activity of a protein of interest, comprising:

expressing a fusion protein comprising the protein of interest linked to a reporter protein in a cell;

contacting the fusion protein with a substrate for the protein of interest;

detecting a signal from the reporter protein; and

detecting degradation of the substrate for the protein of interest,

wherein signal from the reporter protein indicates expression of the fusion protein and increased degradation of the substrate compared to a control indicates that the protein of interest has increased PET degradation activity.

2. The method of claim 1, wherein the reporter protein is a fluorescent protein or portion thereof.

3. The method of claim 2, wherein the fluorescent protein or portion thereof is a green fluorescent protein or portion thereof.

4. The method of claim 1, wherein the reporter protein is a GFP11 split fluorescent protein tag and detecting a signal from the reporter protein comprises contacting the fusion protein with a GFP1-10 split fluorescence protein detector and detecting fluorescence signal.

5. The method of claim 1, wherein expressing the fusion protein comprising the protein of interest linked to a reporter protein in a cell and/or contacting the fusion protein with the substrate for the protein of interest is performed on a solid medium.

6. The method of claim 1, wherein detecting degradation of the substrate for the protein of interest utilizes a colorimetric assay.

7. The method of claim 1, wherein the substrate for the protein of interest comprises bis(2-hydroxyethyl) terephthalate (BHET).

8. The method of claim 1, further comprising:

selecting the protein of interest having increased PET degradation activity; and determining activity of the selected protein of interest to degrade PET.

9. The method of claim 8, wherein activity of the selected protein of interest comprises contacting the selected protein of interest with PET and measuring presence or amount of one or more PET degradation products.

10. The method of claim 9, wherein the one or more PET degradation products is measured by absorbance or high-performance liquid chromatography (HPLC).

11. The method of claim 9, wherein the one or more PET degradation products comprises one or more of bis(2-hydroxyethyl) terephthalate (BHET), mono-(2-hydroxyethyl) terephthalate (MHET), terephthalic acid (TPA) and ethylene glycol (EG).

12. The method of claim 1, wherein expressing the fusion protein comprises expressing a library of fusion proteins comprising one or more amino acid substitutions in the protein of interest compared to a wild type or control protein of interest.

13. The method of claim 12, wherein the control protein of interest is a modified version of the protein of interest compared to the wild type protein of interest.

