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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 June 2019 (20.06.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/118557 A1

(51) International Patent Classification:
C12N 9/18 (2006.01) *A23L 3/3571* (2006.01)

(21) International Application Number:
PCT/US2018/065124

(22) International Filing Date:
12 December 2018 (12.12.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/598,189 13 December 2017 (13.12.2017) US

(71) Applicant: CODEXIS, INC. [US/US]; 200 Penobscot Drive, Redwood City, California 94063 (US).

(72) Inventors: VOLADRI, Rama; 200 Penobscot Drive, Redwood City, California 94063 (US). **ENTWISTLE, David**; 200 Penobscot Drive, Redwood City, California 94063 (US). **MICKLITSCH, Christopher, Michael**; 744 S. Harshaw St., Philadelphia, Pennsylvania 19146 (US). **xMITCHELL, Vesna**; 200 Penobscot Drive, Redwood City, California 94063 (US). **DELLAS, Nikki**; 200 Penobscot Drive, Redwood City, California 94063 (US). **ZHANG, Xiyun**; 34279 Maybird Circle, Fremont, California 94555 (US). **SERO, Antoinette**; 200 Penobscot Drive, Redwood City, California 94063 (US). **DORR, Brent**; 709 Swedeland Road, UW2220, King of Prussia, Pennsylvania 19406 (US). **FUERST, Douglas, E.**; 709 Swedeland Road, UW2220, King of Prussia, Pennsylvania 19406 (US).

(74) Agent: MACKNIGHT, Kamrin, T. et al.; Codexis, Inc., 200 Penobscot Drive, Redwood City, California 94063 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

(54) Title: CARBOXYESTERASE POLYPEPTIDES FOR AMIDE COUPLING

(57) Abstract: The present invention provides engineered carboxyesterase enzymes having improved properties as compared to a naturally occurring wild-type carboxyesterase enzymes, as well as polynucleotides encoding the engineered carboxyesterase enzymes, host cells capable of expressing the engineered carboxyesterase enzymes, and methods of using the engineered carboxyesterase enzymes in amidation reactions.

WO 2019/118557 A1

CARBOXYESTERASE POLYPEPTIDES FOR AMIDE COUPLING

[0001] The present application claims priority to US Prov. Pat. Appln. Ser. No. 62/598,189, filed December 13, 2017, which is hereby incorporated by reference in its entirety for all purposes.

REFERENCE TO SEQUENCE LISTING, TABLE OR COMPUTER PROGRAM

[0002] The Sequence Listing concurrently submitted herewith under 37 C.F.R. §1.821 in a computer readable form (CRF) *via* EFS-Web as file name, CX2-165US1_ST25.txt, is herein incorporated by reference. The electronic copy of the Sequence Listing was created on December 7, 2018, with a file size of 420 Kbytes.

FIELD OF THE INVENTION

[0003] The present invention provides engineered carboxyesterases (E.C. 3.1.1) having improved non-native properties as compared to naturally occurring wild-type (WT) carboxyesterase enzymes, as well as polynucleotides encoding the engineered carboxyesterase enzymes, host cells capable of expressing the engineered carboxyesterase enzymes, and methods of applying the engineered carboxyesterase enzymes to amidation reactions.

BACKGROUND OF THE INVENTION

[0004] Amide bonds are key functional moieties in various synthetic molecules, including polymers (*e.g.*, proteins, nylon), pesticides (*e.g.*, propanil, chlorpropham), and pharmaceuticals (*e.g.*, valsartan, lisdexamfetamine). A recent survey of the prevalence of reaction types employed in the pursuit of novel drug candidates listed N-acyl amidation at approximately 16% among all of those reactions. (Roughley and Jordan, J. Med. Chem., 54: 3451-3479 [2011]). When produced using traditional chemical methods, amide bond formation is a resource intensive transformation. Amide bonds are typically synthesized from carboxylic acids and amines. However, the reaction between these two functional groups does not occur spontaneously at ambient temperatures, with the necessary elimination of water only occurring at high temperatures (*e.g.*, 200°C). These conditions tend to be detrimental to the substrates and products.

[0005] For amidation to occur under more suitable conditions, activation of a carboxylic acid is generally required in order to couple to an amine. Carboxylic activation usually occurs with the aid of a coupling reagent to form an activated ester or anhydride or by transforming the carboxylic acid into the corresponding acid chloride (*i.e.*, through the Schotten-Baumann reaction; See El-Faham and Albericio, Chem. Rev., 111: 6557-6602 [2011]). These reactions are performed with super- or stoichiometric concentrations of expensive coupling reagents that utilize atom-inefficient synthetic routes (See, Pattabiraman and Bode, Nature, 480: 471-479 [2011]). In addition, the reagents, as well as the resulting waste, can be highly toxic and environmentally unfriendly. At least one equivalent of waste is produced per product molecule formed in these reactions, resulting in very low atom economy. Removal of the waste from the reaction mixture is a tedious and expensive process. Thus, more efficient and environmentally friendly means are needed in the art for the production of amide bonds in various settings.

[0006] Lipases have found application on commercial scale for hydrolysis of fatty acid esters, and have been employed for amidation of esters (Seem Faber, Biotransformations in Organic Chemistry, In Special Techniques, Springer-Verlag, New York, NY, [2011]; and Kalkote, et al., Asian J. Biochem., 2: 279-283 [2007]). The formation of amide bonds using enzymes is a highly atom economical process, as there is no need to activate the carboxylic acid as under typical coupling approaches. The use of enzymes in these reactions provides great industrial value, as they are environmentally friendly, occur under mild conditions, and are less expensive than the currently available chemical routes. While eukaryotic proteases and lipases are capable of forming amide bonds (See, Adolfsson et al., Chem. Soc. Rev., 43: 2714-2742 [2014]; Guzman, et al., Elec. J. Biotech., 10(2) [2007]; and Asano et al., J. Biosci. Bioeng., 100(6):662-666 [2005]), these enzymes are typically poorly expressed in prokaryotic expression systems and may perform inferiorly in organic solvents. In contrast, with the aid of directed evolution, the carboxyesterases of the present invention have been easily produced in prokaryotic (*E. coli*) expression systems and have increased tolerance and practicality in organic solvent regimes more suited to this reaction type.

BRIEF DESCRIPTION OF THE FIGURES

[0007] Figure 1 shows a sequence alignment of the polynucleotide sequence encoding the *E. coli* codon optimized for wild-type carboxyesterase enzyme, *Thermobifida fusca* (*T. fusca*) (SEQ ID NO: 1) against each of the polynucleotide sequences that encode the engineered carboxyesterase sequences shown in the Sequence Listing filed herewith (SEQ ID NOs: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131,133, and 135). All of these disclosed polynucleotide sequences are between 98.9-99.9% identical to each other.

[0008] Figure 2 shows a sequence alignment of the polypeptide sequence derived from the wild-type carboxyesterase enzyme, *Thermobifida fusca* (*T. fusca*) (SEQ ID NO: 2) against each of the engineered polypeptide carboxyesterase sequences shown in the Sequence Listing filed herewith (SEQ ID NOs: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, and 136). All of these disclosed polypeptide sequences are between 98.6-99.8% identical to each other.

SUMMARY OF THE INVENTION

[0009] The present invention provides engineered carboxyesterases (E.C. 3.1.1) having improved non-native properties as compared to naturally occurring wild-type (WT) carboxyesterase enzymes, as well as polynucleotides encoding the engineered carboxyesterase enzymes, host cells capable of expressing the engineered carboxyesterase enzymes, and methods of applying the engineered carboxyesterase enzymes to amidation reactions.

[0010] The present invention provides engineered carboxyesterases comprising polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more sequence identity to SEQ ID NO: 2 or a functional fragment thereof, wherein the engineered carboxyesterases comprise at least one substitution or substitution set in their polypeptide sequences, and wherein the amino acid positions of the polypeptide sequences are numbered with reference to SEQ ID NO: 2. In some embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions at positions selected from 39, 39/323, 62, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 147, 153, 153/215, 164, 164/271, 174, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224, 224/268/372, 231, 249, 249/284, 263, 268, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 345, 349, 372, 372/376, 373, 374, 376, 377, 405, 416, 420, 427, 428, 429, 438, and 470, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from: 39/323, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 153, 153/215, 164/271, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224/268/372, 249/284, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 372, 372/376, 373, 376, 377, 405, 420, 427, 428, and 429, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from: 39M/323I, 62H/117G, 63A, 63R, 63T, 63Y, 64A, 64E, 64G, 64I, 64T, 64V, 64W, 65G, 65S, 65T, 65W, 66N, 68L, 68P, 69F, 69G, 69H, 69L, 69V, 69W, 69Y, 70L, 70R, 70T, 70W, 71F, 71G, 71H/263R, 71P, 71R, 71V, 71Y, 77S/184G, 103P, 103R, 103T/147S, 104P, 104Q/429V, 105L, 107D/185W, 107L, 107P, 107S, 108G, 108K, 108Q, 108R, 108S, 108W, 109G/117M, 110A, 110H, 110P, 110S, 111L, 111M, 111R, 111S, 111V, 111W, 113P, 114A, 114H, 114Q, 115H, 115T, 115V, 117A, 117F, 118G/349V, 118I, 118N, 118N/269T, 119G, 119P, 119S, 126C, 153H/215P, 153L, 164R/271T, 174D/282V, 183P, 184F, 184G, 184P, 184S/249T, 184Y, 185A, 185T, 186C, 186G, 186P, 186R, 186T, 187P, 188E, 188G, 190H, 190K, 190L, 190M, 190Q, 190R, 190W, 209E, 209G, 209P, 209S, 209V, 210P, 210T, 210W, 211I, 211L, 211R, 211V, 212A, 212P, 212R, 212S, 213C, 213E, 213L, 213N, 213P, 213Q, 213R/345G, 213S, 213T/271K, 213V, 214K, 214L, 214T, 214V, 215K, 215M, 215P, 215R, 215R/271R, 215W, 216P, 217G, 217L, 217P, 217R, 217R/231V, 217S, 217V, 217W, 224I/268S/372F, 249V/284P, 269N, 269V, 270I, 270I/470M, 270R, 271A, 271K, 271L, 271P, 271Q/416V, 271S, 271T, 276F, 277M, 278H, 278S, 279C, 279E, 279G, 279L/280G/282M, 279V, 280E, 280G, 280S, 281P, 281V, 281Y/374N, 282A, 282C, 282Q, 282R, 282S, 282T, 282W, 283C, 283D, 283K, 283R/429V, 283T, 283V, 283Y, 284C, 284T, 284T/438T, 284V, 285L, 285M, 285P, 286V, 311I, 317C, 317P, 320A, 320F, 320G, 320G/323S, 320S, 320S/323S/372A, 320S/372A/376G, 320S/376G/377V, 320W, 321L, 321S, 323C, 323I, 323R, 323Y, 324A, 372A/376A, 372L, 373G, 376A, 376G, 376L, 376M, 377L, 377W, 377Y, 405D, 420G, 427A, 428V, and 429L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some further embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from: T39M/F323I, R62H/P117G, P63A, P63R, P63T, P63Y, P64A, P64E, P64G, P64I, P64T, P64V, P64W, Y65G, Y65S, Y65T, Y65W, P66N, A68L, A68P, I69F, I69G, I69H, I69L, I69V, I69W, I69Y, G70L, G70R, G70T, G70W, A71F, A71G, A71H/Q263R, A71P, A71R, A71V, A71Y, F77S/E184G, W103P, W103R, W103T/P147S, I104P, I104Q/A429V, H105L, G107D/S185W, G107L, G107P, G107S, A108G, A108K, A108Q, A108R, A108S, A108W, F109G/P117M, T110A, T110H, T110P, T110S, N111L, N111M, N111R, N111S, N111V, N111W, S113P, G114A, G114H, G114Q, S115H, S115T, S115V, P117A, P117F, V118G/A349V, V118I, V118N, V118N/A269T, Y119G, Y119P, Y119S, R126C, R153H/N215P, R153L, W164R/W271T, G174D/L282V, G183P, E184F, E184G, E184P, E184S/A249T, E184Y, S185A, S185T, A186C, A186G, A186P, A186R, A186T, G187P, A188E, A188G, S190H, S190K, S190L, S190M, S190Q, S190R, S190W, L209E, L209G, L209P, L209S, L209V, Q210P, Q210T, Q210W, S211I, S211L, S211R, S211V, G212A, G212P, G212R, G212S, A213C, A213E, A213L, A213N, A213P, A213Q, A213R/S345G, A213S, A213T/W271K, A213V, G214K, G214L, G214T, G214V, N215K, N215M, N215P, N215R, N215R/W271R, N215W, M216P, A217G, A217L, A217P, A217R, A217R/A231V, A217S, A217V, A217W, T224I/P268S/I372F, A249V/F284P, A269N, A269V, V270I, V270I/V470M, V270R, W271A, W271K, W271L, W271P, W271Q/A416V, W271S, W271T, A276F, G277M, G278H, G278S, S279C, S279E, S279G, S279L/V280G/L282M, S279V, V280E, V280G, V280S, L281P, L281V, L281Y/D374N, L282A, L282C, L282Q, L282R, L282S, L282T, L282W, P283C, P283D, P283K, P283R/A429V, P283T, P283V, P283Y, F284C, F284T, F284T/P438T, F284V, A285L, A285M, A285P, P286V, L311I, T317C, T317P, Y320A, Y320F,

Y320G, Y320G/F323S, Y320S, Y320S/F323S/I372A, Y320S/I372A/V376G, Y320S/V376G/F377V, Y320W, R321L, R321S, F323C, F323I, F323R, F323Y, L324A, I372A/V376A, I372L, T373G, V376A, V376G, V376L, V376M, F377L, F377W, F377Y, P405D, P420G, D427A, R428V, and A429L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some embodiments, the engineered carboxyesterases comprise a substitution at position 282, wherein the position is numbered with reference to SEQ ID NO: 2. In some further embodiments, the substitution at position 282 is aliphatic, non-polar, basic, polar, or aromatic. In yet some additional embodiments, the substitution is selected from: X282T, X282G, X282A, X282V, X282M, X282C, X282W, X282Q, X282S, X282T, and X282R.

[0011] The present invention also provides engineered carboxyesterases comprising a polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8 or a functional fragment thereof, wherein the engineered carboxyesterases comprises at least one substitution or substitution set in the polypeptide sequences, wherein the amino acid positions of the polypeptide sequences are numbered with reference to SEQ ID NO: 8. In some embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions at positions selected from: 63, 63/65/108, 63/65/108/189, 63/65/108/377, 63/65/282/285/320/323, 63/65/320/323, 63/108, 63/108/282/285/377, 63/108/285/377, 63/108/320/323, 63/189, 63/212/215, 63/212/215/268/269/343, 63/215, 63/215/269, 63/215/270/271, 63/215/343, 63/268/269/270/429, 63/377, 65/69/70/281/372, 65/69/70/372, 65/69/372, 65/70/372, 65/320, 65/320/323, 68, 68/69/189/214, 68/69/189/214/215, 68/69/189/214/215/271, 68/69/189/214/215/271/281/282/343/381, 68/69/189/214/271/280, 68/69/189/214/372, 68/69/189/214/377/381, 68/69/189/271, 68/69/189/271/280/372/381, 68/69/189/280/281/282/372/377, 68/69/189/281/282/372/377/381, 68/69/189/343/381, 68/69/189/372, 68/69/189/381, 68/69/214/215/271, 68/69/214/343, 68/69/215, 68/69/271, 68/69/282/287, 68/69/343/372, 68/108/377, 68/184, 68/184/189, 68/189/271/372, 68/189/343, 68/214/215/271/281/282/372, 68/215/271/343/372/381, 68/215/377, 68/271/372, 68/377, 69, 69/70, 69/70/331/372, 69/70/372, 69/70/459, 69/108, 69/108/270/372/377, 69/108/281/285, 69/110/215/281, 69/189, 69/189/214/271/281/282/343, 69/189/214/343/372, 69/189/215/343, 69/189/271, 69/189/271/281/282, 69/189/271/343, 69/189/271/343/381, 69/189/280/282/343/372/381, 69/189/281/373, 69/189/282, 69/189/343/381, 69/189/372, 69/189/372/377, 69/189/377, 69/212/213/215/280/281, 69/214/215/271/372/377/381, 69/214/271/282, 69/214/271/343, 69/215, 69/215/269/270/377, 69/215/271/280/281/282, 69/215/271/282, 69/215/271/372, 69/215/285/317, 69/215/323, 69/215/343, 69/215/343/372/381, 69/271, 69/271/372, 69/282, 69/282/343/372, 69/285/373, 69/372, 70, 70/212, 108, 108/189, 108/189/282/285/320, 108/189/320, 108/189/377, 108/215, 108/215/377, 108/269/270, 108/270, 108/282/285/377, 108/285, 108/320/323, 108/377, 126, 126/184/213/280/281/285/320, 126/184/213/372, 126/189/285/372, 126/215, 126/372, 181/215, 189, 189/214, 189/214/215/271/282, 189/215, 189/215/249/277, 189/215/271/281/282/377, 189/215/343/372, 189/270/285, 189/270/372, 189/280/282, 189/320/377, 189/343, 189/343/377, 189/372/377, 189/377, 189/381, 213/215/320, 214/215/271, 214/215/271/377, 214/271, 214/280/282/343/377/381, 215, 215/249/280/281/285/372, 215/271/372, 215/280/281/285/372, 215/281/285/372, 215/281/285/373, 215/281/373, 215/285, 215/285/317, 215/285/346, 215/285/445, 215/320, 215/320/372, 215/323, 215/372, 215/372/377, 215/373, 215/377, 215/381, 249/377, 269/270/281/372/377, 270/377, 271, 271/343, 271/343/372, 271/343/372/381, 280/285/372, 281/372, 282/285/320/323, 285/323, 320, 343/372, 372, 372/377, 372/381, 373, and 377, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8. In some further embodiments, at least one substitution or substitution set in the polypeptide sequence comprises at least one substitution selected from: 63A, 63A/189A, 63A/215R/343V, 63R, 63R/65G/108G, 63R/65G/108G/189L, 63R/65G/108G/377I, 63R/65G/282A/285L/320W/323I, 63R/65G/320W/323I, 63R/108G, 63R/108G/282A/285L/377L, 63R/108G/285L/377I, 63R/108G/320W/323C, 63R/377I, 63T/215R, 63Y, 63Y/189L, 63Y/212P/215R, 63Y/212P/215R/268A/269N/343V, 63Y/215P/269N, 63Y/215R, 63Y/215R/270I/271S, 63Y/268A/269N/270I/429V, 65G/320W, 65G/320W/323I, 65W/69L/372L, 65W/69M/70A/281P/372L, 65W/69W/70L/372L, 65W/70L/372M, 68P, 68P/69L/189E/214R/271Y/280G, 68P/69L/189E/214R/372L, 68P/69L/189I/214R/215P/271Y, 68P/69L/189I/281P/282C/372L/377Y/381L, 68P/69L/189Q/214R, 68P/69L/189Q/271Y/280G/372L/381L, 68P/69L/215P, 68P/69L/271Y, 68P/69L/282C/287I, 68P/69L/343V/372L, 68P/69W/189E/214R/215P/271Y/281P/282G/343V/381L, 68P/69W/189E/280G/281P/282A/372L/377Y, 68P/69W/189E/343V/381L, 68P/69W/189I/214R/215P, 68P/69W/189I/214R/377Y/381L, 68P/69W/189I/271Y, 68P/69W/189I/372L, 68P/69W/189I/381L, 68P/69W/214R/215P/271Y, 68P/69W/214R/343V, 68P/69W/215P, 68P/108G/377L, 68P/184S, 68P/184S/189E, 68P/189I/271Y/372L, 68P/189I/343V, 68P/214R/215P/271Y/281P/282A/372L, 68P/215P/271Y/343V/372L/381L, 68P/215P/377L, 68P/271Y/372L, 68P/377L, 69F/108G/270E/372L/377L, 69F/189L, 69F/215K, 69F/215K/269L/270I/377L, 69F/215R, 69F/285L/373G, 69L, 69L/70L/331Q/372M, 69L/189E/271Y/281P/282A, 69L/189I, 69L/189I/214R/271Y/281P/282A/343V, 69L/189I/271Y/343V/381L, 69L/189I/280G/282G/343V/372L/381L, 69L/189I/282A, 69L/189Q/377Y, 69L/215P/271Y/280G/281P/282C, 69L/215P/271Y/282A, 69L/215P/271Y/372L, 69L/215P/343V/372L/381L, 69L/215R/285P/317P, 69L/271Y, 69L/271Y/372L, 69L/282C/343V/372L, 69L/372L, 69M/70A/372M, 69W, 69W/70L, 69W/70L/372M, 69W/70L/459R, 69W/108S, 69W/189E/214R/343V/372L, 69W/189E/271Y/343V, 69W/189E/372L, 69W/189I, 69W/189I/215P/343V, 69W/189I/271Y, 69W/189I/343V/381L, 69W/189Q/372L/377Y, 69W/212A/213L/215R/280G/281P, 69W/214R/215P/271Y/372L/377Y/381L,

69W/214R/271Y/282A, 69W/214R/271Y/343V, 69W/215K/343V, 69W/215P, 69W/215R, 69W/215R/323Y, 69W/282A, 69W/372M, 69Y/108G/281P/285P, 69Y/110A/215R/281P, 69Y/189L/281P/373G, 70L, 70L/212P, 108G, 108G/189I/282A/285L/320W, 108G/189L, 108G/189L/320W, 108G/189L/377I, 108G/215K, 108G/215P/377L, 108G/269L/270E, 108G/270E, 108G/282A/285L/377L, 108G/285L, 108G/320W/323I, 108G/377I, 108G/377L, 126C, 126C/184S/213S/280G/281P/285L/320G, 126C/184S/213S/372L, 126C/189I/285L/372L, 126C/215P, 126C/372L, 181L/215P, 189E/372L/377Y, 189I, 189I/214R/215P/271Y/282G, 189I/215K, 189I/215P/343V/372L, 189I/215R/249T/277M, 189I/270E/285L, 189I/270E/372L, 189I/280G/282A, 189I/320W/377I, 189I/343V, 189I/377I, 189L, 189Q, 189Q/214R, 189Q/215P/271Y/281P/282C/377Y, 189Q/343V, 189Q/343V/377Y, 189Q/381L, 213S/215P/320G, 214R/215P/271Y, 214R/215P/271Y/377Y, 214R/271Y, 214R/280G/282A/343V/377Y/381L, 215K, 215K/281P/285L/372L, 215K/281P/373G, 215K/285L/317P, 215K/285L/445L, 215K/323Y, 215K/372L, 215K/372L/377L, 215K/373G, 215P, 215P/271Y/372L, 215P/320G, 215P/320G/372L, 215P/372L, 215P/372L/377L, 215P/377L, 215P/381L, 215R, 215R/249T/280G/281P/285L/372L, 215R/280G/281P/285L/372L, 215R/281P/285L/373G, 215R/285P, 215R/320G, 215R/372L, 215W, 215W/285L/346S, 215W/285P, 215W/373G, 249T/377L, 269L/270E/281P/372L/377L, 270E/377L, 271Y, 271Y/343V, 271Y/343V/372L, 271Y/343V/372L/381L, 280G/285L/372L, 281P/372L, 282A/285L/320W/323I, 285L/323I, 320W, 343V/372L, 372L, 372L/377L, 372L/381L, 372M, 373G, and 377L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8. In some further embodiments, at least one substitution or substitution set in the polypeptide sequence comprises at least one substitution selected from: P63A, P63A/M189A, P63A/N215R/A343V, P63R, P63R/Y65G/A108G, P63R/Y65G/A108G/M189L, P63R/Y65G/A108G/F377I, P63R/Y65G/T282A/A285L/Y320W/F323I, P63R/Y65G/Y320W/F323I, P63R/A108G, P63R/A108G/T282A/A285L/F377L, P63R/A108G/A285L/F377I, P63R/A108G/Y320W/F323C, P63R/F377I, P63T/N215R, P63Y, P63Y/M189L, P63Y/G212P/N215R, P63Y/G212P/N215R/P268A/A269N/A343V, P63Y/N215P/A269N, P63Y/N215R, P63Y/N215R/V270I/W271S, P63Y/P268A/A269N/V270I/A429V, Y65G/Y320W, Y65G/Y320W/F323I, Y65W/I69L/I372L, Y65W/I69M/G70A/L281P/I372L, Y65W/I69W/G70L/I372L, Y65W/G70L/I372M, A68P, A68P/I69L/M189E/G214R/W271Y/V280G, A68P/I69L/M189E/G214R/I372L, A68P/I69L/M189I/G214R/N215P/W271Y, A68P/I69L/M189I/L281P/T282C/I372L/F377Y/A381L, A68P/I69L/M189Q/G214R, A68P/I69L/M189Q/W271Y/V280G/I372L/A381L, A68P/I69L/N215P, A68P/I69L/W271Y, A68P/I69L/T282C/V287I, A68P/I69L/A343V/I372L, A68P/I69W/M189E/G214R/N215P/W271Y/L281P/T282G/A343V/A381L, A68P/I69W/M189E/V280G/L281P/T282A/I372L/F377Y, A68P/I69W/M189E/A343V/A381L, A68P/I69W/M189I/G214R/N215P, A68P/I69W/M189I/G214R/F377Y/A381L,

A68P/I69W/M189I/W271Y, A68P/I69W/M189I/I372L, A68P/I69W/M189I/A381L, A68P/I69W/G214R/N215P/W271Y, A68P/I69W/G214R/A343V, A68P/I69W/N215P, A68P/A108G/F377L, A68P/E184S, A68P/E184S/M189E, A68P/M189I/W271Y/I372L, A68P/M189I/A343V, A68P/G214R/N215P/W271Y/L281P/T282A/I372L, A68P/N215P/W271Y/A343V/I372L/A381L, A68P/N215P/F377L, A68P/W271Y/I372L, A68P/F377L, I69F/A108G/V270E/I372L/F377L, I69F/M189L, I69F/N215K, I69F/N215K/A269L/V270I/F377L, I69F/N215R, I69F/A285L/T373G, I69L, I69L/G70L/P331Q/I372M, I69L/M189E/W271Y/L281P/T282A, I69L/M189I, I69L/M189I/G214R/W271Y/L281P/T282A/A343V, I69L/M189I/W271Y/A343V/A381L, I69L/M189I/V280G/T282G/A343V/I372L/A381L, I69L/M189I/T282A, I69L/M189Q/F377Y, I69L/N215P/W271Y/V280G/L281P/T282C, I69L/N215P/W271Y/T282A, I69L/N215P/W271Y/I372L, I69L/N215P/A343V/I372L/A381L, I69L/N215R/A285P/T317P, I69L/W271Y, I69L/W271Y/I372L, I69L/T282C/A343V/I372L, I69L/I372L, I69M/G70A/I372M, I69W, I69W/G70L, I69W/G70L/I372M, I69W/G70L/G459R, I69W/A108S, I69W/M189E/G214R/A343V/I372L, I69W/M189E/W271Y/A343V, I69W/M189E/I372L, I69W/M189I, I69W/M189I/N215P/A343V, I69W/M189I/W271Y, I69W/M189I/A343V/A381L, I69W/M189Q/I372L/F377Y, I69W/G212A/A213L/N215R/V280G/L281P, I69W/G214R/N215P/W271Y/I372L/F377Y/A381L, I69W/G214R/W271Y/T282A, I69W/G214R/W271Y/A343V, I69W/N215K/A343V, I69W/N215P, I69W/N215R, I69W/N215R/F323Y, I69W/T282A, I69W/I372M, I69Y/A108G/L281P/A285P, I69Y/T110A/N215R/L281P, I69Y/M189L/L281P/T373G, G70L, G70L/G212P, A108G, A108G/M189I/T282A/A285L/Y320W, A108G/M189L, A108G/M189L/Y320W, A108G/M189L/F377I, A108G/N215K, A108G/N215P/F377L, A108G/A269L/V270E, A108G/V270E, A108G/T282A/A285L/F377L, A108G/A285L, A108G/Y320W/F323I, A108G/F377I, A108G/F377L, R126C, R126C/E184S/A213S/V280G/L281P/A285L/Y320G, R126C/E184S/A213S/I372L, R126C/M189I/A285L/I372L, R126C/N215P, R126C/I372L, V181L/N215P, M189E/I372L/F377Y, M189I, M189I/G214R/N215P/W271Y/T282G, M189I/N215K, M189I/N215P/A343V/I372L, M189I/N215R/A249T/G277M, M189I/V270E/A285L, M189I/V270E/I372L, M189I/V280G/T282A, M189I/Y320W/F377I, M189I/A343V, M189I/F377I, M189L, M189Q, M189Q/G214R, M189Q/N215P/W271Y/L281P/T282C/F377Y, M189Q/A343V, M189Q/A343V/F377Y, M189Q/A381L, A213S/N215P/Y320G, G214R/N215P/W271Y, G214R/N215P/W271Y/F377Y, G214R/W271Y, G214R/V280G/T282A/A343V/F377Y/A381L, N215K, N215K/L281P/A285L/I372L, N215K/L281P/T373G, N215K/A285L/T317P, N215K/A285L/V445L, N215K/F323Y, N215K/I372L, N215K/I372L/F377L, N215K/T373G, N215P, N215P/W271Y/I372L, N215P/Y320G, N215P/Y320G/I372L, N215P/I372L, N215P/I372L/F377L, N215P/F377L, N215P/A381L, N215R, N215R/A249T/V280G/L281P/A285L/I372L,

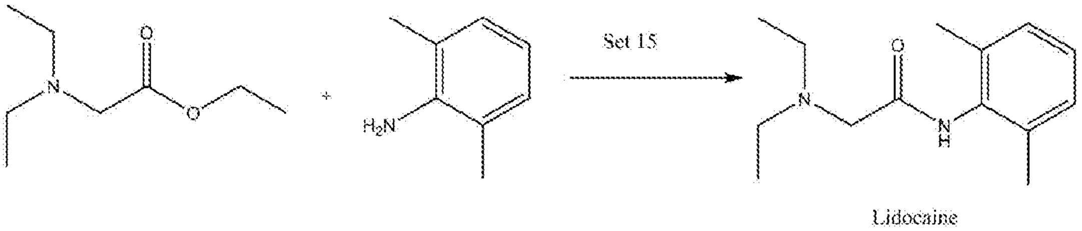
N215R/V280G/L281P/A285L/I372L, N215R/L281P/A285L/T373G, N215R/A285P, N215R/Y320G, N215R/I372L, N215W, N215W/A285L/G346S, N215W/A285P, N215W/T373G, A249T/F377L, A269L/V270E/L281P/I372L/F377L, V270E/F377L, W271Y, W271Y/A343V, W271Y/A343V/I372L, W271Y/A343V/I372L/A381L, V280G/A285L/I372L, L281P/I372L, T282A/A285L/Y320W/F323I, A285L/F323I, Y320W, A343V/I372L, I372L, I372L/F377L, I372L/A381L, I372M, T373G, and F377L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8.

[0012] The present invention also provides engineered carboxyesterases comprising polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 138 or a functional fragment thereof, wherein the engineered carboxyesterases comprise at least one substitution or substitution set in their polypeptide sequences, and wherein the amino acid positions of the polypeptide sequences are numbered with reference to SEQ ID NO: 138.

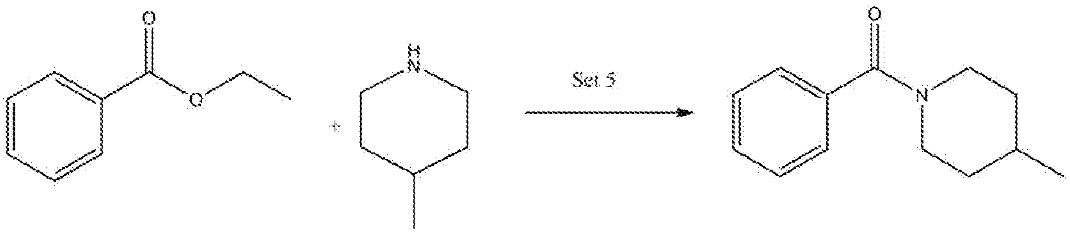
[0013] The present invention also provides engineered carboxyesterases comprising polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, sequence identity to SEQ ID NO: 140 or a functional fragment thereof, wherein the engineered carboxyesterases comprise at least one substitution or substitution set in the polypeptide sequences, and wherein the amino acid positions of the polypeptide sequences are numbered with reference to SEQ ID NO: 140.

[0014] The present invention also provides engineered carboxyesterases comprising polypeptide sequences selected from: SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, and 136.

[0015] The present invention further provides engineered carboxyesterases exhibiting at least one improved property as compared to the wild-type *T. fusca* carboxyesterase of SEQ ID NO:2. In some embodiments, the improved property is selected from: improved amidation activity, solvent tolerance, thermostability, pH stability, regiosteroselectivity, stereoselectivity, substrate scope, and/or reduced substrate or product inhibition, and reduced toxicity to bacterial host cells producing the engineered carboxyesterase. In some additional embodiments, the bacterial host cells comprise *E. coli*. In still some additional embodiments, the engineered carboxyesterases exhibit improved solvent tolerance to at least one solvent selected from: acetone, acetonitrile, toluene, tetrahydrofuran, isopropanol, isopropyl acetate, dimethyl sulfoxide and/or methyl ethyl ketone. In some further embodiments, the engineered carboxyesterases exhibit greater activity than wild-type *T. fusca* carboxyesterase on at least one substrate selected from: aniline, isobutylamine, n-butylamine, t-butylamine, *N'*-t-butoxycarbonyl-benzhydrazide, 4-methylpiperidine, *O*-t-butylhdroxylamine, benzylamine, 2,6-dimethylaniline, (*S*)-(-)-α-methylbenzylamine, (*R*)-(+)-α-methylbenzylamine, methyl phenylacetate, ethyl acetate, ethyl benzoate, 2-pyrazinyl ethyl ester, 4-ethyl-*1H*-indole ester, *N*,*N*-diethylglycyl methyl ester. In some additional embodiments, the engineered carboxyesterases exhibit greater activity than wild-type *T. fusca* carboxyesterase on at least one substrate or substrate set selected from aniline, isobutylamine, n-butylamine, t-butylamine, *N*'-t-butoxycarbonyl-benzhydrazide, 4-methylpiperidine, *O*-t-butylhdroxylamine, benzylamine, 2,6-dimethylaniline, (*S*)-(-)-α-methylbenzylamine, (*R*)-(+)-α-methylbenzylamine, methyl phenylacetate, ethyl acetate, ethyl benzoate, 2-pyrazinyl ethyl ester, 4-ethyl-*1H*-indole ester, *N*,*N*-diethylglycyl methyl ester. In yet some further embodiments, the engineered carboxyesterases exhibit greater activity than wild-type *T. fusca* carboxyesterase in producing at least one product selected from: acetanilide, *N*-n-butyl-benzylacetamide, *N*-[(*S*)-1-phenylethyl]-pyrazinylamide, *N*-[(*S*)-1-phenylethyl]-benzamide, *N*-[(*R*)-1-phenylethyl]-benzamide, *N'*-t-butoxycarbonyl-benzhydrazide, 1-benzoyl-4-methylpiperidine, 2-pyrazinyl-4-methylpiperidine, *N*-isobutyl-benzamide, *N*-t-butyl-benzamide, *N*-t-butylhydroxyl-benzamide, *N*-isobutyl-*1H*-indol-4-amide, *N',N'*-(diethylamino)-*N*-phenylacetamide, *N',N'*-(diethylamino)-*N*-benzylacetamide, *N',N'*-(diethylamino)-*N*-2,6-dimethylphenylacetamide (*i.e.*, lidocaine). In some embodiments, the engineered carboxyesterases of the invention comprises at least one substitution selected from: X343V, X372L, X320W/G, X214R, X282C, X271Y, X65G, wherein the substitutions are numbered with reference to SEQ ID NO:2, and wherein the engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase on a hindered amine for formation of *N',N'*-(diethylamino)-*N*-2,6-dimethylphenylacetamide from ethyl benzoate and 2,6-dimethylaniline, as shown in the following schematic.



[0016] In yet some additional embodiments, the engineered carboxyesterases comprise at least one substitution selected from: X268A, X63A/R, X189Q/I/E, X214R, X282G/C, X381L, and X69W, wherein the substitutions are numbered with reference to SEQ ID NO:2, and wherein the engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase on a secondary amine for formation of 1-benzoyl-4-methyl-piperidine from ethyl benzoate and 4-methyl-piperidine, as shown in the following schematic.



[0017] In yet some additional embodiments, the engineered carboxyesterases provided herein are purified. In still some further embodiments, the engineered carboxyesterases are immobilized. The present invention also provides compositions comprising at least one engineered carboxyesterase provided herein.