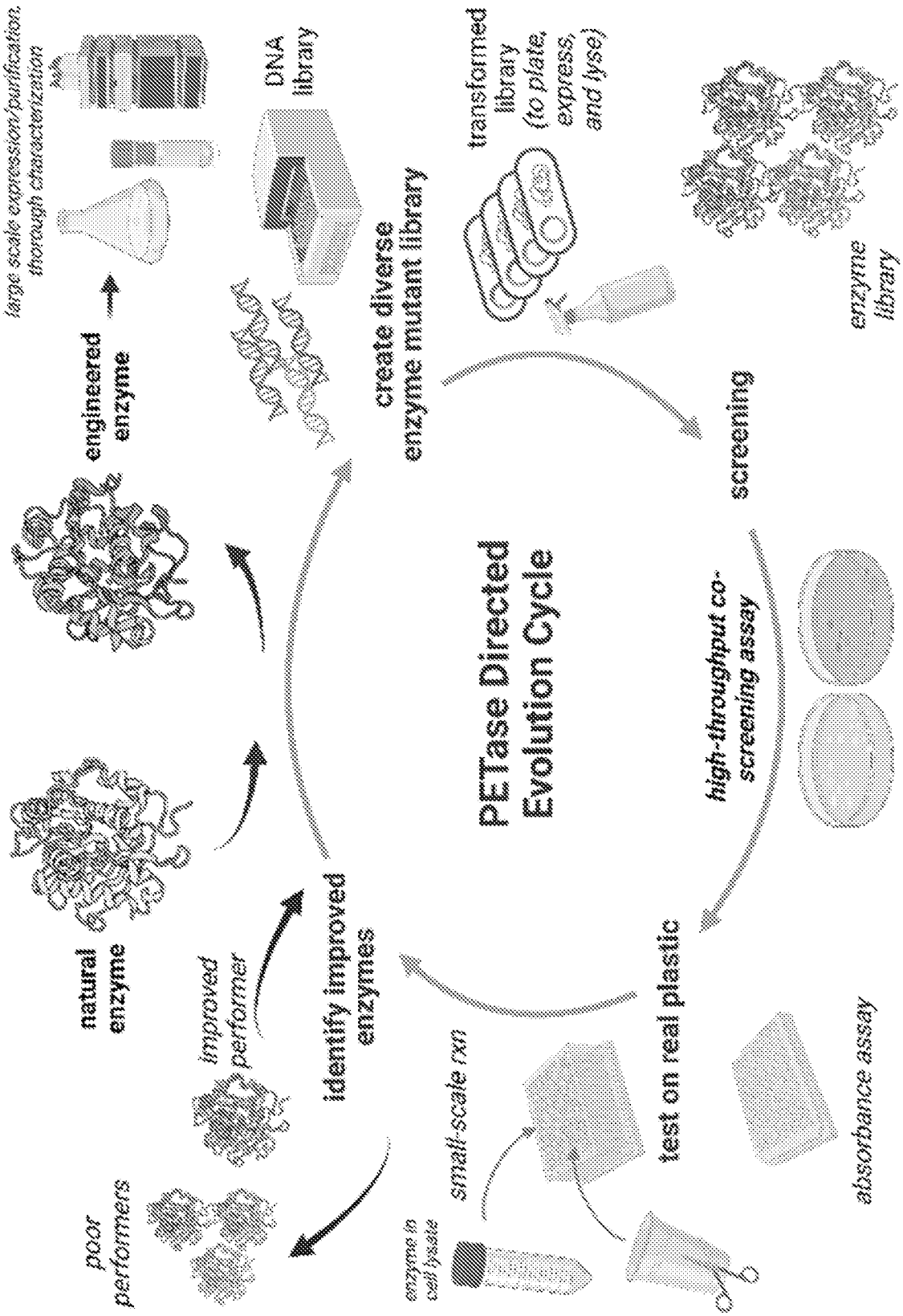
14. The method of claim 1, wherein the protein of interest comprises a leaf-branch compost cutinase or a polyester hydrolase.

15. The method of claim 1, further comprising measuring thermostability of the fusion protein

16. The method of claim 15, wherein measuring thermostability of the fusion protein comprises detecting signal from the reporter protein before and after heat treatment.

17. The method of claim 16, wherein the heat treatment comprises incubating the fusion protein and the substrate at about 65-80°C.

FIG. 1


PETase Directed Evolution Cycle
natural enzyme
improved performer
poor performers
identify improved enzymes
small-scale rxn
enzyme in cell lysate
test on real plastic
high-throughput co-screening assay
absorbance assay
screening
transformed library (to plate, express, and lyse)
enzyme library
create diverse enzyme mutant library
DNA library
engineered enzyme
large scale expression/purification, thorough characterization

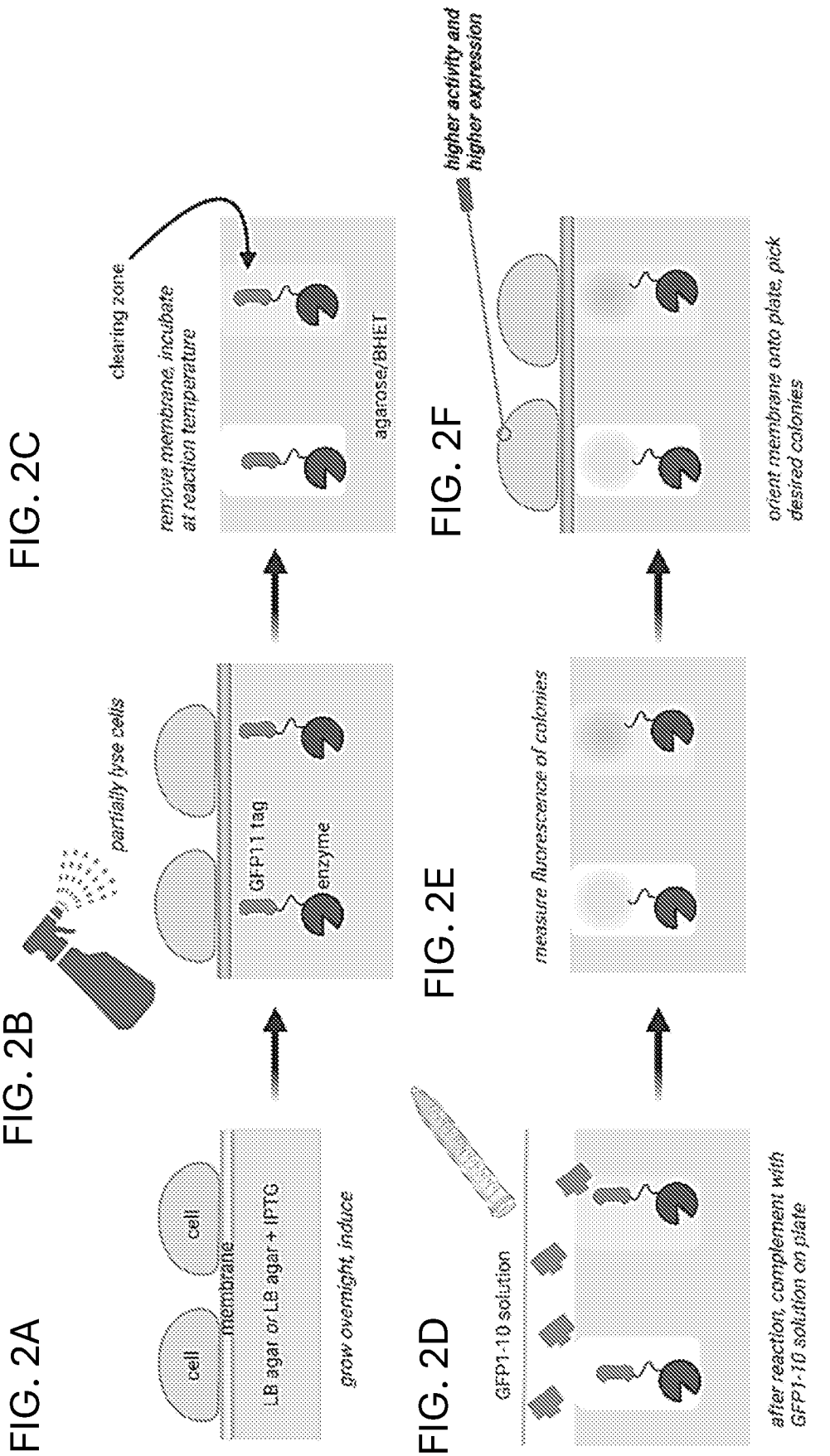


FIG. 2A FIG. 2B FIG. 2C FIG. 2D FIG. 2E FIG. 2F

FIG. 3A

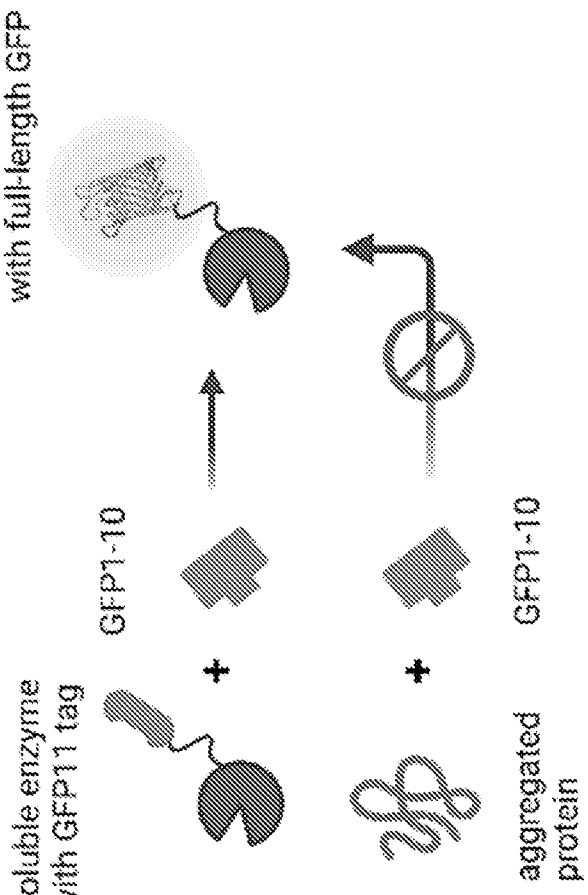

soluble enzyme with GFP11 tag
GFP1-10
soluble enzyme with full-length GFP
aggregated protein
GFP1-10

FIG. 3B