[0018] The present invention also provides polynucleotide sequences encoding at least one engineered carboxyesterase provided herein. In some embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2 or a functional fragment thereof, wherein the engineered carboxyesterase comprises at least one substitution or substitution set in its polypeptide sequence, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising substitutions at positions selected from: 39, 39/323, 62, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 147, 153, 153/215, 164, 164/271, 174, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224, 224/268/372, 231, 249, 249/284, 263, 268, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 345, 349, 372, 372/376, 373, 374, 376, 377, 405, 416, 420, 427, 428, 429, 438, and 470, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising at least one substitution or substitution set selected from: 39/323, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 153, 153/215, 164/271, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224/268/372, 249/284, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 372, 372/376, 373, 376, 377, 405, 420, 427, 428, and 429, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising at least one substitution or substitution set selected from: 39M/323I, 62H/117G, 63A, 63R, 63T, 63Y, 64A, 64E, 64G, 64I, 64T, 64V, 64W, 65G, 65S, 65T, 65W, 66N, 68L, 68P, 69F, 69G, 69H, 69L, 69V, 69W, 69Y, 70L, 70R, 70T, 70W, 71F, 71G, 71H/263R, 71P, 71R, 71V, 71Y, 77S/184G, 103P, 103R, 103T/147S, 104P, 104Q/429V, 105L, 107D/185W, 107L, 107P, 107S, 108G, 108K, 108Q, 108R, 108S, 108W, 109G/117M, 110A, 110H, 110P, 110S, 111L, 111M, 111R, 111S, 111V, 111W, 113P, 114A, 114H, 114Q, 115H, 115T, 115V, 117A, 117F, 118G/349V, 118I, 118N, 118N/269T, 119G, 119P, 119S, 126C, 153H/215P, 153L, 164R/271T, 174D/282V, 183P, 184F, 184G, 184P, 184S/249T, 184Y, 185A, 185T, 186C, 186G, 186P, 186R, 186T, 187P, 188E, 188G, 190H, 190K, 190L, 190M, 190Q, 190R, 190W, 209E, 209G, 209P, 209S, 209V, 210P, 210T, 210W, 211I, 211L, 211R, 211V, 212A, 212P, 212R, 212S, 213C, 213E, 213L, 213N, 213P, 213Q, 213R/345G, 213S, 213T/271K, 213V, 214K, 214L, 214T, 214V, 215K, 215M, 215P, 215R, 215R/271R, 215W, 216P, 217G, 217L, 217P, 217R, 217R/231V, 217S, 217V, 217W, 224I/268S/372F, 249V/284P, 269N, 269V, 270I, 270I/470M, 270R, 271A, 271K, 271L, 271P, 271Q/416V, 271S, 271T, 276F, 277M, 278H, 278S, 279C, 279E, 279G, 279L/280G/282M, 279V, 280E, 280G, 280S, 281P, 281V, 281Y/374N, 282A, 282C, 282Q, 282R, 282S, 282T, 282W, 283C, 283D, 283K, 283R/429V, 283T, 283V, 283Y, 284C, 284T, 284T/438T, 284V, 285L, 285M, 285P, 286V, 311I, 317C, 317P, 320A, 320F, 320G, 320G/323S, 320S, 320S/323S/372A, 320S/372A/376G, 320S/376G/377V, 320W, 321L, 321S, 323C, 323I, 323R, 323Y, 324A, 372A/376A, 372L, 373G, 376A, 376G, 376L, 376M, 377L, 377W, 377Y, 405D, 420G, 427A, 428V, and 429L, wherein the amino acid positions are numbered with reference to SEQ ID NO: 2. In some further embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising at least one substitution or substitution set selected from: T39M/F323I, R62H/P117G, P63A, P63R, P63T, P63Y, P64A, P64E, P64G, P64I, P64T, P64V, P64W, Y65G, Y65S, Y65T, Y65W, P66N, A68L, A68P, I69F, I69G, I69H, I69L, I69V, I69W, I69Y, G70L, G70R, G70T, G70W, A71F, A71G, A71H/Q263R, A71P, A71R, A71V, A71Y, F77S/E184G, W103P, W103R, W103T/P147S, I104P, I104Q/A429V, H105L, G107D/S185W, G107L, G107P, G107S, A108G, A108K, A108Q, A108R, A108S, A108W, F109G/P117M, T110A, T110H, T110P, T110S, N111L, N111M, N111R, N111S, N111V, N111W, S113P, G114A, G114H, G114Q, S115H, S115T, S115V, P117A, P117F, V118G/A349V, V118I, V118N, V118N/A269T, Y119G, Y119P, Y119S, R126C, R153H/N215P, R153L, W164R/W271T, G174D/L282V, G183P, E184F, E184G, E184P, E184S/A249T, E184Y, S185A, S185T, A186C, A186G, A186P, A186R, A186T, G187P, A188E, A188G, S190H, S190K, S190L, S190M, S190Q, S190R, S190W, L209E, L209G, L209P, L209S, L209V, Q210P, Q210T, Q210W, S211I, S211L, S211R, S211V, G212A, G212P, G212R, G212S, A213C, A213E, A213L, A213N, A213P, A213Q, A213R/S345G, A213S, A213T/W271K, A213V, G214K, G214L, G214T, G214V, N215K, N215M, N215P, N215R, N215R/W271R, N215W, M216P,

A217G, A217L, A217P, A217R, A217R/A231V, A217S, A217V, A217W, T224I/P268S/I372F, A249V/F284P, A269N, A269V, V270I, V270I/V470M, V270R, W271A, W271K, W271L, W271P, W271Q/A416V, W271S, W271T, A276F, G277M, G278H, G278S, S279C, S279E, S279G, S279L/V280G/L282M, S279V, V280E, V280G, V280S, L281P, L281V, L281Y/D374N, L282A, L282C, L282Q, L282R, L282S, L282T, L282W, P283C, P283D, P283K, P283R/A429V, P283T, P283V, P283Y, F284C, F284T, F284T/P438T, F284V, A285L, A285M, A285P, P286V, L311I, T317C, T317P, Y320A, Y320F, Y320G, Y320G/F323S, Y320S, Y320S/F323S/I372A, Y320S/I372A/V376G, Y320S/V376G/F377V, Y320W, R321L, R321S, F323C, F323I, F323R, F323Y, L324A, I372A/V376A, I372L, T373G, V376A, V376G, V376L, V376M, F377L, F377W, F377Y, P405D, P420G, D427A, R428V, and A429L, wherein the amino acids are numbered with reference to SEQ ID NO: 2. In some embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising a substitution at position 282, wherein the position is numbered with reference to SEQ ID NO: 2. In some further embodiments, the substitution at position 282 is aliphatic, non-polar, basic, polar, or aromatic. In yet some additional embodiments, the substitution selected from X282T, X282G, X282A, X282V, X282M, X282C, X282W, X282Q, X282S, X282T, and X282R.

[0019] The present invention also provides polynucleotide sequences encoding at least one engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8 or a functional fragment thereof, wherein the engineered carboxyesterase comprises at least one substitution or substitution set in its polypeptide sequence, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8. In some embodiments, the polynucleotide sequences encode engineered carboxyesterases comprising at least one substitution or substitution set at positions selected from: 63, 63/65/108, 63/65/108/189, 63/65/108/377, 63/65/282/285/320/323, 63/65/320/323, 63/108, 63/108/282/285/377, 63/108/285/377, 63/108/320/323, 63/189, 63/212/215, 63/212/215/268/269/343, 63/215, 63/215/269, 63/215/270/271, 63/215/343, 63/268/269/270/429, 63/377, 65/69/70/281/372, 65/69/70/372, 65/69/372, 65/70/372, 65/320, 65/320/323, 68, 68/69/189/214, 68/69/189/214/215, 68/69/189/214/215/271, 68/69/189/214/215/271/281/282/343/381, 68/69/189/214/271/280, 68/69/189/214/372, 68/69/189/214/377/381, 68/69/189/271, 68/69/189/271/280/372/381, 68/69/189/280/281/282/372/377, 68/69/189/281/282/372/377/381, 68/69/189/343/381, 68/69/189/372, 68/69/189/381, 68/69/214/215/271, 68/69/214/343, 68/69/215, 68/69/271, 68/69/282/287, 68/69/343/372, 68/108/377, 68/184, 68/184/189, 68/189/271/372, 68/189/343, 68/214/215/271/281/282/372, 68/215/271/343/372/381, 68/215/377, 68/271/372, 68/377, 69, 69/70, 69/70/331/372, 69/70/372, 69/70/459, 69/108, 69/108/270/372/377, 69/108/281/285, 69/110/215/281, 69/189, 69/189/214/271/281/282/343, 69/189/214/343/372, 69/189/215/343, 69/189/271, 69/189/271/281/282, 69/189/271/343, 69/189/271/343/381, 69/189/280/282/343/372/381, 69/189/281/373, 69/189/282, 69/189/343/381, 69/189/372, 69/189/372/377, 69/189/377, 69/212/213/215/280/281, 69/214/215/271/372/377/381, 69/214/271/282, 69/214/271/343, 69/215, 69/215/269/270/377, 69/215/271/280/281/282, 69/215/271/282, 69/215/271/372, 69/215/285/317, 69/215/323, 69/215/343, 69/215/343/372/381, 69/271, 69/271/372, 69/282, 69/282/343/372, 69/285/373, 69/372, 70, 70/212, 108, 108/189, 108/189/282/285/320, 108/189/320, 108/189/377, 108/215, 108/215/377, 108/269/270, 108/270, 108/282/285/377, 108/285, 108/320/323, 108/377, 126, 126/184/213/280/281/285/320, 126/184/213/372, 126/189/285/372, 126/215, 126/372, 181/215, 189, 189/214, 189/214/215/271/282, 189/215, 189/215/249/277, 189/215/271/281/282/377, 189/215/343/372, 189/270/285, 189/270/372, 189/280/282, 189/320/377, 189/343, 189/343/377, 189/372/377, 189/377, 189/381, 213/215/320, 214/215/271, 214/215/271/377, 214/271, 214/280/282/343/377/381, 215, 215/249/280/281/285/372, 215/271/372, 215/280/281/285/372, 215/281/285/372, 215/281/285/373, 215/281/373, 215/285, 215/285/317, 215/285/346, 215/285/445, 215/320, 215/320/372, 215/323, 215/372, 215/372/377, 215/373, 215/377, 215/381, 249/377, 269/270/281/372/377, 270/377, 271, 271/343, 271/343/372, 271/343/372/381, 280/285/372, 281/372, 282/285/320/323, 285/323, 320, 343/372, 372, 372/377, 372/381, 373, and 377, wherein the amino acid positions are numbered with reference to SEQ ID NO: 8. In some further embodiments, the polynucleotide sequence encodes an engineered carboxyesterase comprising at least one substitution or substitution set selected from: 63A, 63A/189A, 63A/215R/343V, 63R, 63R/65G/108G, 63R/65G/108G/189L, 63R/65G/108G/377I, 63R/65G/282A/285L/320W/323I, 63R/65G/320W/323I, 63R/108G, 63R/108G/282A/285L/377L, 63R/108G/285L/377I, 63R/108G/320W/323C, 63R/377I, 63T/215R, 63Y, 63Y/189L, 63Y/212P/215R, 63Y/212P/215R/268A/269N/343V, 63Y/215P/269N, 63Y/215R, 63Y/215R/270I/271S, 63Y/268A/269N/270I/429V, 65G/320W, 65G/320W/323I, 65W/69L/372L, 65W/69M/70A/281P/372L, 65W/69W/70L/372L, 65W/70L/372M, 68P, 68P/69L/189E/214R/271Y/280G, 68P/69L/189E/214R/372L, 68P/69L/189I/214R/215P/271Y, 68P/69L/189I/281P/282C/372L/377Y/381L, 68P/69L/189Q/214R, 68P/69L/189Q/271Y/280G/372L/381L, 68P/69L/215P, 68P/69L/271Y, 68P/69L/282C/287I, 68P/69L/343V/372L, 68P/69W/189E/214R/215P/271Y/281P/282G/343V/381L, 68P/69W/189E/280G/281P/282A/372L/377Y, 68P/69W/189E/343V/381L, 68P/69W/189I/214R/215P, 68P/69W/189I/214R/377Y/381L, 68P/69W/189I/271Y, 68P/69W/189I/372L, 68P/69W/189I/381L, 68P/69W/214R/215P/271Y, 68P/69W/214R/343V, 68P/69W/215P, 68P/108G/377L, 68P/184S, 68P/184S/189E, 68P/189I/271Y/372L, 68P/189I/343V, 68P/214R/215P/271Y/281P/282A/372L, 68P/215P/271Y/343V/372L/381L, 68P/215P/377L, 68P/271Y/372L, 68P/377L, 69F/108G/270E/372L/377L, 69F/189L, 69F/215K, 69F/215K/269L/270I/377L, 69F/215R, 69F/285L/373G, 69L, 69L/70L/331Q/372M, 69L/189E/271Y/281P/282A, 69L/189I, 69L/189I/214R/271Y/281P/282A/343V,

69L/189I/271Y/343V/381L, 69L/189I/280G/282G/343V/372L/381L, 69L/189I/282A, 69L/189Q/377Y, 69L/215P/271Y/280G/281P/282C, 69L/215P/271Y/282A, 69L/215P/271Y/372L, 69L/215P/343V/372L/381L, 69L/215R/285P/317P, 69L/271Y, 69L/271Y/372L, 69L/282C/343V/372L, 69L/372L, 69M/70A/372M, 69W, 69W/70L, 69W/70L/372M, 69W/70L/459R, 69W/108S, 69W/189E/214R/343V/372L, 69W/189E/271Y/343V, 69W/189E/372L, 69W/189I, 69W/189I/215P/343V, 69W/189I/271Y, 69W/189I/343V/381L, 69W/189Q/372L/377Y, 69W/212A/213L/215R/280G/281P, 69W/214R/215P/271Y/372L/377Y/381L, 69W/214R/271Y/282A, 69W/214R/271Y/343V, 69W/215K/343V, 69W/215P, 69W/215R, 69W/215R/323Y, 69W/282A, 69W/372M, 69Y/108G/281P/285P, 69Y/110A/215R/281P, 69Y/189L/281P/373G, 70L, 70L/212P, 108G, 108G/189I/282A/285L/320W, 108G/189L, 108G/189L/320W, 108G/189L/377I, 108G/215K, 108G/215P/377L, 108G/269L/270E, 108G/270E, 108G/282A/285L/377L, 108G/285L, 108G/320W/323I, 108G/377I, 108G/377L, 126C, 126C/184S/213S/280G/281P/285L/320G, 126C/184S/213S/372L, 126C/189I/285L/372L, 126C/215P, 126C/372L, 181L/215P, 189E/372L/377Y, 189I, 189I/214R/215P/271Y/282G, 189I/215K, 189I/215P/343V/372L, 189I/215R/249T/277M, 189I/270E/285L, 189I/270E/372L, 189I/280G/282A, 189I/320W/377I, 189I/343V, 189I/377I, 189L, 189Q, 189Q/214R, 189Q/215P/271Y/281P/282C/377Y, 189Q/343V, 189Q/343V/377Y, 189Q/381L, 213S/215P/320G, 214R/215P/271Y, 214R/215P/271Y/377Y, 214R/271Y, 214R/280G/282A/343V/377Y/381L, 215K, 215K/281P/285L/372L, 215K/281P/373G, 215K/285L/317P, 215K/285L/445L, 215K/323Y, 215K/372L, 215K/372L/377L, 215K/373G, 215P, 215P/271Y/372L, 215P/320G, 215P/320G/372L, 215P/372L, 215P/372L/377L, 215P/377L, 215P/381L, 215R, 215R/249T/280G/281P/285L/372L, 215R/280G/281P/285L/372L, 215R/281P/285L/373G, 215R/285P, 215R/320G, 215R/372L, 215W, 215W/285L/346S, 215W/285P, 215W/373G, 249T/377L, 269L/270E/281P/372L/377L, 270E/377L, 271Y, 271Y/343V, 271Y/343V/372L, 271Y/343V/372L/381L, 280G/285L/372L, 281P/372L, 282A/285L/320W/323I, 285L/323I, 320W, 343V/372L, 372L, 372L/377L, 372L/381L, 372M, 373G, and 377L, wherein the amino acid positions are numbered with reference to SEQ ID NO: 8. In some further embodiments, the polynucleotide sequence encodes an engineered carboxyesterase comprisng at least one substitution or substitution set selected from: P63A, P63A/M189A, P63A/N215R/A343V, P63R, P63R/Y65G/A108G, P63R/Y65G/A108G/M189L, P63R/Y65G/A108G/F377I, P63R/Y65G/T282A/A285L/Y320W/F323I, P63R/Y65G/Y320W/F323I, P63R/A108G, P63R/A108G/T282A/A285L/F377L, P63R/A108G/A285L/F377I, P63R/A108G/Y320W/F323C, P63R/F377I, P63T/N215R, P63Y, P63Y/M189L, P63Y/G212P/N215R, P63Y/G212P/N215R/P268A/A269N/A343V, P63Y/N215P/A269N, P63Y/N215R, P63Y/N215R/V270I/W271S, P63Y/P268A/A269N/V270I/A429V, Y65G/Y320W, Y65G/Y320W/F323I, Y65W/I69L/I372L, Y65W/I69M/G70A/L281P/I372L, Y65W/I69W/G70L/I372L, Y65W/G70L/I372M, A68P, A68P/I69L/M189E/G214R/W271Y/V280G,

A68P/I69L/M189E/G214R/I372L, A68P/I69L/M189I/G214R/N215P/W271Y, A68P/I69L/M189I/L281P/T282C/I372L/F377Y/A381L, A68P/I69L/M189Q/G214R, A68P/I69L/M189Q/W271Y/V280G/I372L/A381L, A68P/I69L/N215P, A68P/I69L/W271Y, A68P/I69L/T282C/V287I, A68P/I69L/A343V/I372L, A68P/I69W/M189E/G214R/N215P/W271Y/L281P/T282G/A343V/A381L, A68P/I69W/M189E/V280G/L281P/T282A/I372L/F377Y, A68P/I69W/M189E/A343V/A381L, A68P/I69W/M189I/G214R/N215P, A68P/I69W/M189I/G214R/F377Y/A381L, A68P/I69W/M189I/W271Y, A68P/I69W/M189I/I372L, A68P/I69W/M189I/A381L, A68P/I69W/G214R/N215P/W271Y, A68P/I69W/G214R/A343V, A68P/I69W/N215P, A68P/A108G/F377L, A68P/E184S, A68P/E184S/M189E, A68P/M189I/W271Y/I372L, A68P/M189I/A343V, A68P/G214R/N215P/W271Y/L281P/T282A/I372L, A68P/N215P/W271Y/A343V/I372L/A381L, A68P/N215P/F377L, A68P/W271Y/I372L, A68P/F377L, I69F/A108G/V270E/I372L/F377L, I69F/M189L, I69F/N215K, I69F/N215K/A269L/V270I/F377L, I69F/N215R, I69F/A285L/T373G, I69L, I69L/G70L/P331Q/I372M, I69L/M189E/W271Y/L281P/T282A, I69L/M189I, I69L/M189I/G214R/W271Y/L281P/T282A/A343V, I69L/M189I/W271Y/A343V/A381L, I69L/M189I/V280G/T282G/A343V/I372L/A381L, I69L/M189I/T282A, I69L/M189Q/F377Y, I69L/N215P/W271Y/V280G/L281P/T282C, I69L/N215P/W271Y/T282A, I69L/N215P/W271Y/I372L, I69L/N215P/A343V/I372L/A381L, I69L/N215R/A285P/T317P, I69L/W271Y, I69L/W271Y/I372L, I69L/T282C/A343V/I372L, I69L/I372L, I69M/G70A/I372M, I69W, I69W/G70L, I69W/G70L/I372M, I69W/G70L/G459R, I69W/A108S, I69W/M189E/G214R/A343V/I372L, I69W/M189E/W271Y/A343V, I69W/M189E/I372L, I69W/M189I, I69W/M189I/N215P/A343V, I69W/M189I/W271Y, I69W/M189I/A343V/A381L, I69W/M189Q/I372L/F377Y, I69W/G212A/A213L/N215R/V280G/L281P, I69W/G214R/N215P/W271Y/I372L/F377Y/A381L, I69W/G214R/W271Y/T282A, I69W/G214R/W271Y/A343V, I69W/N215K/A343V, I69W/N215P, I69W/N215R, I69W/N215R/F323Y, I69W/T282A, I69W/I372M, I69Y/A108G/L281P/A285P, I69Y/T110A/N215R/L281P, I69Y/M189L/L281P/T373G, G70L, G70L/G212P, A108G, A108G/M189I/T282A/A285L/Y320W, A108G/M189L, A108G/M189L/Y320W, A108G/M189L/F377I, A108G/N215K, A108G/N215P/F377L, A108G/A269L/V270E, A108G/V270E, A108G/T282A/A285L/F377L, A108G/A285L, A108G/Y320W/F323I, A108G/F377I, A108G/F377L, R126C, R126C/E184S/A213S/V280G/L281P/A285L/Y320G, R126C/E184S/A213S/I372L, R126C/M189I/A285L/I372L, R126C/N215P, R126C/I372L, V181L/N215P, M189E/I372L/F377Y, M189I, M189I/G214R/N215P/W271Y/T282G, M189I/N215K, M189I/N215P/A343V/I372L, M189I/N215R/A249T/G277M, M189I/V270E/A285L, M189I/V270E/I372L, M189I/V280G/T282A, M189I/Y320W/F377I, M189I/A343V, M189I/F377I,

M189L, M189Q, M189Q/G214R, M189Q/N215P/W271Y/L281P/T282C/F377Y, M189Q/A343V, M189Q/A343V/F377Y, M189Q/A381L, A213S/N215P/Y320G, G214R/N215P/W271Y, G214R/N215P/W271Y/F377Y, G214R/W271Y, G214R/V280G/T282A/A343V/F377Y/A381L, N215K, N215K/L281P/A285L/I372L, N215K/L281P/T373G, N215K/A285L/T317P, N215K/A285L/V445L, N215K/F323Y, N215K/I372L, N215K/I372L/F377L, N215K/T373G, N215P, N215P/W271Y/I372L, N215P/Y320G, N215P/Y320G/I372L, N215P/I372L, N215P/I372L/F377L, N215P/F377L, N215P/A381L, N215R, N215R/A249T/V280G/L281P/A285L/I372L, N215R/V280G/L281P/A285L/I372L, N215R/L281P/A285L/T373G, N215R/A285P, N215R/Y320G, N215R/I372L, N215W, N215W/A285L/G346S, N215W/A285P, N215W/T373G, A249T/F377L, A269L/V270E/L281P/I372L/F377L, V270E/F377L, W271Y, W271Y/A343V, W271Y/A343V/I372L, W271Y/A343V/I372L/A381L, V280G/A285L/I372L, L281P/I372L, T282A/A285L/Y320W/F323I, A285L/F323I, Y320W, A343V/I372L, I372L, I372L/F377L, I372L/A381L, I372M, T373G, and F377L, wherein the amino acid positions are numbered with reference to SEQ ID NO: 8.

[0020] The present invention also provides polynucleotide sequences encoding at least one engineered carboxyesterase or a functional fragment thereof, the polynucleotide sequence comprising at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131,133, 135, 137, and/or 139.

[0021] The present invention further provides polynucleotide sequences encoding engineered carboxyesterases provided herein, wherein the polynucleotide sequence is operably linked to a control sequence. In some additional embodiments, the polynucleotide sequences are codon optimized.

[0022] The present invention also provides expression vectors comprising at least one polynucleotide sequence encoding an engineered carboxyesterase provided herein. In addition, the present invention provides host cells comprising at least one expression vector provided herein. In some embodiments, the present invention also provides host cells comprising at least one polynucleotide sequence encoding at least one engineered carboxyesterase provided herein.

[0023] The present invention also provides methods of producing an engineered carboxyesterase in a host cell, comprising culturing a host cell comprising an expression vector comprising at least one polynucleotide encoding at least one engineered carboxyesterase, under suitable conditions, such that at least one engineered carboxyesterase is produced. In some embodiments, the methods further comprise recovering at least one engineered carboxyesterase from the culture and/or host cell. In some additional embodiments, the methods further comprise the step of purifying the at least one engineered carboxyesterase.

DESCRIPTION OF THE INVENTION

[0024] The present invention provides engineered carboxyesterases (E.C. 3.1.1) having improved non-native properties as compared to naturally occurring wild-type (WT) carboxyesterase enzymes, as well as polynucleotides encoding the engineered carboxyesterase enzymes, host cells capable of expressing the engineered carboxyesterase enzymes, and methods of applying the engineered carboxyesterase enzymes to amidation reactions.

[0025] Switching enzyme function and enzyme substrate scope is feasible, as it has been observed that enzyme active sites are capable of catalyzing several different chemical reactions *via* one amino acid mutation (See Rauwerdink and Kazluaskas, ACS Cat., 5: 6153-6176 [2015]). To improve carboxyesterase functionality, and substrate scope, wild-type carboxyesterases were subjected to directed evolution. The resultant variants possessed improved capabilities in the generation of amide bonds using a diverse set of amine and carboxyester substrate pairs (See, Scheme 1, below). The engineered carboxyesterases had activity not only in aqueous systems, but also are active in the presence of organic co-solvents and even near-total organic solvent concentrations (*e.g.*, ~98% v/v), as described herein. Further, immobilization of these engineered carboxyesterases facilitates continuous flow operations for amide production, aids in the purification of the final amide product, as well as improves the efficiency and overall cost of amidation operations.

[0026] The present invention provides novel engineered carboxyesterase polypeptides, along with their corresponding polynucleotide sequences and methods of application, which demonstrate general amide bond formation (See, Scheme 1, below). In some embodiments, the engineered polypeptides possess modified properties that broaden the functionality and scope of activity of these enzymes as compared to the naturally occurring wild-type *Thermobifida fusca* (*T. fusca*) carboxyesterase (SEQ ID NO: 2). The improved carboxyesterase properties include, but are not limited to: solvent stability, enzymatic activity, regiospecificity, stereoselectivity, reduced host cell toxicity, thermal stability, pH stability, substrate scope, and/or reduced substrate or product inhibition. The present invention also provides polynucleotides that have been improved to facilitate expression of the desired polypeptides in non-natural host organisms (*e.g.*, *E. coli*).

[0027] In some embodiments, the carboxyesterase polypeptides provided herein possess modified properties that expand the functionality and scope of activity of these enzymes as compared to the naturally occurring wild-type *Geobacillus stearothermophilus* carboxyesterase (SEQ ID NO: 138). In some instances, these polypeptides are carboxyesterase enzymes which are enhanced relative to the wild-type *Mycobacterium tuberculosis* carboxyesterase (SEQ ID NO: 140). Further, in some embodiments, the present invention provides polynucleotides that have been improved to facilitate expression of the desired polypeptides in a non-native host organisms (*e.g.*, *E. coli*).

[0028] The improved properties of the carboxyesterase variants presented are related to the engineered amidation polypeptides containing residue differences at specific residue positions as compared to the reference carboxyesterase sequence of *T. fusca* or another referred engineered amidation polypeptide, such as the sequence of SEQ ID NO: 8. In some embodiments, the residue differences are present at least one of the following amino acid positions: X39, X62, X63, X64, X65, X66, X68, X69, X70, X71, X77, X103, X104, X105, X107, X108, X109, X110, X111, X113, X114, X115, X117, X118, X119, X126, X147, X153, X164, X174, X181, X183, X184, X185, X186, X187, X188, X189, X190, X209, X210, X211, X212, X213, X214, X215, X216, X217, X224, X231, X249, X263, X268, X269, X270, X271, X276, X277, X278, X279, X280, X281, X282, X283, X284, X285, X286, X287, X311, X317, X320, X321, X323, X324, X331, X343, X345, X346, X349, X372, X373, X374, X376, X377, X381, X405, X416, X420, X427, X428, X429, X438, X445, X459, and X470.

[0029] In some embodiments, the engineered carboxyesterases provided herein are characterized as exhibiting increased thermostability as compared to the wild-type polypeptide under the same reaction conditions. The engineered carboxyesterases are capable of mediating amidation conversion (See, Scheme 1, below), as indicated by continued formation of products, at higher temperatures and for longer times than the WT carboxyesterase. In some embodiments, the engineered carboxyesterase polypeptides maintain or have increased activity in the presence of higher concentrations of substrate ester **(I)** and/or amine **(II)**, such as 300 mM isobutylamine. In some embodiments, the engineered carboxyesterase polypeptides maintain or have increased activity under conditions with various pH levels (*e.g.*, pH 9.0), as compared to the WT carboxyesterase. In some embodiments, the engineered polypeptides with increased thermostability, pH stability, and/or substrate stability comprise and amino acid sequence that is at least 80%, 82%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more identical to the reference sequence of SEQ ID NOS: 2 and/or 8.

[0030] In some embodiments, the engineered carboxyesterases are capable of biocatalytic activity improvements for converting the substrate compound(s) to product(s) (See, Scheme 1, below) at least about 1.2-fold, 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 150-fold, 500-fold or more relative to the activity of wild-type carboxyesterase (SEQ ID NO: 2) or a reference engineered carboxyesterase (SEQ ID NO: 8), under suitable reaction conditions. In some embodiments, these improvements in enzyme activity extend to associated increases in thermostability, stereoselectivity, stereospecificity, regiospecificity, solvent stability, pH stability, and/or substrate binding, or reduced substrate and/or product inhibition.

[0031] In some embodiments, the engineered carboxyesterases are characterized by activity on a variety of structurally different carboxyester (I) or amine (II) substrates. In some embodiments, engineered polypeptides are capable of biocatalytically converting esters (*N,N*-diethylamino glycine methyl ester, ethyl benzoate, ethyl acetate, pyrazine-2-carboxylic ethyl ester, *1H*-indole-4-carboxylic ethyl ester, methyl phenylacetate), and amines (n-butylamine, isobutylamine, aniline, benzylamine, 2,6-dimethylaniline, t-butylamine, *N'*-t-butoxycarbonyl-benzhydrazide, 4-methylpiperidine, *O*-tbutylhydroxylamine, 2,6-dimethylaniline,or stereoselective conversion of (*S*)-(-)-α-methylbenzylamine, (*R*)-(+)-α-methylbenzylamine), to their corresponding amide product at a greater rate than the WT polypeptides of SEQ ID NO: 2 and/or the engineered polypeptide SEQ ID NO: 8.

[0032] In some embodiments, the improved engineered variant polypeptide comprises an amino acid sequence corresponding to SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88. 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, or 140. In some embodiments, the carboxyesterase enzymes provided herein are obtained by mutagenizing a gene encoding an engineered carboxyesterase polypeptide that is at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more identical to the amino acid sequence of the naturally-occurring *T. fusca* carboxyesterase (SEQ ID NO: 2).

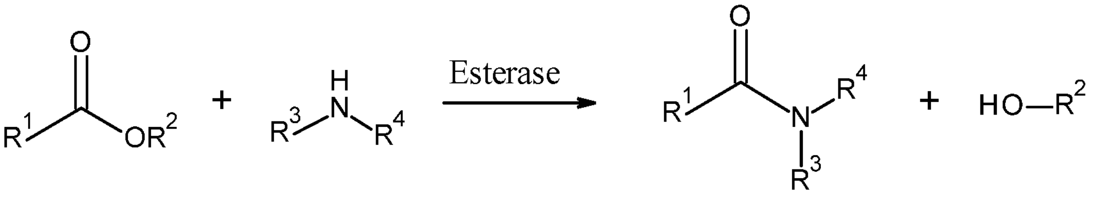
[0033] In some additional embodiments of the present invention, the carboxyesterase polypeptide variants are encoded by polynucleotides or polynucleotides that hybridize to yield such polynucleotides under highly stringent conditions, as provided herein. In some embodiments, the polynucleotides comprise promoters and/or other regulatory elements useful for expression of the encoded engineered carboxyesterase, and can utilize codons optimized for specific expression systems.

[0034] In some embodiments, the polynucleotides encoding the improved carboxyesterase enzymes comprise a sequence selected from SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, and 139.

[0035] In some additional embodiments, the present invention provides host cells comprising the polynucleotides and/or expression vectors provided herein. In some embodiments, the host cells are *T. fusca*, while in some alternative embodiments, they are other organisms (*e.g.*, *E. coli*). The host cells find use in the expression of the encoded polynucleotides to produce the engineered carboxyesterases, and isolation of the engineered carboxyesterases described herein. In some embodiments, the host cells fine use in directly converting substrate(s) to the desired product(s).

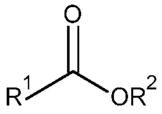
[0036] In some additional embodiments, the present invention provides methods for carrying out reaction Scheme 1 (shown below) using any of the engineered carboxyesterase enzymes provided herein. In some embodiments, the methods comprise contacting or incubating carboxyester (I) and amine (II) substrates with an engineered carboxyesterase polypeptide of the present invention under suitable reaction conditions for the conversion of the substrates to the corresponding amide product, thereby transforming the substrates to the product compounds. Whether carrying out the method with whole cells, cell extracts or purified carboxyesterase enzymes, a single carboxyesterase enzyme can be used or, alternatively, mixtures of at least two carboxyesterase enzymes find use.

Scheme 1



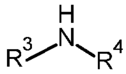
(I) (II) (III) (IV)

[0037] The engineered carboxyesterases of the present invention are capable of converting a diverse set of carboxyester (Scheme 1, **I**) and amine (**II**) substrates to their corresponding amide products (**III**). The scope of carboxyester substrates available can be detailed while considering R^1 and R^2 of formula (**I**). In some embodiments, the functionality of groups at R^1 and R^2 encompasses various components, from a hydrogen atom to alkyl, alkenyl, alkynyl, alkoxy, carboxy, heteroalkyl, heteroalkenyl, heteroalkynyl, carboxyalkyl, aminoalkyl, haloalkyl, alkythioalkyl, cycloalkyl, aryl, arylalkyl, heterocycloalkyl, heteroaryl, and heteroarylalkyl components, excepting that R^2 cannot be a hydrogen. In some additional embodiments, selected functionality at the R^3 and R^4 positions of the amine (**II**) consist of a hydrogen atom, alkyl, alkenyl, alkynyl, alkoxy, carboxy, heteroalkyl, heteroalkenyl, heteroalkynyl, carboxyalkyl, aminoalkyl, haloalkyl, alkythioalkyl, cycloalkyl, aryl, arylalkyl, heterocycloalkyl, heteroaryl, and heteroarylalkyl. In some further embodiments, the selected R^1 and R^2 may be linked to form a 3-membered to 10-membered ring, with the caveat that groups at R^3 and R^4 are separately chosen from alkyl, alkenyl, alkynyl, alkoxy, carboxy, heteroalkyl, heteroalkenyl, heteroalkynyl, carboxyalkyl, aminoalkyl, haloalkyl, alkythioalkyl, cycloalkyl, aryl, arylalkyl, heterocycloalkyl, heteroaryl, heteroarylalkyl, and may also be linked to form a 3-membered to 10-membered ring; and optionally R^1 or R^2 may be linked via a alkyl, alkenyl, alkynyl, alkoxy, carboxy, heteroalkyl, heteroalkenyl, heteroalkynyl, carboxyalkyl, aminoalkyl, haloalkyl, alkythioalkyl, cycloalkyl, aryl, arylalkyl, heterocycloalkyl, heteroaryl, or heteroarylalkyl tether to R^3 or R^4. The amidation process proceeds as the compound (I),



(I)

wherein R^1, and R^2 are as defined above, and a compound of formula (II),



(II)

wherein R^3, and R^4 are as defined above, and an engineered polypeptide having amidative activity under suitable reaction conditions. In some embodiments of the reaction methods provided herein, the WT *T. fusca* carboxyesterase (SEQ ID NO: 2), or a reference engineered polypeptide (*e.g.*, SEQ ID NO: 8), carboxyesterase derivatives presented herein are capable of generating primary and secondary amides matching the generic formula **(III)**. The engineered amidation enzymes provided herein (*e.g.*, the engineered carboxyesterase polypeptides of even numbered sequence identifiers SEQ ID NO: 4 – 136) find use as biocatalysts of the reaction above (Scheme 1).

Definitions

[0038] In reference to the present invention, the technical and scientific terms used in the descriptions herein will have the meanings commonly understood by one of ordinary skill in the art, unless specifically defined otherwise. Accordingly, the following terms are intended to have the following meanings. All U.S patents and published U.S. patent applications, including all sequences disclosed within such patents and patent applications, referred to herein are expressly incorporated by reference. Unless otherwise indicated, the practice of the present invention involves conventional techniques commonly used in molecular biology, fermentation, microbiology, and related fields, which are known to those of skill in the art. Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described. Indeed, it is intended that the present invention not be limited to the particular methodology, protocols, and reagents described herein, as these may vary, depending upon the context in which they are used. The headings provided herein are not limitations of the various aspects or embodiments of the present invention.

[0039] Nonetheless, in order to facilitate understanding of the present invention, a number of terms are defined below. Numeric ranges are inclusive of the numbers defining the range. Thus, every numerical range disclosed herein is intended to encompass every narrower numerical range that falls within such broader numerical range, as if such narrower numerical ranges were all expressly written herein. It is also intended that every maximum (or minimum) numerical limitation disclosed herein includes every lower (or higher) numerical limitation, as if such lower (or higher) numerical limitations were expressly written herein.

[0040] As used herein, the term “comprising” and its cognates are used in their inclusive sense (i.e., equivalent to the term “including” and its corresponding cognates).

[0041] As used herein and in the appended claims, the singular “a”, “an” and “the” include the plural reference unless the context clearly dictates otherwise. Thus, for example, reference to a “host cell” includes a plurality of such host cells.

[0042] Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation and amino acid sequences are written left to right in amino to carboxy orientation, respectively.

[0043] The headings provided herein are not limitations of the various aspects or embodiments of the invention that can be had by reference to the specification as a whole. Accordingly, the terms defined below are more fully defined by reference to the specification as a whole.

[0044] As used herein, “carboxyesterases” are defined as enzymes that naturally have catalytic activity toward the hydrolysis of carboxyesters which results in the formation of an organic acid and an alcohol.

[0045] As used herein, “amidation,” or amide synthesis, refers to the process of generating an amide bond, resulting in a carboxamide (organic amide).

[0046] As used herein, the terms “protein,” “polypeptide,” and “peptide” are used interchangeably herein to denote a polymer of at least two amino acids covalently linked by an amide bond, regardless of length or post-translational modification (e.g., glycosylation, phosphorylation, lipidation, myristilation, ubiquitination, etc.). Included within this definition are D- and L-amino acids, and mixtures of D- and L-amino acids.

[0047] As used herein, “polynucleotide” and “nucleic acid’ refer to two or more nucleosides that are covalently linked together. The polynucleotide may be wholly comprised ribonucleosides (i.e., an RNA), wholly comprised of 2’ deoxyribonucleotides (i.e., a DNA) or mixtures of ribo- and 2’ deoxyribonucleosides. While the nucleosides will typically be linked together via standard phosphodiester linkages, the polynucleotides may include one or more non-standard linkages. The polynucleotide may be single-stranded or double-stranded, or may include both single-stranded regions and double-stranded regions. Moreover, while a polynucleotide will typically be composed of the naturally occurring encoding nucleobases (i.e., adenine, guanine, uracil, thymine, and cytosine), it may include one or more modified and/or synthetic nucleobases (e.g., inosine, xanthine, hypoxanthine, etc.). In one embodiment of the invention, such modified or synthetic nucleobases will be encoding nucleobases.

[0048] As used herein, “coding sequence” refers to that portion of a nucleic acid (e.g., a gene) that encodes an amino acid sequence of a protein.

[0049] As used herein, “naturally occurring,” “wild-type,” and “WT” refer to the form found in nature. For example, a naturally occurring or wild-type polypeptide or polynucleotide sequence is a sequence present in an organism that can be isolated from a source in nature and which has not been intentionally modified by human manipulation.

[0050] As used herein, “non-naturally occurring” or “engineered” or “recombinant” when used in the present invention with reference to (e.g., a cell, nucleic acid, or polypeptide), refers to a material, or a material corresponding to the natural or native form of the material, that has been modified in a manner that would not otherwise exist in nature, or is identical thereto but produced or derived from synthetic materials and/or by manipulation using recombinant techniques. Non-limiting examples include, among others, recombinant cells expressing genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level.

[0051] As used herein, “percentage of sequence identity,” “percent identity,” and “percent identical” refer to comparisons between polynucleotide sequences or polypeptide sequences, and are determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which either the identical nucleic acid base or amino acid residue occurs in both sequences or a nucleic acid base or amino acid residue is aligned with a gap to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Determination of optimal alignment and percent sequence identity is performed using the BLAST and BLAST 2.0 algorithms (See, e.g., Altschul et al., J. Mol. Biol. 215: 403-410 [1990]; and Altschul, et al., Nucleic Acids Res. 3389-3402 [1977]). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information website.

[0052] Briefly, the BLAST analyses involve first identifying high scoring sequence pairs (HSPs) by identifying short words of length within the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as, the neighborhood word score threshold (Altschul, et al., supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (See, e.g., Henikoff and Henikoff, Proc Natl Acad Sci USA 89:10915 [1989]).

[0053] Numerous other algorithms are available and known in the art that function similarly to BLAST in providing percent identity for two sequences. Optimal alignment of sequences for comparison can be conducted using any suitable method known in the art (e.g., by the local homology algorithm of Smith and Waterman, Adv. Appl. Math. 2:482 [1981]; by the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol. 48:443 [1970]; by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. USA 85:2444 [1988]; and/or by computerized implementations of these algorithms [GAP, BESTFIT, FASTA, and TFASTA in the GCG Wisconsin Software Package]), or by visual inspection, using methods commonly known in the art. Additionally, determination of sequence alignment and percent sequence identity can employ the BESTFIT or GAP programs in the GCG Wisconsin Software package (Accelrys, Madison WI), using the default parameters provided.

[0054] As used herein, “reference sequence” refers to a defined sequence to which another sequence is compared. A reference sequence may be a subset of a larger sequence, for example, a segment of a full-length gene or polypeptide sequence. Generally, a reference sequence is at least 20 nucleotide or amino acid residues in length, at least 25 residues in length, at least 50 residues in length, or the full length of the nucleic acid or polypeptide. Since two polynucleotides or polypeptides may each (1) comprise a sequence (i.e., a portion of the complete sequence) that is similar between the two sequences, and (2) may further comprise a sequence that is divergent between the two sequences, sequence comparisons between two (or more) polynucleotides or polypeptide are typically performed by comparing sequences of the two polynucleotides over a comparison window to identify and compare local regions of sequence similarity. The term “reference sequence” is not intended to be limited to wild-type sequences, and can include engineered or altered sequences. For example, in some embodiments, a “reference sequence” can be a previously engineered or altered amino acid sequence.