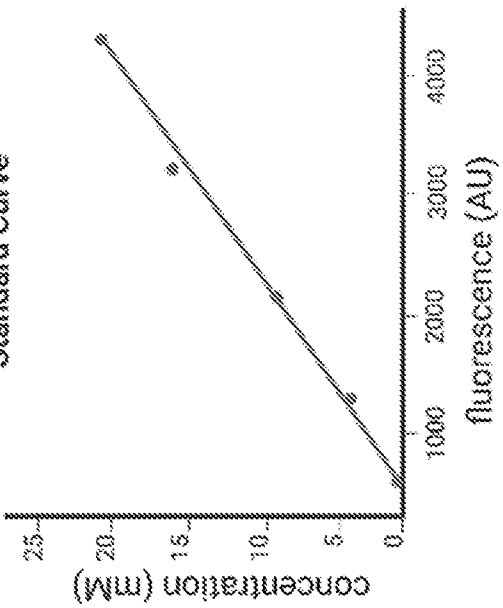

Standard Curve
concentration (mM)
25
20
15
10
5
0
1000
2000
3000
4000
fluorescence (AU)

FIG. 4A

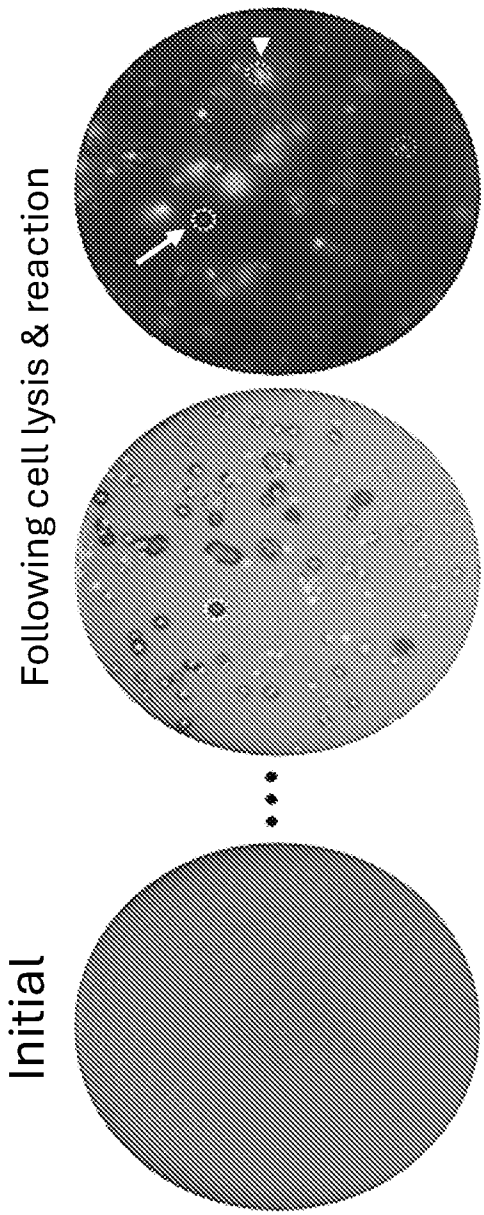

Initial
Following cell lysis & reaction
colorimetric
GFP fluorescence
colorimetric

FIG. 4B

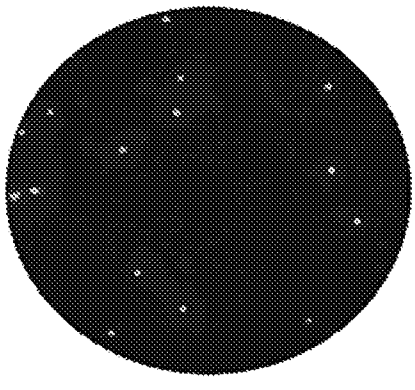

mCherry fluorescence

FIG. 5

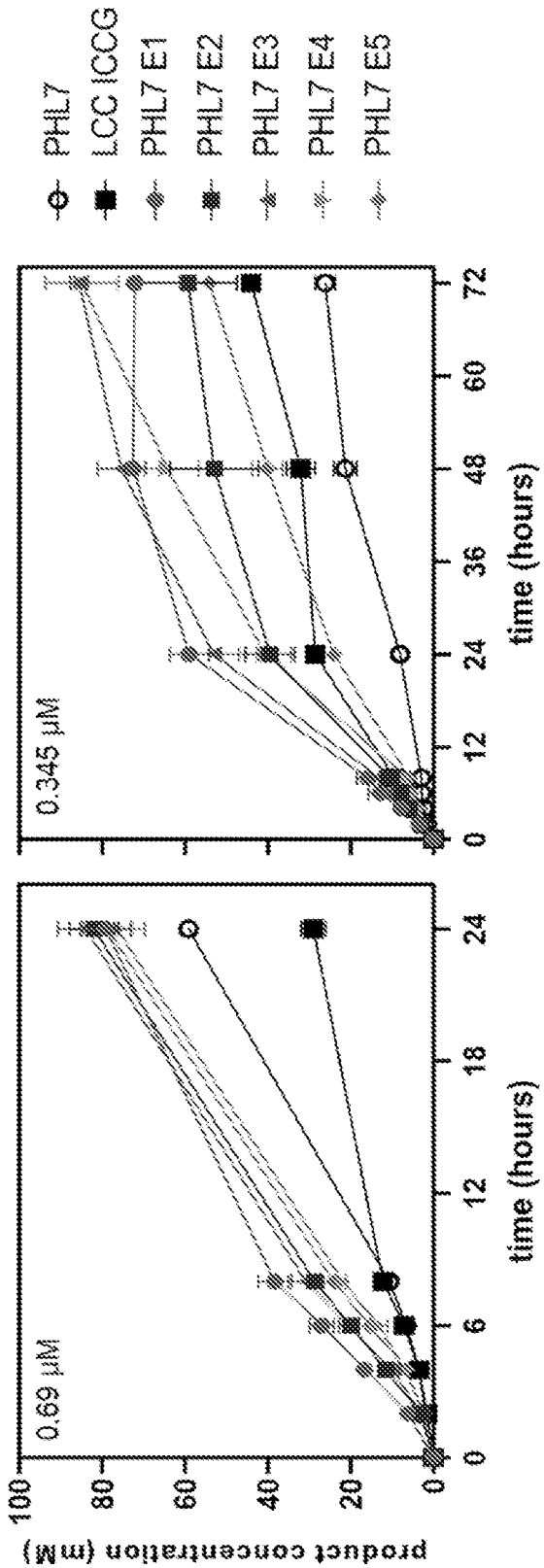
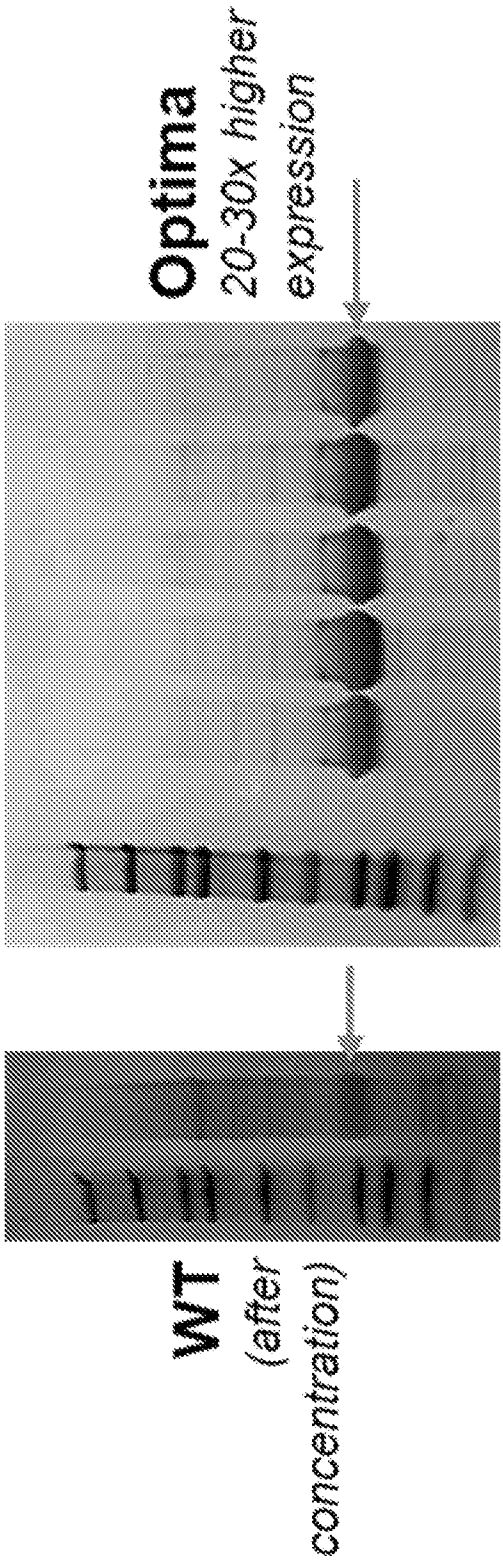