[0055] As used herein, “comparison window” refers to a conceptual segment of at least about 20 contiguous nucleotide positions or amino acids residues wherein a sequence may be compared to a reference sequence of at least 20 contiguous nucleotides or amino acids and wherein the portion of the sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The comparison window can be longer than 20 contiguous residues, and includes, optionally 30, 40, 50, 100, or longer windows.

[0056] As used herein, “corresponding to”, “reference to” or “relative to” when used in the context of the numbering of a given amino acid or polynucleotide sequence refers to the numbering of the residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to the reference sequence. In other words, the residue number or residue position of a given polymer is designated with respect to the reference sequence rather than by the actual numerical position of the residue within the given amino acid or polynucleotide sequence. For example, a given amino acid sequence, such as that of an engineered carboxyesterase, can be aligned to a reference sequence by introducing gaps to optimize residue matches between the two sequences. In these cases, although the gaps are present, the numbering of the residue in the given amino acid or polynucleotide sequence is made with respect to the reference sequence to which it has been aligned. As used herein, a reference to a residue position, such as "Xn" as further described below, is to be construed as referring to "a residue corresponding to", unless specifically denoted otherwise. Thus, for example, "X94" refers to any amino acid at position 94 in a polypeptide sequence (e.g., SEQ ID NOS:2, 4, 10, 26, or 42).

[0057] As used herein, "stereoselectivity" refers to the preferential formation in a chemical or enzymatic reaction of one stereoisomer over another stereoisomer or another set of stereoisomers. Stereoselectivity can be partial, where the formation of a stereoisomer is favored over another, or it may be complete where only one stereoisomer is formed. When the stereoisomers are enantiomers, the stereoselectivity is referred to as enantioselectivity, the fraction (typically reported as a percentage) of one enantiomer in the sum of both enantiomers. It is commonly alternatively reported in the art (typically as a percentage) as the enantiomeric excess (e.e.) calculated therefrom according to the formula [major enantiomer − minor enantiomer]/[major enantiomer + minor enantiomer]. Where the stereoisomers are diastereoisomers, the stereoselectivity is referred to as diastereoselectivity, the fraction (typically reported as a percentage) of one diastereomer in a mixture of two diastereomers, commonly alternatively reported as the diastereomeric excess (d.e.). Enantiomeric excess and diastereomeric excess are types of stereomeric excess. It is also to be understood that stereoselectivity is not limited to single stereoisomers and can be described for sets of stereoisomers.

[0058] As used herein, "highly stereoselective" refers to a chemical or enzymatic reaction that is capable of converting a substrate to its corresponding chiral amide product, with at least about 75% stereomeric excess.

[0059] As used herein, "increased enzymatic activity" and "increased activity" refer to an improved property of an engineered enzyme, which can be represented by an increase in specific activity (*e.g.*, product produced/time/weight protein) or an increase in percent conversion of the substrate to the product (*e.g.*, percent conversion of starting amount of substrate to product in a specified time period using a specified amount of carboxyesterase) as compared to a reference enzyme. Exemplary methods to determine enzyme activity are provided in the Examples. Any property relating to enzyme activity may be affected, including the classical enzyme properties of K_m, V_{max} or k_{cat}, changes of which can lead to increased enzymatic activity. The carboxyesterase activity can be measured by any one of standard assays used for measuring carboxyesterases, such as change in substrate or product concentration. Comparisons of enzyme activities are made using a defined preparation of enzyme, a defined assay under a set condition, and one or more defined substrates, as further described in detail herein. Generally, when enzymes in cell lysates are compared, the numbers of cells and the amount of protein assayed are determined as well as use of identical expression systems and identical host cells to minimize variations in amount of enzyme produced by the host cells and present in the lysates.

[0060] As used herein, "conversion" refers to the enzymatic transformation of a substrate to the corresponding product.

[0061] As used herein "percent conversion" refers to the percent of the substrate that is converted to the product within a period of time under specified conditions. Thus, for example, the "enzymatic activity" or "activity" of a carboxyesterase polypeptide can be expressed as "percent conversion" of the substrate to the product.

[0062] As used herein, "regiospecificity" refers to chemical reactions in which one structural isomer is produced exclusively when other isomers are also theoretically possible.

[0063] As used herein, "thermostable" or "thermal stable" are used interchangeably to refer to a polypeptide that is resistant to inactivation when exposed to a set of temperature conditions (*e.g.*, 40-80°C) for a period of time (*e.g.*, 0.5-24 hrs) compared to the untreated enzyme, thus retaining a certain level of residual activity (*e.g.*, more than 60% to 80% for example) after exposure to elevated temperatures.

[0064] As used herein, "solvent stable" refers to the ability of a polypeptide to maintain similar activity (*e.g.*, more than *e.g.*, 60% to 80%) after exposure to varying concentrations (*e.g.*, 5-99%) of solvent (*e.g.*, isopropyl alcohol, tetrahydrofuran, 2-methyltetrahydrofuran, acetone, toluene, butylacetate, methyl tert-butylether, *etc.*) for a period of time (*e.g.*, 0.5-24 hrs) compared to the untreated enzyme.

[0065] As used herein, "amino acid difference" or "residue difference" refers to a difference in the amino acid residue at a position of a polypeptide sequence relative to the amino acid residue at a corresponding position in a reference sequence. The positions of amino acid differences generally are referred to herein as "Xn", where n refers to the corresponding position in the reference sequence upon which the residue difference is based. For example, a "residue difference at position X40 as compared to SEQ ID NO:2" refers to a difference of the amino acid residue at the polypeptide position corresponding to position 40 of SEQ ID NO:2. Thus, if the reference polypeptide of SEQ ID NO:2 has a histidine at position 40, then a "residue difference at position X40 as compared to SEQ ID NO:2" refers to an amino acid substitution of any residue other than histidine at the position of the polypeptide corresponding to position 40 of SEQ ID NO:2. In most instances herein, the specific amino acid residue difference at a position is indicated as "XnY" where "Xn" specified the corresponding position as described above, and "Y" is the single letter identifier of the amino acid found in the engineered polypeptide (*i.e.*, the different residue than in the reference polypeptide). In some instances, the present invention also provides specific amino acid differences denoted by the conventional notation "AnB", where A is the single letter identifier of the residue in the reference sequence, "n" is the number of the residue position in the reference sequence, and B is the single letter identifier of the residue substitution in the sequence of the engineered polypeptide. In some instances, a polypeptide of the present invention can include at least one amino acid residue differences relative to a reference sequence, which is indicated by a list of the specified positions where residue differences are present relative to the reference sequence. In some embodiments, where more than one amino acid can be used in a specific residue position of a polypeptide, the various amino acid residues that can be used are separated by a "/" (*e.g.*, X192A/G). The present invention includes engineered polypeptide sequences comprising at least one amino acid differences that include either/or both conservative and non-conservative amino acid substitutions. The amino acid sequences of the specific recombinant carbonic anhydrase polypeptides included in the Sequence Listing of the present invention include an initiating methionine (M) residue (*i.e.*, M represents residue position 1). The skilled artisan, however, understands that this initiating methionine residue can be removed by biological processing machinery, such as in a host cell or *in vitro* translation system, to generate a mature protein lacking the initiating methionine residue, but otherwise retaining the enzyme's properties. Consequently, the term "amino acid residue difference relative to SEQ ID NO:2 at position Xn" as used herein may refer to position "Xn" or to the corresponding position (*e.g.*, position (X-1)n) in a reference sequence that has been processed so as to lack the starting methionine.

[0066] As used herein, the phrase "conservative amino acid substitutions" refers to the interchangeability of residues having similar side chains, and thus typically involves substitution of the amino acid in the polypeptide with amino acids within the same or similar defined class of amino acids. By way of example and not limitation, in some embodiments, an amino acid with an aliphatic side chain is substituted with another aliphatic amino acid (*e.g.*, alanine, valine, leucine, and isoleucine); an amino acid with a hydroxyl side chain is substituted with another amino acid with a hydroxyl side chain (*e.g.*, serine and threonine); an amino acids having aromatic side chains is substituted with another amino acid having an aromatic side chain (*e.g.*, phenylalanine, tyrosine, tryptophan, and histidine); an amino acid with a basic side chain is substituted with another amino acid with a basic side chain (*e.g.*, lysine and arginine); an amino acid with an acidic side chain is substituted with another amino acid with an acidic side chain (*e.g.*, aspartic acid or glutamic acid); and/or a hydrophobic or hydrophilic amino acid is replaced with another hydrophobic or hydrophilic amino acid, respectively. The appropriate classification of any amino acid or residue will be apparent to those of skill in the art, especially in light of the detailed invention provided herein. Exemplary conservative substitutions are provided in Table 1.

Table 1. Exemplary Conservative Amino Acid Substitutions

Residue	Possible Conservative Substitutions
A, L, V, I	Other aliphatic (A, L, V, I) Other non-polar (A, L, V, I, G, M)
G, M	Other non-polar (A, L, V, I, G, M)
D, E	Other acidic (D, E)
K, R	Other basic (K, R)
P	none
N, Q, S, T	Other polar
H, Y, W, F	Other aromatic (H, Y, W, F)
C	none

[0067] As used herein, the phrase "non-conservative substitution" refers to substitution of an amino acid in the polypeptide with an amino acid with significantly differing side chain properties. Non-conservative substitutions may use amino acids between, rather than within, the defined groups and affects (a) the structure of the peptide backbone in the area of the substitution (e.g., proline for glycine) (b) the charge or hydrophobicity, or (c) the bulk of the side chain. By way of example and not limitation, an exemplary non-conservative substitution can be an acidic amino acid substituted with a basic or aliphatic amino acid; an aromatic amino acid substituted with a small amino acid; and a hydrophilic amino acid substituted with a hydrophobic amino acid.

[0068] As used herein, "deletion" refers to modification of the polypeptide by removal of one or more amino acids from the reference polypeptide. Deletions can comprise removal of 1 or more amino acids, 2 or more amino acids, 5 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, or up to 20% of the total number of amino acids making up the polypeptide while retaining enzymatic activity and/or retaining the improved properties of an engineered enzyme. Deletions can be directed to the internal portions and/or terminal portions of the polypeptide. In various embodiments, the deletion can comprise a continuous segment or can be discontinuous.

[0069] As used herein, "insertion" refers to modification of the polypeptide by addition of one or more amino acids to the reference polypeptide. In some embodiments, the improved engineered carboxyesterase enzymes comprise insertions of one or more amino acids to the naturally occurring carboxyesterase polypeptide as well as insertions of one or more amino acids to engineered carboxyesterase polypeptides. Insertions can be in the internal portions of the polypeptide, or to the carboxy or amino terminus. Insertions as used herein include fusion proteins as is known in the art.

The insertion can be a contiguous segment of amino acids or separated by one or more of the amino acids in the naturally occurring polypeptide.

[0070] The term "amino acid substitution set" or "substitution set" refers to a group of amino acid substitutions in a polypeptide sequence, as compared to a reference sequence. A substitution set can have 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or more amino acid substitutions. In some embodiments, a substitution set refers to the set of amino acid substitutions that is present in any of the variant carboxyesterases included in the Tables provided in the Examples.

[0071] As used herein, "fragment" refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion, but where the remaining amino acid sequence is identical to the corresponding positions in the sequence. Fragments can typically have about 80%, about 90%, about 95%, about 98%, or about 99% of the full-length carboxyesterase polypeptide, for example, the polypeptide of SEQ ID NO:4. In some embodiments, the fragment is "biologically active" (i.e., it exhibits the same enzymatic activity as the full-length sequence).

[0072] A "functional fragment", or a "biologically active fragment", used interchangeably, herein refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion(s) and/or internal deletions, but where the remaining amino acid sequence is identical to the corresponding positions in the sequence to which it is being compared (e.g., a full-length engineered T. fusca enzyme of the present invention) and that retains substantially all of the activity of the full-length polypeptide.

[0073] As used herein, "isolated polypeptide" refers to a polypeptide which is substantially separated from other contaminants that naturally accompany it (e.g., protein, lipids, and polynucleotides). The term embraces polypeptides which have been removed or purified from their naturally-occurring environment or expression system (e.g., host cell or in vitro synthesis). The improved carboxyesterase enzymes may be present within a cell, present in the cellular medium, or prepared in various forms, such as lysates or isolated preparations. As such, in some embodiments, the engineered carboxyesterase polypeptides of the present invention can be an isolated polypeptide.

[0074] As used herein, "substantially pure polypeptide" refers to a composition in which the polypeptide species is the predominant species present (i.e., on a molar or weight basis it is more abundant than any other individual macromolecular species in the composition), and is generally a substantially purified composition when the object species comprises at least about 50 percent of the macromolecular species present by mole or % weight. Generally, a substantially pure engineered carboxyesterase polypeptide composition will comprise about 60 % or more, about 70% or more, about 80% or more, about 90% or more, about 91% or more, about 92% or more, about 93% or more, about 94% or more, about 95% or more, about 96% or more, about 97% or more, about 98% or more, or about 99% of all macromolecular species by mole or % weight present in the composition. Solvent species, small molecules (<500 Daltons), and elemental ion species are not considered macromolecular species. In some embodiments, the isolated improved carboxyesterase polypeptide is a substantially pure polypeptide composition.

[0075] As used herein, when used with reference to a nucleic acid or polypeptide, the term “heterologous” refers to a sequence that is not normally expressed and secreted by an organism (e.g., a wild-type organism). In some embodiments, the term encompasses a sequence that comprises two or more subsequences which are not found in the same relationship to each other as normally found in nature, or is recombinantly engineered so that its level of expression, or physical relationship to other nucleic acids or other molecules in a cell, or structure, is not normally found in nature. For instance, a heterologous nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged in a manner not found in nature (e.g., a nucleic acid open reading frame (ORF) of the invention operatively linked to a promoter sequence inserted into an expression cassette, such as a vector). In some embodiments, “heterologous polynucleotide” refers to any polynucleotide that is introduced into a host cell by laboratory techniques, and includes polynucleotides that are removed from a host cell, subjected to laboratory manipulation, and then reintroduced into a host cell.

[0076] As used herein, “codon optimized” refers to changes in the codons of the polynucleotide encoding a protein to those preferentially used in a particular organism such that the encoded protein is efficiently expressed in the organism of interest. In some embodiments, the polynucleotides encoding the carboxyesterase enzymes may be codon optimized for optimal production from the host organism selected for expression.

[0077] As used herein, “control sequence” is defined herein to include all components, which are necessary or advantageous for the expression of a polynucleotide and/or polypeptide of the present invention. Each control sequence may be native or foreign to the polynucleotide of interest. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator.

[0078] As used herein, “operably linked” is defined herein as a configuration in which a control sequence is appropriately placed (i.e., in a functional relationship) at a position relative to a polynucleotide of interest such that the control sequence directs or regulates the expression of the polynucleotide and/or polypeptide of interest.

[0079] As used herein, “suitable reaction conditions” refer to those conditions in the biocatalytic reaction solution (e.g., ranges of enzyme loading, substrate loading, temperature, pH, buffers, co-solvents, etc.) under which an carboxyesterase polypeptide of the present invention is capable of converting a substrate compound to a product compound (e.g., conversion of one compound to another compound). Exemplary “suitable reaction conditions” are provided in the present invention and illustrated by the Examples.

[0080] As used herein, “loading,” such as in “compound loading,” “enzyme loading,” or “substrate loading” refers to the concentration or amount of a component in a reaction mixture at the start of the reaction.

[0081] As used herein, “substrate” in the context of a biocatalyst mediated process refers to the compound or molecule acted on by the biocatalyst.

[0082] As used herein “product” in the context of a biocatalyst mediated process refers to the compound or molecule resulting from the action of the biocatalyst.

[0083] As used herein, “equilibration” as used herein refers to the process resulting in a steady state concentration of chemical species in a chemical or enzymatic reaction (e.g., interconversion of two species A and B), including interconversion of stereoisomers, as determined by the forward rate constant and the reverse rate constant of the chemical or enzymatic reaction.

[0084] As used herein, “alkyl” refers to saturated hydrocarbon groups of from 1 to 18 carbon atoms inclusively, either straight chained or branched, more preferably from 1 to 8 carbon atoms inclusively, and most preferably 1 to 6 carbon atoms inclusively. An alkyl with a specified number of carbon atoms is denoted in parenthesis (e.g., (C1-C4) alkyl refers to an alkyl of 1 to 4 carbon atoms).

[0085] As used herein, “alkenyl” refers to groups of from 2 to 12 carbon atoms inclusively, either straight or branched containing at least one double bond but optionally containing more than one double bond.

[0086] As used herein, “alkynyl” refers to groups of from 2 to 12 carbon atoms inclusively, either straight or branched containing at least one triple bond but optionally containing more than one triple bond, and additionally optionally containing one or more double bonded moieties.

[0087] As used herein, “heteroalkyl, “heteroalkenyl,” and heteroalkynyl,” refer to alkyl, alkenyl and alkynyl as defined herein in which one or more of the carbon atoms are each independently replaced with the same or different heteroatoms or heteroatomic groups. Heteroatoms and/or heteroatomic groups which can replace the carbon atoms include, but are not limited to, -O-, -S-, -S-O-, -NRα-, -PH-, -S(O)-, -S(O)2-, -S(O) NRα-, -S(O)2NRα-, and the like, including combinations thereof, where each Rα is independently selected from hydrogen, alkyl, heteroalkyl, cycloalkyl, heterocycloalkyl, aryl, and heteroaryl.

[0088] As used herein, “alkoxy” refers to the group –ORβ wherein R β is an alkyl group is as defined above including optionally substituted alkyl groups as also defined herein.

[0089] As used herein, “aryl” refers to an unsaturated aromatic carbocyclic group of from 6 to 12 carbon atoms inclusively having a single ring (e.g., phenyl) or multiple condensed rings (e.g., naphthyl or anthryl). Exemplary aryls include phenyl, pyridyl, naphthyl and the like.

[0090] As used herein, “amino” refers to the group -NH2. Substituted amino refers to the group –NHRδ, NRδRδ, and NRδRδRδ, where each Rδ is independently selected from substituted or unsubstituted alkyl, cycloalkyl, cycloheteroalkyl, alkoxy, aryl, heteroaryl, heteroarylalkyl, acyl, alkoxycarbonyl, sulfanyl, sulfinyl, sulfonyl, and the like. Typical amino groups include, but are limited to, dimethylamino, diethylamino, trimethylammonium, triethylammonium, methylysulfonylamino, furanyl-oxy-sulfamino, and the like.

[0091] As used herein, "oxo" refers to =O.

[0092] As used herein, "oxy" refers to a divalent group -O-, which may have various substituents to form different oxy groups, including ethers and esters.

[0093] As used herein, "carboxy" refers to -COOH.

[0094] As used herein, "carbonyl" refers to -C(O)-, which may have a variety of substituents to form different carbonyl groups including acids, acid halides, aldehydes, amides, esters, and ketones.

[0095] As used herein, "alkyloxycarbonyl" refers to -C(O)ORε, where Rε is an alkyl group as defined herein, which can be optionally substituted.

[0096] As used herein, "aminocarbonyl" refers to -C(O)NH2. Substituted aminocarbonyl refers to –C(O)NRδRδ, where the amino group NRδRδ is as defined herein.

[0097] As used herein, "halogen" and "halo" refer to fluoro, chloro, bromo and iodo.

[0098] As used herein, "hydroxy" refers to -OH.

[0099] As used herein, "cyano" refers to -CN.

[0100] As used herein, "heteroaryl" refers to an aromatic heterocyclic group of from 1 to 10 carbon atoms inclusively and 1 to 4 heteroatoms inclusively selected from oxygen, nitrogen and sulfur within the ring. Such heteroaryl groups can have a single ring (e.g., pyridyl or furyl) or multiple condensed rings (e.g., indolizinyl or benzothienyl).

[0101] As used herein, "heteroarylalkyl" refers to an alkyl substituted with a heteroaryl (i.e., heteroaryl-alkyl- groups), preferably having from 1 to 6 carbon atoms inclusively in the alkyl moiety and from 5 to 12 ring atoms inclusively in the heteroaryl moiety. Such heteroarylalkyl groups are exemplified by pyridylmethyl and the like.

[0102] As used herein, "heteroarylalkenyl" refers to an alkenyl substituted with a heteroaryl (i.e., heteroaryl-alkenyl- groups), preferably having from 2 to 6 carbon atoms inclusively in the alkenyl moiety and from 5 to 12 ring atoms inclusively in the heteroaryl moiety.

[0103] As used herein, "heteroarylalkynyl" refers to an alkynyl substituted with a heteroaryl (i.e., heteroaryl-alkynyl- groups), preferably having from 2 to 6 carbon atoms inclusively in the alkynyl moiety and from 5 to 12 ring atoms inclusively in the heteroaryl moiety.

[0104] As used herein, "heterocycle," "heterocyclic," and interchangeably "heterocycloalkyl," refer to a saturated or unsaturated group having a single ring or multiple condensed rings, from 2 to 10 carbon ring atoms inclusively and from 1 to 4 hetero ring atoms inclusively selected from nitrogen, sulfur or oxygen within the ring. Such heterocyclic groups can have a single ring (e.g., piperidinyl or tetrahydrofuryl) or multiple condensed rings (e.g., indolinyl, dihydrobenzofuran or quinuclidinyl). Examples of heterocycles include, but are not limited to, furan, thiophene, thiazole, oxazole, pyrrole, imidazole, pyrazole, pyridine, pyrazine, pyrimidine, pyridazine, indolizine, isoindole, indole, indazole, purine, quinolizine, isoquinoline, quinoline, phthalazine, naphthylpyridine, quinoxaline, quinazoline, cinnoline, pteridine, carbazole, carboline, phenanthridine, acridine, phenanthroline, isothiazole, phenazine, isoxazole, phenoxazine, phenothiazine, imidazolidine, imidazoline, piperidine, piperazine, pyrrolidine, indoline and the like.

[0105] As used herein, “membered ring” is meant to embrace any cyclic structure. The number preceding the term “membered” denotes the number of skeletal atoms that constitute the ring. Thus, for example, cyclohexyl, pyridine, pyran and thiopyran are 6-membered rings and cyclopentyl, pyrrole, furan, and thiophene are 5-membered rings.

[0106] Unless otherwise specified, positions occupied by hydrogen in the foregoing groups can be further substituted with substituents exemplified by, but not limited to, hydroxy, oxo, nitro, methoxy, ethoxy, alkoxy, substituted alkoxy, trifluoromethoxy, haloalkoxy, fluoro, chloro, bromo, iodo, halo, methyl, ethyl, propyl, butyl, alkyl, alkenyl, alkynyl, substituted alkyl, trifluoromethyl, haloalkyl, hydroxyalkyl, alkoxyalkyl, thio, alkylthio, acyl, carboxy, alkoxycarbonyl, carboxamido, substituted carboxamido, alkylsulfonyl, alkylsulfinyl, alkylsulfonylamino, sulfonamido, substituted sulfonamido, cyano, amino, substituted amino, alkylamino, dialkylamino, aminoalkyl, acylamino, amidino, amidoximo, hydroxamoyl, phenyl, aryl, substituted aryl, aryloxy, arylalkyl, arylalkenyl, arylalkynyl, pyridyl, imidazolyl, heteroaryl, substituted heteroaryl, heteroaryloxy, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloalkyl, cycloalkenyl, cycloalkylalkyl, substituted cycloalkyl, cycloalkyloxy, pyrrolidinyl, piperidinyl, morpholino, heterocycle, (heterocycle)oxy, and (heterocycle)alkyl; and preferred heteroatoms are oxygen, nitrogen, and sulfur. It is understood that where open valences exist on these substituents they can be further substituted with alkyl, cycloalkyl, aryl, heteroaryl, and/or heterocycle groups, that where these open valences exist on carbon they can be further substituted by halogen and by oxygen-, nitrogen-, or sulfur-bonded substituents, and where multiple such open valences exist, these groups can be joined to form a ring, either by direct formation of a bond or by formation of bonds to a new heteroatom, preferably oxygen, nitrogen, or sulfur. It is further understood that the above substitutions can be made provided that replacing the hydrogen with the substituent does not introduce unacceptable instability to the molecules of the present invention, and is otherwise chemically reasonable.

[0107] As used herein, “optional” and “optionally” means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where the event or circumstance occurs and instances in which it does not. One of ordinary skill in the art would understand that with respect to any molecule described as containing one or more optional substituents, only sterically practical and/or synthetically feasible compounds are meant to be included.

[0108] As used herein, "optionally substituted" refers to all subsequent modifiers in a term or series of chemical groups. For example, in the term "optionally substituted arylalkyl, the "alkyl" portion and the "aryl" portion of the molecule may or may not be substituted, and for the series "optionally substituted alkyl, cycloalkyl, aryl and heteroaryl," the alkyl, cycloalkyl, aryl, and heteroaryl groups, independently of the others, may or may not be substituted.

[0109] As used herein, "protecting group" refers to a group of atoms that mask, reduce or prevent the reactivity of the functional group when attached to a reactive functional group in a molecule. Typically, a protecting group may be selectively removed as desired during the course of a synthesis. Examples of protecting groups are well-known in the art. Functional groups that can have a protecting group include, but are not limited to, hydroxy, amino, and carboxy groups. Representative amino protecting groups include, but are not limited to, formyl, acetyl, trifluoroacetyl, benzyl, benzyloxycarbonyl ("CBZ"), tert-butoxycarbonyl ("Boc"), trimethylsilyl ("TMS"), 2-trimethylsilyl-ethanesulfonyl ("SES"), trityl and substituted trityl groups, allyloxycarbonyl, 9-fluorenylmethyloxycarbonyl ("FMOC"), nitro-veratryloxycarbonyl ("NVOC") and the like. Representative hydroxyl protecting groups include, but are not limited to, those where the hydroxyl group is either acylated (e.g., methyl and ethyl esters, acetate or propionate groups or glycol esters) or alkylated such as benzyl and trityl ethers, as well as alkyl ethers, tetrahydropyranyl ethers, trialkylsilyl ethers (e.g., TMS or TIPPS groups) and allyl ethers. Other protecting groups can be found in the references noted herein.

Engineered Carboxyesterase Polypeptides

[0110] The present invention provides engineered polypeptides having carboxyesterase activity (also referred to herein as "engineered carboxyesterase polypeptides") useful for amidation reactions. Accordingly, in one aspect, the present invention provides engineered polypeptides having carboxyesterase activity which are capable of converting substrate compound(s) to product compound(s) as shown in Table 3.1 in Example 3. Further, the present invention provides polynucleotides encoding the engineered polypeptides, associated vectors and host cells comprising the polynucleotides, methods for making the engineered polypeptides, and methods for using the engineered polypeptides, including suitable reaction conditions.

[0111] The engineered polypeptides of the present invention are non-naturally occurring carboxyesterases engineered to have improved enzyme properties (*e.g.*, increased stereoselectivity) as compared to the wild-type carboxyesterase polypeptide of *T. fusca* (GenBank Acc. No. WP_011292850.1; SEQ ID NO: 2). In some embodiments, various engineered carboxyesterase polypeptides provided herein exhibit improved enzyme properties as compared to other engineered reference carboxyesterase polypeptides provided herein. In some embodiments, the engineered polypeptides of the present invention are non-naturally occurring carboxyesterases engineered to have improved enzyme properties (*e.g.*, increased thermostability) as compared to the wild-type carboxyesterase polypeptide of *G. stearothermophilus* (GenBank Acc. No. WP_033015113; SEQ ID NO: 138). In some further embodiments, the engineered polypeptides are non-naturally occurring carboxyesterases engineered to have improved enzyme properties (*e.g.*, increased thermostability) as compared to the wild-type carboxyesterase polypeptide of *M. tuberculosis* (GenBank Acc. No. WP_003407276; SEQ ID NO: 140).

[0112] In some embodiments, the engineered carboxyesterase variants provided herein comprise polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2, a reference engineered carboxyesterase (SEQ ID NO: 8), or a functional fragment thereof, wherein the engineered carboxyesterase comprises at least one substitution or substitution set in the polypeptide sequence, and wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some embodiments, the engineered polypeptides having carboxyesterase activity comprise polypeptides having the amino acid substitutions provided herein (See, *e.g.*, Tables 8.1 and 11.).

[0113] The present invention provides engineered carboxyesterases comprising polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2 or a functional fragment thereof, wherein the engineered carboxyesterases comprise at least one substitution or substitution set in their polypeptide sequences, and wherein the amino acid positions of the polypeptide sequences are numbered with reference to SEQ ID NO: 2. In some embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions at positions selected from: 39, 39/323, 62, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 147, 153, 153/215, 164, 164/271, 174, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224, 224/268/372, 231, 249, 249/284, 263, 268, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 345, 349, 372, 372/376, 373, 374, 376, 377, 405, 416, 420, 427, 428, 429, 438, and 470, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from: 39/323, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 153, 153/215, 164/271, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224/268/372, 249/284, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 372, 372/376, 373, 376, 377, 405, 420, 427, 428, and 429, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from: 39M/323I, 62H/117G, 63A, 63R, 63T, 63Y, 64A, 64E, 64G, 64I, 64T, 64V, 64W, 65G, 65S, 65T, 65W, 66N, 68L, 68P, 69F, 69G, 69H, 69L, 69V, 69W, 69Y, 70L, 70R, 70T, 70W, 71F, 71G, 71H/263R, 71P, 71R, 71V, 71Y, 77S/184G, 103P, 103R, 103T/147S, 104P, 104Q/429V, 105L, 107D/185W, 107L, 107P, 107S, 108G, 108K, 108Q, 108R, 108S, 108W, 109G/117M, 110A, 110H, 110P, 110S, 111L, 111M, 111R, 111S, 111V, 111W, 113P, 114A, 114H, 114Q, 115H, 115T, 115V, 117A, 117F, 118G/349V, 118I, 118N, 118N/269T, 119G, 119P, 119S, 126C, 153H/215P, 153L, 164R/271T, 174D/282V, 183P, 184F, 184G, 184P, 184S/249T, 184Y, 185A, 185T, 186C, 186G, 186P, 186R, 186T, 187P, 188E, 188G, 190H, 190K, 190L, 190M, 190Q, 190R, 190W, 209E, 209G, 209P, 209S, 209V, 210P, 210T, 210W, 211I, 211L, 211R, 211V, 212A, 212P, 212R, 212S, 213C, 213E, 213L, 213N, 213P, 213Q, 213R/345G, 213S, 213T/271K, 213V, 214K, 214L, 214T, 214V, 215K, 215M, 215P, 215R, 215R/271R, 215W, 216P, 217G, 217L, 217P, 217R, 217R/231V, 217S, 217V, 217W, 224I/268S/372F, 249V/284P, 269N, 269V, 270I, 270I/470M, 270R, 271A, 271K, 271L, 271P, 271Q/416V, 271S, 271T, 276F, 277M, 278H, 278S, 279C, 279E, 279G, 279L/280G/282M, 279V, 280E, 280G, 280S, 281P, 281V, 281Y/374N, 282A, 282C, 282Q, 282R, 282S, 282T, 282W, 283C, 283D, 283K, 283R/429V, 283T, 283V, 283Y, 284C, 284T, 284T/438T, 284V, 285L, 285M, 285P, 286V, 311I, 317C, 317P, 320A, 320F, 320G, 320G/323S, 320S, 320S/323S/372A, 320S/372A/376G, 320S/376G/377V, 320W, 321L, 321S, 323C, 323I, 323R, 323Y, 324A, 372A/376A, 372L, 373G, 376A, 376G, 376L, 376M, 377L, 377W, 377Y, 405D, 420G, 427A, 428V, and 429L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some further embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from: T39M/F323I, R62H/P117G, P63A, P63R, P63T, P63Y, P64A, P64E, P64G, P64I, P64T, P64V, P64W, Y65G, Y65S, Y65T, Y65W, P66N, A68L, A68P, I69F, I69G, I69H, I69L, I69V, I69W, I69Y, G70L, G70R, G70T, G70W, A71F, A71G, A71H/Q263R, A71P, A71R, A71V, A71Y, F77S/E184G, W103P, W103R, W103T/P147S, I104P, I104Q/A429V, H105L, G107D/S185W, G107L, G107P, G107S, A108G, A108K, A108Q, A108R, A108S, A108W, F109G/P117M, T110A, T110H, T110P, T110S, N111L, N111M, N111R, N111S, N111V, N111W, S113P, G114A, G114H, G114Q, S115H, S115T, S115V, P117A, P117F, V118G/A349V, V118I, V118N, V118N/A269T, Y119G, Y119P, Y119S, R126C, R153H/N215P, R153L, W164R/W271T, G174D/L282V, G183P, E184F, E184G, E184P, E184S/A249T, E184Y, S185A, S185T, A186C, A186G, A186P, A186R, A186T, G187P, A188E, A188G, S190H, S190K, S190L, S190M, S190Q, S190R, S190W, L209E, L209G, L209P, L209S, L209V, Q210P, Q210T, Q210W, S211I, S211L, S211R, S211V, G212A, G212P, G212R, G212S, A213C, A213E, A213L, A213N, A213P, A213Q,

A213R/S345G, A213S, A213T/W271K, A213V, G214K, G214L, G214T, G214V, N215K, N215M, N215P, N215R, N215R/W271R, N215W, M216P, A217G, A217L, A217P, A217R, A217R/A231V, A217S, A217V, A217W, T224I/P268S/I372F, A249V/F284P, A269N, A269V, V270I, V270I/V470M, V270R, W271A, W271K, W271L, W271P, W271Q/A416V, W271S, W271T, A276F, G277M, G278H, G278S, S279C, S279E, S279G, S279L/V280G/L282M, S279V, V280E, V280G, V280S, L281P, L281V, L281Y/D374N, L282A, L282C, L282Q, L282R, L282S, L282T, L282W, P283C, P283D, P283K, P283R/A429V, P283T, P283V, P283Y, F284C, F284T, F284T/P438T, F284V, A285L, A285M, A285P, P286V, L311I, T317C, T317P, Y320A, Y320F, Y320G, Y320G/F323S, Y320S, Y320S/F323S/I372A, Y320S/I372A/V376G, Y320S/V376G/F377V, Y320W, R321L, R321S, F323C, F323I, F323R, F323Y, L324A, I372A/V376A, I372L, T373G, V376A, V376G, V376L, V376M, F377L, F377W, F377Y, P405D, P420G, D427A, R428V, and A429L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some embodiments, the engineered carboxyesterases comprise a substitution at position 282, wherein the position is numbered with reference to SEQ ID NO: 2. In some further embodiments, the substitution at position 282 is aliphatic, non-polar, basic, polar, or aromatic. In yet some additional embodiments, the substitution selected from: X282T, X282G, X282A, X282V, X282M, X282C, X282W, X282Q, X282S, X282T, and X282R.

[0114] The present invention also provides engineered carboxyesterases comprising a polypeptide sequences having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8 or a functional fragment thereof, wherein the engineered carboxyesterases comprises at least one substitution or substitution set in the polypeptide sequences, wherein the amino acid positions of the polypeptide sequences are numbered with reference to SEQ ID NO: 8. In some embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions at positions selected from: 63, 63/65/108, 63/65/108/189, 63/65/108/377, 63/65/282/285/320/323, 63/65/320/323, 63/108, 63/108/282/285/377, 63/108/285/377, 63/108/320/323, 63/189, 63/212/215, 63/212/215/268/269/343, 63/215, 63/215/269, 63/215/270/271, 63/215/343, 63/268/269/270/429, 63/377, 65/69/70/281/372, 65/69/70/372, 65/69/372, 65/70/372, 65/320, 65/320/323, 68, 68/69/189/214, 68/69/189/214/215, 68/69/189/214/215/271, 68/69/189/214/215/271/281/282/343/381, 68/69/189/214/271/280, 68/69/189/214/372, 68/69/189/214/377/381, 68/69/189/271, 68/69/189/271/280/372/381, 68/69/189/280/281/282/372/377, 68/69/189/281/282/372/377/381, 68/69/189/343/381, 68/69/189/372, 68/69/189/381, 68/69/214/215/271, 68/69/214/343, 68/69/215, 68/69/271, 68/69/282/287, 68/69/343/372, 68/108/377, 68/184, 68/184/189, 68/189/271/372, 68/189/343, 68/214/215/271/281/282/372, 68/215/271/343/372/381, 68/215/377, 68/271/372, 68/377, 69, 69/70, 69/70/331/372, 69/70/372, 69/70/459, 69/108, 69/108/270/372/377, 69/108/281/285, 69/110/215/281, 69/189, 69/189/214/271/281/282/343, 69/189/214/343/372, 69/189/215/343, 69/189/271, 69/189/271/281/282, 69/189/271/343, 69/189/271/343/381, 69/189/280/282/343/372/381, 69/189/281/373, 69/189/282, 69/189/343/381, 69/189/372, 69/189/372/377, 69/189/377, 69/212/213/215/280/281, 69/214/215/271/372/377/381, 69/214/271/282, 69/214/271/343, 69/215, 69/215/269/270/377, 69/215/271/280/281/282, 69/215/271/282, 69/215/271/372, 69/215/285/317, 69/215/323, 69/215/343, 69/215/343/372/381, 69/271, 69/271/372, 69/282, 69/282/343/372, 69/285/373, 69/372, 70, 70/212, 108, 108/189, 108/189/282/285/320, 108/189/320, 108/189/377, 108/215, 108/215/377, 108/269/270, 108/270, 108/282/285/377, 108/285, 108/320/323, 108/377, 126, 126/184/213/280/281/285/320, 126/184/213/372, 126/189/285/372, 126/215, 126/372, 181/215, 189, 189/214, 189/214/215/271/282, 189/215, 189/215/249/277, 189/215/271/281/282/377, 189/215/343/372, 189/270/285, 189/270/372, 189/280/282, 189/320/377, 189/343, 189/343/377, 189/372/377, 189/377, 189/381, 213/215/320, 214/215/271, 214/215/271/377, 214/271, 214/280/282/343/377/381, 215, 215/249/280/281/285/372, 215/271/372, 215/280/281/285/372, 215/281/285/372, 215/281/285/373, 215/281/373, 215/285, 215/285/317, 215/285/346, 215/285/445, 215/320, 215/320/372, 215/323, 215/372, 215/372/377, 215/373, 215/377, 215/381, 249/377, 269/270/281/372/377, 270/377, 271, 271/343, 271/343/372, 271/343/372/381, 280/285/372, 281/372, 282/285/320/323, 285/323, 320, 343/372, 372, 372/377, 372/381, 373, and 377, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8. In some further embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from 63A, 63A/189A, 63A/215R/343V, 63R, 63R/65G/108G, 63R/65G/108G/189L, 63R/65G/108G/377I, 63R/65G/282A/285L/320W/323I, 63R/65G/320W/323I, 63R/108G, 63R/108G/282A/285L/377L, 63R/108G/285L/377I, 63R/108G/320W/323C, 63R/377I, 63T/215R, 63Y, 63Y/189L, 63Y/212P/215R, 63Y/212P/215R/268A/269N/343V, 63Y/215P/269N, 63Y/215R, 63Y/215R/270I/271S, 63Y/268A/269N/270I/429V, 65G/320W, 65G/320W/323I, 65W/69L/372L, 65W/69M/70A/281P/372L, 65W/69W/70L/372L, 65W/70L/372M, 68P, 68P/69L/189E/214R/271Y/280G, 68P/69L/189E/214R/372L, 68P/69L/189I/214R/215P/271Y, 68P/69L/189I/281P/282C/372L/377Y/381L, 68P/69L/189Q/214R, 68P/69L/189Q/271Y/280G/372L/381L, 68P/69L/215P, 68P/69L/271Y, 68P/69L/282C/287I, 68P/69L/343V/372L, 68P/69W/189E/214R/215P/271Y/281P/282G/343V/381L, 68P/69W/189E/280G/281P/282A/372L/377Y, 68P/69W/189E/343V/381L, 68P/69W/189I/214R/215P, 68P/69W/189I/214R/377Y/381L, 68P/69W/189I/271Y, 68P/69W/189I/372L, 68P/69W/189I/381L, 68P/69W/214R/215P/271Y, 68P/69W/214R/343V, 68P/69W/215P, 68P/108G/377L, 68P/184S, 68P/184S/189E, 68P/189I/271Y/372L, 68P/189I/343V, 68P/214R/215P/271Y/281P/282A/372L, 68P/215P/271Y/343V/372L/381L, 68P/215P/377L, 68P/271Y/372L, 68P/377L, 69F/108G/270E/372L/377L, 69F/189L, 69F/215K, 69F/215K/269L/270I/377L, 69F/215R, 69F/285L/373G, 69L, 69L/70L/331Q/372M, 69L/189E/271Y/281P/282A, 69L/189I, 69L/189I/214R/271Y/281P/282A/343V,

69L/189I/271Y/343V/381L, 69L/189I/280G/282G/343V/372L/381L, 69L/189I/282A, 69L/189Q/377Y, 69L/215P/271Y/280G/281P/282C, 69L/215P/271Y/282A, 69L/215P/271Y/372L, 69L/215P/343V/372L/381L, 69L/215R/285P/317P, 69L/271Y, 69L/271Y/372L, 69L/282C/343V/372L, 69L/372L, 69M/70A/372M, 69W, 69W/70L, 69W/70L/372M, 69W/70L/459R, 69W/108S, 69W/189E/214R/343V/372L, 69W/189E/271Y/343V, 69W/189E/372L, 69W/189I, 69W/189I/215P/343V, 69W/189I/271Y, 69W/189I/343V/381L, 69W/189Q/372L/377Y, 69W/212A/213L/215R/280G/281P, 69W/214R/215P/271Y/372L/377Y/381L, 69W/214R/271Y/282A, 69W/214R/271Y/343V, 69W/215K/343V, 69W/215P, 69W/215R, 69W/215R/323Y, 69W/282A, 69W/372M, 69Y/108G/281P/285P, 69Y/110A/215R/281P, 69Y/189L/281P/373G, 70L, 70L/212P, 108G, 108G/189I/282A/285L/320W, 108G/189L, 108G/189L/320W, 108G/189L/377I, 108G/215K, 108G/215P/377L, 108G/269L/270E, 108G/270E, 108G/282A/285L/377L, 108G/285L, 108G/320W/323I, 108G/377I, 108G/377L, 126C, 126C/184S/213S/280G/281P/285L/320G, 126C/184S/213S/372L, 126C/189I/285L/372L, 126C/215P, 126C/372L, 181L/215P, 189E/372L/377Y, 189I, 189I/214R/215P/271Y/282G, 189I/215K, 189I/215P/343V/372L, 189I/215R/249T/277M, 189I/270E/285L, 189I/270E/372L, 189I/280G/282A, 189I/320W/377I, 189I/343V, 189I/377I, 189L, 189Q, 189Q/214R, 189Q/215P/271Y/281P/282C/377Y, 189Q/343V, 189Q/343V/377Y, 189Q/381L, 213S/215P/320G, 214R/215P/271Y, 214R/215P/271Y/377Y, 214R/271Y, 214R/280G/282A/343V/377Y/381L, 215K, 215K/281P/285L/372L, 215K/281P/373G, 215K/285L/317P, 215K/285L/445L, 215K/323Y, 215K/372L, 215K/372L/377L, 215K/373G, 215P, 215P/271Y/372L, 215P/320G, 215P/320G/372L, 215P/372L, 215P/372L/377L, 215P/377L, 215P/381L, 215R, 215R/249T/280G/281P/285L/372L, 215R/280G/281P/285L/372L, 215R/281P/285L/373G, 215R/285P, 215R/320G, 215R/372L, 215W, 215W/285L/346S, 215W/285P, 215W/373G, 249T/377L, 269L/270E/281P/372L/377L, 270E/377L, 271Y, 271Y/343V, 271Y/343V/372L, 271Y/343V/372L/381L, 280G/285L/372L, 281P/372L, 282A/285L/320W/323I, 285L/323I, 320W, 343V/372L, 372L, 372L/377L, 372L/381L, 372M, 373G, and 377L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8. In some further embodiments, at least one substitution or substitution set in the polypeptide sequence comprises substitutions selected from P63A, P63A/M189A, P63A/N215R/A343V, P63R, P63R/Y65G/A108G, P63R/Y65G/A108G/M189L, P63R/Y65G/A108G/F377I, P63R/Y65G/T282A/A285L/Y320W/F323I, P63R/Y65G/Y320W/F323I, P63R/A108G, P63R/A108G/T282A/A285L/F377L, P63R/A108G/A285L/F377I, P63R/A108G/Y320W/F323C, P63R/F377I, P63T/N215R, P63Y, P63Y/M189L, P63Y/G212P/N215R, P63Y/G212P/N215R/P268A/A269N/A343V, P63Y/N215P/A269N, P63Y/N215R, P63Y/N215R/V270I/W271S, P63Y/P268A/A269N/V270I/A429V, Y65G/Y320W, Y65G/Y320W/F323I, Y65W/I69L/I372L, Y65W/I69M/G70A/L281P/I372L, Y65W/I69W/G70L/I372L, Y65W/G70L/I372M, A68P, A68P/I69L/M189E/G214R/W271Y/V280G,

A68P/I69L/M189E/G214R/I372L, A68P/I69L/M189I/G214R/N215P/W271Y, A68P/I69L/M189I/L281P/T282C/I372L/F377Y/A381L, A68P/I69L/M189Q/G214R, A68P/I69L/M189Q/W271Y/V280G/I372L/A381L, A68P/I69L/N215P, A68P/I69L/W271Y, A68P/I69L/T282C/V287I, A68P/I69L/A343V/I372L, A68P/I69W/M189E/G214R/N215P/W271Y/L281P/T282G/A343V/A381L, A68P/I69W/M189E/V280G/L281P/T282A/I372L/F377Y, A68P/I69W/M189E/A343V/A381L, A68P/I69W/M189I/G214R/N215P, A68P/I69W/M189I/G214R/F377Y/A381L, A68P/I69W/M189I/W271Y, A68P/I69W/M189I/I372L, A68P/I69W/M189I/A381L, A68P/I69W/G214R/N215P/W271Y, A68P/I69W/G214R/A343V, A68P/I69W/N215P, A68P/A108G/F377L, A68P/E184S, A68P/E184S/M189E, A68P/M189I/W271Y/I372L, A68P/M189I/A343V, A68P/G214R/N215P/W271Y/L281P/T282A/I372L, A68P/N215P/W271Y/A343V/I372L/A381L, A68P/N215P/F377L, A68P/W271Y/I372L, A68P/F377L, I69F/A108G/V270E/I372L/F377L, I69F/M189L, I69F/N215K, I69F/N215K/A269L/V270I/F377L, I69F/N215R, I69F/A285L/T373G, I69L, I69L/G70L/P331Q/I372M, I69L/M189E/W271Y/L281P/T282A, I69L/M189I, I69L/M189I/G214R/W271Y/L281P/T282A/A343V, I69L/M189I/W271Y/A343V/A381L, I69L/M189I/V280G/T282G/A343V/I372L/A381L, I69L/M189I/T282A, I69L/M189Q/F377Y, I69L/N215P/W271Y/V280G/L281P/T282C, I69L/N215P/W271Y/T282A, I69L/N215P/W271Y/I372L, I69L/N215P/A343V/I372L/A381L, I69L/N215R/A285P/T317P, I69L/W271Y, I69L/W271Y/I372L, I69L/T282C/A343V/I372L, I69L/I372L, I69M/G70A/I372M, I69W, I69W/G70L, I69W/G70L/I372M, I69W/G70L/G459R, I69W/A108S, I69W/M189E/G214R/A343V/I372L, I69W/M189E/W271Y/A343V, I69W/M189E/I372L, I69W/M189I, I69W/M189I/N215P/A343V, I69W/M189I/W271Y, I69W/M189I/A343V/A381L, I69W/M189Q/I372L/F377Y, I69W/G212A/A213L/N215R/V280G/L281P, I69W/G214R/N215P/W271Y/I372L/F377Y/A381L, I69W/G214R/W271Y/T282A, I69W/G214R/W271Y/A343V, I69W/N215K/A343V, I69W/N215P, I69W/N215R, I69W/N215R/F323Y, I69W/T282A, I69W/I372M, I69Y/A108G/L281P/A285P, I69Y/T110A/N215R/L281P, I69Y/M189L/L281P/T373G, G70L, G70L/G212P, A108G, A108G/M189I/T282A/A285L/Y320W, A108G/M189L, A108G/M189L/Y320W, A108G/M189L/F377I, A108G/N215K, A108G/N215P/F377L, A108G/A269L/V270E, A108G/V270E, A108G/T282A/A285L/F377L, A108G/A285L, A108G/Y320W/F323I, A108G/F377I, A108G/F377L, R126C, R126C/E184S/A213S/V280G/L281P/A285L/Y320G, R126C/E184S/A213S/I372L, R126C/M189I/A285L/I372L, R126C/N215P, R126C/I372L, V181L/N215P, M189E/I372L/F377Y, M189I, M189I/G214R/N215P/W271Y/T282G, M189I/N215K, M189I/N215P/A343V/I372L, M189I/N215R/A249T/G277M, M189I/V270E/A285L, M189I/V270E/I372L, M189I/V280G/T282A, M189I/Y320W/F377I, M189I/A343V, M189I/F377I,

M189L, M189Q, M189Q/G214R, M189Q/N215P/W271Y/L281P/T282C/F377Y, M189Q/A343V, M189Q/A343V/F377Y, M189Q/A381L, A213S/N215P/Y320G, G214R/N215P/W271Y, G214R/N215P/W271Y/F377Y, G214R/W271Y, G214R/V280G/T282A/A343V/F377Y/A381L, N215K, N215K/L281P/A285L/I372L, N215K/L281P/T373G, N215K/A285L/T317P, N215K/A285L/V445L, N215K/F323Y, N215K/I372L, N215K/I372L/F377L, N215K/T373G, N215P, N215P/W271Y/I372L, N215P/Y320G, N215P/Y320G/I372L, N215P/I372L, N215P/I372L/F377L, N215P/F377L, N215P/A381L, N215R, N215R/A249T/V280G/L281P/A285L/I372L, N215R/V280G/L281P/A285L/I372L, N215R/L281P/A285L/T373G, N215R/A285P, N215R/Y320G, N215R/I372L, N215W, N215W/A285L/G346S, N215W/A285P, N215W/T373G, A249T/F377L, A269L/V270E/L281P/I372L/F377L, V270E/F377L, W271Y, W271Y/A343V, W271Y/A343V/I372L, W271Y/A343V/I372L/A381L, V280G/A285L/I372L, L281P/I372L, T282A/A285L/Y320W/F323I, A285L/F323I, Y320W, A343V/I372L, I372L, I372L/F377L, I372L/A381L, I372M, T373G, and F377L, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8.

[0115] In some embodiments, the present invention also provides engineered carboxyesterase polypeptides that comprise a fragment of any of the engineered carboxyesterase polypeptides described herein that retains the functional carboxyesterase activity and/or improved property of that engineered carboxyesterase polypeptide. Accordingly, in some embodiments, the present invention provides a polypeptide fragment having carboxyesterase activity (*e.g.*, capable of converting substrate to product under suitable reaction conditions), wherein the fragment comprises at least about 80%, 90%, 95%, 98%, or 99% of a full-length amino acid sequence of an engineered polypeptide of the present invention, such as an exemplary engineered polypeptide of having the even-numbered sequence identifiers of SEQ ID NOS: 2 - 136.

[0116] In some embodiments, the engineered carboxyesterase polypeptide of the invention comprises an amino acid sequence comprising at least one deletion, addition, and/or substitution, as compared to any one of the engineered carboxyesterase polypeptide sequences described herein, such as the exemplary engineered polypeptide sequences having the even-numbered sequence identifiers of SEQ ID NOS: 2 - 136. Thus, for each and every embodiment of the engineered carboxyesterase polypeptides of the invention, the amino acid sequence can comprise deletions, additions, and/or substitutions of one or more amino acids, 2 or more amino acids, 3 or more amino acids, 4 or more amino acids, 5 or more amino acids, 6 or more amino acids, 8 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, up to 10% of the total number of amino acids, up to 20% of the total number of amino acids, or up to 30% of the total number of amino acids of the carboxyesterase polypeptides, where the associated functional activity and/or improved properties of the engineered carboxyesterase described herein is maintained. In some embodiments, the deletions, additions, and/or substitutions can comprise 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-15, 1-20, 1-21, 1-22, 1-23, 1-24, 1-25, 1-30, 1-35, 1-40, 1-45, 1-50, 1-55, or 1-60 additions, and/or substitutions of the amino acid residues. In some embodiments, the number of deletions, additions, and/or substitutions can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 30, 35, 40, 45, 50, 55, or 60 of the amino acid residues. In some embodiments, the deletions, additions, and/or substitutions can comprise deletions, additions, and/or substitutions of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 24, 25 or 30 amino acid residues.

[0117] In some embodiments, the present invention provides an engineered carboxyesterase polypeptide having an amino acid sequence comprising an insertion as compared to any one of the engineered carboxyesterase polypeptide sequences described herein, such as the exemplary engineered polypeptide sequences having the even-numbered sequence identifiers of SEQ ID NO: 2 - 136. Thus, for each and every embodiment of the carboxyesterase polypeptides of the invention, the insertions can comprise one or more amino acids, 2 or more amino acids, 3 or more amino acids, 4 or more amino acids, 5 or more amino acids, 6 or more amino acids, 8 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, where the associated functional activity and/or improved properties of the engineered carboxyesterase described herein is maintained. The insertions can be to amino or carboxy terminus, or internal portions of the carboxyesterase polypeptide.

[0118] In some embodiments, the polypeptides of the present invention are in the form of fusion polypeptides in which the engineered polypeptides are fused to other polypeptides, such as, by way of example and not limitation, antibody tags (e.g., myc epitope), purification sequences (e.g., His tags for binding to metals), and cell localization signals (e.g., secretion signals). Thus, the engineered polypeptides described herein can be used with or without fusions to other polypeptides.

[0119] The engineered carboxyesterase polypeptides described herein are not restricted to the genetically encoded amino acids. Thus, in addition to the genetically encoded amino acids, the polypeptides described herein may be comprised, either in whole or in part, of naturally-occurring and/or synthetic non-encoded amino acids. Certain commonly encountered non-encoded amino acids of which the polypeptides described herein may be comprised include, but are not limited to: the D-stereoisomers of the genetically-encoded amino acids; 2,3-diaminopropionic acid (Dpr); α-aminoisobutyric acid (Aib); ε-aminohexanoic acid (Aha); δ-aminovaleric acid (Ava); N-methylglycine or sarcosine (MeGly or Sar); ornithine (Orn); citrulline (Cit); t-butylalanine (Bua); t-butylglycine (Bug); N-methylisoleucine (MeIle); phenylglycine (Phg); cyclohexylalanine (Cha); norleucine (Nle); naphthylalanine (Nal); 2-chlorophenylalanine (Ocf); 3-chlorophenylalanine (Mcf); 4-chlorophenylalanine (Pcf); 2-fluorophenylalanine (Off); 3-fluorophenylalanine (Mff); 4-fluorophenylalanine (Pff); 2-bromophenylalanine (Obf); 3-bromophenylalanine (Mbf); 4-bromophenylalanine (Pbf); 2-methylphenylalanine (Omf); 3-methylphenylalanine (Mmf); 4-

methylphenylalanine (Pmf); 2-nitrophenylalanine (Onf); 3-nitrophenylalanine (Mnf); 4-nitrophenylalanine (Pnf); 2-cyanophenylalanine (Ocf); 3-cyanophenylalanine (Mcf); 4-cyanophenylalanine (Pcf); 2-trifluoromethylphenylalanine (Otf); 3-trifluoromethylphenylalanine (Mtf); 4-trifluoromethylphenylalanine (Ptf); 4-aminophenylalanine (Paf); 4-iodophenylalanine (Pif); 4-aminomethylphenylalanine (Pamf); 2,4-dichlorophenylalanine (Opef); 3,4-dichlorophenylalanine (Mpcf); 2,4-difluorophenylalanine (Opff); 3,4-difluorophenylalanine (Mpff); pyrid-2-ylalanine (2pAla); pyrid-3-ylalanine (3pAla); pyrid-4-ylalanine (4pAla); naphth-1-ylalanine (1nAla); naphth-2-ylalanine (2nAla); thiazolylalanine (taAla); benzothienylalanine (bAla); thienylalanine (tAla); furylalanine (fAla); homophenylalanine (hPhe); homotyrosine (hTyr); homotryptophan (hTrp); pentafluorophenylalanine (5ff); styrylkalanine (sAla); authrylalanine (aAla); 3,3-diphenylalanine (Dfa); 3-amino-5-phenypentanoic acid (Afp); penicillamine (Pen); 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (Tic); β-2-thienylalanine (Thi); methionine sulfoxide (Mso); N(w)-nitroarginine (nArg); homolysine (hLys); phosphonomethylphenylalanine (pmPhe); phosphoserine (pSer); phosphothreonine (pThr); homoaspartic acid (hAsp); homoglutamic acid (hGlu); 1-aminocyclopent-(2 or 3)-ene-4 carboxylic acid; pipecolic acid (PA), azetidine-3-carboxylic acid (ACA); 1-aminocyclopentane-3-carboxylic acid; allylglycine (aOly); propargylglycine (pgGly); homoalanine (hAla); norvaline (nVal); homoleucine (hLeu), homovaline (hVal); homoisoleucine (hIle); homoarginine (hArg); N-acetyl lysine (AcLys); 2,4-diaminobutyric acid (Dbu); 2,3-diaminobutyric acid (Dab); N-methylvaline (MeVal); homocysteine (hCys); homoserine (hSer); hydroxyproline (Hyp) and homoproline (hPro). Additional non-encoded amino acids of which the polypeptides described herein may be comprised will be apparent to those of skill in the art. These amino acids may be in either the L- or D-configuration.

[0120] Those of skill in the art will recognize that amino acids or residues bearing side chain protecting groups may also comprise the polypeptides described herein. Non-limiting examples of such protected amino acids, which in this case belong to the aromatic category, include (protecting groups listed in parentheses), but are not limited to: Arg(tos), Cys(methylbenzyl), Cys (nitropyridinesulfenyl), Glu(δ-benzylester), Gln(xanthyl), Asn(N-δ-xanthyl), His(bom), His(benzyl), His(tos), Lys(fmoc), Lys(tos), Ser(O-benzyl), Thr (O-benzyl) and Tyr(O-benzyl).

[0121] Non-encoding amino acids that are conformationally constrained of which the polypeptides described herein may be composed include, but are not limited to, N-methyl amino acids (L-configuration); 1-aminocyclopent-(2 or 3)-ene-4-carboxylic acid; pipecolic acid; azetidine-3-carboxylic acid; homoproline (hPro); and 1-aminocyclopentane-3-carboxylic acid.

[0122] As will be apparent to the skilled artisan, the foregoing residue positions and the specific amino acid residues for each residue position can be used individually or in various combinations to synthesize carboxyesterase polypeptides having desired improved properties, including, among others, enzyme activity, substrate/product preference, stereoselectivity, substrate/product tolerance, and stability under various conditions, such as increased temperature, solvent, and/or pH.

[0123] The engineered carboxyesterase polypeptides of the present invention were generated by directed evolution of SEQ ID NO: 2 for efficient amidation of substrates of interest to products of interest, under certain industrially relevant conditions and have one or more residue differences as compared to a reference carboxyesterase polypeptide. These residue differences are associated with improvements in various enzyme properties, particularly increased activity, increased solvent tolerance, and reduced toxicity to host cells (*e.g.*, *E. coli*). In some additional embodiments, the variant carboxyesterases also exhibited increased stereoselectivity, increased stability, and tolerance of increased substrate and/or product concentration (*e.g.*, decreased product inhibition). Accordingly, in some embodiments, the engineered polypeptides having carboxyesterase activity are capable of converting the substrate compound(s) to product(s) with an activity that is increased at least about 1.2 fold, 1.5 fold, 2 fold, 3 fold, 4 fold, 5 fold, 10 fold, 20 fold, 30 fold, 40 fold, 50 fold, 100 fold, 200 fold, 500 fold, 1000 fold, or more relative to the activity of wild-type *T. fusca* carboxyesterase polypeptide (*e.g.*, SEQ ID NO: 2), under suitable reaction conditions. In some embodiments, the engineered polypeptides having carboxyesterase activity are capable of converting substrate to product with a percent conversion of at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99%, in a reaction time of about 48 h, about 36 h, about 24 h, or even a shorter length of time, under suitable reaction conditions. In some embodiments, the engineered polypeptides having carboxyesterase activity are capable of converting substrate to product diastereomeric excess of at least 90%, 95%, 97%, 98%, 99%, or greater, under suitable reaction conditions.

[0124] In some embodiments, the engineered polypeptides having carboxyesterase activity are capable of converting substrate to product with increased tolerance for the presence of the substrate relative to the substrate tolerance of a reference polypeptide (*e.g.*, SEQ ID NO: 2), under suitable reaction conditions. Accordingly, in some embodiments the engineered polypeptides are capable of converting the substrate of substrate to product in the presence of a substrate loading concentration of at least about 1 g/L, 5 g/L, 10 g/L, 20 g/L, about 30 g/L, about 40 g/L, about 50 g/L, about 70 g/L, about 75 g/L, about 100 g/L, with a percent conversion of at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 94%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99%, in a reaction time of about 72h, about 48h, about 36h, about 24 h, or even shorter length of time, under suitable reaction conditions.

[0125] Some suitable reaction conditions under which the above-described improved properties of the engineered polypeptides can be determined with respect to concentrations or amounts of polypeptide, substrate, buffer, co-solvent, pH, and/or conditions including temperature and reaction time are provided herein. In some embodiments, the suitable reaction conditions comprise the assay conditions described below and in the Examples.

[0126] In some embodiments, the engineered polypeptides can be provided on a solid support, such as a membrane, resin, solid carrier, or other solid phase material. A solid support can be composed of organic polymers such as polystyrene, polyethylene, polypropylene, polyfluoroethylene, polyethyleneoxy, and polyacrylamide, as well as co-polymers and grafts thereof. A solid support can also be inorganic, such as glass, silica, controlled pore glass (CPG), reverse phase silica or metal, such as gold or platinum. The configuration of a solid support can be in the form of beads, spheres, particles, granules, a gel, a membrane or a surface. Surfaces can be planar, substantially planar, or non-planar. Solid supports can be porous or non-porous, and can have swelling or non-swelling characteristics. A solid support can be configured in the form of a well, depression, or other container, vessel, feature, or location.

[0127] In some embodiments, the engineered polypeptides having carboxyesterase activity are bound or immobilized on the solid support such that they retain at least a portion of their improved properties relative to a reference polypeptide (*e.g.*, SEQ ID NO: 2). In such embodiments, the immobilized polypeptides can facilitate the biocatalytic conversion of the substrate compound to the desired product, and after the reaction is complete are easily retained (*e.g.*, by retaining beads on which polypeptide is immobilized) and then reused or recycled in subsequent reactions. Such immobilized enzyme processes allow for further efficiency and cost reduction. Accordingly, it is further contemplated that any of the methods of using the engineered carboxyesterase polypeptides of the present invention can be carried out using the same carboxyesterase polypeptides bound or immobilized on a solid support.

[0128] The engineered carboxyesterase polypeptide can be bound non-covalently or covalently. Various methods for conjugation and immobilization of enzymes to solid supports (*e.g.*, resins, membranes, beads, glass, *etc.*) are well known in the art. Other methods for conjugation and immobilization of enzymes to solid supports (*e.g.*, resins, membranes, beads, glass, *etc.*) are well known in the art (See, *e.g.*, Yi et al., Proc. Biochem., 42: 895-898 [2007]; Martin et al., Appl. Microbiol. Biotechnol., 76: 843-851 [2007]; Koszelewski et al., J. Mol. Cat. B: Enz., 63: 39-44 [2010]; Truppo et al., Org. Proc. Res. Develop., published online: dx.doi.org/10.1021/op200157c; and Mateo et al., Biotechnol. Prog., 18:629-34 [2002], *etc.*). Solid supports useful for immobilizing the engineered carboxyesterase polypeptides of the present invention include, but are not limited to, beads or resins comprising polymethacrylate with epoxide functional groups, polymethacrylate with amino epoxide functional groups, styrene/DVB copolymer or polymethacrylate with octadecyl functional groups. Exemplary solid supports useful for immobilizing the engineered carboxyesterases of the present invention include, but are not limited to, chitosan beads, Eupergit C, and SEPABEADs (Mitsubishi), including the following different types of SEPABEAD: EC-EP, EC-HFA/S, EXA252, EXE119 and EXE120.

[0129] In some embodiments, the engineered carboxyesterase polypeptides can be provided in the form of an array in which the polypeptides are arranged in positionally distinct locations. In some embodiments, the positionally distinct locations are wells in a solid support such as a 96-well plate. A plurality of supports can be configured on an array at various locations, addressable for robotic delivery of reagents, or by detection methods and/or instruments. Such arrays can be used to test a variety of substrate compounds for conversion by the polypeptides.

[0130] In some embodiments, the engineered polypeptides described herein can be provided in the form of kits. The polypeptides in the kits may be present individually or as a plurality of polypeptides. The kits can further include reagents for carrying out enzymatic reactions, substrates for assessing the activity of polypeptides, as well as reagents for detecting the products. The kits can also include reagent dispensers and instructions for use of the kits. In some embodiments, the kits of the present invention include arrays comprising a plurality of different engineered carboxyesterase polypeptides at different addressable position, wherein the different polypeptides are different variants of a reference sequence each having at least one different improved enzyme property. Such arrays comprising a plurality of engineered polypeptides and methods of their use are known (See, *e.g.*, WO2009/008908A2).

Polynucleotides Encoding Engineered Carboxyesterases

[0131] In another aspect, the present invention provides polynucleotides encoding the engineered carboxyesterase enzymes. The polynucleotides may be operatively linked to one or more heterologous regulatory sequences that control gene expression to create a recombinant polynucleotide capable of expressing the polypeptide. Expression constructs containing a heterologous polynucleotide encoding the engineered carboxyesterase can be introduced into appropriate host cells to express the corresponding carboxyesterase polypeptide.

[0132] Because of the knowledge of the codons corresponding to the various amino acids, availability of a protein sequence provides a description of all the polynucleotides capable of encoding the subject. The degeneracy of the genetic code, where the same amino acids are encoded by alternative or synonymous codons allows an extremely large number of nucleic acids to be made, all of which encode the improved carboxyesterase enzymes disclosed herein. Thus, having identified a particular amino acid sequence, those skilled in the art could make any number of different nucleic acids by simply modifying the sequence of one or more codons in a way which does not change the amino acid sequence of the protein. In this regard, the present invention specifically contemplates each and every possible variation of polynucleotides that could be made by selecting combinations based on the possible codon choices, and all such variations are to be considered specifically disclosed for any polypeptide disclosed herein, including the amino acid sequences presented in the Tables in the Examples. In various embodiments, the codons are preferably selected to fit the host cell in which the protein is being produced. For example, preferred codons used in bacteria are used to express the gene in bacteria; preferred codons used in yeast are used for expression in yeast; and preferred codons used in mammals are used for expression in mammalian cells. By way of example, the polynucleotide of SEQ ID NO: 1 has been codon optimized for expression in *E. coli*, but otherwise encodes the naturally occurring carboxyesterase of *T. fusca*.

[0133] In some embodiments, the polynucleotide encodes an engineered carboxyesterase polypeptide comprising an amino acid sequence selected from SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104,106, 108, 110, 112, 114, 116,118, 120, 122, 124, 126, 128, 130, 132, 134, or 136.

[0134] In some embodiments, the polynucleotides encoding the engineered carboxyesterases or a functional fragment thereof, are selected from polynucleotide sequences comprising at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, and/or 139.

[0135] In some embodiments, the polynucleotides are capable of hybridizing under highly stringent conditions to a polynucleotide comprising SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, and/or 139.

[0136] In some embodiments, the polynucleotides encode the polypeptides described herein but have about 80% or more sequence identity, about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more sequence identity at the nucleotide level to a reference polynucleotide encoding the engineered carboxyesterase. In some embodiments, the reference polynucleotide comprises SEQ ID NO: 1, while in some other embodiments, the reference polynucloeotide comprises SEQ ID NO:137. In some further embodiments, the reference polynucleotide sequence comprises SEQ ID NO:139. In some additional embodiments, the engineered carboxyesterase sequences comprise sequences that comprise positions identified to be beneficial, as described in the Examples.

[0137] The present invention also provides polynucleotide sequences encoding at least one engineered carboxyesterase provided herein. In some embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2 or a functional fragment thereof, wherein the engineered carboxyesterase comprises at least one substitution or substitution set in its polypeptide sequence, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising substitutions at positions selected from: 39, 39/323, 62, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 147, 153, 153/215, 164, 164/271, 174, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224, 224/268/372, 231, 249, 249/284, 263, 268, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 345, 349, 372, 372/376, 373, 374, 376, 377, 405, 416, 420, 427, 428, 429, 438, and 470, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising at least one substitution or substitution set selected from: 39/323, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 153, 153/215, 164/271, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224/268/372, 249/284, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 372, 372/376, 373, 376, 377, 405, 420, 427, 428, and 429, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 2. In some additional embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising at least one substitution or substitution set selected from: 39M/323I, 62H/117G, 63A, 63R, 63T, 63Y, 64A, 64E, 64G, 64I, 64T, 64V, 64W, 65G, 65S, 65T, 65W, 66N, 68L, 68P, 69F, 69G, 69H, 69L, 69V, 69W, 69Y, 70L, 70R, 70T, 70W, 71F, 71G, 71H/263R, 71P, 71R, 71V, 71Y, 77S/184G, 103P, 103R, 103T/147S, 104P, 104Q/429V, 105L, 107D/185W, 107L, 107P, 107S, 108G, 108K, 108Q, 108R, 108S, 108W, 109G/117M, 110A, 110H, 110P, 110S, 111L, 111M, 111R, 111S, 111V, 111W, 113P, 114A, 114H, 114Q, 115H, 115T, 115V, 117A, 117F, 118G/349V, 118I, 118N, 118N/269T, 119G, 119P, 119S, 126C, 153H/215P, 153L, 164R/271T, 174D/282V, 183P, 184F, 184G, 184P, 184S/249T, 184Y, 185A, 185T, 186C, 186G, 186P, 186R, 186T, 187P, 188E, 188G, 190H, 190K, 190L, 190M, 190Q, 190R, 190W, 209E, 209G, 209P, 209S, 209V, 210P, 210T, 210W, 211I, 211L, 211R, 211V, 212A, 212P, 212R, 212S, 213C, 213E, 213L, 213N, 213P, 213Q, 213R/345G, 213S, 213T/271K, 213V, 214K, 214L, 214T, 214V, 215K, 215M, 215P, 215R, 215R/271R, 215W, 216P, 217G, 217L, 217P, 217R, 217R/231V, 217S, 217V, 217W, 224I/268S/372F, 249V/284P, 269N, 269V, 270I, 270I/470M, 270R, 271A, 271K, 271L, 271P, 271Q/416V, 271S, 271T, 276F, 277M, 278H, 278S, 279C, 279E, 279G, 279L/280G/282M, 279V, 280E, 280G, 280S, 281P, 281V, 281Y/374N, 282A, 282C, 282Q, 282R, 282S, 282T, 282W, 283C, 283D, 283K, 283R/429V, 283T, 283V, 283Y, 284C, 284T, 284T/438T, 284V, 285L, 285M, 285P, 286V, 311I, 317C, 317P, 320A, 320F, 320G, 320G/323S, 320S, 320S/323S/372A, 320S/372A/376G, 320S/376G/377V, 320W, 321L, 321S, 323C, 323I, 323R, 323Y, 324A, 372A/376A, 372L, 373G, 376A, 376G, 376L, 376M, 377L, 377W, 377Y, 405D, 420G, 427A, 428V, and 429L, wherein the amino acid positions are numbered with reference to SEQ ID NO: 2. In some further embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising at least one substitution or substitution set selected from: T39M/F323I, R62H/P117G, P63A, P63R, P63T, P63Y, P64A, P64E, P64G, P64I, P64T, P64V, P64W, Y65G, Y65S, Y65T, Y65W, P66N, A68L, A68P, I69F, I69G, I69H, I69L, I69V, I69W, I69Y, G70L, G70R, G70T, G70W, A71F, A71G, A71H/Q263R, A71P, A71R, A71V, A71Y, F77S/E184G, W103P, W103R, W103T/P147S, I104P, I104Q/A429V, H105L, G107D/S185W, G107L, G107P, G107S, A108G, A108K, A108Q, A108R, A108S, A108W, F109G/P117M, T110A, T110H, T110P, T110S, N111L, N111M, N111R, N111S, N111V, N111W, S113P, G114A, G114H, G114Q, S115H, S115T, S115V, P117A, P117F, V118G/A349V, V118I, V118N, V118N/A269T, Y119G, Y119P, Y119S, R126C, R153H/N215P, R153L, W164R/W271T, G174D/L282V, G183P, E184F, E184G, E184P, E184S/A249T, E184Y, S185A, S185T, A186C, A186G, A186P, A186R, A186T, G187P, A188E, A188G, S190H, S190K, S190L, S190M, S190Q, S190R, S190W, L209E, L209G, L209P, L209S, L209V, Q210P, Q210T, Q210W, S211I, S211L, S211R, S211V, G212A, G212P, G212R, G212S, A213C, A213E, A213L, A213N, A213P, A213Q, A213R/S345G, A213S, A213T/W271K, A213V, G214K, G214L, G214T, G214V, N215K, N215M, N215P, N215R, N215R/W271R, N215W, M216P, A217G, A217L, A217P, A217R, A217R/A231V, A217S, A217V, A217W, T224I/P268S/I372F, A249V/F284P, A269N, A269V, V270I, V270I/V470M, V270R, W271A, W271K, W271L, W271P, W271Q/A416V, W271S, W271T, A276F, G277M, G278H, G278S, S279C, S279E, S279G, S279L/V280G/L282M, S279V, V280E, V280G, V280S, L281P, L281V, L281Y/D374N, L282A, L282C, L282Q, L282R, L282S, L282T, L282W, P283C, P283D, P283K, P283R/A429V, P283T, P283V, P283Y, F284C, F284T, F284T/P438T, F284V, A285L, A285M, A285P, P286V, L311I, T317C, T317P, Y320A, Y320F, Y320G, Y320G/F323S, Y320S, Y320S/F323S/I372A, Y320S/I372A/V376G, Y320S/V376G/F377V, Y320W, R321L, R321S, F323C, F323I, F323R, F323Y, L324A, I372A/V376A, I372L, T373G, V376A, V376G, V376L, V376M, F377L, F377W, F377Y, P405D, P420G, D427A, R428V, and A429L, wherein the amino acids are numbered with reference to SEQ ID NO: 2. In some embodiments, the polynucleotide sequences encode at least one engineered carboxyesterase comprising a substitution at position 282, wherein the position is numbered with reference to SEQ ID NO: 2. In some further embodiments, the substitution at position 282 is aliphatic, non-polar, basic, polar, or aromatic. In yet some additional embodiments, the substitution is selected from X282T, X282G, X282A, X282V, X282M, X282C, X282W, X282Q, X282S, X282T, and X282R.

[0138] The present invention also provides polynucleotide sequences encoding at least one engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8 or a functional fragment thereof, wherein the engineered carboxyesterase comprises at least one substitution or substitution set in its polypeptide sequence, wherein the amino acid positions of the polypeptide sequence are numbered with reference to SEQ ID NO: 8. In some embodiments, the polynucleotide sequences encode engineered carboxyesterases comprising at least one substitution or substitution set at positions selected from: 63, 63/65/108, 63/65/108/189, 63/65/108/377, 63/65/282/285/320/323, 63/65/320/323, 63/108, 63/108/282/285/377, 63/108/285/377, 63/108/320/323, 63/189, 63/212/215, 63/212/215/268/269/343, 63/215, 63/215/269, 63/215/270/271, 63/215/343, 63/268/269/270/429, 63/377, 65/69/70/281/372, 65/69/70/372, 65/69/372, 65/70/372, 65/320, 65/320/323, 68, 68/69/189/214, 68/69/189/214/215, 68/69/189/214/215/271, 68/69/189/214/215/271/281/282/343/381, 68/69/189/214/271/280, 68/69/189/214/372, 68/69/189/214/377/381, 68/69/189/271, 68/69/189/271/280/372/381, 68/69/189/280/281/282/372/377, 68/69/189/281/282/372/377/381, 68/69/189/343/381, 68/69/189/372, 68/69/189/381, 68/69/214/215/271, 68/69/214/343, 68/69/215, 68/69/271, 68/69/282/287, 68/69/343/372, 68/108/377, 68/184, 68/184/189, 68/189/271/372, 68/189/343, 68/214/215/271/281/282/372, 68/215/271/343/372/381, 68/215/377, 68/271/372, 68/377, 69, 69/70, 69/70/331/372, 69/70/372, 69/70/459, 69/108, 69/108/270/372/377, 69/108/281/285, 69/110/215/281, 69/189, 69/189/214/271/281/282/343, 69/189/214/343/372, 69/189/215/343, 69/189/271, 69/189/271/281/282, 69/189/271/343, 69/189/271/343/381, 69/189/280/282/343/372/381, 69/189/281/373, 69/189/282, 69/189/343/381, 69/189/372, 69/189/372/377, 69/189/377, 69/212/213/215/280/281, 69/214/215/271/372/377/381, 69/214/271/282, 69/214/271/343, 69/215, 69/215/269/270/377, 69/215/271/280/281/282, 69/215/271/282, 69/215/271/372, 69/215/285/317, 69/215/323, 69/215/343, 69/215/343/372/381, 69/271, 69/271/372, 69/282, 69/282/343/372, 69/285/373, 69/372, 70, 70/212, 108, 108/189, 108/189/282/285/320, 108/189/320, 108/189/377, 108/215, 108/215/377, 108/269/270, 108/270, 108/282/285/377, 108/285, 108/320/323, 108/377, 126, 126/184/213/280/281/285/320, 126/184/213/372, 126/189/285/372, 126/215, 126/372, 181/215, 189, 189/214, 189/214/215/271/282, 189/215, 189/215/249/277, 189/215/271/281/282/377, 189/215/343/372, 189/270/285, 189/270/372, 189/280/282, 189/320/377, 189/343, 189/343/377, 189/372/377, 189/377, 189/381, 213/215/320, 214/215/271, 214/215/271/377, 214/271, 214/280/282/343/377/381, 215, 215/249/280/281/285/372, 215/271/372, 215/280/281/285/372, 215/281/285/372, 215/281/285/373, 215/281/373, 215/285, 215/285/317, 215/285/346, 215/285/445, 215/320, 215/320/372, 215/323, 215/372, 215/372/377, 215/373, 215/377, 215/381, 249/377, 269/270/281/372/377, 270/377, 271, 271/343, 271/343/372, 271/343/372/381, 280/285/372, 281/372, 282/285/320/323, 285/323, 320, 343/372, 372, 372/377, 372/381, 373, and 377, wherein the amino acid positions are numbered with reference to SEQ ID NO: 8. In some further embodiments, the polynucleotide sequence encodes an engineered carboxyesterase comprising at least one substitution or substitution set selected from: 63A, 63A/189A, 63A/215R/343V, 63R, 63R/65G/108G, 63R/65G/108G/189L, 63R/65G/108G/377I, 63R/65G/282A/285L/320W/323I, 63R/65G/320W/323I, 63R/108G, 63R/108G/282A/285L/377L, 63R/108G/285L/377I, 63R/108G/320W/323C, 63R/377I, 63T/215R, 63Y, 63Y/189L, 63Y/212P/215R, 63Y/212P/215R/268A/269N/343V, 63Y/215P/269N, 63Y/215R, 63Y/215R/270I/271S, 63Y/268A/269N/270I/429V, 65G/320W, 65G/320W/323I, 65W/69L/372L, 65W/69M/70A/281P/372L, 65W/69W/70L/372L, 65W/70L/372M, 68P, 68P/69L/189E/214R/271Y/280G, 68P/69L/189E/214R/372L, 68P/69L/189I/214R/215P/271Y, 68P/69L/189I/281P/282C/372L/377Y/381L, 68P/69L/189Q/214R, 68P/69L/189Q/271Y/280G/372L/381L, 68P/69L/215P, 68P/69L/271Y, 68P/69L/282C/287I, 68P/69L/343V/372L, 68P/69W/189E/214R/215P/271Y/281P/282G/343V/381L, 68P/69W/189E/280G/281P/282A/372L/377Y, 68P/69W/189E/343V/381L, 68P/69W/189I/214R/215P, 68P/69W/189I/214R/377Y/381L, 68P/69W/189I/271Y, 68P/69W/189I/372L, 68P/69W/189I/381L, 68P/69W/214R/215P/271Y, 68P/69W/214R/343V, 68P/69W/215P, 68P/108G/377L, 68P/184S, 68P/184S/189E, 68P/189I/271Y/372L, 68P/189I/343V, 68P/214R/215P/271Y/281P/282A/372L, 68P/215P/271Y/343V/372L/381L, 68P/215P/377L, 68P/271Y/372L, 68P/377L, 69F/108G/270E/372L/377L, 69F/189L, 69F/215K, 69F/215K/269L/270I/377L, 69F/215R, 69F/285L/373G, 69L, 69L/70L/331Q/372M, 69L/189E/271Y/281P/282A, 69L/189I, 69L/189I/214R/271Y/281P/282A/343V, 69L/189I/271Y/343V/381L, 69L/189I/280G/282G/343V/372L/381L, 69L/189I/282A, 69L/189Q/377Y, 69L/215P/271Y/280G/281P/282C, 69L/215P/271Y/282A, 69L/215P/271Y/372L, 69L/215P/343V/372L/381L, 69L/215R/285P/317P, 69L/271Y, 69L/271Y/372L, 69L/282C/343V/372L, 69L/372L, 69M/70A/372M, 69W, 69W/70L, 69W/70L/372M, 69W/70L/459R, 69W/108S, 69W/189E/214R/343V/372L, 69W/189E/271Y/343V, 69W/189E/372L, 69W/189I, 69W/189I/215P/343V, 69W/189I/271Y, 69W/189I/343V/381L, 69W/189Q/372L/377Y, 69W/212A/213L/215R/280G/281P, 69W/214R/215P/271Y/372L/377Y/381L, 69W/214R/271Y/282A, 69W/214R/271Y/343V, 69W/215K/343V, 69W/215P, 69W/215R, 69W/215R/323Y, 69W/282A, 69W/372M, 69Y/108G/281P/285P, 69Y/110A/215R/281P, 69Y/189L/281P/373G, 70L, 70L/212P, 108G, 108G/189I/282A/285L/320W, 108G/189L, 108G/189L/320W, 108G/189L/377I, 108G/215K, 108G/215P/377L, 108G/269L/270E, 108G/270E, 108G/282A/285L/377L, 108G/285L, 108G/320W/323I, 108G/377I, 108G/377L, 126C,

126C/184S/213S/280G/281P/285L/320G, 126C/184S/213S/372L, 126C/189I/285L/372L, 126C/215P, 126C/372L, 181L/215P, 189E/372L/377Y, 189I, 189I/214R/215P/271Y/282G, 189I/215K, 189I/215P/343V/372L, 189I/215R/249T/277M, 189I/270E/285L, 189I/270E/372L, 189I/280G/282A, 189I/320W/377I, 189I/343V, 189I/377I, 189L, 189Q, 189Q/214R, 189Q/215P/271Y/281P/282C/377Y, 189Q/343V, 189Q/343V/377Y, 189Q/381L, 213S/215P/320G, 214R/215P/271Y, 214R/215P/271Y/377Y, 214R/271Y, 214R/280G/282A/343V/377Y/381L, 215K, 215K/281P/285L/372L, 215K/281P/373G, 215K/285L/317P, 215K/285L/445L, 215K/323Y, 215K/372L, 215K/372L/377L, 215K/373G, 215P, 215P/271Y/372L, 215P/320G, 215P/320G/372L, 215P/372L, 215P/372L/377L, 215P/377L, 215P/381L, 215R, 215R/249T/280G/281P/285L/372L, 215R/280G/281P/285L/372L, 215R/281P/285L/373G, 215R/285P, 215R/320G, 215R/372L, 215W, 215W/285L/346S, 215W/285P, 215W/373G, 249T/377L, 269L/270E/281P/372L/377L, 270E/377L, 271Y, 271Y/343V, 271Y/343V/372L, 271Y/343V/372L/381L, 280G/285L/372L, 281P/372L, 282A/285L/320W/323I, 285L/323I, 320W, 343V/372L, 372L, 372L/377L, 372L/381L, 372M, 373G, and 377L, wherein the amino acid positions are numbered with reference to SEQ ID NO: 8. In some further embodiments, the polynucleotide sequence encodes an engineered carboxyesterase comprising at least one substitution or substitution set selected from P63A, P63A/M189A, P63A/N215R/A343V, P63R, P63R/Y65G/A108G, P63R/Y65G/A108G/M189L, P63R/Y65G/A108G/F377I, P63R/Y65G/T282A/A285L/Y320W/F323I, P63R/Y65G/Y320W/F323I, P63R/A108G, P63R/A108G/T282A/A285L/F377L, P63R/A108G/A285L/F377I, P63R/A108G/Y320W/F323C, P63R/F377I, P63T/N215R, P63Y, P63Y/M189L, P63Y/G212P/N215R, P63Y/G212P/N215R/P268A/A269N/A343V, P63Y/N215P/A269N, P63Y/N215R, P63Y/N215R/V270I/W271S, P63Y/P268A/A269N/V270I/A429V, Y65G/Y320W, Y65G/Y320W/F323I, Y65W/I69L/I372L, Y65W/I69M/G70A/L281P/I372L, Y65W/I69W/G70L/I372L, Y65W/G70L/I372M, A68P, A68P/I69L/M189E/G214R/W271Y/V280G, A68P/I69L/M189E/G214R/I372L, A68P/I69L/M189I/G214R/N215P/W271Y, A68P/I69L/M189I/L281P/T282C/I372L/F377Y/A381L, A68P/I69L/M189Q/G214R, A68P/I69L/M189Q/W271Y/V280G/I372L/A381L, A68P/I69L/N215P, A68P/I69L/W271Y, A68P/I69L/T282C/V287I, A68P/I69L/A343V/I372L, A68P/I69W/M189E/G214R/N215P/W271Y/L281P/T282G/A343V/A381L, A68P/I69W/M189E/V280G/L281P/T282A/I372L/F377Y, A68P/I69W/M189E/A343V/A381L, A68P/I69W/M189I/G214R/N215P, A68P/I69W/M189I/G214R/F377Y/A381L, A68P/I69W/M189I/W271Y, A68P/I69W/M189I/I372L, A68P/I69W/M189I/A381L, A68P/I69W/G214R/N215P/W271Y, A68P/I69W/G214R/A343V, A68P/I69W/N215P, A68P/A108G/F377L, A68P/E184S, A68P/E184S/M189E, A68P/M189I/W271Y/I372L, A68P/M189I/A343V, A68P/G214R/N215P/W271Y/L281P/T282A/I372L, A68P/N215P/W271Y/A343V/I372L/A381L, A68P/N215P/F377L, A68P/W271Y/I372L,

A68P/F377L, I69F/A108G/V270E/I372L/F377L, I69F/M189L, I69F/N215K, I69F/N215K/A269L/V270I/F377L, I69F/N215R, I69F/A285L/T373G, I69L, I69L/G70L/P331Q/I372M, I69L/M189E/W271Y/L281P/T282A, I69L/M189I, I69L/M189I/G214R/W271Y/L281P/T282A/A343V, I69L/M189I/W271Y/A343V/A381L, I69L/M189I/V280G/T282G/A343V/I372L/A381L, I69L/M189I/T282A, I69L/M189Q/F377Y, I69L/N215P/W271Y/V280G/L281P/T282C, I69L/N215P/W271Y/T282A, I69L/N215P/W271Y/I372L, I69L/N215P/A343V/I372L/A381L, I69L/N215R/A285P/T317P, I69L/W271Y, I69L/W271Y/I372L, I69L/T282C/A343V/I372L, I69L/I372L, I69M/G70A/I372M, I69W, I69W/G70L, I69W/G70L/I372M, I69W/G70L/G459R, I69W/A108S, I69W/M189E/G214R/A343V/I372L, I69W/M189E/W271Y/A343V, I69W/M189E/I372L, I69W/M189I, I69W/M189I/N215P/A343V, I69W/M189I/W271Y, I69W/M189I/A343V/A381L, I69W/M189Q/I372L/F377Y, I69W/G212A/A213L/N215R/V280G/L281P, I69W/G214R/N215P/W271Y/I372L/F377Y/A381L, I69W/G214R/W271Y/T282A, I69W/G214R/W271Y/A343V, I69W/N215K/A343V, I69W/N215P, I69W/N215R, I69W/N215R/F323Y, I69W/T282A, I69W/I372M, I69Y/A108G/L281P/A285P, I69Y/T110A/N215R/L281P, I69Y/M189L/L281P/T373G, G70L, G70L/G212P, A108G, A108G/M189I/T282A/A285L/Y320W, A108G/M189L, A108G/M189L/Y320W, A108G/M189L/F377I, A108G/N215K, A108G/N215P/F377L, A108G/A269L/V270E, A108G/V270E, A108G/T282A/A285L/F377L, A108G/A285L, A108G/Y320W/F323I, A108G/F377I, A108G/F377L, R126C, R126C/E184S/A213S/V280G/L281P/A285L/Y320G, R126C/E184S/A213S/I372L, R126C/M189I/A285L/I372L, R126C/N215P, R126C/I372L, V181L/N215P, M189E/I372L/F377Y, M189I, M189I/G214R/N215P/W271Y/T282G, M189I/N215K, M189I/N215P/A343V/I372L, M189I/N215R/A249T/G277M, M189I/V270E/A285L, M189I/V270E/I372L, M189I/V280G/T282A, M189I/Y320W/F377I, M189I/A343V, M189I/F377I, M189L, M189Q, M189Q/G214R, M189Q/N215P/W271Y/L281P/T282C/F377Y, M189Q/A343V, M189Q/A343V/F377Y, M189Q/A381L, A213S/N215P/Y320G, G214R/N215P/W271Y, G214R/N215P/W271Y/F377Y, G214R/W271Y, G214R/V280G/T282A/A343V/F377Y/A381L, N215K, N215K/L281P/A285L/I372L, N215K/L281P/T373G, N215K/A285L/T317P, N215K/A285L/V445L, N215K/F323Y, N215K/I372L, N215K/I372L/F377L, N215K/T373G, N215P, N215P/W271Y/I372L, N215P/Y320G, N215P/Y320G/I372L, N215P/I372L, N215P/I372L/F377L, N215P/F377L, N215P/A381L, N215R, N215R/A249T/V280G/L281P/A285L/I372L, N215R/V280G/L281P/A285L/I372L, N215R/L281P/A285L/T373G, N215R/A285P, N215R/Y320G, N215R/I372L, N215W, N215W/A285L/G346S, N215W/A285P, N215W/T373G, A249T/F377L, A269L/V270E/L281P/I372L/F377L, V270E/F377L, W271Y, W271Y/A343V, W271Y/A343V/I372L, W271Y/A343V/I372L/A381L, V280G/A285L/I372L, L281P/I372L, T282A/A285L/Y320W/F323I, A285L/F323I, Y320W, A343V/I372L, I372L, I372L/F377L, I372L/A381L, I372M, T373G, and F377L, wherein the amino acid positions are numbered with reference to SEQ ID NO: 8.

[0139] An isolated polynucleotide encoding an improved carboxyesterase polypeptide may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the isolated polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides and nucleic acid sequences utilizing recombinant DNA methods are well known in the art.

[0140] For bacterial host cells, suitable promoters for directing transcription of the nucleic acid constructs of the present invention, include the promoters obtained from the *E. coli* lac operon, *Streptomyces coelicolor* agarase gene (dagA), *Bacillus subtilis* levansucrase gene (sacB), *Bacillus licheniformis* alpha-amylase gene (amyL), *Bacillus stearothermophilus* maltogenic amylase gene (amyM), *Bacillus amyloliquefaciens* alpha-amylase gene (amyQ), *Bacillus licheniformis* penicillinase gene (penP), *Bacillus subtilis* xylA and xylB genes, and prokaryotic beta-lactamase gene (See, *e.g.*, Villa-Kamaroff et al., Proc. Natl. Acad. Sci. USA 75: 3727-3731 [1978]), as well as the *tac* promoter (See, *e.g.*, DeBoer et al., Proc. Natl Acad. Sci. USA 80: 21-25 [1983]). Additional suitable promoters are known to those in the art.

[0141] For filamentous fungal host cells, suitable promoters for directing the transcription of the nucleic acid constructs of the present invention include promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (glaA), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

[0142] In a yeast host, useful promoters include, but are not limited to those from the genes for Saccharomyces cerevisiae enolase (ENO-1), Saccharomyces cerevisiae galactokinase (GAL1), Saccharomyces cerevisiae alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP), and Saccharomyces cerevisiae 3-phosphoglycerate kinase, as well as other useful promoters for yeast host cells (See, e.g., Romanos, et al., Yeast 8:423-488 [1992]).

[0143] The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleic acid sequence encoding the polypeptide. Any terminator that is functional in the host cell of choice may be used in the present invention.

[0144] For example, exemplary transcription terminators for filamentous fungal host cells can be obtained from the genes for Aspergillus oryzae TAKA amylase, Aspergillus niger glucoamylase, Aspergillus nidulans anthranilate synthase, Aspergillus niger alpha-glucosidase, and Fusarium oxysporum trypsin-like protease.

[0145] Exemplary terminators for yeast host cells can be obtained from the genes for Saccharomyces cerevisiae enolase, Saccharomyces cerevisiae cytochrome C (CYC1), and Saccharomyces cerevisiae glyceraldehyde-3-phosphate dehydrogenase, as well as other useful terminators for yeast host cells known in the art (See, e.g,. Romanos et al., supra).

[0146] The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA that is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleic acid sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used. Exemplary leaders for filamentous fungal host cells are obtained from the genes for Aspergillus oryzae TAKA amylase and Aspergillus nidulans triose phosphate isomerase. Suitable leaders for yeast host cells are obtained from the genes for Saccharomyces cerevisiae enolase (ENO-1), Saccharomyces cerevisiae 3-phosphoglycerate kinase, Saccharomyces cerevisiae alpha-factor, and Saccharomyces cerevisiae alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

[0147] The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3' terminus of the nucleic acid sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention. Exemplary polyadenylation sequences for filamentous fungal host cells can be from the genes for Aspergillus oryzae TAKA amylase, Aspergillus niger glucoamylase, Aspergillus nidulans anthranilate synthase, Fusarium oxysporum trypsin-like protease, and Aspergillus niger alpha-glucosidase., as well as additional useful polyadenylation sequences for yeast host cells known in the art (See, e.g., Guo et al., Mol. Cell. Biol., 15:5983-5990 [1995]).

[0148] The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region that encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region that is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region.

[0149] Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

[0150] Effective signal peptide coding regions for bacterial host cells are the signal peptide coding regions obtained from the genes for Bacillus NClB 11837 maltogenic amylase, Bacillus stearothermophilus alpha-amylase, Bacillus licheniformis subtilisin, Bacillus licheniformis beta-lactamase, Bacillus stearothermophilus neutral proteases (nprT, nprS, nprM), and Bacillus subtilis prsA,, as well as additional signal peptides known in the art (See, e.g., Simonen et al., Microbiol. Rev., 57: 109-137 [1993]).

[0151] Effective signal peptide coding regions for filamentous fungal host cells include, but are not limited to the signal peptide coding regions obtained from the genes for Aspergillus oryzae TAKA amylase, Aspergillus niger neutral amylase, Aspergillus niger glucoamylase, Rhizomucor miehei aspartic proteinase, Humicola insolens cellulase, and Humicola lanuginosa lipase. Useful signal peptides for yeast host cells can be from the genes for Saccharomyces cerevisiae alpha-factor and Saccharomyces cerevisiae invertase, as well as additional useful signal peptide coding regions (See, e.g., Romanos et al., 1992, supra).

[0152] The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the genes for Bacillus subtilis alkaline protease (aprE), Bacillus subtilis neutral protease (nprT), Saccharomyces cerevisiae alpha-factor, Rhizomucor miehei aspartic proteinase, and Myceliophthora thermophila lactase (WO 95/33836).

[0153] Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

[0154] It may also be desirable to add regulatory sequences, which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. In prokaryotic host cells, suitable regulatory sequences include the lac, tac, and trp operator systems. In yeast host cells, suitable regulatory systems include, as examples, the ADH2 system or GAL1 system. In filamentous fungi, suitable regulatory sequences include the TAKA alpha-amylase promoter, Aspergillus niger glucoamylase promoter, and Aspergillus oryzae glucoamylase promoter.

[0155] Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene, which is amplified in the presence of methotrexate, and the metallothionein genes, which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the carboxyesterase polypeptide of the present invention would be operably linked with the regulatory sequence.

[0156] Thus, in some embodiments, the present invention is also directed to a recombinant expression vector comprising a polynucleotide encoding an engineered carboxyesterase polypeptide or a variant thereof, and one or more expression regulating regions such as a promoter and a terminator, a replication origin, etc., depending on the type of hosts into which they are to be introduced. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present invention may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0157] The recombinant expression vector may be any vector (e.g., a plasmid or virus), which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the polynucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

[0158] The expression vector may be an autonomously replicating vector (i.e., a vector that exists as an extrachromosomal entity), the replication of which is independent of chromosomal replication, (e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome). The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

[0159] The expression vector of the present invention preferably contains one or more selectable markers, which permit easy selection of transformed cells. A selectable marker can be a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like. Examples of bacterial selectable markers are the dal genes from Bacillus subtilis or Bacillus licheniformis, or markers, which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol, or tetracycline resistance. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3.

[0160] Selectable markers for use in a filamentous fungal host cell include, but are not limited to, amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hph (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase), sC (sulfate adenyltransferase), and trpC (anthranilate synthase), as well as equivalents thereof. Embodiments for use in an Aspergillus cell include the amdS and pyrG genes of Aspergillus nidulans or Aspergillus oryzae and the bar gene of Streptomyces hygroscopicus.

[0161] The expression vectors of the present invention can contain an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome. For integration into the host cell genome, the vector may rely on the nucleic acid sequence encoding the polypeptide or any other element of the vector for integration of the vector into the genome by homologous or nonhomologous recombination.

[0162] Alternatively, the expression vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, preferably 400 to 10,000 base pairs, and most preferably 800 to 10,000 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[0163] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Non-limiting examples of bacterial origins of replication are P15A ori or the origins of replication of plasmids pBR322, pUC19, pACYCl77 (which plasmid has the P15A ori), or pACYC184 permitting replication in *E. coli*, and pUB110, pE194, or pTA1060, permitting replication in *Bacillus*. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes it's functioning temperature-sensitive in the host cell (See, *e.g.*, Ehrlich, Proc. Natl. Acad. Sci. USA 75:1433 [1978]).

[0164] More than one copy of a nucleic acid sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleic acid sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0165] Many of the expression vectors for use in the present invention are commercially available. Suitable commercial expression vectors include, but are not limited to p3xFLAGTM expression vectors (Sigma-Aldrich), which include a CMV promoter and hGH polyadenylation site for expression in mammalian host cells and a pBR322 origin of replication and ampicillin resistance markers for amplification in *E. coli*. Other commercially available suitable expression vectors include but are not limited to the pBluescriptII SK(-) and pBK-CMV vectors (Stratagene), and plasmids derived from pBR322 (Gibco BRL), pUC (Gibco BRL), pREP4, pCEP4 (Invitrogen) or pPoly (See, Lathe et al., Gene 57:193-201 [1987]).

[0166] The skilled person will appreciate that, upon production of an enzyme, in particular, depending upon the cell line used and the particular amino acid sequence of the enzyme, post-translational modifications may occur. For example, such post-translational modifications may include the cleavage of certain leader sequences, the addition of various sugar moieties in various glycosylation and phosphorylation patterns, deamidation, oxidation, disulfide bond scrambling, isomerisation, C-terminal lysine clipping, and N-terminal glutamine cyclisation. The present invention encompasses the use of engineered carboxyesterase enzymes that have been subjected to, or have undergone, one or more post-translational modifications. Thus, the engineered carboxyesterases of the invention includes one which has undergone a post-translational modification, such as described herein.

[0167] Deamidation is an enzymatic reaction primarily converting asparagine (N) to iso-aspartic acid (iso-aspartate) and aspartic acid (aspartate) (D) at approximately 3:1 ratio. This deamidation reaction is, therefore, related to isomerization of aspartate (D) to iso-aspartate. The deamidation of asparagine and the isomerisation of aspartate, both involve the intermediate succinimide. To a much lesser degree, deamidation can occur with glutamine residues in a similar manner.

[0168] Oxidation can occur during production and storage (*i.e.,* in the presence of oxidizing conditions) and results in a covalent modification of a protein, induced either directly by reactive oxygen species, or indirectly by reaction with secondary by-products of oxidative stress. Oxidation happens primarily with methionine residues, but may occur at tryptophan and free cysteine residues.

[0169] Disulfide bond scrambling can occur during production and basic storage conditions. Under certain circumstances, disulfide bonds can break or form incorrectly, resulting in unpaired cysteine residues (-SH). These free (unpaired) sulfhydryls (-SH) can promote shuffling.

[0170] N-terminal glutamine (Q) and glutamate (glutamic acid) (E) in the engineered carboxyesterases are likely to form pyroglutamate (pGlu) *via* cyclization. Most pGlu formation happens in manufacturing, but it can be formed non-enzymatically, depending upon pH and temperature of processing and storage conditions.

[0171] C-terminal lysine clipping is an enzymatic reaction catalyzed by carboxypeptidases, and is commonly observed in enzymes. Variants of this process include removal of lysine from the enzymes from the recombinant host cell.

[0172] In the present invention, the post-translational modifications and changes in primary amino acid sequence described above do not result in significant changes in the activity of the engineered carboxyesterase enzymes.

Host Cells for Expression of Carboxyesterase Polypeptides

[0173] In another aspect, the present invention provides a host cell comprising a polynucleotide encoding an improved carboxyesterase polypeptide of the present invention, the polynucleotide being operatively linked to one or more control sequences for expression of the carboxyesterase enzyme in the host cell. Host cells for use in expressing the carboxyesterase polypeptides encoded by the expression vectors of the present invention are well known in the art and include but are not limited to, bacterial cells, such as E. coli, Geobacillus stearothermophilus, Lactobacillus kefir, Lactobacillus brevis, Lactobacillus minor, Mycobacterium tuberculosis, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells (e.g., Saccharomyces cerevisiae or Pichia pastoris (ATCC Accession No. 201178)); insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, BHK, 293, and Bowes melanoma cells; and plant cells. Appropriate culture media and growth conditions for the above-described host cells are well known in the art.

[0174] Polynucleotides for expression of the carboxyesterase may be introduced into cells by various methods known in the art. Techniques include among others, electroporation, biolistic particle bombardment, liposome mediated transfection, calcium chloride transfection, and protoplast fusion. Various methods for introducing polynucleotides into cells will be apparent to the skilled artisan.

[0175] Escherichia coli W3110 is a host strain that finds use in the present invention, although it is not intended that the present invention be limited to this specific host strain. The expression vector was created by operatively linking a polynucleotide encoding an improved carboxyesterase into the plasmid pCK110900 operatively linked to the lac promoter under control of the lacI repressor. The expression vector also contained the P15a origin of replication and the chloramphenicol resistance gene. Cells containing the subject polynucleotide in Escherichia coli W3110 can be isolated by subjecting the cells to chloramphenicol selection.

Methods of Generating Engineered Carboxyesterase Polypeptides.

[0176] In some embodiments, to make the improved carboxyesterase polynucleotides and polypeptides of the present invention, the naturally-occurring carboxyesterase enzyme that catalyzes the amidation reaction is obtained (or derived) from *T. fusca.* In some embodiments, the parent polynucleotide sequence is codon optimized to enhance expression of the carboxyesterase in a specified host cell. As an illustration, the parental polynucleotide sequence encoding the wild-type carboxyesterase polypeptide of *T. fusca* was constructed from oligonucleotides prepared based upon the known polypeptide sequence of *T. fusca* carboxyesterase sequence available in Genbank database (Genbank accession no. WP_011292850.1). The parental polynucleotide sequence, designated as SEQ ID NO: 1, was codon optimized for expression in *E. coli* and the codon-optimized polynucleotide cloned into an expression vector, placing the expression of the carboxyesterase gene under the control of the lac promoter and lacI repressor gene. Clones expressing the active carboxyesterase in *E. coli* were identified and the genes sequenced to confirm their identity. The codon-optimized polynucleotide sequence designated SEQ ID NO: 1 was the parent sequence utilized as the starting point for most experiments and library construction of engineered carboxyesterases evolved from the original wild-type carboxyesterase.

[0177] In other embodiments, to make the improved carboxyesterase polynucleotides and polypeptides of the present invention, the naturally-occurring carboxyesterase enzyme that catalyzes the amidation reaction is obtained (or derived) from *G. stearothermophilus.* In some embodiments, the parent polynucleotide sequence is codon optimized to enhance expression of the carboxyesterase in a specified host cell. As an illustration, the parental polynucleotide sequence encoding the wild-type carboxyesterase polypeptide of *G. stearothermophilus* was constructed from oligonucleotides prepared based upon the known polypeptide sequence of *G. stearothermophilus* carboxyesterase sequence available in Genbank database (Genbank accession no. WP_033015113). The parental polynucleotide sequence, designated as SEQ ID NO: 137, was codon optimized for expression in *E. coli* and the codon-optimized polynucleotide cloned into an expression vector, placing the expression of the carboxyesterase gene under the control of the lac promoter and lacI repressor gene. Clones expressing the active carboxyesterase in *E. coli* were identified and the genes sequenced to confirm their identity. The polynucleotide sequence designated SEQ ID NO: 137 was the parent sequence utilized as the starting point for most experiments and library construction of engineered carboxyesterases evolved from the original wild-type carboxyesterase.

[0178] In some embodiments, to make the improved carboxyesterase polynucleotides and polypeptides of the present invention, the naturally-occurring carboxyesterase enzyme that catalyzes the amidation reaction was obtained or derived from *M. tuberculosis.* In some embodiments, the parent polynucleotide sequence was codon optimized to enhance expression of the carboxyesterase in a specified host cell. As an illustration, the parental polynucleotide sequence encoding the wild-type carboxyesterase polypeptide of *M. tuberculosis* was constructed from oligonucleotides prepared based upon the known polypeptide sequence of *M. tuberculosis* carboxyesterase sequence available in Genbank database (Genbank accession no. WP_003407276). The parental polynucleotide sequence, designated as SEQ ID NO: 139, was codon optimized for expression in *E. coli* and the codon-optimized polynucleotide cloned into an expression vector, placing the expression of the carboxyesterase gene under the control of the lac promoter and lacI repressor gene. Clones expressing the active carboxyesterase in *E. coli* were identified and the genes sequenced to confirm their identity. The codon-optimized polynucleotide sequence designated SEQ ID NO: 139, was the parent sequence utilized as the starting point for most experiments and library construction of engineered carboxyesterases evolved from the original wild-type carboxyesterase, as described herein.

[0179] In some embodiments, engineered carboxyesterases are obtained by subjecting the polynucleotide encoding the naturally occurring carboxyesterase or a codon-optimized version of the polynucleotide encoding naturally-occurring carboxyesterase to mutagenesis and/or directed evolution methods, as discussed above. Mutagenesis may be performed in accordance with any of the techniques known in the art, including random and site-specific mutagenesis. Directed evolution can be performed with any of the techniques known in the art to screen for improved promoter variants including shuffling. Mutagenesis and directed evolution methods are well known in the art (See, *e.g.*, US Patent Nos. 5,605,793, 5,811,238, 5,830,721, 5,834,252, 5,837,458, 5,928,905, 6,096,548, 6,117,679, 6,132,970, 6,165,793, 6,180,406, 6,251,674, 6,265,201, 6,277,638, 6,287,861, 6,287,862, 6,291,242, 6,297,053, 6,303,344, 6,309,883, 6,319,713, 6,319,714, 6,323,030, 6,326,204, 6,335,160, 6,335,198, 6,344,356, 6,352,859, 6,355,484, 6,358,740, 6,358,742, 6,365,377, 6,365,408, 6,368,861, 6,372,497, 6,337,186, 6,376,246, 6,379,964, 6,387,702, 6,391,552, 6,391,640, 6,395,547, 6,406,855, 6,406,910, 6,413,745, 6,413,774, 6,420,175, 6,423,542, 6,426,224, 6,436,675, 6,444,468, 6,455,253, 6,479,652, 6,482,647, 6,483,011, 6,484,105, 6,489,146, 6,500,617, 6,500,639, 6,506,602, 6,506,603, 6,518,065, 6,519,065, 6,521,453, 6,528,311, 6,537,746, 6,573,098, 6,576,467, 6,579,678, 6,586,182, 6,602,986, 6,605,430, 6,613,514, 6,653,072, 6,686,515, 6,703,240, 6,716,631, 6,825,001, 6,902,922, 6,917,882, 6,946,296, 6,961,664, 6,995,017, 7,024,312, 7,058,515, 7,105,297, 7,148,054, 7,220,566, 7,288,375, 7,384,387, 7,421,347, 7,430,477, 7,462,469, 7,534,564, 7,620,500, 7,620,502, 7,629,170, 7,702,464, 7,747,391, 7,747,393, 7,751,986, 7,776,598, 7,783,428, 7,795,030, 7,853,410, 7,868,138, 7,783,428, 7,873,477, 7,873,499, 7,904,249, 7,957,912, 7,981,614, 8,014,961, 8,029,988, 8,048,674, 8,058,001, 8,076,138, 8,108,150, 8,170,806, 8,224,580, 8,377,681, 8,383,346, 8,457,903, 8,504,498, 8,589,085, 8,762,066, 8,768,871, 9,593,326, 9,665,694, 9,684,771, and all related non-US counterparts; Ling *et al.*, Anal. Biochem., 254(2):157-78 [1997]; Dale *et al.*, Meth. Mol. Biol., 57:369-74 [1996]; Smith, Ann. Rev. Genet., 19:423-462 [1985]; Botstein *et al.*, Science, 229:1193-1201 [1985]; Carter, Biochem. J., 237:1-7 [1986]; Kramer *et al.*, Cell, 38:879-887 [1984]; Wells *et al.*, Gene, 34:315-323 [1985]; Minshull *et al.*, Curr. Op. Chem. Biol., 3:284-290 [1999]; Christians *et al.*, Nat. Biotechnol., 17:259-264 [1999]; Crameri *et al.*, Nature, 391:288-291 [1998]; Crameri, *et al.*, Nat. Biotechnol., 15:436-438 [1997]; Zhang *et al.*, Proc. Nat. Acad. Sci. U.S.A., 94:4504-4509 [1997]; Crameri *et al.*, Nat. Biotechnol., 14:315-319 [1996]; Stemmer, Nature, 370:389-391 [1994]; Stemmer, Proc. Nat. Acad. Sci. USA, 91:10747-10751 [1994]; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767; and WO 2009/152336. It is not intended that the present invention be limited to any particular methods, as various methods find use in the art.

[0180] In some embodiments, where the improved enzyme property desired is thermal stability, enzyme activity may be measured after subjecting the enzyme preparations to a defined temperature and measuring the amount of enzyme activity remaining after heat treatments. Clones containing a polynucleotide encoding a carboxyesterase are then isolated, sequenced to identify the nucleotide sequence changes (if any), and used to express the enzyme in a host cell.

[0181] Where the sequence of the engineered polypeptide is known, the polynucleotides encoding the enzyme can be prepared by standard solid-phase methods, according to known synthetic methods. In some embodiments, fragments of up to about 100 bases can be individually synthesized, then joined (*e.g.*, by enzymatic or chemical ligation methods or polymerase mediated methods) to form any desired continuous sequence. For example, polynucleotides and oligonucleotides of the invention can be prepared by chemical synthesis (*e.g.*, using the classical phosphoramidite method described by Beaucage et al., Tet. Lett., 22:1859-69 [1981], or the method described by Matthes et al., EMBO J., 3:801-05 [1984], as it is typically practiced in automated synthetic methods). According to the phosphoramidite method, oligonucleotides are synthesized (*e.g.*, in an automatic DNA synthesizer), purified, annealed, ligated and cloned in appropriate vectors. In addition, essentially any nucleic acid can be obtained from any of a variety of commercial sources (*e.g.*, The Midland Certified Reagent Company, Midland, TX, The Great American Gene Company, Ramona, CA, ExpressGen Inc. Chicago, IL, Operon Technologies Inc., Alameda, CA, and many others).

[0182] Engineered carboxyesterase enzymes expressed in a host cell can be recovered from the cells and or the culture medium using any one or more of the well-known techniques for protein purification, including, among others, lysozyme treatment, sonication, filtration, salting-out, ultra-centrifugation, and chromatography. Suitable solutions for lysing and the high efficiency extraction of proteins from bacteria, such as *E. coli*, are commercially available under the trade name CelLytic B (Sigma-Aldrich).

[0183] Chromatographic techniques for isolation of the carboxyesterase polypeptide include, among others, reverse phase chromatography high performance liquid chromatography (HPLC), ion exchange chromatography, gel electrophoresis, and affinity chromatography. Conditions for purifying a particular enzyme will depend, in part, on factors such as net charge, hydrophobicity, hydrophilicity, molecular weight, molecular shape, *etc.*, and will be apparent to those having skill in the art.

[0184] In some embodiments, affinity techniques may be used to isolate the improved carboxyesterase enzymes. For affinity chromatography purification, any antibody which specifically binds the carboxyesterase polypeptide may be used. For the production of antibodies, various host animals, including but not limited to rabbits, mice, rats, *etc.*, may be immunized by injection with the carboxyesterase. The carboxyesterase polypeptide may be attached to a suitable carrier, such as BSA, by means of a side chain functional group or linkers attached to a side chain functional group.

Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (Bacillus Calmette Guerin) and *Corynebacterium parvum*.

[0185] The carboxyesterases may be prepared and used in the form of cells expressing the enzymes, as crude extracts, or as isolated or purified preparations. The carboxyesterases may be prepared as lyophilizates, in powder form (*e.g.*, acetone powders), or prepared as enzyme solutions. In some embodiments, the carboxyesterases can be in the form of substantially pure preparations.

[0186] In some embodiments, the carboxyesterase polypeptides can be attached to a solid substrate. The substrate can be a solid phase, surface, and/or membrane. A solid support can be composed of organic polymers such as polystyrene, polyethylene, polypropylene, polyfluoroethylene, polyethyleneoxy, and polyacrylamide, as well as co-polymers and grafts thereof. A solid support can also be inorganic, such as glass, silica, controlled pore glass (CPG), reverse phase silica or metal, such as gold or platinum. The configuration of the substrate can be in the form of beads, spheres, particles, granules, a gel, a membrane or a surface. Surfaces can be planar, substantially planar, or non-planar. Solid supports can be porous or non-porous, and can have swelling or non-swelling characteristics. A solid support can be configured in the form of a well, depression, or other container, vessel, feature, or location. A plurality of supports can be configured on an array at various locations, addressable for robotic delivery of reagents, or by detection methods and/or instruments.

Methods of Using the Engineered Carboxyesterase Enzymes and Compounds Prepared Therewith

[0187] Whole cells transformed with gene(s) encoding the engineered carboxyesterase enzyme and/or the optional cofactor regeneration enzymes, or cell extracts and/or lysates thereof, may be employed in a variety of different forms, including solid (e.g., lyophilized, spray-dried, and the like) or semisolid (e.g., a crude paste).

[0188] The cell extracts or cell lysates may be partially purified by precipitation (ammonium sulfate, polyethyleneimine, heat treatment or the like), followed by a desalting procedure prior to lyophilization (e.g., ultrafiltration, dialysis, and the like). Any of the cell preparations may be stabilized by crosslinking using known crosslinking agents, such as, for example, glutaraldehyde or immobilization to a solid phase (e.g., Eupergit C, and the like).

[0189] The solid reactants (e.g., enzyme, salts, etc.) may be provided to the reaction in a variety of different forms, including powder (e.g., lyophilized, spray dried, and the like), solution, emulsion, suspension, and the like. The reactants can be readily lyophilized or spray dried using methods and equipment that are known to those having ordinary skill in the art. For example, the protein solution can be frozen at -80°C in small aliquots, then added to a prechilled lyophilization chamber, followed by the application of a vacuum. After the removal of water from the samples, the temperature is typically raised to 4°C for two hours before release of the vacuum and retrieval of the lyophilized samples.

[0190] The quantities of reactants used in the amidation reaction will generally vary depending on the quantities of product desired, and concomitantly the amount of carboxyesterase substrate employed. The following guidelines can be used to determine the amounts of carboxyesterase, and/or amine. Generally, ester and amine substrates can be employed at a concentration of about 5 to 200 grams/liter using from about 50 mg to about 5 g of carboxyesterase.

[0191] Those having ordinary skill in the art will readily understand how to vary these quantities to tailor them to the desired level of productivity and scale of production.

[0192] The order of addition of reactants is not critical. The reactants may be added together at the same time to a solvent (e.g., monophasic solvent, biphasic aqueous co-solvent system, and the like), or alternatively, some of the reactants may be added separately, and some together at different time points. For example, the carboxyesterase and the carboxyesterase substrate may be added first to the solvent.

[0193] For improved mixing efficiency when an aqueous co-solvent system is used, the carboxyesterase and amine may be added and mixed into the aqueous phase first. The organic phase may then be added and mixed in, followed by addition of the carboxyesterase substrate. Alternatively, the carboxyesterase substrate may be premixed in the organic phase, prior to addition to the aqueous phase

[0194] Suitable conditions for carrying out the carboxyesterase-catalyzed amidation reactions described herein include a wide variety of conditions which can be readily optimized by routine experimentation that includes, but is not limited to, contacting the engineered carboxyesterase enzyme and substrates at an experimental pH and temperature and detecting product, for example, using the methods described in the Examples provided herein.

[0195] The carboxyesterase catalyzed amidation is typically carried out at a temperature in the range of from about 15°C to about 75°C. For some embodiments, the reaction is carried out at a temperature in the range of from about 20°C to about 55°C. In still other embodiments, it is carried out at a temperature in the range of from about 20°C to about 45°C. The reaction may also be carried out under ambient conditions.

[0196] The amidation reaction is generally allowed to proceed until essentially complete, or near complete, coupling of substrates is obtained. Amide formation (product) can be monitored using known methods by detecting substrates and/or product. Suitable methods include, but are not limited to, gas chromatography, HPLC, and the like. Conversion yields of the amide product generated in the reaction mixture are generally greater than about 50%, may also be greater than about 60%, may also be greater than about 70%, may also be greater than about 80%, may also be greater than 90%, and are often greater than about 97%.

EXAMPLES

[0197] Various features and embodiments of the present invention are illustrated in the following representative examples, which are intended to be illustrative, and not limiting.

[0198] In the experimental disclosure below, the following abbreviations apply: ppm (parts per million); M (molar); mM (millimolar), uM and μM (micromolar); nM (nanomolar); mol (moles); gm and g (gram); mg (milligrams); ug and μg (micrograms); L and l (liter); ml and mL (milliliter); cm (centimeters); mm (millimeters); um and μm (micrometers); sec. (seconds); min(s) (minute(s)); h(s) and hr(s) (hour(s)); U (units); MW (molecular weight); rpm (rotations per minute); °C (degrees Centigrade); RT (room temperature); MWD (multiple wavelength detector); CDS (coding sequence); DNA (deoxyribonucleic acid); RNA (ribonucleic acid); RP-HPLC (reversed-phased high performance liquid chromatography); FIOP (fold improvement over positive control); HTP (high throughput); LB (Luria broth); TFA (trifluoroacetic acid); MeCN (acetonitrile); TEoA (triethanolamine); THF (tetrahydrofuran); Sigma-Aldrich (Sigma-Aldrich, St. Louis, MO); Millipore (Millipore, Corp., Billerica MA); Difco (Difco Laboratories, BD Diagnostic Systems, Detroit, MI); Daicel (Daicel, West Chester, PA); Genetix (Genetix USA, Inc., Beaverton, OR); Molecular Devices (Molecular Devices, LLC, Sunnyvale, CA); Applied Biosystems (Applied Biosystems, part of Life Technologies, Corp., Grand Island, NY), Agilent (Agilent Technologies, Inc., Santa Clara, CA); Thermo Scientific (part of Thermo Fisher Scientific, Waltham, MA); Corning (Corning, Inc., Palo Alto, CA); and Bio-Rad (Bio-Rad Laboratories, Hercules, CA); Phenomenex (Phenomenex, Inc., Torrence, CA); Epicentre (Epicentre, Madison, WI) .

EXAMPLE 1

Wild-type Carboxyesterase Gene Acquisition and Construction of Expression Vectors

[0199] This Example describes the acquisition of the codon optimized recombinant polynucleotides encoding wild-type carboxyesterases (SEQ ID NOS: 2, 138, and 140) from which genes encoding engineered carboxyesterases in the following examples were derived, and expression vectors and host cells suitable for such engineering.

[0200] The codon optimized versions of the wild-type genes (SEQ ID NO: 1, 137, and 139) encoding the wild-type carboxyesterases (SEQ ID NO: 2, 138, and 140) of *T. fusca*, *G. stearothermophilus*, and *M. tuberculosis*, respectively, were synthesized for expression in *E. coli*. The codon optimized gene was cloned into expression vector pCK11 0900 (See, *e.g.*, US Pat. No. 9,714,437 and US Pat. Appln. Publn. No. 2006/0195947, both of which are incorporated herein by reference in their entireties and for all purposes), under the control of a lac promoter. The expression vector also contained the P15a origin of replication and the chloramphenicol resistance gene. These sequence-verified vectors were transformed into a *E. coli* W3110 strain for expression. The polynucleotides (odd numbered SEQ ID NOS: 1-135) encoding the engineered carboxyesterases (even numbered SEQ ID NOS: 2-136) of the present invention were likewise cloned into vector pCK11 0900 for expression in a derivative of *E. coli* W3110 strain. Directed evolution techniques generally known to those skilled in the art were used to generate the libraries of the engineered carboxyesterases.

EXAMPLE 2

HTP Production and Analysis of Wild-type Carboxyesterase Polypeptides

[0201] HTP lysates were prepared by taking the codon-optimized carboxyesterase genes (described in Example 1) which were then transformed into E.coli W3110 and plated on Luria-Bertani (LB) agar medium containing 1% glucose and 30 μg/mL chloramphenicol (CAM). After incubation for at least 16 h at 30 °C, colonies were picked using a Q-botobotic colony picker (Genetix) into a 96-well shallow well microtiter plate containing 200 μL of LB, 1% glucose, and 30 μg/mL CAM. Cells were grown 18 - 20 h at 30°C, with shaking at 200 rpm. Then, 20 μL of these cultures were then transferred to 360 μL of Terrific Broth (TB) and 30 μg/mL CAM in a deep well plate. After incubation at 30°C with shaking at 250 rpm for 2.5 h (OD600 0.6-0.8), recombinant gene expression was induced by adding isopropyl thioglycoside (IPTG) to a final concentration of 1 mM. The plates were then incubated at 30 °C, with shaking at 250 rpm for 18 - 21 h.

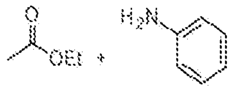
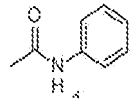
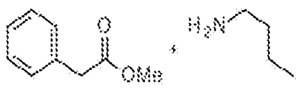
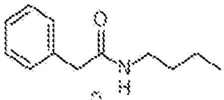
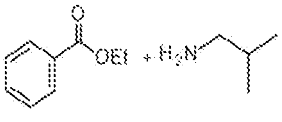
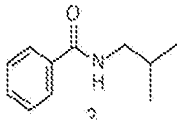
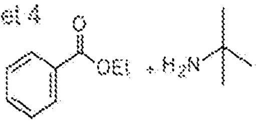
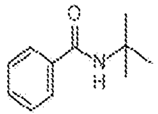
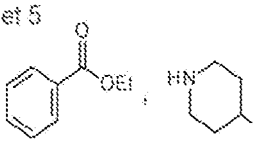
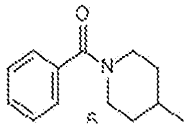
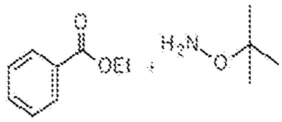
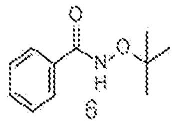
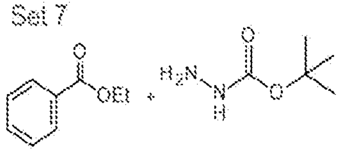
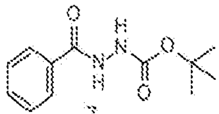
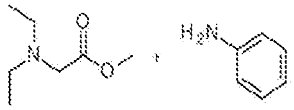
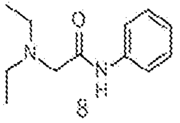
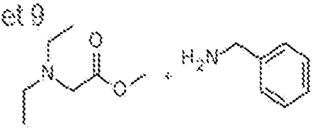
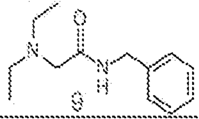
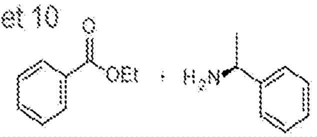
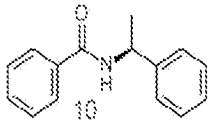
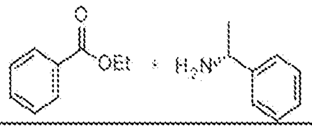
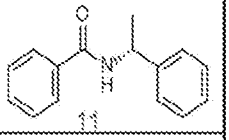
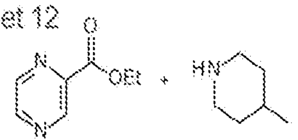
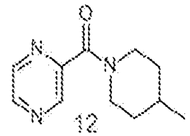
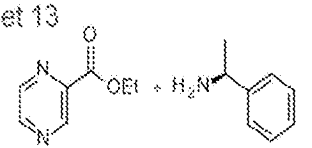
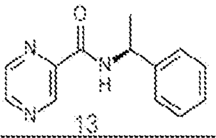
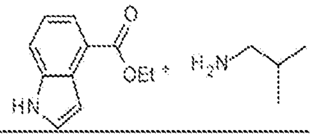
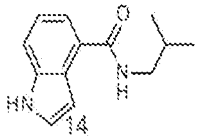
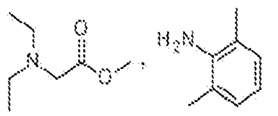
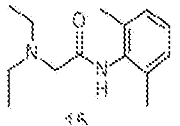
[0202] Cell cultures were pelleted at 3500 x g for 20 min, and the media discarded. Cell pellets were lysed in 200 μL of 0.2 M TEoA, pH 7.5 with 1 g/L lysozyme, 0.5 g/L polymixin B sulfate, and 0.5 μL OmniCleave™ endonuclease (Epicentre) by shaking at RT for 2 h. Samples were centrifuged at 3500 x g for 20 min to clarify cellular debris, and the supernatant was used to carry out the transformations described in Examples 3 and 4.

EXAMPLE 3

Activity of Wild-type Carboxyesterases on Substrate Set #1

[0203] The wild-type carboxyesterase polypeptides were generated as described in Example 2. To analyze for amidation activity on substrate set #1 (Table 3.1), 20 μL supernatant were added to a mixture of 10 μL aniline, 10 μL ethyl acetate, 100 μL of 0.1 M sodium phosphate buffer, pH 7.0. Reactions were incubated at 20 °C and shaken at 300 rpm for 18 h. Samples were quenched by diluting 2-fold in MeCN. Analysis of reactions was performed by RP-HPLC as described in Example 12.1.

Table 3.1 Substrate Sets Used to Evaluate Carboxyesterase Amidation Activity

	Products
Set 1 	 1
Set 2 	 2
Set 3 	 3
Set 4 	 4
Set 5 	 5
Set 6 	 6
Set 7 	 7
Set 8 	 8
Set 9 	 9
Set 10 	 10
Set 11 	 11
Set 12 	 12
Set 13 	 13
Set 14 	 14
Set 15 	 15

EXAMPLE 4

Activity of Wild-type Carboxyesterases on Substrate Sets #2-7

[0204] Wild-type carboxyesterase polypeptides were generated according to the methods described in Example 2. To analyze for amidation activity on substrate set #2 (See, Table 3.1), 100 μL supernatant was added to a mixture of 10 μL methyl phenylacetate, 10 μL n-butylamine, 10 μL dimethyl sulfoxide (DMSO), and 20 μL of 0.1 M sodium phosphate buffer, pH 7.0. Reactions were incubated at RT and shaken at 300 rpm for 18 h. Samples were quenched by diluting with an equal volume of MeCN. Analysis of substrates and products was performed by RP-HPLC using the methods described in Example 12.1.

[0205] To analyze for amidation activity on substrate sets #3-7 (See, Table 3.1), 120 µL of a mixture of 75 - 100 g/L amine and 15 g/L ethyl benzoate in 0.1 M potassium phosphate buffer, pH 7.5, with or without 50% isopropanol, were added to 100 µL supernatant, produced as described in Example 2. Reactions were incubated at 55 °C and shaken at 400 rpm for 18 h. Samples were quenched by diluting 3-fold in MeCN and then centrifuged for 5 - 8 min at 4000 x g. The resultant supernatant was analyzed by RP-HPLC using the methods described in Example 12.2. In one experiment using substrate set #3, 120 µL of a mixture of 5 g/L isobutylamine and 15 g/L ethyl benzoate in 0.1 M potassium phosphate buffer, pH 7.5, was added to 100 µL supernatant, produced as described in Example 2. The results for all of these substrate sets are provided in Table 4.1 with the substrate sets indicated by "reaction" numbers (i.e., "R1," "R2," "R3," etc.).

Table 4.1 Solvent Tolerance and Substrate Scoping of Wild-type Carboxyesterases (SEQ ID NOS: 2, 138, and 140)[1]

Conditions (KPi Buffer, pH 7.5)	R1	R2	R3	R4	R5	R6	R7
T. fusca	‡		†		+++++	+++	+++
G. stearothermophilus	##	+++++	#	+		†	
M. tuberculosis	#	++++	+	+++		++++	
Conditions (KPi Buffer, pH 7.5 / IPA (50:50 v/v%))	R1	R2	R3	R4	R5	R6	R7
T. fusca			†				++++
G. stearothermophilus			##			#	++
M. tuberculosis			##	+	+		+++

[1] Levels of increased activity were determined relative to the reference peak area of the negative control powder. In this Table, "+" = > 0, < 20; "++" = > 20, < 50; "+++" = > 50, < 100; "++++" = > 100, < 500; "+++++" = > 500, < 1000; "#" = > 1000, < 2000; "##" = > 2000, < 5000; "†" = >5000, < 10000; "‡" = > 10000, < 15000.

EXAMPLE 5

Production of Wild-type Carboxyesterase Shake Flask Powders

[0206] The wild-type carboxyesterases (SEQ ID NOS: 2, 138, and 140) were produced in shake flasks for further characterization. The *E. coli* transformants containing the plasmid encoding WT carboxyesterases were grown on Luria-Bertani (LB) agar medium containing 1% glucose and 30 µg/mL chloramphenicol (CAM). After incubation for at least 16 h at 30 °C, single colonies were picked into 6 mL of LB, 1% glucose, and 30 µg/mL CAM. Cells were grown 18 - 20 h at 30 °C, with shaking at 250 rpm. This culture was then transferred into Terrific Broth (TB) and 30 µg/mL CAM at a final OD_{600} of 0.2 and a final volume of 250 mL. After incubation of the flasks at 25 °C or 30 °C with shaking at 250 rpm for 2.5 h (OD_{600} 0.6-0.8), recombinant gene expression was induced by isopropyl thioglycoside (IPTG) to a final concentration of 1 mM. The flask was then incubated at 30 °C with shaking at 250 rpm for 18 - 21 h. Cells were pelleted at 3500 x g for 20 min, and the supernatant was discarded. The cell pellet was washed 3x in 30 mL ice cold 50 mM sodium phosphate pH 7.5, resuspended in 12 mL of the same buffer, and lysed using a cell disruptor at 18-20 kpsi. Lysates were clarified at 10000 x g for 30 min, and clarified supernatants were lyophilized to an off-white powder.

Table 5.1 Wild-type Carboxyesterase (SEQ ID NO: 2, 138, and 140) Growth Evaluation

	SEQ ID NO:	Growth OD_{600}[1]	Growth OD_{600}	Harvest OD_{600}	Harvest OD_{600}	[2]Pellet Mass	Pellet Mass	[3]Powder Mass	Powder Mass
Shake Flask		25°C	30°C	25°C	30°C	25°C	30°C	25°C	30°C
Neg. Ctrl.	-	+++	+++	++	++	++++	+++	+	++
T. fusca	2	++++	++++	+++	++++	+++	++++	++	++
G. stearothermophilus	138	++++	++++	+	N/A	++	N/A	++	N/A
M. tuberculosis	140	++++	++++	+++	+++	++++	++++	+++	+++

[1] OD 600 rankings as follows ("+" = >10, <15: "++"= >15, <20: "+++" = >20, <25; "++++" = >25, <30);

[2] Mass pellet rankings as follows ("+" = > 3 g, <4 g; "++"= > 4 g, <5 g; "+++"= > 5 g, <6 g; "++++"= > 6 g, <7 g);

[3] Mass of powder ranking as follows ("+" = > 0.2 g, <0.5 g; "++"= > 0.5 g, <1.0 g; "+++"= > 1.0 g, < 1.5 g)

EXAMPLE 6

Solvent Tolerance Evaluation of Wild-Type Carboxyesterase Polypeptides

[0207] In this Example, experiments conducted to determine the solvent tolerance of the wild-type carboxyesterases are described. In these experiments, the tolerance to organic solvent of the lyophilized shake flask powders prepared in Example 5 was determined, by testing activity on substrate sets #3 and #5. Reactions were conducted in a 96 well plate (reaction volume 220 µL) with 10 g/L enzyme powder, 30 - 34 g/L amine, 43 g/L ester, and 10% DMSO, or 50% isopropanol, or 15 - 25% MeCN or 25% THF in 0.1 M potassium phosphate buffer, pH 7.5. Reactions were heated to 50°C with shaking at 400 rpm for 18 - 21 h. Reactions were quenched by diluting 3-fold in MeCN. Reaction samples were analyzed by RP-HPLC using methods described in Example 12.2.

Table 6.1 Solvent Tolerance Evaluation of Wild-type Carboxyesterases (SEQ ID NO: 2, 138, and 140)

Substrate Set	Co-solvent		*T. fusca*	*G. stearothermophilus*	*M. tuberculosis*
#3	Buffer Only		+++++	++	++
	MeCN	15%	++	++	++
		25%	++	++	++
	THF	15%	++	-	++
		25%	++	++	++
	DMSO	15%	++++	+	-
		25%	+++	++	++
#5	Buffer Only		++	-	-
	MeCN	15%	+	+	++
		25%	+	+	-
	THF	15%	++	-	-
		25%	+	-	-
	DMSO	15%	++	-	-
		25%	++	-	-

[1] Levels of increased activity were determined relative to the reference negative control powder (Table 5.1) and defined as follows: “-” = no activity; “+” > 1-fold activity; “++” >1-fold activity, < 2-fold activity; “+++” > 2-fold, <3-fold activity; "++++" > 4-fold, <5-fold activity; “+++++” > than 5-fold activity.

EXAMPLE 7

Production and Analysis of Engineered *T. fusca* Carboxyesterase Polypeptide Libraries

[0208] Plasmid libraries containing evolved *T. fusca* carboxyesterase genes were transformed in *E. coli* according to the methods described in Example 1, and produced following the methods described in Example 2. The cell lysates were used to carry out the activity assessments described in Examples 8 and 9.

EXAMPLE 8

Amidation Activity of Engineered *T. fusca* Carboxyesterase Polypeptides on Substrate Sets #3, 8 - 12

[0209] *T. fusca* carboxyesterase variants were generated according to Example 7. To analyze for amidation activity on substrate set #3 (See, Table 3.1), 20 μL of supernatant produced as described in

Example 7, were added to a mixture of 25 µL MeCN with 3 M ethyl benzoate, 110 µL of 0.6 M isobutylamine in 0.2 M TEoA (pH adjusted to 9.5), and 65 µL 0.2 M TEoA, pH 8.5. Reactions were incubated at 50 °C and shaken at 300 rpm for 18 h. Samples were quenched by diluting 4-fold in MeCN. Analysis of reactions was performed by RP-HPLC using the methods described in Example 12.2.

[0210] To evaluate for amidation activity on substrate sets #3, 8-12 (See, Table 3.1), 20 - 50 µL supernatant produced as described in Example 7, were added to a mixture of 25 µL MeCN with 2 - 3.2 M amine, 110 µL of 0.4 - 0.6 M ester suspension in 0.2 M TEoA (pH adjusted to 8), and 35 - 65 µL 0.2 M TEoA, pH 8.5 - 9.0. Reactions were incubated at 50 °C and shaken at 300 rpm for 16 – 18 h. Samples were quenched by diluting 4-fold in MeCN. Analysis of reactions was performed by RP-HPLC, as described in Examples 13.3, 13.4, and 13.5. In one experiment using substrate set #8, 50 µL supernatant produced as described in Example 7, were added to a mixture of 25 µL MeCN with 3 M glycine methyl ester, 110 µL of 0.6 M aniline in 0.2 M TEoA (pH adjusted to 8), and 35 µL 0.2 M TEoA, pH 8.5. Reactions were incubated at 50 °C and shaken at 300 rpm for 18 h. Samples were quenched by diluting 4-fold in MeCN. Analysis of reactions was performed by RP-HPLC using the methods described in Example 12.3. The amino acid substitutions of these variants are indicated relative to SEQ ID NO: 2.

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
1		A108W	***					
2	3/4	L282Q	***					
3	5/6	L282A	***	*		*		
4	7/8	L282T	***			*	*****	
5		A108K	***					
6	9/10	A285L	**	*				
7		I69F	**					
8	11/12	L282C	**	*				
9		L282W	**					
10		L282R	**					
11		F323Y	**					

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
12		I69W	*					
13	13/14	A108G	*	***				
14		I69Y	*					
15		T110A	*					
16		A108Q	*					
17		A285P	*					
18		T373G	*					
19	15/16	N215R	*	*				
20		T317P	*					
21		I69L	*					
22		N215K	*	*				
23		A108R	*					
24	17/18	L281P	*				*****	
25	19/20	N215W	*					*
26		L209P	*					
27		P283K	*					
28		F377Y	*					
29		A213P	*					
30		G70W	*					
31		A249V/F284P	*					
32	21/22	G212P	*	*				
33		A186G	*					
34		I372A/V376A	*					
35		A71R	*	*				
36		Y320S	*					
37	135/ 136	I372L	*					
38		A71Y	*					

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
39		A108S	*					
40		S115T	*					
41		A71G	*					
42		A71F	*					
43		S190H	*					
44	23/24	V280G	*				*****	
45		P64W	*					
46		I69V	*					
47		L209E	*					
48		P117A	*					
49		A68L	*					
50		P64I	*					
51		L209S	*					
52		Y65W	*					
53	25/26	N215P	*	*				
54		P63A	*					
55		S279L/V280G/ L282M	*					
56		T110S	*					
57		S115H	*					
58		V376M	*					
59		V376L	*					
60		G70L	*					
61		G214T	*					
62		G212A	*					
63		V376A	*					
64		F284T	*					
65		P405D	*					
66		S185T	*					

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
67		G70R	*					
68		V280E	*					
69	27/28	P283D	*				*****	
70		Y320S/I372A/ V376G	*					
71		G70T	*					
72		P64E	*					
73		Y320F	*					
74		R428V	*					
75		Y320A	*					
76		R321S	*					
77		W271K	*					
78		W271P	*					
79		Y65T	*					
80		V376G	*					
81		F377W	*					
82		V118I	*					
83		F323R	*					
84		A217W	*					
85		G214K	*					
86		A429L	*					
87		T110H	*					
88		T110P	*					
89		R153L	*					
90	29/30	Y320G	*				***	
91		Y320S/V376G/ F377V	*					
92		G114H	*					
93		V280S	*					

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
94	31/32	R321L	*				***	
95		Y320S/F323S/I372A	*					
96		T224I/P268S/I372F	*					
97		V118N	*					
98		A217V	*					
99		A71P	*					
100		L209V	*					
101		Y320G/F323S	*					
102		W271L	*					
103		P420G	*					
104	33/34	W271T	*					*
105		P283C	*					
106	35/36	P63R	*	*				
107	37/38	Y320W	*	*				
108	39/40	S190K	*					*
109		R126C					****	
110		L282S						*
111		F284C					*****	
112		F284V						*
113		A269N		*				
114		W271A					**	
115		P283T					****	
116		A71H/Q263R					*	
117	41/42	N215R/W271R			*		*	
118		A217G		*				
119		A213S					**	
120		G277M					***	

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
121		L281V					***	
122	43/44	P283V					****	
123		D427A					*	
124		P63Y		*				
125		V270I/V470M			*			
126		A269V					*	
127		S279G					****	
128		A217S					*	
129	45/46	A68P				*	*	
130		P63T		*				
131		A285M		*				*
132	47/48	P283Y					*****	
133		L311I		*				
134		S279C					****	
135		V270I		*				
136		V270R					**	
137		T317C					**	
138		W271S		*			*	
139		S279V					****	
140		A217R			*			
141		S190L					*	
142		P64A					*	
143		A71V					*	
144		P117F					*	
145	49/50	E184S/A249T					*****	
146	51/52	F284T/P438T					*****	
147		L209G					*	
148		G278H					***	

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
149		L324A					****	
150		S190M					*	
151		P64G					*	
152		A276F					****	
153		P64V					*	
154		P64T					*	
155		P66N					*	
156		A217L					*	
157		I69H					*	
158		M216P					*	
159		A213N					*	
160		A217R/A231V					*	
161		A213T/W271K					*	
162		V118G/A349V					*	
163		N215M					*	
164		A188G					*	
165		S190Q					*	
166		T39M/F323I		*				
167		G278S					****	
168		V118N/A269T					*	
169		A213C					*	
170		Y65S					*	
171		P283R/A429V						*
172		A213V					*	
173	53/54	A213L			*		*	
174		A186C					*	
175		E184F			*			
176		A213Q					*	

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
177		I104Q/A429V					*	
178		A217P					*	
179		N111W					*	
180	55/56	F377L		**				
181	57/58	E184G			*		*****	
182		G214V					*	
183		F323C		*				
184	59/60	R153H/N215P			*		*	
185		W164R/W271T					**	
186		G212R					*	
187		P286V					**	
188		F323I		*				
189		N111M					*	
190		I69G					*	
191	61/62	G214L					*	
192		G212S					*	
193		W271Q/A416V					*	
194	63/64	S190W			*		*	
195		Q210T			*			
196		G114Q					*	
197		N111V					*	
198		Y119S					*	
199		N111L					*	
200		A213E					*	
201		S211I						*
202		A186T					*	
203		F109G/P117M					*	
204		Y119G					*	

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
205		S211V					*	
206		L281Y/D374N					*	
207		E184P					*****	
208		Y119P					*	
209		A213R/S345G					*	
210		W103T/P147S					*	
211		W103R						*
212		N111S					*	
213		W103P					*	
214		Q210W					*	
215		I104P					*	
216		S211L					*	
217		S190R					*	
218		G183P					*	
219		Q210P					*	
220		A188E					*	
221		H105L					*	
222		G107P					*	
223		S113P					*	
224		G114A					*	
225	65/66	F77S/E184G					*****	
226		S279E					****	
227		G107D/S185W					**	
228		S211R					*	
229		S185A					*	
230		A186R					*	
231		G187P					*	
232		A186P					*	

Table 8.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 2 on Substrate Sets 3, 8, 9, 10, 11, and 12								
Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 8	Set 9	Set 10	Set 11	Set 12
233		R62H/P117G					*	
234		N111R					*	
235		S115V					*	
236		G107L			*			
237		G107S			*			
238		Y65G		*				
239		E184Y						*
240		G174D/L282V						*
[1]Levels of increased activity were determined relative to the reference polypeptide and defined as follows: < 1.5x; “ * ” = > 1.5x, < 3.5x; “ ** ” = > 3.5x, < 5.5x; “ *** ” = > 5.5x, < 7.5x; “ **** ” = > 7.5x, < 9.5x; “ ***** ” = > 9.5x, <10.5x.								

EXAMPLE 9

Evaluation of Shake Flask Powders of Engineered *T. fusca* Carboxyesterase Polypeptides

[0211] Powders of evolved *T. fusca* carboxyesterases were prepared in shake flask scale quantities following the methods described in Example 5. The amidation activity of the lyophilized shake flask powders was assessed by testing their activity on substrate set #3. First, 1.0 mL reaction mixtures were prepared with 10 g/L enzyme powder, 0.3 M isobutylamine, 0.3 M ethyl benzoate in toluene, followed by addition of 20 μL of 0.2 M TEoA buffer, pH 8.5. Reactions were heated to 50 °C with shaking at 500 rpm for 60 h. Reactions were quenched by diluting 100 μL reaction mixture into 1.4 mL acetone. Reaction samples were analyzed by RP-HPLC using the methods provided in Table 12.2.

Table 9.1: Amide Formation by *T. fusca* Carboxyesterase Variants	
Variant No.:	**Amide Formation (FIOP)[1]** Relative to SEQ ID NO: 2
4	++++
3	++++
12	+++
31	+++

28	++
24	++
37	++

[1] Levels of increased activity or selectivity were determined relative to the reference polypeptide of SEQ ID NO: 2 and defined as follows: “++” > than 1.2-fold but less than 2.5-fold increase; “+++” > than 2.5-fold but less than 5-fold; "++++" > than 5-fold but less than 10-fold.

EXAMPLE 10

Production and Analysis of Combinatorial *T. fusca* Carboxyesterase Libraries

[0212] Plasmid libraries obtained through combinatorial shuffling on a T. fusca carboxyesterase variant (SEQ ID NO: 8) were transformed in to E. coli W3110 according to the methods described in Example 1. The HTP lysates produced according to the methods described in Example 2, were used to carry out the activity assessments described in Example 11.

EXAMPLE 11

Activity of Engineered *T. fusca* Carboxyesterase Polypeptides on Selected Substrate Sets 3, 5, 8, 12-15

[0213] *T. fusca* carboxyesterase variants generated from combinatorial libraries were analyzed for amidation activity on substrate sets #3, 5, 8, and 12-15 (See, Table 3.1). To analyze for amidation activity on substrate set #3 (See, Table 3.1), 10 µL supernatant produced as described in Example 7, and were added to a mixture of 25 µL MeCN with 3 M ethyl benzoate, 110 µL of 0.6 M isobutylamine in 0.2 M TEoA (pH adjusted to 8), and 75 µL 0.2 M TEoA, pH 8.5. Reactions were incubated at 50 °C and shaken at 300 rpm for 18 h. Samples were quenched by diluting 4-fold in MeCN. To analyze for amidation activity on substrate sets #5, 8, and 12-15, 20 - 50 µL supernatant were produced as described in Example 7, and added to a mixture of 25 µL MeCN with 2 - 3.2 M amine, 110 µL of 0.4 - 0.6 M ester suspension in 0.2 M TEoA (pH adjusted to 8), and 35 - 65 µL 0.2 M TEoA, pH 8.5 - 9. Reactions were incubated at 50 °C and shaken at 300 rpm for 16-18 h. Samples were quenched by diluting 4-fold in MeCN. Analysis of reactions was performed by RP-HPLC, using the methods described in Examples 12.2 (Set 3), 12.3 (Set 8), 12.5 (Set 12, and 13), 12.6 (Set 14), and 12.7 (Set 5 and 15). In one experiment using substrate set #8, 20 µL supernatant produced as described in Example 7 were added to a mixture of 25 µL MeCN with 3 M glycine methyl ester, 110 µL of 0.6 M aniline in 0.2 M TEoA (pH adjusted to 9.5), and 65 µL 0.2 M TEoA, pH 8.5. Reactions were incubated at 50 °C and shaken at 300 rpm for 18 h. Samples were quenched by diluting 4-fold in MeCN. Analysis of reactions was performed by RP-HPLC, using the methods describe in Example 12.3.

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
241	67/68	Y65W/I69L/I372L	****						
242		Y65W/I69W/G70L/I372L	***						
243		I69W/G70L/G459R	***						
244		I69W/G70L	***						
245		G70L	**						
246	69/70	A68P/I69L/A343V/I372L	*		*			*	
247		I69L/G70L/P331Q/I372M	*		*				
248	71/72	I69L/W271Y/I372L	*	*	*				
249	73/74	I372L	*	*	*	*			
250	75/76	P63R/A108G	*					*	*
251		R126C/I372L	*						
252		Y320W	*						*
253		I69L/I372L	*						
254		I69W/G212A/A213L/N215R/ V280G/L281P	*						
255		I69W/G70L/I372M	*						
256	77/78	Y65W/I69M/G70A/L281P/ I372L	*		*	*			
257		A108G	*					*	*
258		R126C	*						
259	79/80	I69W/I372M	*	*	**			*	
260		I69M/G70A/I372M	*		*				*
261		I69L/T282C/A343V/I372L	*						
262		P63R	*						*
263	81/82	V280G/A285L/I372L	*			*			*
264		A68P/I69L/W271Y	*		*				
265	83/84	W271Y/A343V/I372L	*	*		*	*		
266	85/86	I69L	*		*		*		*
267	87/88	L281P/I372L	*	*	*			*	

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15									
Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
268		I69L/W271Y			*				
269		I372M		*	*				
270	89/90	P63R/Y65G/A108G						*	***
271		A343V/I372L			*				
272	91/92	A68P/W271Y/I372L		**					**
273	93/94	A108G/A285L						*	**
274		Y65W/G70L/I372M		*		*			
275		I372L/A381L		*					**
276		I69W/N215P							***
277		I69W		*****	*				
278		A108G/V270E						*	*
279		W271Y/A343V/I372L/A381L		*****					
280		P63Y				*			
281		A108G/F377I							****
282		P63A						*	
283		A68P/I69L/T282C/V287I							*
284		T373G				*			*
285		I69W/T282A			*				
286	95/96	A108G/N215K				*			*****
287		N215P/I372L				*			*
288		Y65G/Y320W							*
289		I69Y/A108G/L281P/A285P						*	
290		P63R/A108G/Y320W/F323C							*
291		I69L/N215P/W271Y/I372L				**			
292	97/98	A68P		**			*		*
293		P63Y/P268A/A269N/V270I/A429V		*		*			
294		W271Y		*					

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
295		W271Y/A343V		*****					
296		A68P/I69W/N215P							*****
297		R126C/E184S/A213S/I372L				*			
298		A108G/M189L/Y320W							***
299		I69F/A108G/V270E/I372L/F377L							*****
300		I69L/N215P/W271Y/V280G/L281P/T282C							***
301	99/100	N215R/I372L				*			**
302		P63R/F377I							****
303		A108G/Y320W/F323I							*
304		I372L/F377L							*****
305		N215P/W271Y/I372L							***
306		I69L/N215P/A343V/I372L/A381L				*			
307		R126C/N215P							*
308	101/102	N215K/L281P/A285L/I372L				**			**
309		A68P/F377L							**
310		A68P/I69W/G214R/N215P/W271Y				*	*		
311		A68P/A108G/F377L							*****
312		F377L							**
313	103/104	P63T/N215R				*	*		**
314		A68P/I69W/M189I/A381L					**		
315		A68P/E184S				*			
316	105/106	I69W/N215R				*			*****
317		I69L/N215P/W271Y/T282A							**
318		I69W/A108S			*				
319		N215P/Y320G				*			*

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
320		N215R/V280G/L281P/A285L/I372L				*			*
321		V181L/N215P							*
322	107/108	P63A/N215R/A343V				*	*		***
323		N215P/Y320G/I372L				*			
324		M189I/F377I							***
325	109/110	P63Y/G212P/N215R				*	*		**
326		A108G/F377L							***
327		P63Y/N215P/A269N				*	*		
328		I69W/N215K/A343V					*		
329		A68P/I69W/M189E/G214R/N215P/W271Y/L281P/T282G/A343V/A381L		*					
330	111/112	P63Y/G212P/N215R/P268A/A269N/A343V				*			**
331		I69F/N215K				*			*****
332		G214R/N215P/W271Y							***
333		I69F/A285L/T373G				*			*
334		N215R/A249T/V280G/L281P/A285L/I372L				*			*
335		A68P/I69L/N215P					*		
336		A68P/I69W/M189I/I372L					**		
337		R126C/M189I/A285L/I372L							**
338		N215P				*			**
339		A249T/F377L							*
340		N215K/A285L/V445L				*			***
341	113/114	P63Y/N215R				*	*		*
342		P63A/M189A							*
343		N215K				*			***
344		N215K/F323Y				*			*
345		A285L/F323I							*

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
346		A68P/I69W/M189I/W271Y					*		
347		I69W/M189I/A343V/A381L		*					
348		P63Y/N215R/V270I/W271S							*
349		N215R				*	*		*
350		I69W/M189E/I372L		*					
351		N215P/I372L/F377L				*			*****
352		A68P/N215P/W271Y/A343V/I372L/A381L				*	*		
353		N215R/Y320G				*			***
354		I69W/M189I/W271Y							***
355		I69L/M189I/V280G/T282G/A343V/I372L/A381L		*****					
356		P63R/A108G/A285L/F377I							*****
357		I69W/M189E/W271Y/A343V		*					
358		I69W/M189I					*		
359		A68P/G214R/N215P/W271Y/L281P/T282A/I372L							*
360		M189I/Y320W/F377I							****
361		N215K/L281P/T373G				*			*
362		N215R/A285P							****
363		M189I/V270E/A285L							**
364		P63R/Y65G/A108G/F377I							*****
365		I69W/N215R/F323Y				*			*
366		M189Q		*****					
367		M189I/N215P/A343V/I372L				*			
368		M189Q/A343V		*****					
369		P63R/Y65G/Y320W/F323I							**
370		I69W/G214R/N215P/W271Y/I372L/F377Y/A381L		*****					

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
371		A108G/T282A/A285L/F377L							*****
372		N215P/A381L				*	*		
373		A108G/M189I/T282A/A285L/Y320W							**
374		T282A/A285L/Y320W/F323I							**
375		P63R/A108G/T282A/A285L/F377L							*****
376		I69W/G214R/W271Y/T282A							**
377		M189Q/A381L		****					
378	115/116	A68P/I69W/M189I/G214R/N215P		*****		*	*		
379		Y65G/Y320W/F323I							***
380	117/118	N215K/T373G				*			****
381		I69W/M189Q/I372L/F377Y		*****					
382		M189E/I372L/F377Y		*****					
383		A108G/M189L							**
384		A68P/I69W/M189E/A343V/A381L					*		
385		I69W/M189E/G214R/A343V/I372L		*****					
386		I69L/M189I/T282A					*		
387		A68P/M189I/W271Y/I372L					*		
388		G214R/N215P/W271Y/F377Y				*			
389	119/120	N215K/I372L/F377L				*			*****
390	121/122	M189I/A343V		*****			*		
391	123/124	N215R/L281P/A285L/T373G				***			*
392	125/126	N215W/A285L/G346S				*			**
393		G214R/W271Y							*

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
394		R126C/E184S/A213S/V280G/L281P/A285L/Y320G				**			
395		A68P/I69L/M189I/L281P/T282C/I372L/F377Y/A381L				***			
396		A68P/M189I/A343V					*		
397		I69L/M189I/W271Y/A343V/A381L					**		
398		M189L							**
399		P63Y/M189L							*
400		I69L/M189I					**		
401		A68P/I69W/G214R/A343V					*		
402		N215W				*			*
403		M189I/V280G/T282A					*		
404		N215W/A285P							**
405		M189I/N215K							*
406		A68P/I69L/M189I/G214R/N215P/W271Y					*		
407		A68P/I69L/M189Q/W271Y/V280G/I372L/A381L		**					
408	127/128	M189I		**			*		*
409		P63R/Y65G/T282A/A285L/Y320W/F323I							*
410		I69W/M189I/N215P/A343V					*		
411		A108G/M189L/F377I							**
412		I69F/M189L							**
413		N215K/A285L/T317P				*			*
414		I69Y/T110A/N215R/L281P				*			
415		I69L/N215R/A285P/T317P							*
416		I69W/G214R/W271Y/A343V					*		
417		A213S/N215P/Y320G				*			
418		I69Y/M189L/L281P/T373G							*

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15									
Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
419		N215W/T373G				*			*
420		M189I/G214R/N215P/W271Y/T282G					*		
421		P63R/Y65G/A108G/M189L							*
422	129/130	I69F/N215K/A269L/V270I/F377L							*****
423		A68P/N215P/F377L							****
424		N215P/F377L							****
425		I69L/M189I/G214R/W271Y/L281P/T282A/A343V							**
426		M189I/N215R/A249T/G277M							*
427		G70L/G212P							*
428		A68P/E184S/M189E				**			
429		M189Q/G214R		*****					
430		M189Q/A343V/F377Y		*****					
431	131/132	M189Q/N215P/W271Y/L281P/T282C/F377Y		*****					
432		G214R/V280G/T282A/A343V/F377Y/A381L		**					
433		A68P/I69L/M189E/G214R/W271Y/V280G		**					
434		A68P/I69W/M189E/V280G/L281P/T282A/I372L/F377Y		**					
435		A68P/I69L/M189E/G214R/I372L		*					
436		I69L/M189E/W271Y/L281P/T282A		*					
437		I69L/M189Q/F377Y		*					
438		A68P/I69L/M189Q/G214R		*					
439		A68P/I69W/M189I/G214R/F377Y/A381L		*					

Table 11.1: Amidation[1] by *T. fusca* Carboxyesterase Variants Relative to SEQ ID NO: 8 on Substrate Sets 3, 5, 8, and 12-15

Variant No.	SEQ ID NO: (nt/aa)	Amino Acid Mutations	Set 3	Set 5	Set 8	Set 12	Set 13	Set 14	Set 15
440	133/134	A108G/N215P/F377L							*****
441		I69F/N215R				*			***
442		N215K/I372L				*			***
443		V270E/F377L							*
444		A269L/V270E/L281P/I372L/ F377L							*
445		M189I/V270E/I372L							*
446		A108G/A269L/V270E						*	

[1] Levels of increased activity were determined relative to the reference polypeptide and defined as follows: < 1.5x; " * " = > 1.5x, < 3.5x; " ** " = > 3.5x, < 5.5x; " *** " = > 5.5x, < 7.5x; " **** " = > 7.5x, < 9.5x; " ***** " = > 9.5x, <10.5x.

EXAMPLE 12

Analytical Detection of Produced Amides and Precursor Substrates

[0214] Data described in the above Examples were collected using analytical methods in Tables 12.1 through 12.7. The methods provided herein all find use in analyzing the products from the *T. fusca* carboxyesterase variants produced using the present invention. However, it is not intended that the present invention be limited to the methods described herein for the analysis of the products provided herein and/or produced using the methods provided herein. Indeed, any suitable method finds use in the present invention. Product peak elution was confirmed either by confirmation with a commercially available standard or by LC/MS/MS analysis.

Table 12.1: Analytical Method

Instrument	Agilent 1100 Series HPLC
Column	Agilent XDB C-18, 4.6 x 100 mm, 5 µm
Mobile Phase	Gradient I (20% Methanol; 80% Water) Isocratic Gradient
Flow Rate	1.000 mL/min
Run Time	10.0 min
Elution order	Substrate Set #1 – Aniline, acetanilide (**1**) Substrate Set #2 – amide (**2**), methyl phenylacetate

Table 12.1: Analytical Method	
Column Temperature	RT
Injection Volume	10 µL
Detection	UV 254 nm; Detector: MWD

Table 12.2: Analytical Method	
Instrument	Agilent 1100 Series HPLC
Column	Phenomenex Luna, C-18 4.6 x 150 mm, 5 µm
Mobile Phase	Gradient I (A=0.1% TFA in water, B=0.1% TFA in MeCN) Time(min) % A 0.000 100 7.000 5
Flow Rate	1.5 mL/min
Run Time	8.000 min
Elution order	Substrate Set #3 – isobutylbenzamide (**3**), ethyl benzoate Substrate Set #4 – t-butylbenzamide (**4**), ethyl benzoate Substrate Set #6 – t-butoxy-benzamide (**6**), ethyl benzoate Substrate Set #7 – N'-t-butoxycarbonyl-benzhydrazide (**7**), ethyl benzoate
Column Temperature	25 °C
Injection Volume	10 µL
Detection	UV 254 nm; Detector: MWD

Table 12.3: Analytical Method	
Instrument	Agilent 1100 Series HPLC
Column	Phenomenex Luna, C-18 4.6 x 150 mm, 5 µm
Mobile Phase	Gradient I (A=0.1% TFA in water, B=0.1% TFA in MeCN) Time(min) % A 0.000 90 2.750 45 3.150 5
Flow Rate	1.5 mL/min

Table 12.3: Analytical Method	
Run Time	5.000 min
Elution order	Substrate Set #8 – aniline, (amide **8**)
Column Temperature	30 °C
Injection Volume	10 µL
Detection	UV 220 and 260 nm; Detector: MWD

Table 12.4: Analytical Method	
Instrument	Agilent 1100 Series HPLC
Column	Phenomenex Luna, C-18 4.6 x 150 mm, 5 µm
Mobile Phase	Gradient I (A=0.1% TFA in water, B=0.1% TFA in MeCN) Time(min) % A 0.000 90 2.750 30 3.150 5
Flow Rate	1.5 mL/min
Run Time	5.000 min
Elution order	Substrate Set #9 – benzylamine, (amide **9**)
Column Temperature	30 °C
Injection Volume	10 µL
Detection	UV 230 and 260 nm; Detector: MWD

Table 12.5: Analytical Method	
Instrument	Agilent 1100 Series HPLC
Column	Phenomenex Luna, C-18 4.6 x 150 mm, 5 µm
Mobile Phase	Gradient I (A=0.1% TFA in water, B=0.1% TFA in MeCN) Time(min) %A 0.000 95 1.100 75 5.900 30
Flow Rate	1.5 mL/min
Run Time	8.000 min

Table 12.5: Analytical Method

Elution order	Substrate Set #10 – S-phenylethylamine, (amide **10**), ethyl benzoate Substrate Set #11 – R-phenylethylamine, (amide **11**), ethyl benzoate Substrate Set #12 – 4-methylpiperidine, pyrazine ethyl ester, (amide **12**) Substrate Set #13 – S-phenylethylamine, pyrazine ethylester, (amide **13**)
Column Temperature	25 °C
Injection Volume	10 µL
Detection	UV 230 and 260 nm; Detector: MWD

Table 12.6: Analytical Method

Instrument	Agilent 1100 Series HPLC
Column	Phenomenex Luna, C-18 4.6 x 150 mm, 5 µm
Mobile Phase	Gradient I (A=0.1% TFA in water, B=0.1% TFA in MeCN) Time(min) %A 0.000 70 2.750 5 3.750 5
Flow Rate	1.5 mL/min
Run Time	5.800 min
Elution order	Substrate Set #14 – (amide **14**), ethyl-4-indole ester
Column Temperature	25 °C
Injection Volume	10 µL
Detection	UV 230 and 260 nm; Detector: MWD

Table 12.7: Analytical Method

Instrument	Agilent 1200 Series HPLC with CTC-PAL Autosampler; AB Sciex 4000 Q-Trap MS
Column	Agilent Eclipse, C-18 4.6 x 50 mm, 1.6 µm
Mobile Phase	Gradient I (A: 0.1% formic acid in water; B: 0.1% formic acid in MeCN) Time(min) % A 0.000 90 1.000 90 3.000 5 4.000 5

Table 12.7: Analytical Method	
Flow Rate	0.6 mL/min
Run Time	5.000 min
MRM Target Mass	MRM: 204.3 → 105.4; (amide **5**) MRM: 235.4 → 86.4; (amide **15**)
Column Temperature	RT
Injection Volume	10 µL (Samples were quenched by 2-fold dilution in 1:1 MeCN : water)
Detection Parameters	LC/MS/MS analysis Source dependent parameters: CUR: 40; IS: 5500; TEM: 550 °C; GS1: 40; GS2: 40; DP: 120; EP: 10; CE: 27; CXP: 14.

[0215] While various specific embodiments have been illustrated and described, it will be appreciated that various changes can be made without departing from the spirit and scope of the invention(s).

CLAIMS

1. An engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2 or a functional fragment thereof, wherein said engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence, and wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 2.

2. The engineered carboxyesterase of Claim 1, wherein said at least one substitution or substitution set in said polypeptide sequence comprises substitutions at positions selected from: 39, 39/323, 62, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 147, 153, 153/215, 164, 164/271, 174, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224, 224/268/372, 231, 249, 249/284, 263, 268, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 345, 349, 372, 372/376, 373, 374, 376, 377, 405, 416, 420, 427, 428, 429, 438, and 470, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 2.

3. An engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8 or a functional fragment thereof, wherein said engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 8.

4. The engineered carboxyesterase of Claim 3, wherein said at least one substitution or substitution set in said polypeptide sequence comprises substitutions at positions selected from: 63, 63/65/108, 63/65/108/189, 63/65/108/377, 63/65/282/285/320/323, 63/65/320/323, 63/108, 63/108/282/285/377, 63/108/285/377, 63/108/320/323, 63/189, 63/212/215, 63/212/215/268/269/343, 63/215, 63/215/269, 63/215/270/271, 63/215/343, 63/268/269/270/429, 63/377, 65/69/70/281/372, 65/69/70/372, 65/69/372, 65/70/372, 65/320, 65/320/323, 68, 68/69/189/214, 68/69/189/214/215, 68/69/189/214/215/271, 68/69/189/214/215/271/281/282/343/381, 68/69/189/214/271/280, 68/69/189/214/372, 68/69/189/214/377/381, 68/69/189/271, 68/69/189/271/280/372/381,

68/69/189/280/281/282/372/377, 68/69/189/281/282/372/377/381, 68/69/189/343/381, 68/69/189/372, 68/69/189/381, 68/69/214/215/271, 68/69/214/343, 68/69/215, 68/69/271, 68/69/282/287, 68/69/343/372, 68/108/377, 68/184, 68/184/189, 68/189/271/372, 68/189/343, 68/214/215/271/281/282/372, 68/215/271/343/372/381, 68/215/377, 68/271/372, 68/377, 69, 69/70, 69/70/331/372, 69/70/372, 69/70/459, 69/108, 69/108/270/372/377, 69/108/281/285, 69/110/215/281, 69/189, 69/189/214/271/281/282/343, 69/189/214/343/372, 69/189/215/343, 69/189/271, 69/189/271/281/282, 69/189/271/343, 69/189/271/343/381, 69/189/280/282/343/372/381, 69/189/281/373, 69/189/282, 69/189/343/381, 69/189/372, 69/189/372/377, 69/189/377, 69/212/213/215/280/281, 69/214/215/271/372/377/381, 69/214/271/282, 69/214/271/343, 69/215, 69/215/269/270/377, 69/215/271/280/281/282, 69/215/271/282, 69/215/271/372, 69/215/285/317, 69/215/323, 69/215/343, 69/215/343/372/381, 69/271, 69/271/372, 69/282, 69/282/343/372, 69/285/373, 69/372, 70, 70/212, 108, 108/189, 108/189/282/285/320, 108/189/320, 108/189/377, 108/215, 108/215/377, 108/269/270, 108/270, 108/282/285/377, 108/285, 108/320/323, 108/377, 126, 126/184/213/280/281/285/320, 126/184/213/372, 126/189/285/372, 126/215, 126/372, 181/215, 189, 189/214, 189/214/215/271/282, 189/215, 189/215/249/277, 189/215/271/281/282/377, 189/215/343/372, 189/270/285, 189/270/372, 189/280/282, 189/320/377, 189/343, 189/343/377, 189/372/377, 189/377, 189/381, 213/215/320, 214/215/271, 214/215/271/377, 214/271, 214/280/282/343/377/381, 215, 215/249/280/281/285/372, 215/271/372, 215/280/281/285/372, 215/281/285/372, 215/281/285/373, 215/281/373, 215/285, 215/285/317, 215/285/346, 215/285/445, 215/320, 215/320/372, 215/323, 215/372, 215/372/377, 215/373, 215/377, 215/381, 249/377, 269/270/281/372/377, 270/377, 271, 271/343, 271/343/372, 271/343/372/381, 280/285/372, 281/372, 282/285/320/323, 285/323, 320, 343/372, 372, 372/377, 372/381, 373, and 377, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 8.

5. An engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 138 or a functional fragment thereof, wherein said engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence, and wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 138.

6. An engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 140 or a functional fragment thereof, wherein said engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence,

and wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 140.

7. The engineered carboxyesterase of Claim 1, wherein said engineered carboxyesterase comprises an amino acid substitution at position 282, wherein said position is numbered with reference to SEQ ID NO: 2.

8. The engineered carboxyesterase of Claim 7, wherein said amino acid at position 282 is selected from: aliphatic, non-polar, basic, polar, and aromatic amino acids.

9. The engineered carboxyesterase of Claim 7, wherein said engineered carboxyesterase comprises an amino acid substitution selected from: X282T, X282G, X282A, X282V, X282M, X282C, X282W, X282Q, X282S, X282T, and X282R.

10. The engineered carboxyesterase of Claim 1, wherein said polypeptide sequence is selected from: SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, and 136.

11. The engineered carboxyesterase of any of Claims 1-4 and 7-10, wherein said engineered carboxyesterase comprises at least one improved property compared to the wild-type *T. fusca* carboxyesterase of SEQ ID NO:2.

12. The engineered carboxyesterase of Claim 11, wherein said improved property is selected from: improved amidation activity, solvent tolerance, thermostability, pH stability, regioselectivity, stereoselectivity, substrate scope, and/or reduced substrate or product inhibition, and reduced toxicity to bacterial host cells producing said engineered carboxyesterase.

13. The engineered carboxyesterase of Claim 12, wherein said bacterial host cells comprise *E. coli*.

14. The engineered carboxyesterase of Claim 12, wherein said solvent is selected from: acetone, acetonitrile, toluene, tetrahydrofuran, isopropanol, isopropyl acetate, dimethyl sulfoxide, and methyl ethyl ketone.

15. The engineered carboxyesterase of any of Claims 1-14, wherein said engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase on at least one substrate selected from: aniline; isobutylamine; n-butylamine; t-butylamine; *N*'-t-butoxycarbonyl-benzhydrazide; 4-methylpiperidine; *O*-t-butylhdroxylamine; benzylamine; 2,6-dimethylaniline; (*S*)-(-)-α-methylbenzylamine; (*R*)-(+)-α-methylbenzylamine; methyl phenylacetate; ethyl acetate; ethyl benzoate; 2-pyrazinyl ethyl ester; 4-ethyl-*1H*-indole ester; and *N,N*-diethylglycyl methyl ester.

16. The engineered carboxyesterase of any of Claims 1-15, wherein said engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase on at least one substrate or substrate set selected from: aniline; isobutylamine; n-butylamine; t-butylamine; *N*'-t-butoxycarbonyl-benzhydrazide; 4-methylpiperidine; *O*-t-butylhdroxylamine; benzylamine; 2,6-dimethylaniline; (*S*)-(-)-α-methylbenzylamine; (*R*)-(+)-α-methylbenzylamine; methyl phenylacetate; ethyl acetate; ethyl benzoate; 2-pyrazinyl ethyl ester; 4-ethyl-*1H*-indole ester; and *N,N*-diethylglycyl methyl ester.

17. The engineered carboxyesterase of any of Claims 1-16, wherein said engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase in producing at least one product selected from: acetanilide; *N*-n-butyl-benzylacetamide; *N*-[(*S*)-1-phenylethyl]-pyrazinylamide; *N*-[(*S*)-1-phenylethyl]-benzamide; *N*-[(*R*)-1-phenylethyl]-benzamide; *N'*-t-butoxycarbonyl-benzhydrazide; 1-benzoyl-4-methylpiperidine; 2-pyrazinyl-4-methylpiperidine; *N*-isobutyl-benzamide; *N*-t-butyl-benzamide; *N*-t-butylhydroxyl-benzamide; *N*-isobutyl-*1H*-indol-4-amide; *N',N'*-(diethylamino)-*N*-phenylacetamide; *N',N'*-(diethylamino)-*N*-benzylacetamide; and *N',N'*-(diethylamino)-*N*-2,6-dimethylphenylacetamide.

18. The engineered carboxyesterase of Claim 1, comprising at least one substitution selected from: X343V, X372L, X320W/G, X214R, X282C, X271Y, X65G, wherein said engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase on a hindered amine for formation of *N',N'*-(diethylamino)-*N*-2,6-dimethylphenylacetamide from ethyl benzoate; and 2,6-dimethylaniline.

19. The engineered carboxyesterase of Claim 1, comprising at least one substitution selected from: X268A, X63A/R, X189Q/I/E, X214R, X282G/C, X381L, X69W, wherein said engineered carboxyesterase exhibits greater activity than wild-type *T. fusca* carboxyesterase on a secondary amine for formation of 1-benzoyl-4-methyl-piperidine from ethyl benzoate and 4-methyl-piperidine.

20. The engineered carboxyesterase of any of Claims 1 - 19, wherein said engineered carboxyesterase is purified.

21. The engineered carboxyesterase of any of Claims 1 - 20, wherein said engineered carboxyesterase is immobilized.

22. A composition comprising at least one engineered carboxyesterase of any of Claims 1-21.

23. A polynucleotide sequence encoding at least one engineered carboxyesterase of any of Claims 1–19.

24. A polynucleotide sequence encoding at least one engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2 or a functional fragment thereof, wherein said engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 2.

25. The polynucleotide sequence of Claim 24, wherein said encoded at least one engineered carboxyesterase comprises at least one substitution or substitution set at positions selected from: 39, 39/323, 62, 62/117, 63, 64, 65, 66, 68, 69, 70, 71, 71/263, 77, 77/184, 103, 103/147, 104, 104/429, 105, 107, 107/185, 108, 109, 109/117, 110, 111, 113, 114, 115, 117, 118, 118/269, 118/349, 119, 126, 147, 153, 153/215, 164, 164/271, 174, 174/282, 183, 184, 184/249, 185, 186, 187, 188, 190, 209, 210, 211, 212, 213, 213/271, 213/345, 214, 215, 215/271, 216, 217, 217/231, 224, 224/268/372, 231, 249, 249/284, 263, 268, 269, 270, 270/470, 271, 271/416, 276, 277, 278, 279, 279/280/282, 280, 281, 281/374, 282, 283, 283/429, 284, 284/438, 285, 286, 311, 317, 320, 320/323, 320/323/372, 320/372/376, 320/376/377, 321, 323, 324, 345, 349, 372, 372/376, 373, 374, 376, 377, 405, 416, 420, 427, 428, 429, 438, and 470, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 2.

26. A polynucleotide sequence encoding at least one engineered carboxyesterase comprising a polypeptide sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8 or a functional fragment thereof, wherein said engineered carboxyesterase comprises at least one

substitution or substitution set in said polypeptide sequence, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 8.

27. The polynucleotide sequence of Claim 26, wherein said encoded engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence at positions selected from: 63, 63/65/108, 63/65/108/189, 63/65/108/377, 63/65/282/285/320/323, 63/65/320/323, 63/108, 63/108/282/285/377, 63/108/285/377, 63/108/320/323, 63/189, 63/212/215, 63/212/215/268/269/343, 63/215, 63/215/269, 63/215/270/271, 63/215/343, 63/268/269/270/429, 63/377, 65/69/70/281/372, 65/69/70/372, 65/69/372, 65/70/372, 65/320, 65/320/323, 68, 68/69/189/214, 68/69/189/214/215, 68/69/189/214/215/271, 68/69/189/214/215/271/281/282/343/381, 68/69/189/214/271/280, 68/69/189/214/372, 68/69/189/214/377/381, 68/69/189/271, 68/69/189/271/280/372/381, 68/69/189/280/281/282/372/377, 68/69/189/281/282/372/377/381, 68/69/189/343/381, 68/69/189/372, 68/69/189/381, 68/69/214/215/271, 68/69/214/343, 68/69/215, 68/69/271, 68/69/282/287, 68/69/343/372, 68/108/377, 68/184, 68/184/189, 68/189/271/372, 68/189/343, 68/214/215/271/281/282/372, 68/215/271/343/372/381, 68/215/377, 68/271/372, 68/377, 69, 69/70, 69/70/331/372, 69/70/372, 69/70/459, 69/108, 69/108/270/372/377, 69/108/281/285, 69/110/215/281, 69/189, 69/189/214/271/281/282/343, 69/189/214/343/372, 69/189/215/343, 69/189/271, 69/189/271/281/282, 69/189/271/343, 69/189/271/343/381, 69/189/280/282/343/372/381, 69/189/281/373, 69/189/282, 69/189/343/381, 69/189/372, 69/189/372/377, 69/189/377, 69/212/213/215/280/281, 69/214/215/271/372/377/381, 69/214/271/282, 69/214/271/343, 69/215, 69/215/269/270/377, 69/215/271/280/281/282, 69/215/271/282, 69/215/271/372, 69/215/285/317, 69/215/323, 69/215/343, 69/215/343/372/381, 69/271, 69/271/372, 69/282, 69/282/343/372, 69/285/373, 69/372, 70, 70/212, 108, 108/189, 108/189/282/285/320, 108/189/320, 108/189/377, 108/215, 108/215/377, 108/269/270, 108/270, 108/282/285/377, 108/285, 108/320/323, 108/377, 126, 126/184/213/280/281/285/320, 126/184/213/372, 126/189/285/372, 126/215, 126/372, 181/215, 189, 189/214, 189/214/215/271/282, 189/215, 189/215/249/277, 189/215/271/281/282/377, 189/215/343/372, 189/270/285, 189/270/372, 189/280/282, 189/320/377, 189/343, 189/343/377, 189/372/377, 189/377, 189/381, 213/215/320, 214/215/271, 214/215/271/377, 214/271, 214/280/282/343/377/381, 215, 215/249/280/281/285/372, 215/271/372, 215/280/281/285/372, 215/281/285/372, 215/281/285/373, 215/281/373, 215/285, 215/285/317, 215/285/346, 215/285/445, 215/320, 215/320/372, 215/323, 215/372, 215/372/377, 215/373, 215/377, 215/381, 249/377, 269/270/281/372/377, 270/377, 271, 271/343, 271/343/372, 271/343/372/381, 280/285/372, 281/372, 282/285/320/323, 285/323, 320, 343/372, 372, 372/377, 372/381, 373, and 377, wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 8.

28. A polynucleotide sequence encoding at least one engineered carboxyesterase or a functional fragment thereof, said polynucleotide sequence comprising at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%, sequence identity to SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131,133, 135, 137, and/or 139.

29. The polynucleotide sequence of any of Claims 23-28, wherein said polynucleotide sequence is operably linked to a control sequence.

30. The polynucleotide sequence of any of Claims 23-29, wherein said polynucleotide sequence is codon optimized.

31. An expression vector comprising at least one polynucleotide sequence of any of Claims 23-30.

32. A host cell comprising at least one expression vector of Claim 31.

33. A host cell comprising at least one polynucleotide sequence of any of Claims 23-30.

34. The host cell of Claim 32 and/or 33, wherein said host cell is *E. coli*.

35. A method of producing an engineered carboxyesterase in a host cell, comprising culturing the host cell of any of Claims 32 -34, under suitable conditions, such that at least one engineered carboxyesterase is produced.

36. The method of Claim 35, further comprising recovering at least one engineered carboxyesterase from the culture and/or host cell.

37. The method of Claim 35 and/or 36, further comprising the step of purifying said at least one engineered carboxyesterase.

Figure 1. Pairwise alignment of the Polynucleotide Sequence Encoding the *E. coli* codon optimized for the Wild-Type Carboxyesterase Enzyme, *Thermobifida fusca* (*T. fusca*) (SEQ ID NO: 1) against the polynucleotide sequences of *T. fusca* variants

```
                       10         20         30         40         50         60         70
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1    1  atggagatcg tgattcgcac gggcagcgga gatgtgcgtg gtagtaaaga gaacggcatt gcagtgtttc  70
SEQIDNO:3    1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:5    1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:7    1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:9    1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:11   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:13   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:15   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:17   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:19   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:21   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:23   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:25   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:27   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:29   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:31   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:33   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:35   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:37   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:39   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:41   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:43   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:45   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:47   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:49   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:51   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:53   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:55   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:57   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:59   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:61   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:63   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:65   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:67   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:69   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:71   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:73   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:75   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:77   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:79   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:81   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:83   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:85   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:87   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:89   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:91   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:93   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:95   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:97   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:99   1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:101  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:103  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:105  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:107  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:109  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:111  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:113  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:115  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:117  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:119  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:121  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:123  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:125  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:127  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:129  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:131  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:133  1  .......... .......... .......... .......... .......... .......... ..........  70
SEQIDNO:135  1  .......... .......... .......... .......... .......... .......... ..........  70
```

Figure 1, Continued

```
                             80         90        100        110        120        130        140
                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1      71    gtggcatccc ctatgcggaa ccaccggttg gcgctcaccg ctttacggcg ccacgtccac cacgtccatg  140
SEQIDNO:3      71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:5      71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:7      71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:9      71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:11     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:13     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:15     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:17     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:19     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:21     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:23     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:25     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:27     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:29     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:31     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:33     71    .......... .......... .......... .t........ .......... .......... ..........  140
SEQIDNO:35     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:37     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:39     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:41     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:43     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:45     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:47     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:49     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:51     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:53     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:55     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:57     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:59     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:61     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:63     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:65     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:67     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:69     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:71     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:73     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:75     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:77     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:79     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:81     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:83     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:85     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:87     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:89     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:91     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:93     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:95     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:97     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:99     71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:101    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:103    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:105    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:107    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:109    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:111    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:113    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:115    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:117    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:119    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:121    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:123    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:125    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:127    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:129    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:131    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:133    71    .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:135    71    .......... .......... .......... .......... .......... .......... ..........  140
```

Figure 1, Continued

```
                            150        160        170        180        190        200        210
                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1      141   ggatggtgtt cgtgacgcta ctgaattctc ggcgactgca ccacgtccac catatccaga ggcaattggt  210
SEQIDNO:3      141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:5      141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:7      141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:9      141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:11     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:13     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:15     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:17     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:19     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:21     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:23     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:25     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:27     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:29     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:31     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:33     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:35     141   .......... .......... .......... .......... .......gg. .......... ..........  210
SEQIDNO:37     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:39     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:41     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:43     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:45     141   .......... .......... .......... .......... .......... .......... .c.t......  210
SEQIDNO:47     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:49     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:51     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:53     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:55     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:57     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:59     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:61     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:63     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:65     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:67     141   .......... .......... .......... .......... .......... ...gg..... ....c.....  210
SEQIDNO:69     141   .......... .......... .......... .......... .......... .......... .c.gc.g...  210
SEQIDNO:71     141   .......... .......... .......... .......... .......... .......... ....c.g...  210
SEQIDNO:73     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:75     141   .......... .......... .......... .......... .......gg. ....c..... ..........  210
SEQIDNO:77     141   .......... .......... .......... .......... .......... ...gg..... ......g.c.  210
SEQIDNO:79     141   .......... .......... .......... .......... .......... .......... ....tgg...  210
SEQIDNO:81     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:83     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:85     141   .......... .......... .......... .......... .......... .......... ....c.g...  210
SEQIDNO:87     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:89     141   .......... .......... .......... .......... .......gg. ..gga..... ..........  210
SEQIDNO:91     141   .......... .......... .......... .......... .......... .......... .c.g......  210
SEQIDNO:93     141   .......... .......... .......... .......... ........g. ....c..... ..........  210
SEQIDNO:95     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:97     141   .......... .......... .......... .......... .......... .......... .c.g......  210
SEQIDNO:99     141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:101    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:103    141   .......... .......... .......... .......... ......a.g. .......... ..........  210
SEQIDNO:105    141   .......... .......... .......... .......... .......... .......... ....tgg...  210
SEQIDNO:107    141   .......... .......... .......... .......... ......g.g. .......... ..........  210
SEQIDNO:109    141   .......... .......... .......... .......... ......tat. .......... ..........  210
SEQIDNO:111    141   .......... .......... .......... .......... ......tat. .......... ..........  210
SEQIDNO:113    141   .......... .......... .......... .......... ......tat. .......... ..........  210
SEQIDNO:115    141   .......... .......... .......... .......... .......... .......... .c.gtgg...  210
SEQIDNO:117    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:119    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:121    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:123    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:125    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:127    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:129    141   .......... .......... .......... .......... .......... .......... ....t.....  210
SEQIDNO:131    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:133    141   .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:135    141   .......... .......... .......... .......... .......... .......... ..........  210
```

Figure 1, Continued

```
                           220        230        240        250        260        270        280
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     211   gcgttgctga tcgaacgctt cattcctggc gacgattacc tgaccctgaa cgtatggact ccggacccga  280
SEQIDNO:3     211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:5     211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:7     211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:9     211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:11    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:13    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:15    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:17    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:19    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:21    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:23    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:25    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:27    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:29    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:31    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:33    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:35    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:37    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:39    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:41    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:43    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:45    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:47    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:49    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:51    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:53    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:55    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:57    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:59    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:61    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:63    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:65    211   .......... .........c .......... .......... .......... .......... ..........  280
SEQIDNO:67    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:69    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:71    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:73    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:75    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:77    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:79    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:81    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:83    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:85    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:87    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:89    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:91    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:93    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:95    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:97    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:99    211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:101   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:103   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:105   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:107   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:109   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:111   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:113   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:115   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:117   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:119   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:121   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:123   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:125   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:127   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:129   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:131   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:133   211   .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:135   211   .......... .......... .......... .......... .......... .......... ..........  280
```

Figure 1, Continued

```
                        290        300        310        320        330        340        350
                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1    281 atgcagttgg tctgccagtc atggtgtgga ttcatggtgg tgcctttact aacggtagtg gtagcgaacc 350
SEQIDNO:3    281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:5    281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:7    281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:9    281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:11   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:13   281 .......... .......... .......... .......... ..gg...... .......... .......... 350
SEQIDNO:15   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:17   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:19   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:21   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:23   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:25   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:27   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:29   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:31   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:33   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:35   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:37   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:39   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:41   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:43   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:45   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:47   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:49   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:51   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:53   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:55   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:57   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:59   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:61   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:63   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:65   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:67   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:69   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:71   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:73   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:75   281 .......... .......... .......... .......... ..g....... .......... .......... 350
SEQIDNO:77   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:79   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:81   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:83   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:85   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:87   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:89   281 .......... .......... .......... .......... ..g....... .......... .......... 350
SEQIDNO:91   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:93   281 .......... .......... .......... .......... ..g....... .......... .......... 350
SEQIDNO:95   281 .......... .......... .......... .......... ..gg...... .......... .......... 350
SEQIDNO:97   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:99   281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:101  281 .......... .......... .......... .......... ...a...... .......... .......... 350
SEQIDNO:103  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:105  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:107  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:109  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:111  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:113  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:115  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:117  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:119  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:121  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:123  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:125  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:127  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:129  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:131  281 .......... .......... .......... .......... .......... .......... .......... 350
SEQIDNO:133  281 .......... .......... .......... .......... ..gg...... .......... .......... 350
SEQIDNO:135  281 .......... .......... .......... .......... .......... .......... .......... 350
```

Figure 1, Continued

```
                              360        370        380        390        400        410        420
                      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1       351   ggtgtatgac ggtgcagcat ttgcacgtga tggtgtggtg tttgtgagct tcaactatcg tctgggcatc  420
SEQIDNO:3       351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:5       351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:7       351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:9       351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:11      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:13      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:15      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:17      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:19      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:21      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:23      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:25      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:27      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:29      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:31      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:33      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:35      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:37      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:39      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:41      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:43      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:45      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:47      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:49      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:51      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:53      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:55      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:57      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:59      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:61      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:63      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:65      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:67      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:69      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:71      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:73      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:75      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:77      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:79      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:81      351   .......... .......... .......c.. .......... .......... .......... ..........  420
SEQIDNO:83      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:85      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:87      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:89      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:91      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:93      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:95      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:97      351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:99      351   .......... .......... .......c.. .......... .......... .......... ..........  420
SEQIDNO:101     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:103     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:105     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:107     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:109     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:111     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:113     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:115     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:117     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:119     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:121     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:123     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:125     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:127     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:129     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:131     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:133     351   .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:135     351   .......... .......... .......... .......... .......... .......... ..........  420
```

Figure 1, Continued

```
                          430        440        450        460        470        480        490
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1    421  attggctttg cagatctgcc agatgcacca agtaatcgcg gcctgttaga tcaaatcgcc gcactggaat  490
SEQIDNO:3    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:5    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:7    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:9    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:11   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:13   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:15   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:17   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:19   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:21   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:23   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:25   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:27   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:29   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:31   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:33   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:35   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:37   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:39   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:41   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:43   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:45   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:47   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:49   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:51   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:53   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:55   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:57   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:59   421  .......... .......... .......... .......a.. .......... .......... ..........  490
SEQIDNO:61   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:63   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:65   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:67   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:69   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:71   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:73   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:75   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:77   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:79   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:81   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:83   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:85   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:87   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:89   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:91   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:93   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:95   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:97   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:99   421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:101  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:103  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:105  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:107  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:109  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:111  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:113  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:115  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:117  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:119  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:121  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:123  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:125  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:127  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:129  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:131  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:133  421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:135  421  .......... .......... .......... .......... .......... .......... ..........  490
```

Figure 1, Continued

```
                            500        510        520        530        540        550        560
                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1       491  gggttcgcga caacattgca cgctttgggg gtgacccagg taatgtgacc gtgtttggcg aaagtgcggg  560
SEQIDNO:3       491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:5       491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:7       491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:9       491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:11      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:13      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:15      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:17      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:19      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:21      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:23      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:25      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:27      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:29      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:31      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:33      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:35      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:37      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:39      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:41      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:43      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:45      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:47      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:49      491  .......... .......... .......... .......... .......... .........t ct........  560
SEQIDNO:51      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:53      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:55      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:57      491  .......... .......... .......... .......... .......... .......... gg........  560
SEQIDNO:59      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:61      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:63      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:65      491  .......... .......... .......... .......... .......... .......... gg........  560
SEQIDNO:67      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:69      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:71      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:73      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:75      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:77      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:79      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:81      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:83      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:85      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:87      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:89      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:91      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:93      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:95      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:97      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:99      491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:101     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:103     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:105     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:107     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:109     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:111     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:113     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:115     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:117     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:119     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:121     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:123     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:125     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:127     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:129     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:131     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:133     491  .......... .......... .......... .......... .......... .......... ..........  560
SEQIDNO:135     491  .......... .......... .......... .......... .......... .......... ..........  560
```

Figure 1, Continued

```
                            570        580        590        600        610        620        630
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     561   cgcgatgagc gtgtgcaccc tgatggcaac cccgcgtgca cgtggtctgt ttcgtcgtgc tatcttacag   630
SEQIDNO:3     561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:5     561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:7     561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:9     561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:11    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:13    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:15    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:17    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:19    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:21    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:23    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:25    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:27    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:29    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:31    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:33    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:35    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:37    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:39    561   ........ag .......... .......... .......... .......... .......... ..........   630
SEQIDNO:41    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:43    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:45    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:47    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:49    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:51    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:53    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:55    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:57    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:59    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:61    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:63    561   .......t.g .......... .......... .......... .......... .......... ..........   630
SEQIDNO:65    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:67    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:69    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:71    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:73    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:75    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:77    561   .......tca .......... .......... .......... .......... .......... ..........   630
SEQIDNO:79    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:81    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:83    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:85    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:87    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:89    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:91    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:93    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:95    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:97    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:99    561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:101   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:103   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:105   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:107   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:109   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:111   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:113   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:115   561   ......t... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:117   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:119   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:121   561   ......t... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:123   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:125   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:127   561   ......t... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:129   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:131   561   ....ca.... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:133   561   .......... .......... .......... .......... .......... .......... ..........   630
SEQIDNO:135   561   .......... .......... .......... .......... .......... .......... ..........   630
```

Figure 1, Continued

```
                          640        650        660        670        680        690        700
                   ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     631  agcggcgcag gcaacatggc agtcgctgca gaagatgcta ccacaatcgc agctgtgatt gcccatcgtt  700
SEQIDNO:3     631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:5     631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:7     631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:9     631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:11    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:13    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:15    631  .......... ...gg..... .......... .......... .......... .......... ..........  700
SEQIDNO:17    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:19    631  .......... ..tgg..... .......... .......... .......... .......... ..........  700
SEQIDNO:21    631  ...cct.... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:23    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:25    631  .......... ..ccg..... .......... .......... .......... .......... ..........  700
SEQIDNO:27    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:29    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:31    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:33    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:35    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:37    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:39    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:41    631  .......... ...gg..... .......... .......... .......... .......... ..........  700
SEQIDNO:43    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:45    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:47    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:49    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:51    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:53    631  ......ctg. .......... .......... .......... .......... .......... ..........  700
SEQIDNO:55    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:57    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:59    631  .......... ..cct..... .......... .......... .......... .......... ..........  700
SEQIDNO:61    631  .........c tt........ .......... .......... .......... .......... ..........  700
SEQIDNO:63    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:65    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:67    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:69    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:71    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:73    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:75    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:77    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:79    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:81    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:83    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:85    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:87    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:89    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:91    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:93    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:95    631  .......... ....g..... .......... .......... .......... .......... ..........  700
SEQIDNO:97    631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:99    631  ........g. ..cga..... .......... .......... .......... .......... ..........  700
SEQIDNO:101   631  .......... ....g..... .......... .......... .......... .......... ..........  700
SEQIDNO:103   631  .....a.... ..cg...... .......... .......... .......... .......... ..........  700
SEQIDNO:105   631  .......... ...gg..... .......... .......... .......... .......... ..........  700
SEQIDNO:107   631  .....a.... ..cg...... .......... .......... .......... .......... ..........  700
SEQIDNO:109   631  ...cc..... ..cg...... .......... .......... .......... .......... ..........  700
SEQIDNO:111   631  ...cc..... ..cg...... .......... .......... .......... .......... ..........  700
SEQIDNO:113   631  .....a.... ..cg...... .......... .......... .......... .......... ..........  700
SEQIDNO:115   631  .........c .tccg..... .......... .......... .......... .......... ..........  700
SEQIDNO:117   631  .......... ....a..... .......... .......... .......... .......... ..........  700
SEQIDNO:119   631  .......... ....g..... .......... .......... .......... .......... ..........  700
SEQIDNO:121   631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:123   631  .......... ...gg..... .......... .......... .......... .......... ..........  700
SEQIDNO:125   631  .......... ..tgg..... .......... .......... .......... .......... ..........  700
SEQIDNO:127   631  .......... .......... .......... .......... .......... .......... ..........  700
SEQIDNO:129   631  .......... ....g..... .......... .......... .......... .......... ..........  700
SEQIDNO:131   631  .......... ..ccg..... .......... .......... .......... .......... ..........  700
SEQIDNO:133   631  .......... ..ccg..... .......... .......... .......... .......... ..........  700
SEQIDNO:135   631  .......... .......... .......... .......... .......... .......... ..........  700
```

Figure 1, Continued

```
                            710        720        730        740        750        760        770
                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1      701   taggtgttga gccaactgca gcagcattag cgcatgttcc agttgcaceg ttactggatg ttcagcagca   770
SEQIDNO:3      701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:5      701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:7      701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:9      701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:11     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:13     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:15     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:17     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:19     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:21     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:23     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:25     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:27     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:29     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:31     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:33     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:35     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:37     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:39     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:41     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:43     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:45     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:47     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:49     701   .......... .......... .......... .......... ....a..... .......... ..........   770
SEQIDNO:51     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:53     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:55     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:57     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:59     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:61     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:63     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:65     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:67     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:69     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:71     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:73     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:75     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:77     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:79     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:81     701   .......... .......... .......... .......... ......c... .......... ..........   770
SEQIDNO:83     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:85     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:87     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:89     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:91     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:93     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:95     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:97     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:99     701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:101    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:103    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:105    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:107    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:109    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:111    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:113    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:115    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:117    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:119    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:121    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:123    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:125    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:127    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:129    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:131    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:133    701   .......... .......... .......... .......... .......... .......... ..........   770
SEQIDNO:135    701   .......... .......... .......... .......... .......... .......... ..........   770
```

Figure 1, Continued

```
                           780        790        800        810        820        830        840
                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     771    agttgcccaa gagattcaag gcgcgccaga tccagcagtt tggggtgaac gtattgcggg cggcagtgtg  840
SEQIDNO:3     771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:5     771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:7     771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:9     771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:11    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:13    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:15    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:17    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:19    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:21    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:23    771    .......... .......... .......... .......... .......... .......... ........g.  840
SEQIDNO:25    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:27    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:29    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:31    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:33    771    .......... .......... .......... .......... act....... .......... ..........  840
SEQIDNO:35    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:37    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:39    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:41    771    .......... .......... .......... .......... a......... .......... ..........  840
SEQIDNO:43    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:45    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:47    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:49    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:51    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:53    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:55    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:57    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:59    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:61    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:63    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:65    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:67    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:69    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:71    771    .......... .......... .......... .......... .at....... .......... ..........  840
SEQIDNO:73    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:75    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:77    771    .......... .......... .......... .......... .......... .......... .........t  840
SEQIDNO:79    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:81    771    .......... .......... .......... .......... .......... .......... ........gt  840
SEQIDNO:83    771    .......... .......... .......... .......... .at....... .......... ..........  840
SEQIDNO:85    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:87    771    .......... .......... .......... .......... .......... .......... .........t  840
SEQIDNO:89    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:91    771    .......... .......... .......... .......... .at....... .......... ..........  840
SEQIDNO:93    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:95    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:97    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:99    771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:101   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:103   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:105   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:107   771    .......... .......... .......... ...g...... .......... .......... ..........  840
SEQIDNO:109   771    .......... .......... .......... ...g...... .......... .......... ..........  840
SEQIDNO:111   771    .......... .......... .......... .g.gaat... .......... .......... ..........  840
SEQIDNO:113   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:115   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:117   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:119   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:121   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:123   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:125   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:127   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:129   771    .......... .......... .......... ....tt.a.c .......... .......... ..........  840
SEQIDNO:131   771    .......... .......... .......... .......... .at....... .......... ..........  840
SEQIDNO:133   771    .......... .......... .......... .......... .......... .......... ..........  840
SEQIDNO:135   771    .......... .......... .......... .......... .......... .......... ..........  840
```

Figure 1, Continued

```
                          850        860        870        880        890        900        910
                   ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     841  ctgctgccat ttgctcccgt cattgatggg gaactgctga gccaacgtcc agcggaagcc attgcgggcg  910
SEQIDNO:3     841  ....a..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:5     841  ...gct.... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:7     841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:9     841  .......... ..ctg..... .......... .......... .......... .......... ..........  910
SEQIDNO:11    841  ...tgt.... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:13    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:15    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:17    841  .ct....... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:19    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:21    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:23    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:25    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:27    841  ......gat. .......... .......... .......... .......... .......... ..........  910
SEQIDNO:29    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:31    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:33    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:35    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:37    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:39    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:41    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:43    841  ......gtt. .......... .......... .......... .......... .......... ..........  910
SEQIDNO:45    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:47    841  ......tat. .......... .......... .......... .......... .......... ..........  910
SEQIDNO:49    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:51    841  .........a c......... .......... .......... .......... .......... ..........  910
SEQIDNO:53    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:55    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:57    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:59    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:61    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:63    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:65    841  .......... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:67    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:69    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:71    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:73    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:75    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:77    841  .c.ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:79    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:81    841  ...ac..... ..ctg..... .......... .......... .......... .......... ..........  910
SEQIDNO:83    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:85    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:87    841  .c.ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:89    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:91    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:93    841  ...acc.... ..tta..... .......... .......... .......... .......... ..........  910
SEQIDNO:95    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:97    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:99    841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:101   841  .c.ac..... ..tta..... .......... .......... .......... .......... ..........  910
SEQIDNO:103   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:105   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:107   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:109   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:111   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:113   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:115   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:117   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:119   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:121   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:123   841  .c.ac..... ..ctg..... .......... .......... .......... .......... ..........  910
SEQIDNO:125   841  ...ac..... ..ctg..... .......... .......... .......... .......... ..........  910
SEQIDNO:127   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:129   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:131   841  .c.tgc.... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:133   841  ...ac..... .......... .......... .......... .......... .......... ..........  910
SEQIDNO:135   841  .......... .......... .......... .......... .......... .......... ..........  910
```

Figure 1, Continued

```
                                920        930        940        950        960        970        980
                         ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1        911     gcgcaggtca tgatgtagac ctcttgtttg gcactaccac cgatgaatcc cgtctgtttc tggcaccaac   980
SEQIDNO:3        911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:5        911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:7        911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:9        911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:11       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:13       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:15       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:17       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:19       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:21       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:23       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:25       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:27       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:29       911     .......... .......... .......... .......... .......ggt .......... ..........   980
SEQIDNO:31       911     .......... .......... .......... .......... .......... ttg....... ..........   980
SEQIDNO:33       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:35       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:37       911     .......... .......... .......... .......... ........gg .......... ..........   980
SEQIDNO:39       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:41       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:43       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:45       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:47       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:49       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:51       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:53       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:55       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:57       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:59       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:61       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:63       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:65       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:67       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:69       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:71       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:73       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:75       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:77       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:79       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:81       911     .......... .......... .......... .......... .........t .......... ..........   980
SEQIDNO:83       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:85       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:87       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:89       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:91       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:93       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:95       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:97       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:99       911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:101      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:103      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:105      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:107      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:109      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:111      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:113      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:115      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:117      911     .......... .......... .......... .......... g......... .......... ..........   980
SEQIDNO:119      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:121      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:123      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:125      911     .......... .......... .......... .......... g......... .......... ..........   980
SEQIDNO:127      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:129      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:131      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:133      911     .......... .......... .......... .......... .......... .......... ..........   980
SEQIDNO:135      911     .......... .......... .......... .......... .......... .......... ..........   980
```

Figure 1, Continued

```
                              990        1000       1010       1020       1030       1040       1050
                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1        981    tgggctgctg ccgttcatca caggggacta tgttaccact cacttagcca agagcggttt agatgcagat  1050
SEQIDNO:3        981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:5        981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:7        981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:9        981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:11       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:13       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:15       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:17       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:19       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:21       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:23       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:25       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:27       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:29       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:31       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:33       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:35       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:37       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:39       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:41       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:43       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:45       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:47       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:49       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:51       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:53       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:55       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:57       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:59       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:61       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:63       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:65       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:67       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:69       981    .......... .......... .......... .......... .......tg. .......... ..........  1050
SEQIDNO:71       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:73       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:75       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:77       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:79       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:81       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:83       981    .......... .......... .......... .......... .......tg. .......... ..........  1050
SEQIDNO:85       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:87       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:89       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:91       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:93       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:95       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:97       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:99       981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:101      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:103      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:105      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:107      981    .......... .......... .......... .......... .......tg. .......... ..........  1050
SEQIDNO:109      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:111      981    .......... .......... .......... .......... .......tg. .......... ..........  1050
SEQIDNO:113      981    .......... .......... .......... .......... ........g. .......... ..........  1050
SEQIDNO:115      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:117      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:119      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:121      981    .......... .......... .......... .......... .......tg. .......... ..........  1050
SEQIDNO:123      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:125      981    .......... .......... .......... .......... .......... .....a.... ..........  1050
SEQIDNO:127      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:129      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:131      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:133      981    .......... .......... .......... .......... .......... .......... ..........  1050
SEQIDNO:135      981    .......... .......... .......... .......... .......... .......... ..........  1050
```

Figure 1, Continued

```
                            1060       1070       1080       1090       1100       1110       1120
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1      1051 gcggccaaag cgtataccgc ggaaggtcgt ggtgaagaac cgggcgacat cttggccagc atcatcaccg  1120
SEQIDNO:3      1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:5      1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:7      1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:9      1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:11     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:13     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:15     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:17     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:19     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:21     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:23     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:25     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:27     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:29     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:31     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:33     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:35     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:37     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:39     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:41     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:43     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:45     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:47     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:49     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:51     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:53     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:55     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:57     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:59     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:61     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:63     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:65     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:67     1051 .......... .......... .......... .......... .......... .......... ...c.g....  1120
SEQIDNO:69     1051 .......... .......... .......... .......... .......... .......... ...c.g....  1120
SEQIDNO:71     1051 .......... .......... .......... .......... .......... .......... ...c.g....  1120
SEQIDNO:73     1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
SEQIDNO:75     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:77     1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
SEQIDNO:79     1051 .......... .......... .......... .......... .......... .......... .....g....  1120
SEQIDNO:81     1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
SEQIDNO:83     1051 .......... .......... .......... .......... .......... .......... ...c.g....  1120
SEQIDNO:85     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:87     1051 .......... .......... .......... .......... .......... .......... ...c.g....  1120
SEQIDNO:89     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:91     1051 .......... .......... .......... .......... .......... .......... ...c.g....  1120
SEQIDNO:93     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:95     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:97     1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:99     1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
SEQIDNO:101    1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
SEQIDNO:103    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:105    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:107    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:109    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:111    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:113    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:115    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:117    1051 .......... .......... .......... .......... .......... .......... ......ggg.  1120
SEQIDNO:119    1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
SEQIDNO:121    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:123    1051 .......... .......... .......... .......... .......... .......... ......ggg.  1120
SEQIDNO:125    1051 .......... .......... .......... .......... .......... .......... ........t.  1120
SEQIDNO:127    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:129    1051 .......... .......... .......... .......... .......... .......... .....t....  1120
SEQIDNO:131    1051 .......... .......... .......... .......... .......... .......... ..........  1120
SEQIDNO:133    1051 .......... .......... .......... .......... .......... .......... .....t....  1120
SEQIDNO:135    1051 .......... .......... .......... .......... .......... .......... ...c.t....  1120
```

Figure 1, Continued

```
                              1130       1140       1150       1160       1170       1180       1190
                      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1       1121 accaggtgtt tcgtattccg gcgctgcgta ttgcagaatc ccgtgttgat gcgcctgcac gtacccttgg  1190
SEQIDNO:3       1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:5       1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:7       1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:9       1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:11      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:13      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:15      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:17      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:19      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:21      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:23      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:25      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:27      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:29      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:31      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:33      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:35      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:37      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:39      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:41      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:43      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:45      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:47      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:49      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:51      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:53      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:55      1121 ........c. .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:57      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:59      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:61      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:63      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:65      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:67      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:69      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:71      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:73      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:75      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:77      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:79      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:81      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:83      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:85      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:87      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:89      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:91      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:93      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:95      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:97      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:99      1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:101     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:103     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:105     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:107     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:109     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:111     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:113     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:115     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:117     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:119     1121 .......... g......... .......... .......... .......... .......... ..........  1190
SEQIDNO:121     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:123     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:125     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:127     1121 .......... .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:129     1121 .......... g......... .......... .......... .......... .......... ..........  1190
SEQIDNO:131     1121 .........a .......... .......... .......... .......... .......... ..........  1190
SEQIDNO:133     1121 .......... g......... .......... .......... .......... .......... ..........  1190
SEQIDNO:135     1121 .......... .......... .......... .......... .......... .......... ..........  1190
```

Figure 1, Continued

```
                        1200       1210       1220       1230       1240       1250       1260
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1      1191 ctacgaattc gcgtggcgta cgccacaatt ggacggcatt ctgggtgctt gtcatgcggt ggaacttccg  1260
SEQIDNO:3      1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:5      1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:7      1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:9      1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:11     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:13     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:15     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:17     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:19     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:21     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:23     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:25     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:27     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:29     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:31     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:33     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:35     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:37     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:39     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:41     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:43     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:45     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:47     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:49     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:51     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:53     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:55     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:57     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:59     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:61     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:63     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:65     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:67     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:69     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:71     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:73     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:75     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:77     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:79     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:81     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:83     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:85     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:87     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:89     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:91     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:93     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:95     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:97     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:99     1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:101    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:103    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:105    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:107    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:109    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:111    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:113    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:115    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:117    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:119    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:121    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:123    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:125    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:127    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:129    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:131    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:133    1191 .......... .......... .......... .......... .......... .......... ..........  1260
SEQIDNO:135    1191 .......... .......... .......... .......... .......... .......... ..........  1260
```

Figure 1, Continued

```
                         1270       1280       1290       1300       1310       1320       1330
                   ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     1261 ttcgtgtttc gcaccttaga ccgtgcggca tcgctggtgg gcacaaaccc gccagaagaa ctggcagaaa  1330
SEQIDNO:3     1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:5     1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:7     1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:9     1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:11    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:13    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:15    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:17    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:19    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:21    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:23    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:25    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:27    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:29    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:31    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:33    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:35    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:37    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:39    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:41    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:43    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:45    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:47    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:49    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:51    1261 .......... .......... .......... .......... .......... .a........ ..........  1330
SEQIDNO:53    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:55    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:57    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:59    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:61    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:63    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:65    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:67    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:69    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:71    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:73    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:75    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:77    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:79    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:81    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:83    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:85    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:87    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:89    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:91    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:93    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:95    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:97    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:99    1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:101   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:103   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:105   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:107   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:109   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:111   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:113   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:115   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:117   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:119   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:121   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:123   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:125   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:127   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:129   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:131   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:133   1261 .......... .......... .......... .......... .......... .......... ..........  1330
SEQIDNO:135   1261 .......... .......... .......... .......... .......... .......... ..........  1330
```

Figure 1, Continued

```
                          1340       1350       1360       1370       1380       1390       1400
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1      1331 ccgtacacaa tgcctgggtg cgcttcgcga catcaggtga tccagggtgg ccggcgtgga atccggaaac  1400
SEQIDNO:3      1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:5      1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:7      1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:9      1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:11     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:13     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:15     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:17     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:19     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:21     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:23     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:25     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:27     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:29     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:31     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:33     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:35     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:37     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:39     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:41     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:43     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:45     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:47     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:49     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:51     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:53     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:55     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:57     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:59     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:61     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:63     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:65     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:67     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:69     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:71     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:73     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:75     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:77     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:79     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:81     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:83     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:85     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:87     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:89     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:91     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:93     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:95     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:97     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:99     1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:101    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:103    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:105    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:107    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:109    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:111    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:113    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:115    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:117    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:119    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:121    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:123    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:125    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:127    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:129    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:131    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:133    1331 .......... .......... .......... .......... .......... .......... ..........  1400
SEQIDNO:135    1331 .......... .......... .......... .......... .......... .......... ..........  1400
```

Figure 1, Continued

```
                        1410       1420       1430       1440       1450       1460       1470
                   ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:1     1401 ccgcagcgtg atgcgctttg atcacccggt ttccgagatg gttaccgacc cgtacccggc gacccgtgca  1470
SEQIDNO:3     1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:5     1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:7     1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:9     1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:11    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:13    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:15    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:17    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:19    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:21    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:23    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:25    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:27    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:29    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:31    1401 .......... .......... .......... .......... .......... .......a.. ..........  1470
SEQIDNO:33    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:35    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:37    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:39    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:41    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:43    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:45    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:47    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:49    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:51    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:53    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:55    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:57    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:59    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:61    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:63    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:65    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:67    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:69    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:71    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:73    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:75    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:77    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:79    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:81    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:83    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:85    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:87    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:89    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:91    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:93    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:95    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:97    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:99    1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:101   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:103   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:105   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:107   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:109   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:111   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:113   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:115   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:117   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:119   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:121   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:123   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:125   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:127   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:129   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:131   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:133   1401 .......... .......... .......... .......... .......... .......... ..........  1470
SEQIDNO:135   1401 .......... .......... .......... .......... .......... .......... ..........  1470
```

Figure 1, Continued

```
                           1480       1490
                    ....|....| ....|....| .
SEQIDNO:1     1471 ctgtgggatg gtgtgccatt g 1491
SEQIDNO:3     1471 .......... .......... . 1491
SEQIDNO:5     1471 .......... .......... . 1491
SEQIDNO:7     1471 .......... .......... . 1491
SEQIDNO:9     1471 .......... .......... . 1491
SEQIDNO:11    1471 .......... .......... . 1491
SEQIDNO:13    1471 .......... .......... . 1491
SEQIDNO:15    1471 .......... .......... . 1491
SEQIDNO:17    1471 .......... .......... . 1491
SEQIDNO:19    1471 .......... .......... . 1491
SEQIDNO:21    1471 .......... .......... . 1491
SEQIDNO:23    1471 .......... .......... . 1491
SEQIDNO:25    1471 .......... .......... . 1491
SEQIDNO:27    1471 .......... .......... . 1491
SEQIDNO:29    1471 .......... .......... . 1491
SEQIDNO:31    1471 .......... .......... . 1491
SEQIDNO:33    1471 .......... .......... . 1491
SEQIDNO:35    1471 .......... .......... . 1491
SEQIDNO:37    1471 .......... .......... . 1491
SEQIDNO:39    1471 .......... .......... . 1491
SEQIDNO:41    1471 .......... .......... . 1491
SEQIDNO:43    1471 .......... .......... . 1491
SEQIDNO:45    1471 .......... .......... . 1491
SEQIDNO:47    1471 .......... .......... . 1491
SEQIDNO:49    1471 .......... .......... . 1491
SEQIDNO:51    1471 .......... .......... . 1491
SEQIDNO:53    1471 .......... .......... . 1491
SEQIDNO:55    1471 .......... .......... . 1491
SEQIDNO:57    1471 .......... .......... . 1491
SEQIDNO:59    1471 .......... .......... . 1491
SEQIDNO:61    1471 .......... .......... . 1491
SEQIDNO:63    1471 .......... .......... . 1491
SEQIDNO:65    1471 .......... .......... . 1491
SEQIDNO:67    1471 .......... .......... . 1491
SEQIDNO:69    1471 .......... .......... . 1491
SEQIDNO:71    1471 .......... .......... . 1491
SEQIDNO:73    1471 .......... .......... . 1491
SEQIDNO:75    1471 .......... .......... . 1491
SEQIDNO:77    1471 .......... .......... . 1491
SEQIDNO:79    1471 .......... .......... . 1491
SEQIDNO:81    1471 .......... .......... . 1491
SEQIDNO:83    1471 .......... .......... . 1491
SEQIDNO:85    1471 .......... .......... . 1491
SEQIDNO:87    1471 .......... .......... . 1491
SEQIDNO:89    1471 .......... .......... . 1491
SEQIDNO:91    1471 .......... .......... . 1491
SEQIDNO:93    1471 .......... .......... . 1491
SEQIDNO:95    1471 .......... .......... . 1491
SEQIDNO:97    1471 .......... .......... . 1491
SEQIDNO:99    1471 .......... .......... . 1491
SEQIDNO:101   1471 .......... .......... . 1491
SEQIDNO:103   1471 .......... .......... . 1491
SEQIDNO:105   1471 .......... .......... . 1491
SEQIDNO:107   1471 .......... .......... . 1491
SEQIDNO:109   1471 .......... .......... . 1491
SEQIDNO:111   1471 .......... .......... . 1491
SEQIDNO:113   1471 .......... .......... . 1491
SEQIDNO:115   1471 .......... .......... . 1491
SEQIDNO:117   1471 .......... .......... . 1491
SEQIDNO:119   1471 .......... .......... . 1491
SEQIDNO:121   1471 .......... .......... . 1491
SEQIDNO:123   1471 .......... .......... . 1491
SEQIDNO:125   1471 .......... .......... . 1491
SEQIDNO:127   1471 .......... .......... . 1491
SEQIDNO:129   1471 .......... .......... . 1491
SEQIDNO:131   1471 .......... .......... . 1491
SEQIDNO:133   1471 .......... .......... . 1491
SEQIDNO:135   1471 .......... .......... . 1491
```

Figure 2. Pairwise alignment of the Polypeptide Sequence Encoding the Wild-Type Carboxyesterase enzyme, *Thermobifida fusca* (*T. fusca*) (SEQ ID NO: 2) against the polypeptide sequences of *T. fusca* variants

```
                              10         20         30         40         50         60         70
                      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2        1    MEIVIRTGSG DVRGSKENGI AVFRGIPYAE PPVGAHRFTA PRPPRPWDGV RDATEFSATA PRPPYPEAIG   70
SEQIDNO:4        1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:6        1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:8        1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:10       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:12       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:14       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:16       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:18       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:20       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:22       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:24       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:26       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:28       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:30       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:32       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:34       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:36       1    .......... .......... .......... .......... .......... .......... ..R.......   70
SEQIDNO:38       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:40       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:42       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:44       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:46       1    .......... .......... .......... .......... .......... .......... .......P..   70
SEQIDNO:48       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:50       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:52       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:54       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:56       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:58       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:60       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:62       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:64       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:66       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:68       1    .......... .......... .......... .......... .......... .......... ....W...L.   70
SEQIDNO:70       1    .......... .......... .......... .......... .......... .......... .......PL.   70
SEQIDNO:72       1    .......... .......... .......... .......... .......... .......... ........L.   70
SEQIDNO:74       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:76       1    .......... .......... .......... .......... .......... .......... ..R.......   70
SEQIDNO:78       1    .......... .......... .......... .......... .......... .......... ....W...MA   70
SEQIDNO:80       1    .......... .......... .......... .......... .......... .......... ........W.   70
SEQIDNO:82       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:84       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:86       1    .......... .......... .......... .......... .......... .......... ........L.   70
SEQIDNO:88       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:90       1    .......... .......... .......... .......... .......... .......... ..R.G.....   70
SEQIDNO:92       1    .......... .......... .......... .......... .......... .......... .......P..   70
SEQIDNO:94       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:96       1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:98       1    .......... .......... .......... .......... .......... .......... .......P..   70
SEQIDNO:100      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:102      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:104      1    .......... .......... .......... .......... .......... .......... ..I.......   70
SEQIDNO:106      1    .......... .......... .......... .......... .......... .......... ........W.   70
SEQIDNO:108      1    .......... .......... .......... .......... .......... .......... ..A.......   70
SEQIDNO:110      1    .......... .......... .......... .......... .......... .......... ..Y.......   70
SEQIDNO:112      1    .......... .......... .......... .......... .......... .......... ..Y.......   70
SEQIDNO:114      1    .......... .......... .......... .......... .......... .......... ..Y.......   70
SEQIDNO:116      1    .......... .......... .......... .......... .......... .......... .......PW.   70
SEQIDNO:118      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:120      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:122      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:124      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:126      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:128      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:130      1    .......... .......... .......... .......... .......... .......... ........F.   70
SEQIDNO:132      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:134      1    .......... .......... .......... .......... .......... .......... ..........   70
SEQIDNO:136      1    .......... .......... .......... .......... .......... .......... ..........   70
```

Figure 2, Continued

```
                          80         90        100        110        120        130        140
                  ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2     71  ALLIERFIPG DDYLILNVWI PDPNAVGLPV MVWIHGGAFT NGSGSEPVYD GAAFARDGVV FVSFNYRLGI  140
SEQIDNO:4     71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:6     71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:8     71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:10    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:12    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:14    71  .......... .......... .......... .......G.. .......... .......... ..........  140
SEQIDNO:16    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:18    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:20    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:22    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:24    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:26    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:28    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:30    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:32    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:34    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:36    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:38    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:40    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:42    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:44    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:46    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:48    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:50    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:52    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:54    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:56    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:58    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:60    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:62    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:64    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:66    71  ......S... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:68    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:70    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:72    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:74    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:76    71  .......... .......... .......... .......G.. .......... .......... ..........  140
SEQIDNO:78    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:80    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:82    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:84    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:86    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:88    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:90    71  .......... .......... .......... .......G.. .......... .......... ..........  140
SEQIDNO:92    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:94    71  .......... .......... .......... .......G.. .......... .......... ..........  140
SEQIDNO:96    71  .......... .......... .......... .......G.. .......... .......... ..........  140
SEQIDNO:98    71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:100   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:102   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:104   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:106   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:108   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:110   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:112   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:114   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:116   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:118   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:120   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:122   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:124   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:126   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:128   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:130   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:132   71  .......... .......... .......... .......... .......... .......... ..........  140
SEQIDNO:134   71  .......... .......... .......... .......G.. .......... .......... ..........  140
SEQIDNO:136   71  .......... .......... .......... .......... .......... .......... ..........  140
```

Figure 2, Continued

```
                           150        160        170        180        190        200        210
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2       141 IGFADLPDAP SNRGLLDQIA ALEWVRDNIA RFGGDPGNVT VFGESAGAMS VCTLMAIPRA RGLFRRAILQ  210
SEQIDNO:4       141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:6       141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:8       141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:10      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:12      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:14      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:16      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:18      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:20      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:22      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:24      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:26      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:28      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:30      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:32      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:34      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:36      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:38      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:40      141 .......... .......... .......... .......... .........K .......... ..........  210
SEQIDNO:42      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:44      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:46      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:48      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:50      141 .......... .......... .......... .......... ...S...... .......... ..........  210
SEQIDNO:52      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:54      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:56      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:58      141 .......... .......... .......... .......... ...G...... .......... ..........  210
SEQIDNO:60      141 .......... ..H....... .......... .......... .......... .......... ..........  210
SEQIDNO:62      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:64      141 .......... .......... .......... .......... .........W .......... ..........  210
SEQIDNO:66      141 .......... .......... .......... .......... ...G...... .......... ..........  210
SEQIDNO:68      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:70      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:72      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:74      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:76      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:78      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:80      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:82      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:84      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:86      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:88      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:90      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:92      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:94      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:96      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:98      141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:100     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:102     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:104     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:106     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:108     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:110     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:112     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:114     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:116     141 .......... .......... .......... .......... ........I. .......... ..........  210
SEQIDNO:118     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:120     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:122     141 .......... .......... .......... .......... ........I. .......... ..........  210
SEQIDNO:124     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:126     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:128     141 .......... .......... .......... .......... ........I. .......... ..........  210
SEQIDNO:130     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:132     141 .......... .......... .......... .......... ........Q. .......... ..........  210
SEQIDNO:134     141 .......... .......... .......... .......... .......... .......... ..........  210
SEQIDNO:136     141 .......... .......... .......... .......... .......... .......... ..........  210
```

Figure 2, Continued

```
                          220        230        240        250        260        270        280
                   ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2      211 SGAGNMAVAA EDATTIAAVI AHRLGVEPTA AALAHVPVAQ LLDVQQQVAQ EIQGAPDPAV WGERIAGGSV  280
SEQIDNO:4      211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:6      211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:8      211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:10     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:12     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:14     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:16     211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:18     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:20     211 ....W..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:22     211 .P........ .......... .......... .......... .......... .......... ..........  280
SEQIDNO:24     211 .......... .......... .......... .......... .......... .......... .........G  280
SEQIDNO:26     211 ....P..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:28     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:30     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:32     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:34     211 .......... .......... .......... .......... .......... .......... Y.........  280
SEQIDNO:36     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:38     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:40     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:42     211 ....R..... .......... .......... .......... .......... .......... R.........  280
SEQIDNO:44     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:46     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:48     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:50     211 .......... .......... .......... ........T. .......... .......... ..........  280
SEQIDNO:52     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:54     211 ..L....... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:56     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:58     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:60     211 ....P..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:62     211 ...L...... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:64     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:66     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:68     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:70     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:72     211 .......... .......... .......... .......... .......... .......... Y.........  280
SEQIDNO:74     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:76     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:78     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:80     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:82     211 .......... .......... .......... .......... .......... .......... .........G  280
SEQIDNO:84     211 .......... .......... .......... .......... .......... .......... Y.........  280
SEQIDNO:86     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:88     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:90     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:92     211 .......... .......... .......... .......... .......... .......... Y.........  280
SEQIDNO:94     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:96     211 ....K..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:98     211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:100    211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:102    211 ....K..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:104    211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:106    211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:108    211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:110    211 .P..R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:112    211 .P..R..... .......... .......... .......... .......... .......AN. ..........  280
SEQIDNO:114    211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:116    211 ...RP..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:118    211 ....K..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:120    211 ....K..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:122    211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:124    211 ....R..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:126    211 ....W..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:128    211 .......... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:130    211 ....K..... .......... .......... .......... .......... ........LI ..........  280
SEQIDNO:132    211 ....P..... .......... .......... .......... .......... .......... Y.........  280
SEQIDNO:134    211 ....P..... .......... .......... .......... .......... .......... ..........  280
SEQIDNO:136    211 .......... .......... .......... .......... .......... .......... ..........  280
```

Figure 2, Continued

```
                            290        300        310        320        330        340        350
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2       281 LLPFAPVIDG ELLSQRPAEA IAGGAGHDVD LLFGTTTDEY RLFLAPTGLL PFITGDYVTT HLAKSGLDAD  350
SEQIDNO:4       281 .Q........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:6       281 .A........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:8       281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:10      281 ....L..... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:12      281 .C........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:14      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:16      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:18      281 P......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:20      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:22      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:24      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:26      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:28      281 ..D....... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:30      281 .......... .......... .......... .........G .......... .......... ..........  350
SEQIDNO:32      281 .......... .......... .......... .......... L......... .......... ..........  350
SEQIDNO:34      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:36      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:38      281 .......... .......... .......... .........W .......... .......... ..........  350
SEQIDNO:40      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:42      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:44      281 ..V....... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:46      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:48      281 ..Y....... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:50      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:52      281 ...T...... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:54      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:56      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:58      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:60      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:62      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:64      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:66      281 .......... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:68      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:70      281 .T........ .......... .......... .......... .......... .......... ..V.......  350
SEQIDNO:72      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:74      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:76      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:78      281 PT........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:80      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:82      281 .T..L..... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:84      281 .T........ .......... .......... .......... .......... .......... ..V.......  350
SEQIDNO:86      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:88      281 PT........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:90      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:92      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:94      281 .T..L..... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:96      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:98      281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:100     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:102     281 PT..L..... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:104     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:106     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:108     281 .T........ .......... .......... .......... .......... .......... ..V.......  350
SEQIDNO:110     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:112     281 .T........ .......... .......... .......... .......... .......... ..V.......  350
SEQIDNO:114     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:116     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:118     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:120     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:122     281 .T........ .......... .......... .......... .......... .......... ..V.......  350
SEQIDNO:124     281 PT..L..... .......... .......... .......... .......... .......... ..........  350
SEQIDNO:126     281 .T..L..... .......... .......... .......... .......... .......... .....S....  350
SEQIDNO:128     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:130     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:132     281 PC........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:134     281 .T........ .......... .......... .......... .......... .......... ..........  350
SEQIDNO:136     281 .......... .......... .......... .......... .......... .......... ..........  350
```

Figure 2, Continued

```
                           360        370        380        390        400        410        420
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2       351 AAKAYTAEGR GEEPGDILAS IITDQVFRIP ALRIAESRVD APARTFGYEF ANRTPQLDGI LGACHAVELP  420
SEQIDNO:4       351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:6       351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:8       351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:10      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:12      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:14      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:16      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:18      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:20      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:22      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:24      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:26      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:28      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:30      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:32      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:34      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:36      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:38      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:40      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:42      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:44      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:46      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:48      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:50      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:52      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:54      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:56      351 .......... .......... ......L... .......... .......... .......... ..........  420
SEQIDNO:58      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:60      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:62      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:64      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:66      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:68      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:70      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:72      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:74      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:76      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:78      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:80      351 .......... .......... .M........ .......... .......... .......... ..........  420
SEQIDNO:82      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:84      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:86      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:88      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:90      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:92      351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:94      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:96      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:98      351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:100     351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:102     351 .......... .......... .L........ .......... .......... .......... ..........  420
SEQIDNO:104     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:106     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:108     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:110     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:112     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:114     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:116     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:118     351 .......... .......... ..G....... .......... .......... .......... ..........  420
SEQIDNO:120     351 .......... .......... .L....L... .......... .......... .......... ..........  420
SEQIDNO:122     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:124     351 .......... .......... ..G....... .......... .......... .......... ..........  420
SEQIDNO:126     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:128     351 .......... .......... .......... .......... .......... .......... ..........  420
SEQIDNO:130     351 .......... .......... ......L... .......... .......... .......... ..........  420
SEQIDNO:132     351 .......... .......... ......Y... .......... .......... .......... ..........  420
SEQIDNO:134     351 .......... .......... ......L... .......... .......... .......... ..........  420
SEQIDNO:136     351 .......... .......... .L........ .......... .......... .......... ..........  420
```

Figure 2, Continued

```
                           430        440        450        460        470        480        490
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
SEQIDNO:2      421  FVFRILDRAA SLVGINPPEE LAEIVHNAWV RFATSGDPGW PAWNPEIRSV MREDHPVSEM VIDPYPATRA  490
SEQIDNO:4      421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:6      421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:8      421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:10     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:12     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:14     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:16     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:18     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:20     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:22     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:24     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:26     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:28     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:30     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:32     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:34     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:36     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:38     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:40     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:42     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:44     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:46     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:48     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:50     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:52     421  .......... .......T.. .......... .......... .......... .......... ..........  490
SEQIDNO:54     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:56     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:58     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:60     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:62     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:64     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:66     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:68     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:70     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:72     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:74     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:76     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:78     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:80     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:82     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:84     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:86     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:88     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:90     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:92     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:94     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:96     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:98     421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:100    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:102    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:104    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:106    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:108    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:110    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:112    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:114    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:116    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:118    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:120    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:122    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:124    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:126    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:128    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:130    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:132    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:134    421  .......... .......... .......... .......... .......... .......... ..........  490
SEQIDNO:136    421  .......... .......... .......... .......... .......... .......... ..........  490
```

Figure 2, Continued

```
                    ....|..
SEQIDNO:2      491 LWDGVPL 497
SEQIDNO:4      491 ....... 497
SEQIDNO:6      491 ....... 497
SEQIDNO:8      491 ....... 497
SEQIDNO:10     491 ....... 497
SEQIDNO:12     491 ....... 497
SEQIDNO:14     491 ....... 497
SEQIDNO:16     491 ....... 497
SEQIDNO:18     491 ....... 497
SEQIDNO:20     491 ....... 497
SEQIDNO:22     491 ....... 497
SEQIDNO:24     491 ....... 497
SEQIDNO:26     491 ....... 497
SEQIDNO:28     491 ....... 497
SEQIDNO:30     491 ....... 497
SEQIDNO:32     491 ....... 497
SEQIDNO:34     491 ....... 497
SEQIDNO:36     491 ....... 497
SEQIDNO:38     491 ....... 497
SEQIDNO:40     491 ....... 497
SEQIDNO:42     491 ....... 497
SEQIDNO:44     491 ....... 497
SEQIDNO:46     491 ....... 497
SEQIDNO:48     491 ....... 497
SEQIDNO:50     491 ....... 497
SEQIDNO:52     491 ....... 497
SEQIDNO:54     491 ....... 497
SEQIDNO:56     491 ....... 497
SEQIDNO:58     491 ....... 497
SEQIDNO:60     491 ....... 497
SEQIDNO:62     491 ....... 497
SEQIDNO:64     491 ....... 497
SEQIDNO:66     491 ....... 497
SEQIDNO:68     491 ....... 497
SEQIDNO:70     491 ....... 497
SEQIDNO:72     491 ....... 497
SEQIDNO:74     491 ....... 497
SEQIDNO:76     491 ....... 497
SEQIDNO:78     491 ....... 497
SEQIDNO:80     491 ....... 497
SEQIDNO:82     491 ....... 497
SEQIDNO:84     491 ....... 497
SEQIDNO:86     491 ....... 497
SEQIDNO:88     491 ....... 497
SEQIDNO:90     491 ....... 497
SEQIDNO:92     491 ....... 497
SEQIDNO:94     491 ....... 497
SEQIDNO:96     491 ....... 497
SEQIDNO:98     491 ....... 497
SEQIDNO:100    491 ....... 497
SEQIDNO:102    491 ....... 497
SEQIDNO:104    491 ....... 497
SEQIDNO:106    491 ....... 497
SEQIDNO:108    491 ....... 497
SEQIDNO:110    491 ....... 497
SEQIDNO:112    491 ....... 497
SEQIDNO:114    491 ....... 497
SEQIDNO:116    491 ....... 497
SEQIDNO:118    491 ....... 497
SEQIDNO:120    491 ....... 497
SEQIDNO:122    491 ....... 497
SEQIDNO:124    491 ....... 497
SEQIDNO:126    491 ....... 497
SEQIDNO:128    491 ....... 497
SEQIDNO:130    491 ....... 497
SEQIDNO:132    491 ....... 497
SEQIDNO:134    491 ....... 497
SEQIDNO:136    491 ....... 497
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 18/65124

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12N 9/18, A23L 3/3571 (2019.01)
CPC - A23L 3/3571, C12N 9/18, C12N 9/14

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 data base consulted during the international search (name of data base 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 ----- A	BILLIG et al. Hydrolysis of cyclic poly(ethylene terephthalate) trimers by a carboxylesterase from Thermobifida fusca KW3, Applied Microbiology and Biotechnology, 14 May 2010, Vol 87, No 5, pp 1753-1764, Especially Abstract pg 1753, 1754, 1757	1, 24 ------------- (11-14)/1 ------------- 2, (11-14)/2, 25
Y	US 2011/0137002 A1 (HAUER et al.) 9 June 2011 (09.06.2011) Especially para [0035], [0094], [0120]	(11-14)/1
Y	OESER et al. High level expression of a hydrophobic poly(ethylene terephthalate)-hydrolyzing carboxylesterase from Thermobifida fusca KW3 in Escherichia coli BL21(DE3), Journal of Biotechnology, 1 April 2010, Vol 146, No 3, pp 100-104, Especially p 101, 102, Fig. 1	(12-14)/1
Y	US 2007/0015909 A1 (CASH et al.) 18 January 2007 (18.01.2007) Especially para [0087]	(12-14)/1

☐ 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
"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 7 May 2019	Date of mailing of the international search report 23 MAY 2019
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 18/65124

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 15-17, 20-23, 29-37
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

-----please see continuation in extra sheet-----

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1, 2, 11-14(in part), 24, 25 limited to SEQ ID NO: 2 with a substitution at position 39

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Continuation of Box No. III Observations where unity of invention is lacking

Group I+, Claims 1-14, 18, 19, and 24-28, directed to an engineered carboxyesterase and a polynucleotide encoding the same comprising at least one substitution. The carboxyesterase will be searched to the extent that the polypeptide sequence encompass SEQ ID NO: 2 with a single substitution at position 39. It is believed that claims 1, 2, 11-14(in part), 24, 25 and 28 encompass this first named invention, and thus these claims will be searched without fee to the extent that the carboxyesterase polypeptide sequence encompass SEQ ID NO: 2 with a single substitution at position 39. Additional carboxyesterase polypeptide sequence(s) will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected carboxyesterase polypeptide sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be the carboxyesterase polypeptide sequence comprising SEQ ID NO: 2 with substitutions at positions 39 and 323 (claims 1, 2, 11-14(in part), 24, 25, 28).

[NOTE: None of the claimed peptide or polynucleotide sequences (SEQ ID Nos: 3-136) are SEQ ID NO: 2 with a single substitution at position 39, as shown in Figure 2, or corresponding nucleotide sequence SEQ ID NO: 1 with a single substitution nts 117-119 (corresponding to encoded position 39), as shown in Figure 1.]

The inventions listed as Group I+ do not relate to a single special technical feature under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special technical features

The inventions of Group I+ each include the special technical feature of a unique amino acid sequence. Each amino acid sequence encodes a unique peptide, and is considered a distinct technical feature.

No technical features are shared between the carboxyesterase amino acid sequences of Group I+ and, accordingly, these groups lack unity a priori.

Additionally, even if Groups I+ were considered to share the technical features of including: an engineered carboxyesterase from a Thermobifida fusca comprising a polypeptide sequence with at least one substitution, or a functional fragment thereof, and a polynucleotide sequence encoding said engineered carboxyesterase, these shared technical features are previously disclosed by the article entitled "Hydrolysis of cyclic poly(ethylene terephthalate) trimers by a carboxylesterase from Thermobifida fusca KW3" by Billig et al., (Appl Microbiol Biotechnol, August 2010, Vol 87, No 5, pp 1753-1764), (hereinafter 'Billig').

Billig teaches an engineered carboxyesterase from a Thermobifida fusca comprising a polypeptide sequence having at least 80% sequence identity to SEQ ID NO: 2 (pg 1754 col 1 para 3, "Recently, we reported the recombinant production of the carboxylesterase TfCa (T. fusca carboxylesterase) from T. fusca KW3 following optimization for codon usage in E. coli (Oeser et al. 2010). In this study, we have further characterized this unusual enzyme with PET-hydrolyzing activity that is distinct from cutinases."; pg 1753, col 1, last para, "the protein sequence data in the UniProt knowledgebase under the accession number P86325."; Note, UniProtKB-P86325 (hereinafter 'P86325') teaches a 497 amino acid polypeptide with amino acids 1-497 comprising 99.6% sequence identity to SEQ ID NO:2), wherein said engineered carboxyesterase comprises at least one substitution or substitution set in said polypeptide sequence, and wherein the amino acid positions of said polypeptide sequence are numbered with reference to SEQ ID NO: 2 (pg 1760 col 2 last para - pg 1762 col 1 para 1 "A comparison of the specific activities of the wild-type enzyme and enzyme variants of TfCa where the Glu of the active site and Glu and Ala of the G-E-S-A-G motif were substituted with Asp, His, and Met is shown in Fig. 7. Substitution of Glu184Gln, one variable amino acid of the serine motif, resulted in an 84% decrease in the specific activity compared with the wild-type enzyme. Substitution of Glu319Asp in the acid position of the active site resulted in a 75% decrease in thespecific activity. Combination of both substitutions (E184Q and E319D) resulted in an 80% decrease of its specific activity compared with wild-type TfCa. A mutant TfCa carrying a double mutation (E184H/A186M) retained only 5% specific activity compared with wild-type TfCa. Introduction of a third substitution at the acid position (Glu to Asp) of the catalytic triad created a mutant (E184H/A186M/E319D).").

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Group I+ inventions lack unity under PCT Rule 13 because they do not share the same or corresponding special technical feature.

NOTE claims 15-17, 20-23, 29-37 are held unsearchable because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).