PHL7
LCC ICCG
PHL7 E1
PHL7 E2
PHL7 E3
PHL7 E4
PHL7 E5
0.345 µM
0.69 µM
product concentration (mM)
time (hours)
time (hours)
0
20
40
60
80
100
0
6
12
18
24
0
12
24
36
48
60
72

FIG. 6



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2024/037855

A. CLASSIFICATION OF SUBJECT MATTER

IPC - INV. G01N 33/52 (2024.01)
ADD. C12N 9/14; G01N 21/64 (2024.01)

CPC - INV. Y02W 30/62; G01N 33/52
ADD. C07K 2319/60; C12N 9/14; G01N 21/6428; G01N 2021/6439

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See Search History document

Electronic database consulted during the international search (name of database and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	CN 114636682 A (SHANDONG UNIV) 17 June 2022; Page 2, 3, 4 ; Abstract, claim 1,2	1-3, 5-6, 8-10, 14 --- 4, 7, 11-13, 15-17
Y	WO 2021/253126 A1 (9412-1126 QUéBEC INC. (CA)) 23 December 2021; [0063]	4
Y	JP 2023058381 A (NAT INSTITUTES OF NATURAL SCIENCE et al.) 25 April 2023; [0010], [0011], [0023], [0055]	7, 11-13, 15-17
A	LIU. "Dual Fluorescence Assay Enables High-Throughput Screening for Poly(ethylene terephthalate) Hydrolases" 1-10. Chem Sus Chem. Web. 08 March 2023; <Entire document>; <DOI: 10.1002/cssc.202202019>	1-17
A	CN 112159478 A (JIANGNAN UNIV) 01 January 2021; Entire document	1-17
P,X	WO 2024/122636 A1 (KIRIN HOLDING CO LTD) 13 June 2024; Entire Document	1-17

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 September 2024 (17.09.2024)	OCT 09 2024
Name and mailing address of the ISA/ Mail Stop PCT, Attn: P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/037855

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed.

   b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)),

      ☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments: