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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 November 2011 (03.11.2011)

PCT

(10) International Publication Number
WO 2011/137369 A1

(51) International Patent Classification:
C12N 1/21 (2006.01) *C12P 13/00* (2006.01)
C12N 9/88 (2006.01)

(21) International Application Number:
PCT/US2011/034612

(22) International Filing Date:
29 April 2011 (29.04.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/329,204 29 April 2010 (29.04.2010) US

(71) Applicant *(for all designated States except US)*: **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 1111 Franklin Street, Oakland, CA 94607-5200 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LIAO, James, C.** [US/US]; 10573 Willworth Avenue, Los Angeles, CA 90024 (US). **CHO, Kwang, Myung** [KR/US]; 10600 Wilshire Blvd., #428, Los Angeles, CA 90024 (US).

(74) Agent: **KALE, Tamara, A.**; Christensen O'connor Johnson Kindness, 1420 5th Avenue Suite 2800, Seattle, WA 98101-2372 (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, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).

Published:

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

(54) **Title**: PRODUCTION OF GAMMA-AMINOBUTYRIC ACID BY RECOMBINANT MICROORGANISMS

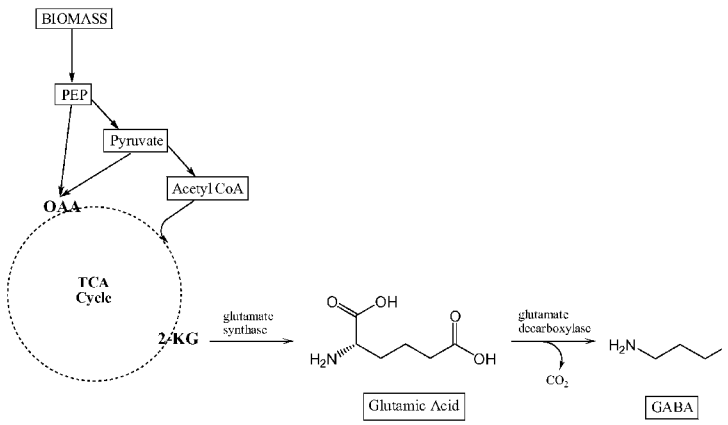

FIGURE 1

(57) **Abstract**: Provided herein are metabolically-modified microorganisms useful for producing gamma-aminobutyric acid.

WO 2011/137369 A1

PRODUCTION OF GAMMA-AMINOBUTYRIC ACID BY RECOMBINANT MICROORGANISMS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application Serial No. 61/329,204, filed April 29, 2010, the disclosure of which is incorporated herein by reference.

STATEMENT REGARDING SEQUENCE LISTING

The sequence listing associated with this application is provided in text format in lieu of a paper copy and is hereby incorporated by reference into the specification. The name of the text file containing the sequence listing is "36585_FINAL_SEQ.txt". The text file is 19.7 KB; was created on 29 April 2011; and is being submitted via EFS-Web with the filing of the specification.

TECHNICAL FIELD

Metabolically-modified microorganisms and methods of producing such organisms are provided. Also provided are methods of producing gamma-aminobutyric acid by contacting a suitable substrate with a metabolically-modified microorganism and enzymatic preparations there from.

BACKGROUND

Gamma-aminobutyric acid is a compound that serves both industrial purposes in the production of various synthetics including, for example, Nylon 4, as well as biological roles such as in nerve cell communications.

SUMMARY

The disclosure provides methods and recombinant microorganisms for the production of gamma-aminobutyric acid.

Accordingly, provided herein is a recombinant microorganism derived from a parental microorganism that produces gamma-aminobutyric acid (GABA, also termed 4-aminobutyric acid) and over-expresses a polypeptide having glutamate decarboxylase activity.

Also provided is a method of producing gamma-aminobutyric acid, comprising: (a) culturing a recombinant cell or an extract of a cell of claim 1 with a suitable carbon source, wherein the cell or cell extract converts the carbon source to gamma-

aminobutyric acid; and (b) isolating the gamma-aminobutyric acid. A suitable carbon source may be a sugar, such as glucose, sucrose, or a combination thereof.

Also provided is a method of producing a recombinant microorganism that produces gamma-aminobutyric acid, comprising transforming a microorganism with a nucleic acid sequence encoding a polypeptide having glutamate decarboxylase activity.

The disclosure further provides a recombinant microorganism or microorganism culture (*e.g.*, a plurality of recombinant organisms with the same or different enzymes) that produces gamma-aminobutyric acid comprising a recombinant metabolic pathway as set forth in Figure 1. In one embodiment, the microorganism comprises expression of a non-natural polypeptide of the organism or over-expression of an endogenous polypeptide of the organism wherein the polypeptide has glutamate decarboxylase activity. In yet another embodiment, the microorganism comprises a pathway that produces glutamic acid. In one embodiment, a recombinant bacteria produces gamma-aminobutyric acid at greater than about 1 g/L (*e.g.*, 2 g/L, 3 g/L, 4 g/L or more).

The disclosure provides for the efficient production of GABA, which can be used for the monomer of Nylon 4 as well as pharmaceutical intermediates alone. Nylon 4 is a biodegradable engineering plastic with much superior mechanical strength and heat resistant properties to current the biopolymer, polylactic acid (PLA). The melting temperature of Nylon 4 is known to be 262°C, which is much higher than that of PLA of 184°C. Because of these better characteristics in both mechanical strength and processing, it could have much wider spectrum of application than PLA. In addition, it is quite compatible with various kinds of pre-existing co-monomers to improve its properties.

In one embodiment, the disclosure provides a recombinant bacterium, *Corynebacterium glutamicum*, by introducing a glutamate decarboxylase gene (*gadA*) from another bacteria. This recombinant microorganism successfully produced GABA directly from glucose. Glutamic acid is one of the major amino acids produced by direct fermentation of *Corynebacterium glutamicum*.

The details of one or more embodiments of the disclosure are set forth in the accompanying drawings and the description below. Other features, objects, and advantages will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF THE DRAWINGS

The accompanying drawing, which is incorporated into and constitutes a part of this specification, illustrates one or more embodiments of the disclosure and, together with the detailed description, serves to explain the principles and implementations of the invention.

Figure 1 shows an exemplary metabolic pathway for the production of gamma-aminobutyric acid.

DETAILED DESCRIPTION

As used herein and in the appended claims, the singular forms "a," "and," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a polynucleotide" includes a plurality of such polynucleotides and reference to "the microorganism" includes reference to one or more microorganisms, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice of the disclosed methods and compositions, the exemplary methods, devices and materials are described herein.

Any publications discussed above and throughout the text are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior disclosure.

By "about" is meant a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight, or length that varies by as much as 30, 25, 20, 25, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1% to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight, or length.

The disclosure provides metabolically engineered microorganisms comprising a biochemical pathway for the production of gamma-aminobutyric acid from a suitable substrate. A metabolically engineered microorganism of the disclosure comprises one or more recombinant polynucleotides within the genome of the organism or external to the genome within the organism. The microorganism can comprise a reduction, disruption, or knockout of a gene found in the wild-type organism and/or introduction of a heterologous polynucleotide.

The disclosure also includes metabolically engineered biosynthetic pathways that utilize an organism's native pathways, in part, to provide metabolic intermediates for the production of gamma-aminobutyric acid.

Accordingly, provided herein is a recombinant microorganism derived from a parental microorganism that produces gamma-aminobutyric acid (GABA, also termed 4-aminobutyric acid) and over-expresses a polypeptide having glutamate decarboxylase activity. The polypeptide may be endogenous or heterologous to the parental microorganism. In some embodiments, the microorganism is derived from a genera selected from Brevibacterium, Arthrobacter, Microbacterium, and Corynebacterium. In some embodiments, the microorganism is derived from *Corynebacterium glutamicum.*

As discussed herein, the polypeptide having glutamate decarboxylase activity may be gadA or gadB. The polypeptide having glutamate decarboxylase activity may comprise an amino acid sequence of SEQ ID NO:2, or a conservative-substituted variant of or homologous polypeptide to SEQ ID NO:2 having at least 60% identity to SEQ ID NO:2 and having glutamate decarboxylase activity. In some embodiments, the polypeptide having glutamate decarboxylase activity is obtained from a species selected from *E. coli, Shigella dysenteriae, Shigella flexneri, Escherichia albertii, Yersinia ruckeri, Edwardsiella tarda, Brucella suis, Brucella ovis, Akkermansia muciniphila, Desulfovibrio desulfuricans, Blastopirellula marina,* and *Photobacterium damselae.* The polypeptide having glutamate decarboxylase activity may be encoded by a gene having at least 70-100% identity to polynucleotide SEQ ID NO:1. The polynucleotide having glutamate decarboxylase activity may comprise a gadA polynucleotide selected from: (a) a deoxyribonucleic acid (DNA) comprising the nucleotide sequence of SEQ ID NO:1; and (b) a DNA that hybridizes with the nucleotide sequence of SEQ ID NO:1 or a complement thereof and which encodes a polypeptide having glutamate decarboxylase activity.

A recombinant microorganism may be further modified to enhance glutamic acid/GABA antiporter activity compared to a parental microorganism. In some embodiments, glutamic acid/GABA antiporter activity is produced by a gadC polypeptide.

Also provided is a method of producing gamma-aminobutyric acid, comprising: (a) culturing a recombinant cell or an extract of a cell of claim 1 with a suitable carbon source, wherein the cell or cell extract converts the carbon source to gamma-

aminobutyric acid; and (b) isolating the gamma-aminobutyric acid. A suitable carbon source may be a sugar, such as glucose, sucrose, or a combination thereof.

Also provided is a method of producing a recombinant microorganism that produces gamma-aminobutyric acid, comprising transforming a microorganism with a nucleic acid sequence encoding a polypeptide having glutamate decarboxylase activity.

In one embodiment, the disclosure provides a recombinant microorganism comprising elevated expression of at least one target enzyme as compared to a parental microorganism or encodes an enzyme not found in the parental organism. In another, or further embodiment, the microorganism comprises a reduction, disruption, or knockout of at least one gene encoding an enzyme that competes for a metabolite necessary for the production of a desired metabolic intermediate. The recombinant microorganism produces at least one metabolite involved in a biosynthetic pathway for the production of gamma-aminobutyric acid. In general, the recombinant microorganism comprises at least one recombinant metabolic pathway that comprises a target enzyme and can further include a reduction in activity or expression of an enzyme in a competitive biosynthetic pathway. The pathway acts to modify a substrate or metabolic intermediate in the production of gamma-aminobutyric acid. The target enzyme is encoded by, and expressed from, a polynucleotide derived from a suitable biological source. In some embodiments, the polynucleotide comprises a gene derived from a bacterial or yeast source and recombinantly engineered into the microorganism of the disclosure.

As used herein, the term "metabolically engineered" or "metabolic engineering" involves rational pathway design and assembly of biosynthetic genes, genes associated with operons, and control elements of such polynucleotides, for the production of a desired metabolite, such as gamma-aminobutyric acid and the like as set forth in Figure 1 in a microorganism. "Metabolically engineered" can further include optimization of metabolic flux by regulation and optimization of transcription, translation, protein stability and protein functionality using genetic engineering and appropriate culture condition including the reduction of, disruption, or knocking out of, a competing metabolic pathway that competes for an intermediate leading to a desired pathway. A biosynthetic gene can be heterologous to the host microorganism, either by virtue of being foreign to the host, or being modified by mutagenesis, recombination, and/or association with a heterologous expression control sequence in an endogenous host cell.

In one embodiment, where the polynucleotide is xenogenetic to the host organism, the polynucleotide can be codon optimized.

The term "biosynthetic pathway", also referred to as "metabolic pathway", refers to a set of anabolic or catabolic biochemical reactions for converting (transmuting) one chemical species into another. Gene products belong to the same "metabolic pathway" if they, in parallel or in series, act on the same substrate, produce the same product, or act on or produce a metabolic intermediate (*i.e.*, metabolite) between the same substrate and metabolite end product.

The term "substrate" or "suitable substrate" refers to any substance or compound that is converted or meant to be converted into another compound by the action of an enzyme. The term includes not only a single compound, but also combinations of compounds, such as solutions, mixtures and other materials which contain at least one substrate, or derivatives thereof. Further, the term "substrate" encompasses not only compounds that provide a carbon source suitable for use as a starting material, such as any biomass derived sugar, but also intermediate and end product metabolites used in a pathway associated with a metabolically engineered microorganism as described herein. A "biomass derived sugar" includes, but is not limited to, molecules such as glucose, sucrose, mannose, xylose, and arabinose. The term biomass derived sugar encompasses suitable carbon substrates ordinarily used by microorganisms in either D or L form.

Recombinant microorganisms provided herein can express target enzymes involved in pathways for the production of gamma-aminobutyric acid by using a suitable carbon substrate.

Accordingly, metabolically "engineered" or "modified" microorganisms are produced via the introduction of genetic material into a host or parental microorganism of choice thereby modifying or altering the cellular physiology and biochemistry of the microorganism. Through the introduction of genetic material the parental microorganism acquires new properties, *e.g.*, the ability to produce a new, or greater quantities of, an intracellular metabolite. In an illustrative embodiment, the introduction of genetic material into a parental microorganism results in a new or modified ability to produce gamma-aminobutyric acid. The genetic material introduced into the parental microorganism contains gene(s), or parts of genes, coding for one or more of the enzymes involved in a biosynthetic pathway for the production of gamma-aminobutyric acid and

can also include additional elements for the expression and/or regulation of expression of these genes, *e.g.* promoter sequences.

An engineered or modified microorganism can also include in the alternative or in addition to the introduction of a genetic material into a host or parental microorganism, the disruption, deletion, or knocking out of a gene or polynucleotide to alter the cellular physiology and biochemistry of the microorganism. Through the reduction, disruption or knocking out of a gene or polynucleotide the microorganism acquires new or improved properties (*e.g.*, the ability to produced a new or greater quantities of an intracellular metabolite, improve the flux of a metabolite down a desired pathway, and/or reduce the production of undesirable by-products).

The disclosure demonstrates that the expression of one or more heterologous polynucleotide(s) or over-expression of one or more endogenous polynucleotide(s) encoding a polypeptide having glutamate decarboxylase activity can promote the production of gamma-aminobutyric acid. In one embodiment, the microorganism can further include expression or over-expression of a polypeptide having glutamate synthase activity. In yet another embodiment, the microorganism comprises an endogenous glutamate production pathway. For example, useful organisms having glutamate production pathways include those bacterial or yeast species from genera such as Brevibacterium, Arthrobacter, Microbacterium, and Corynebacterium.

For example, the disclosure demonstrates that with over-expression or heterologous expression of a polypeptide having glutamate decarboxylase activity in a microorganism comprising a glutamate synthesis pathway that gamma-aminobutyric acid can be effectively produced. In one embodiment, the polypeptide having glutamate decarboxylase activity is selected from the group consisting of GadA and GadB. In yet another embodiment, the polypeptide having glutamate decarboxylase activity has at least 60%-99% sequence identity to the glutamate decarboxylase consisting of SEQ ID NO:2. In another embodiment, the polypeptide having glutamate decarboxylase activity is obtained from a species selected from the group consisting of: *Escherichia coli*, *Shigella dysenteriae*, *Shigella flexneri*, *Escherichia albertii*, *Yersinia ruckeri*, *Edwardsiella tarda*, *Brucella suis*, *Brucella ovis*, *Akkermansia muciniphila*, *Desulfovibrio desulfuricans*, *Blastopirellula marina*, and *Photobacterium damselae*. In yet another embodiment, the glutamate decarboxylase comprises a sequence of SEQ ID NO:2:

MDQKLLTDFR SELLDSRFGA KAISTIAESK RFPLHEMRDD VAFQIINDEL YLDGNARQNL
ATFCQTWDDE NVHKLMDLSI NKNWIDKEEY PQSAAIDLRC VNMVADLWHA PAPKNGQAVG
TNTIGSSEAC MLGGMAMKWR WRKRMEAAGK PTDKPNLVCG PVQICWHKFA RYWDVELREI
PMRPGQLFMD PKRMIEACDE NTIGVVPTFG VTYTGNYEFP QPLHDALDKF QADTGIDIDM
HIDAASGGFL APFVAPDIVW DFRLPRVKSI SASGHKFGLA PLGCGWVIWR DEEALPQELV
FNVDYLGGQI GTFAINFSRP AGQVIAQYYE FLRLGREGYT KVQNASYQVA AYLADEIAKL
GPYEFICTGR PDEGIPAVCF KLKDGEDPGY TLYDLSERLR LRGWQVPAFT LGGEATDIVV
MRIMCRRGFE MDFAELLLED YKASLKYLSD HPKLQGIAQQ NSFKHT

In yet another embodiment, the glutamate decarboxylase is encoded by a polynucleotide having at least 70%-100 identity to a polynucleotide consisting of SEQ ID NO:1. For example, the polynucleotide comprises a *gadA* polynucleotide selected from the group consisting of: (a) a DNA comprising the nucleotide sequence of SEQ ID NO:1; and (b) a DNA that hybridizes with a nucleotide sequence consisting of SEQ ID NO:1 or a complement thereof and which encodes a polypeptide having glutamate decarboxylase activity.

SEQ ID NO:1:

atgGACCAGA AGCTGTTAAC GGATTTCCGC TCAGAACTAC TCGATTCACG TTTTGGCGCA
AAGGCCATTT CTACTATCGC GGAGTCAAAA CGATTTCCGC TGCACGAAAT GCGCGATGAT
GTCGCATTTC AGATTATCAA TGATGAATTA TATCTTGATG GCAACGCTCG TCAGAACCTG
GCCACTTTCT GCCAGACCTG GGACGACGAA AACGTCCATA AATTGATGGA TTTGTCGATC
AATAAAAACT GGATCGACAA AGAAGAATAT CCGCAATCCG CAGCCATCGA CCTGCGTTGC
GTAAATATGG TTGCCGATCT GTGGCATGCG CCTGCGCCGA AAAATGGTCA GGCCGTTGGC
ACCAACACCA TTGGTTCTTC CGAGGCCTGT ATGCTCGGCG GGATGGCGAT GAAATGGCGT
TGGCGCAAGC GTATGGAAGC TGCAGGCAAA CCAACGGATA AACCAAACCT GGTGTGCGGT
CCGGTACAAA TCTGCTGGCA TAAATTCGCC CGCTACTGGG ATGTGGAGCT GCGTGAGATC
CCTATGCGCC CCGGTCAGTT GTTTATGGAC CCGAAACGCA TGATTGAAGC CTGTGACGAA
AACACCATCG GCGTGGTGCC GACTTTCGGC GTGACCTACA CCGGTAACTA TGAGTTCCCA
CAACCGCTGC ACGATGCGCT GGATAAATTC CAGGCCGACA CCGGTATCGA CATCGACATG
CACATCGACG CTGCCAGCGG TGGCTTCCTG GCACCGTTCG TCGCCCCGGA TATCGTCTGG
GACTTCCGCC TGCCGCGTGT GAAATCGATC AGTGCTTCAG GCCATAAATT CGGTCTGGCT
CCGCTGGGCT GCGGCTGGGT TATCTGGCGT GACGAAGAAG CGCTGCCGCA GGAACTGGTG

TTCAACGTTG ACTACCTGGG TGGTCAAATT GGTACTTTTG CCATCAACTT CTCCCGCCCG
GCGGGTCAGG TAATTGCACA GTACTATGAA TTCCTGCGCC TCGGTCGTGA AGGCTATACC
AAAGTACAGA ACGCCTCTTA CCAGGTTGCC GCTTATCTGG CGGATGAAAT CGCCAAACTG
GGGCCGTATG AGTTCATCTG TACGGGTCGC CCGGACGAAG GCATCCCGGC GGTTTGCTTC
AAACTGAAAG ATGGTGAAGA TCCGGGATAC ACCCTGTACG ACCTCTCTGA ACGTCTGCGT
CTGCGCGGCT GGCAGGTTCC GGCCTTCACT CTCGGCGGTG AAGCCACCGA CATCGTGGTG
ATGCGCATTA TGTGTCGTCG CGGCTTCGAA ATGGACTTTG CTGAACTGTT GCTGGAAGAC
TACAAAGCCT CCCTGAAATA TCTCAGCGAT CACCCGAAAC TGCAGGGTAT TGCCCAGCAG
AACAGCTTTA AACACACCTG A

In yet another embodiment, a recombinant microorganism of the disclosure is further modified to enhance glutamic acid/GABA antiporter activity. In one embodiment, the glutamic acid/GABA antiporter activity is produced by a GadC polypeptide.

As used herein, "hybridization", "hybridizes" or "capable of hybridizing" is understood to mean the forming of a double or triple stranded molecule or a molecule with partial double or triple stranded nature. The term "hybridization", "hybridize(s)" or "capable of hybridizing" encompasses the terms "stringent condition(s)" or "high stringency" and the terms "low stringency" or "low stringency condition(s)."

As used herein "stringent condition(s)" or "high stringency" are those conditions that allow hybridization between or within one or more nucleic acid strand(s) containing complementary sequence(s), but precludes hybridization of random sequences. Stringent conditions tolerate little, if any, mismatch between a nucleic acid and a target strand. Such conditions are well known to those of ordinary skill in the art, and are preferred for applications requiring high selectivity. Non-limiting applications include isolating a nucleic acid, such as a gene or a nucleic acid segment thereof.

Stringent conditions may comprise low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.15 M NaCl at temperatures of about 50°C to about 70°C. It is understood that the temperature and ionic strength of a desired stringency are determined in part by the length of the particular nucleic acid(s), the length and nucleobase content of the target sequence(s), the charge composition of the nucleic acid(s), and to the presence or concentration of formamide, tetramethylammonium chloride or other solvent(s) in a hybridization mixture.

It is also understood that these ranges, compositions, and conditions for hybridization are mentioned by way of non-limiting examples only, and that the desired

stringency for a particular hybridization reaction is often determined empirically by comparison to one or more positive or negative controls. Depending on the application envisioned it is preferred to employ varying conditions of hybridization to achieve varying degrees of selectivity of a nucleic acid towards a target sequence. In a non-limiting example, identification or isolation of a related target nucleic acid that does not hybridize to a nucleic acid under stringent conditions may be achieved by hybridization at low temperature and/or high ionic strength. For example, a medium stringency condition could be provided by about 0.1 to 0.25 M NaCl at temperatures of about 37ºC to about 55ºC. Under these conditions, hybridization may occur even though the sequences of probe and target strand are not perfectly complementary, but are mismatched at one or more positions. In another example, a low stringency condition could be provided by about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20ºC to about 55ºC. Of course, it is within the skill of one in the art to further modify the low or high stringency conditions to suite a particular application.

As described herein, enhancement of glutamate decarboxylase activity can be achieved by enhancing the expression of a polynucleotide encoding glutamate decarboxylase. The expression can be enhanced by increasing expression of an endogenous gene by modification of an expression regulatory region such as a promoter; or by introduction of a plasmid containing a polynucleotide encoding a polypeptide having glutamate decarboxylase activity or the like.

The term "glutamate decarboxylase activity" refers to an activity of glutamate decarboxylase (GAD) that irreversibly catalyzes the reaction to produce gamma-aminobutyric acid by decarboxylation at the alpha-position of L-glutamic acid (EC: 4.1.1.15), and includes any number of different polypeptides and polynucleotides encoding such polypeptides (as described herein). A recombinant microorganism of the disclosure comprises a number of glutamate decarboxylase molecules per cell that is increased relative to a parental microorganism. Where a recombinant microorganism of the disclosure is derived from a parental strain that normally does not produce glutamic acid in any substantial amount, the microorganism can be modified to over-express or express a glutamate synthase or other enzyme(s) required for sufficient production of glutamic acid.

Escherichia coli is known to have a gene encoding glutamate decarboxylase termed the *gadA* gene encoding GadA protein (SEQ ID NO:2) and a *gadB* gene encoding

GadB protein. GadA and GadB are known to be isozymes having a very high homology (having 99% amino acid homology) (Smith *et al.*, *J. Bacteriol.* 174:5820-5826 (1992)).

The increased expression of a polynucleotide encoding glutamate decarboxylase compared to a parental strain such as a wild-type or unmodified strain can be confirmed by comparing the mRNA level of glutamate decarboxylase with that of a wild-type or an unmodified strain. Examples of a method of confirming the amount of expression can include Northern hybridization and RT-PCR (Molecular Cloning (Cold Spring Harbor Laboratory Press, Cold Spring Harbor (USA), 2001)). The amount of expression can be any level so long as it is increased as compared to a wild-type or a parental microorganism, and for example, the level is preferably increased by about 1.5-fold.

In another embodiment, the glutamic acid/GABA antiporter is known to be involved in acid resistance of cells grown in neutral pH at the time of addition of glutamic acid, and is presumed to play a role in maintaining the pH in cells in cooperation with GAD by simultaneously exporting intercellular glutamic acid and importing extracellular GABA (Hersh *et al.*, *J. Bacteriol.* 178:3978-3981 (1996)). *Escherichia coli* is known to have a *gadC* gene encoding GadC protein, which functions as a glutamic acid/GABA antiporter. The glutamic acid/GABA antiporter activity can be determined by, for example, the method described in Smith *et al.*, *J. Bacteriol.* 178:3978-3981 (1996).

Therefore, the "glutamate decarboxylase" in accordance with the presently disclosed subject matter can refer to GadA protein or GadB protein, and the "glutamic acid/GABA antiporter" can refer to a GadC protein.

As described in *J. Bacteriol.* 178:3978-3981 (1996), *gadB* gene and *gadC* gene form an operon on the genome of *E. coli*, so the operon can be used to simultaneously enhance expression of *gadB* and *gadC* genes. Hereinafter, the operon including the *gadB* and *gadC* genes is referred to as the *gadBC* operon, in some cases.

The *gadA* gene includes, for example, the *gadA* gene of an *Escherichia* bacterium, and homologs thereof. Examples of the *gadA* gene of *Escherichia coli* can include a gene encoding the GadA polypeptide which sequence is set forth in GenBank Accession No. AAC76542 [GI: 1789934] and set forth as (SEQ ID NO:2).

The *gadB* gene includes for example, the *gadB* gene of an *Escherichia* bacterium, and homologs thereof. Examples of the *gadB* gene of *Escherichia coli* can include a gene encoding the polypeptide having the amino acid sequence set forth in GenBank Accession No. AAC74566 [GI: 1787769] (SEQ ID NO:3).

Examples of the homologs of *gadA* and *gadB* genes can include *gadA* genes of a *Shigella* bacterium, *Vibrio* bacterium, and the like registered in GenBank. In addition, the *gadA* gene can be obtained by cloning, based on homology to any of the above-mentioned genes, from a *Streptomyces* bacterium such as *Streptomyces coelicolor*, a *Mycobacterium* such as *Mycobacterium tuberculosis*, or lactic acid bacterium such as *Lactococcus* or *Lactobacillus*. The homologs may be given different gene names however, they have homology to the *gadA* gene of an *Escherichia* bacterium. Further, the homologs of the *gadA* and *gadB* genes can be obtained by selecting genes having high homologies from a known database based on the above-mentioned sequence information. The homologs of amino acid sequences and nucleotide sequences can be determined by using, for example, the algorithm BLAST (Karlin andAtlschul, *Proc. Natl. Acad. Sci. USA* 90:5873-5877 (1993)) or FASTA (Pearson, *Methods Enzymol.* 183:63-98 (1990)). Based on the algorithm BLAST, programs called BLASTN and BLASTX have been developed and are available on the World Wide Web.

The *gadC* gene in accordance with the presently disclosed subject matter can include the GadC protein having the sequence as set forth in GenBank Accession No. AAC74565 [GI: 1787768] (SEQ ID NO:4) and homologs thereof.

As described elsewhere herein, the *gadA* and *gadB* polynucleotides are not limited to naturally occurring sequences but can be mutant or artificially modified polynucleotide molecules that encode a polypeptide having identity to a glutamate decarboxylase of the disclosure and having glutamate decarboxylase activity. Such modifications can include substitutions, deletions, insertions, and additions of one or several amino acids at one or a plurality of positions, so long as the function of the resulting polypeptide has glutamate decarboxylase activity.

Microorganisms provided herein are modified to produce metabolites in quantities not available in the parental microorganism. A "metabolite" refers to any substance produced by metabolism or a substance necessary for or taking part in a particular metabolic process. A metabolite can be an organic compound that is a starting material (*e.g.*, glutamic acid), an intermediate or an end product (*e.g.*, gamma-aminobutyric acid) of metabolism. Metabolites can be used to construct more complex molecules, or they can be broken down into simpler ones. Intermediate metabolites can be synthesized from other metabolites, perhaps used to make more complex substances, or broken down into simpler compounds, often with the release of chemical energy.

In addition, it can be advantageous to decrease or remove expression of enzymes that compete for a metabolite used in the production of gamma-aminobutyric acid.

The disclosure identifies specific genes and enzymes useful in the methods, compositions, and organisms of the disclosure; however it will be recognized that absolute identity to such genes or polypeptides is not necessary. For example, changes in a particular gene or polynucleotide comprising a sequence encoding a polypeptide or enzyme can be performed and screened for activity. Typically such changes comprise conservative mutations and/or silent mutations. Such modified or mutated polynucleotides and polypeptides can be screened for expression of a functional enzyme activity using methods known in the art.

Due to the inherent degeneracy of the genetic code, other polynucleotides which encode substantially the same or a functionally equivalent polypeptide can also be used to clone and express the polynucleotides encoding such enzymes.

As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence to enhance its expression in a particular host. The genetic code is redundant with 64 possible codons, but most organisms typically use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons. Codons can be substituted to reflect the preferred codon usage of the host, a process sometimes called "codon optimization" or "controlling for species codon bias."

Optimized coding sequences containing codons preferred by a particular prokaryotic or eukaryotic host (see also, Murray *et al. Nucl. Acids Res.* 17:477-508 (1989)) can be prepared, for example, to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced from a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, typical stop codons for *S. cerevisiae* and mammals are UAA and UGA, respectively. The typical stop codon for monocotyledonous plants is UGA, whereas insects and *E. coli* commonly use UAA as the stop codon (Dalphin *et al. Nucl. Acids Res.* 24:216-218 (1996)). Methodology for optimizing a nucleotide sequence for expression in a plant is provided, for example, in U.S. Pat. No. 6,015,891, and the references cited therein.

Those of skill in the art will recognize that, due to the degenerate nature of the genetic code, a variety of DNA compounds differing in their nucleotide sequences can be

used to encode a given amino acid sequence (*e.g.*, enzyme) of the disclosure. The native DNA sequence encoding the biosynthetic enzymes described above are referenced herein merely to illustrate an embodiment of the disclosure, and the disclosure includes DNA compounds of any sequence that encode the amino acid sequences of the polypeptides and proteins of the enzymes utilized in the methods of the disclosure. In similar fashion, a polypeptide can typically tolerate one or more amino acid substitutions, deletions, and/or insertions in its amino acid sequence without loss or significant loss of a desired activity. The disclosure includes such polypeptides with different amino acid sequences than the specific proteins described herein so long as the modified or variant polypeptides have the enzymatic anabolic or catabolic activity of the reference polypeptide. Furthermore, the amino acid sequences encoded by the DNA sequences shown herein merely illustrate embodiments of the disclosure.

In addition, homologs of enzymes useful for generating metabolites are encompassed by the microorganisms and methods provided herein. The term "homologs" used with respect to an original enzyme or gene of a first family or species refers to distinct enzymes or genes of a second family or species which are determined by functional, structural or genomic analyses to be an enzyme or gene of the second family or species which corresponds to the original enzyme or gene of the first family or species. Most often, homologs will have functional, structural, or genomic similarities. Techniques are known by which homologs of an enzyme or gene can readily be cloned using genetic probes and PCR. Identity of cloned sequences as a homolog can be confirmed using functional assays and/or by genomic mapping of the genes.

A protein has "homology" or is "homologous" to a second protein if the nucleic acid sequence that encodes the protein has a similar sequence to the nucleic acid sequence that encodes the second protein. Alternatively, a protein has homology to a second protein if the two proteins have "similar" amino acid sequences. (Thus, the term "homologous proteins" is defined to mean that the two proteins have similar amino acid sequences).

As used herein, two proteins (or a region of the proteins) are substantially homologous when the amino acid sequences have at least about 30%, 40%, 50% 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity. To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*,

gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In one embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, typically at least 40%, more typically at least 50%, even more typically at least 60%, and even more typically at least 70%, 80%, 90%, or 100% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

When "homologous" is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (*e.g.*, charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative substitution, the percent sequence identity or degree of homology can be adjusted upwards to correct for the conservative nature of the substitution. Methods for making such an adjustment are well known to those of skill in the art (see, *e.g.*, Pearson *et al.*, *Meth. Mol. Biol.* 25:365-389 (1994), hereby incorporated herein by reference).

These following groups contain amino acid families that are considered conservative substitutions for one another and include amino acids with basic side chains (*e.g.*, lysine (K), arginine (R), histidine (H)), acidic side chains (*e.g.*, aspartic acid (D), glutamic acid (E)), uncharged polar side chains (*e.g.*, glycine (G), asparagine (N), glutamine (Q), serine (S), threonine (T), tyrosine (Y), cysteine (C)), nonpolar side chains (*e.g.*, alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), phenylalanine (F), methionine (M), tryptophan (W)), beta-branched side chains (*e.g.*, threonine (T), valine (V), isoleucine (I)) and aromatic side chains (*e.g.*, tyrosine (Y), phenylalanine (F),

tryptophan (W), histidine (H)). Another grouping of amino acids considered conservative substitutions by the skilled artisan is as follows: 1) Serine (S), Threonine (T); 2) Aspartic Acid (D), Glutamic Acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Sequence homology for polypeptides, which is also referred to as percent sequence identity, is typically measured using sequence analysis software. See, *e.g.*, the Sequence Analysis Software Package of the Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 910 University Avenue, Madison, Wis. 53705. Protein analysis software matches similar sequences using measure of homology assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild type protein and a mutein thereof. See, *e.g.*, GCG Version 6.1.

A typical algorithm used for comparing a molecule sequence to a database containing a large number of sequences from different organisms is the computer program BLAST (Altschul, *J. Mol. Biol.* 215:403-441 (1990); Gish, *Nature Genet.* 3:266-272 (1993); Madden, *Meth. Enzymol.* 266:131-141 (1996); Altschul, *Nucl. Acids Res.* 25:3389-3402 (1997); Zhang, *Genome Res.* 7:649-656 (1997)), especially blastp or tblastn (Altschul, *Nucl. Acids Res.* 25:3389-3402 (1997)). Typical parameters for BLASTp are: Expectation value: 10 (default); Filter: seg (default); Cost to open a gap: 11 (default); Cost to extend a gap: 1 (default); Max. alignments: 100 (default); Word size: 11 (default); No. of descriptions: 100 (default); Penalty Matrix: BLOWSUM62.

When searching a database containing sequences from a large number of different organisms, it is typical to compare amino acid sequences. Database searching using amino acid sequences can be measured by algorithms other than blastp known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson, *Meth. Enzymol.* 183:63-98 (1990), hereby incorporated herein by reference). For example, percent sequence identity between amino acid sequences can be determined using FASTA with

its default parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1, hereby incorporated herein by reference.

The disclosure provides accession numbers for various genes, homologs and variants useful in the generation of recombinant microorganism described herein. It is to be understood that homologs and variants described herein are exemplary and non-limiting. Additional homologs, variants and sequences are available to those of skill in the art using various databases including, for example, the National Center for Biotechnology Information (NCBI) access to which is available on the World-Wide-Web.

It is understood that a range of microorganisms can be modified to include a recombinant metabolic pathway suitable for the production of gamma-aminobutyric acid. It is also understood that various microorganisms can act as "sources" for genetic material encoding target enzymes suitable for use in a recombinant microorganism provided herein. The term "microorganism" includes prokaryotic and eukaryotic microbial species from the Domains Archaea, Bacteria and Eucarya, the latter including yeast and filamentous fungi, protozoa, algae, or higher Protista. The terms "microbial cells" and "microbes" are used interchangeably with the term microorganism.

The term "prokaryotes" is art recognized and refers to cells which contain no nucleus or other cell organelles. The prokaryotes are generally classified in one of two domains, the Bacteria and the Archaea. The definitive difference between organisms of the Archaea and Bacteria domains is based on fundamental differences in the nucleotide base sequence in the 16S ribosomal RNA.

The term "Archaea" refers to a categorization of organisms of the division Mendosicutes, typically found in unusual environments and distinguished from the rest of the prokaryotes by several criteria, including the number of ribosomal proteins and the lack of muramic acid in cell walls. On the basis of ssrRNA analysis, the Archaea comprise two phylogenetically-distinct groups: Crenarchaeota and Euryarchaeota. On the basis of their physiology, the Archaea can be organized into three types: methanogens (prokaryotes that produce methane); extreme halophiles (prokaryotes that live at very high concentrations of salt ((NaCl)); and extreme (hyper) thermophilus (prokaryotes that live at very high temperatures). Besides the unifying archaeal features that distinguish them from Bacteria (*i.e.*, no murein in cell wall, ester-linked membrane lipids, and the like.), these prokaryotes exhibit unique structural or biochemical attributes which adapt them to their particular habitats. The Crenarchaeota comprise mainly hyperthermophilic

sulfur-dependent prokaryotes and the Euryarchaeota contains the methanogens and extreme halophiles.

"Bacteria", or "eubacteria", refers to a domain of prokaryotic organisms. Bacteria include at least 11 distinct groups as follows: (1) Gram-positive (gram$^+$) bacteria, of which there are two major subdivisions: (1) high G+C group (Actinomycetes, Mycobacteria, Micrococcus, others) (2) low G+C group (Bacillus, Clostridia, Lactobacillus, Staphylococci, Streptococci, Mycoplasmas); (2) Proteobacteria, *e.g.*, Purple photosynthetic +non-photosynthetic Gram-negative bacteria (includes most "common" Gram-negative bacteria); (3) Cyanobacteria, *e.g.*, oxygenic phototrophs; (4) Spirochetes and related species; (5) Planctomyces; (6) Bacteroides, Flavobacteria; (7) Chlamydia; (8) Green sulfur bacteria; (9) Green non-sulfur bacteria (also anaerobic phototrophs); (10) Radioresistant micrococci and relatives; and (11) Thermotoga and Thermosipho thermophiles.

"Gram-negative bacteria" include cocci, nonenteric rods, and enteric rods. The genera of Gram-negative bacteria include, for example, Neisseria, Spirillum, Pasteurella, Brucella, Yersinia, Francisella, Haemophilus, Bordetella, Escherichia, Salmonella, Shigella, Klebsiella, Proteus, Vibrio, Pseudomonas, Bacteroides, Acetobacter, Aerobacter, Agrobacterium, Azotobacter, Spirilla, Serratia, Vibrio, Rhizobium, Chlamydia, Rickettsia, Treponema, and Fusobacterium.

"Gram positive bacteria" include cocci, nonsporulating rods, and sporulating rods. The genera of gram positive bacteria include, for example, Actinomyces, Bacillus, Clostridium, Corynebacterium, Erysipelothrix, Lactobacillus, Listeria, Mycobacterium, Myxococcus, Nocardia, Staphylococcus, Streptococcus, and Streptomyces.

The term "recombinant microorganism" and "recombinant host cell" are used interchangeably herein and refer to microorganisms that have been genetically modified to express or over-express endogenous polynucleotides, or to express non-endogenous polynucleotides, such as those included in a vector, or which have a reduction in expression of an endogenous gene. The polynucleotide generally encodes a target enzyme involved in a metabolic pathway for producing a desired metabolite as described above. Accordingly, recombinant microorganisms described herein have been genetically engineered to express or over-express target enzymes not previously expressed or over-expressed by a parental microorganism. It is understood that the terms "recombinant microorganism" and "recombinant host cell" refer not only to the particular

recombinant microorganism but to the progeny or potential progeny of such a microorganism.

A "parental microorganism" refers to a cell used to generate, or derive, a recombinant microorganism. The term "parental microorganism" describes a cell that occurs in nature, *i.e.* a "wild-type" cell that has not been genetically modified. The term "parental microorganism" also describes a cell that has been genetically modified but which does not express or over-express a target enzyme *e.g.*, an enzyme involved in the biosynthetic pathway for the production of a desired metabolite. For example, a wild-type microorganism can be genetically modified to express or over-express a first target enzyme such as a glutamate synthase. This microorganism can act as a parental microorganism in the generation of a microorganism modified to express or over-express a second target enzyme *e.g.*, glutamate decarboxylase. In turn, the microorganism modified to express or over-express *e.g.*, a glutamate decarboxylase can be modified to express or over-express a third target enzyme. Accordingly, a parental microorganism functions as a reference cell for successive genetic modification events. Each modification event can be accomplished by introducing a nucleic acid molecule into the reference cell. The introduction facilitates the expression or over-expression of a target enzyme. It is understood that the term "facilitates" encompasses the activation of endogenous polynucleotides encoding a target enzyme through genetic modification of *e.g.*, a promoter sequence in a parental microorganism. It is further understood that the term "facilitates" encompasses the introduction of exogenous polynucleotides encoding a target enzyme into a parental microorganism.

In another embodiment a method of producing a recombinant microorganism that converts a suitable carbon substrate to gamma-aminobutyric acid is provided. The method includes transforming a microorganism with one or more recombinant polynucleotides encoding polypeptides that include, for example, glutamate decarboxylase. Polynucleotides that encode enzymes useful for generating metabolites including homologs, variants, fragments, related fusion proteins, or functional equivalents thereof, are used in recombinant nucleic acid molecules that direct the expression of such polypeptides in appropriate an host cell, such as bacterial or yeast cells. It is understood that the addition of polynucleotide sequences which do not alter the encoded activity of a polynucleotide, such as the addition of a non-functional or non-coding sequence, is a conservative variation of the basic nucleic acid. The "activity" of an enzyme is a measure

of its ability to catalyze a reaction resulting in a metabolite, *i.e.*, to "function", and may be expressed as the rate at which the metabolite of the reaction is produced. For example, enzyme activity can be represented as the amount of metabolite produced per unit of time or per unit of enzyme (*e.g.*, concentration or weight), or in terms of affinity or dissociation constants.

A "protein" or "polypeptide", which terms are used interchangeably herein, comprises one or more chains of chemical building blocks called amino acids that are linked together by peptide bonds. An "enzyme" means any substance, composed wholly or largely of protein, that catalyzes or promotes, more or less specifically, one or more chemical or biochemical reactions. A "native" or "wild-type" protein, enzyme, polynucleotide, gene, or cell, means a protein, enzyme, polynucleotide, gene, or cell that occurs in nature.

It is understood that the polynucleotides described above include "genes" and that the nucleic acid molecules described above include "vectors" or "plasmids." Accordingly, the term "gene", also called a "structural gene", refers to a polynucleotide that codes for a particular sequence of amino acids, which comprise all or part of one or more proteins or enzymes, and may include regulatory (non-transcribed) nucleotide (DNA) sequences, such as promoter sequences, which determine for example the conditions under which the gene is expressed. The transcribed region of the gene may include untranslated regions, including introns, 5'-untranslated region (UTR), and 3'-UTR, as well as the coding sequence. The term "nucleic acid" or "recombinant nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term "expression" with respect to a gene sequence refers to transcription of the gene and, as appropriate, translation of the resulting mRNA transcript to a protein. Thus, as will be clear from the context, expression of a protein results from transcription and translation of the open reading frame sequence.

The term "operon" refers to two or more genes which are transcribed as a single transcriptional unit from a common promoter. In some embodiments, the genes comprising the operon are contiguous genes. It is understood that transcription of an entire operon can be modified (*i.e.*, increased, decreased, or eliminated) by modifying the common promoter. Alternatively, any gene or combination of genes in an operon can be modified to alter the function or activity of the encoded polypeptides. The modification

can result in an increase in the activity of the encoded polypeptides. Further, the modification can impart new activities on the encoded polypeptides. Exemplary new activities include the use of alternative substrates and/or the ability to function in alternative environmental conditions.

A "vector" is any means by which a nucleic acid can be propagated and/or transferred between organisms, cells, or cellular components. Vectors include viruses, bacteriophage, pro-viruses, plasmids, phagemids, transposons, and artificial chromosomes such as YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), and PLACs (plant artificial chromosomes), and the like, that are "episomes," that is, that replicate autonomously or can integrate into a chromosome of a host cell. A vector can also be a naked RNA polynucleotide, a naked DNA polynucleotide, a polynucleotide composed of both DNA and RNA within the same strand, a poly-lysine -conjugated DNA or RNA, a peptide-conjugated DNA or RNA, a liposome-conjugated DNA, or the like, that are not episomal in nature, or it can be an organism which comprises one or more of the above polynucleotide constructs such as an agrobacterium or a bacterium.

"Transformation" refers to the process by which a vector is introduced into a host cell. Transformation (or transduction, or transfection), can be achieved by any one of a number of means including electroporation, microinjection, biolistics (or particle bombardment-mediated delivery), or agrobacterium mediated transformation. These means of transformation are well known to the skilled artisan.

The disclosure provides nucleic acid molecules in the form of recombinant DNA expression vectors or plasmids, as described in more detail below, that encode one or more target enzymes. Generally, such vectors can either replicate in the cytoplasm of the host microorganism or integrate into the chromosomal DNA of the host microorganism. In either case, the vector can be a stable vector (*i.e.*, the vector remains present over many cell divisions, even if only with selective pressure) or a transient vector (*i.e.*, the vector is gradually lost by a host microorganism with increasing numbers of cell divisions). The disclosure provides DNA molecules in isolated (*i.e.*, not pure, but existing in a preparation in an abundance and/or concentration not found in nature) and purified (*i.e.*, substantially free of contaminating materials or substantially free of materials with which the corresponding DNA would be found in nature) forms.

Provided herein are methods for the heterologous expression of one or more of the biosynthetic genes involved in gamma-aminobutyric acid biosynthesis and recombinant DNA expression vectors useful in the method. Thus, included within the scope of the disclosure are recombinant expression vectors that include such nucleic acids. The term expression vector refers to a nucleic acid that can be introduced into a host microorganism or cell-free transcription and translation system. An expression vector can be maintained permanently or transiently in a microorganism, whether as part of the chromosomal or other DNA in the microorganism or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a promoter that drives expression of an RNA, which typically is translated into a polypeptide in the microorganism or cell extract. For efficient translation of RNA into protein, the expression vector also typically contains a ribosome-binding site sequence positioned upstream of the start codon of the coding sequence of the gene to be expressed. Other elements, such as enhancers, secretion signal sequences, transcription termination sequences, and one or more marker genes by which host microorganisms containing the vector can be identified and/or selected, may also be present in an expression vector. Selectable markers, *i.e.*, genes that confer antibiotic resistance or sensitivity, are used and confer a selectable phenotype on transformed cells when the cells are grown in an appropriate selective medium.

The various components of an expression vector can vary widely, depending on the intended use of the vector and the host cell(s) in which the vector is intended to replicate or drive expression. Expression vector components suitable for the expression of genes and maintenance of vectors in *E. coli*, yeast, Streptomyces, and other commonly used host cells are widely known and commercially available. For example, suitable promoters for inclusion in the expression vectors of the disclosure include those that function in eukaryotic or prokaryotic host microorganisms. Promoters can comprise regulatory sequences that allow for regulation of expression relative to the growth of the host microorganism or that cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus. For *E. coli* and certain other bacterial host cells, promoters derived from genes for biosynthetic enzymes, antibiotic-resistance conferring enzymes, and phage proteins can be used and include, for example, the galactose, lactose (lac), maltose, tryptophan (trp), beta-lactamase (bla), bacteriophage lambda PL, and T5 promoters. In addition, synthetic promoters, such as the tac promoter

(U.S. Pat. No. 4,551,433), can also be used. For *E. coli* expression vectors, it is useful to include an *E. coli* origin of replication, such as from pUC, p1P, p1, and pBR.

Thus, recombinant expression vectors contain at least one expression system, which, in turn, is composed of at least a portion of polyketide synthase (PKS) and/or other biosynthetic gene coding sequences operably linked to a promoter and optionally termination sequences that operate to effect expression of the coding sequence in compatible host cells. The host cells are modified by transformation with the recombinant DNA expression vectors of the disclosure to contain the expression system sequences either as extrachromosomal elements or integrated into the chromosome.

A nucleic acid of the disclosure can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques and those procedures described in the Examples section below. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

It is also understood that an isolated nucleic acid molecule encoding a polypeptide homologous to the enzymes described herein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence encoding the particular polypeptide, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into the polynucleotide by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. In contrast to those positions where it may be desirable to make a non-conservative amino acid substitutions (see above), in some positions it is preferable to make conservative amino acid substitutions. As noted above, a conservative amino acid substitution is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art as discussed above.

In another embodiment a method for producing gamma-aminobutyric acid is provided. The method includes culturing a recombinant microorganism as provided herein in the presence of a suitable substrate and under conditions suitable for the conversion of the substrate to gamma-aminobutyric acid or an intermediate which may be further converted to gamma-aminobutyric acid. The gamma-aminobutyric acid produced

by a microorganism provided herein can be detected by any method known to the skilled artisan. Such methods include mass spectrometry. Culture conditions suitable for the growth and maintenance can be modified to accommodate the requirements of each microorganism.

As previously discussed, general texts which describe molecular biological techniques useful herein, including the use of vectors, promoters and many other relevant topics, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology Volume 152, (Academic Press, Inc., San Diego, Calif.) ("Berger"); Sambrook and Russell, Molecular Cloning--A Laboratory Manual, 3rd ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 2001 ("Sambrook") and Current Protocols in Molecular Biology, Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2011) ("Ausubel"). Examples of protocols sufficient to direct persons of skill through *in vitro* amplification methods, including the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Q-replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA), *e.g.*, for the production of the homologous nucleic acids of the disclosure are found in Berger, Sambrook, and Ausubel, as well as in Mullis *et al.* U.S. Pat. No. 4,683,202 (1987); Innis *et al.*, eds., PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, Calif. (1990)) ("Innis"); Arnheim and Levenson, C&EN 68:36-47 (1990); Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA* 86:1173 (1989); Guatelli *et al.*, *Proc. Natl. Acad. Sci. USA* 87:1874 (1990); Lomell *et al.*, *J. Clin. Chem.* 35:1826 (1989); Landegren *et al.*, *Science* 241:1077-1080 (1988); Van Brunt, *Biotechnology* 8:291-294 (1990); Wu and Wallace, *Gene* 4:560 (1989); Barringer *et al. Gene* 89:117 (1990); and Sooknanan and Malek, *Biotechnology* 13:563-564 (1995). Improved methods for cloning *in vitro* amplified nucleic acids are described in Wallace *et al.*, U.S. Pat. No. 5,426,039. Improved methods for amplifying large nucleic acids by PCR are summarized in Cheng *et al.*, *Nature* 369:684-685 (1994) and the references cited therein, in which PCR amplicons of up to 40 kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. See, *e.g.*, Ausubel, Sambrook and Berger, all *supra*.

Appropriate culture conditions are conditions of culture medium pH, ionic strength, nutritive content, and the like; temperature; oxygen/CO_2/nitrogen content;

humidity; and other culture conditions that permit production of the compound by the host microorganism, *i.e.*, by the metabolic action of the microorganism. Appropriate culture conditions are well known for microorganisms that can serve as host cells.

EXAMPLE

The following example demonstrates the construction of a recombinant microorganism, *Corynebacterium glutamicum*, that produces gamma-aminobutyric acid by transformation with a polynucleotide sequence that encodes a glutamate decarboxylase. The polynucleotide sequence that encodes that glutamate decarboxylase was obtained from *Escherichia coli*.

Cloning of Glutamate Decarboxylase Genes

The *E. coli* glutamate decarboxylase (*gadA*, SEQ ID NO:1) was amplified from genomic DNA of *E. coli* MG1655 with synthesized primer F1 (5'-ggccgctggagatggaccagaagctgtta-3'; SEQ ID NO:5) and R1 (5'-ccaatgcattcaggtgtgtttaaagct; SEQ ID NO:6). The PCR products were recovered using a commercial DNA cleaning and concentration product. such as Zymo Kit (Zymo Research Corp., Irvine, CA) and then digested with XhoI and NsiI. The purified fragment was ligated into XhoI and NsiI sites of pKS140 plasmid (Smith *et al.*, *Appl. Microbiol. Biotechnol.* 87:1045-1055 (2010)), which has the strong and constitutive *eftu* promoter, which transcribes the *eftu* gene (Cg0587). The resulting plasmid was designated pN4-1 and transformed into *Corynebacterium glutamicum* ATCC13032 to test production of GABA directly from carbon sources.

Enzyme Assay for Glutamate Decarboxylase

The *Corynebacterium glutamicum* (pN4-1) was grown in 5 ml CGXIII medium (Menkel *et al.*, *Appl. Environ. Microbiol.* 55:684-688 (1989)) for 24 hours. Crude extracts were prepared by concentrating the cultures by 5-fold in 0.1 M phosphate buffer (pH 7.1), and lysing them with 0.1 mM glass beads using bead beater. Enzyme assay for the conversion of glutamic acid to GABA was done by the reaction mixture consisting of 100 µl of crude extracts, PLP 50 µM, glutamic acid 50 mM in phosphate buffer (pH 6.2). After 30 minutes incubation at 30°C, products were analyzed using HPLC with the amino acid analysis method described by Zhang *et al.* (*Proc. Natl. Acad. Sci. USA* 107:6234-6239 (2010)).

Production of GABA

Seed culture of the production strain *Corynebacterium glutamicum* 13032 (pN4-1) was grown in 5 ml CGXIII medium containing 25 µg/ml at 30°C, 250 rpm overnight. Seed culture (90.5 ml) was inoculated into 20 ml of the production medium in a 250 ml flask. The GABA production medium contained glucose (40 g/L), $MgSO_4$ (0.5 g/L), $FeSO_4$ (20 mg/L), $MnSO_4$ (20mg/L), Biotin (1 µg/L), Thiamin (300 µg/L), Yeast extract (2.5 g/L), $(NH_4)_2HPO_4$ (1 g/l), Urea (8 g/L), KH_2PO_4 (1 g/L), K_2HPO_4 (0.5 g/L) and 0.25 mM PLP. After 48 hours cultivation at 30°C, 250 rpm, GABA production was analyzed using HPLC. The amount of GABA produced by the various *Corynebacterium glutamicum* strains is depicted in Table 1.

Tabl26e 1. GABA production from direct fermentation

Strain	GABA Produced (g/L)
C. glutamicum 13032	0
C. glutamicum 13032 (pKS140)	0
C. glutamicum 13032 (pN4-1)	1.3

A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.

CLAIMS

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. A recombinant microorganism derived from a parental microorganism that produces gamma-aminobutyric acid (GABA) and over-expresses a polypeptide having glutamate decarboxylase activity.

2. The recombinant microorganism of claim 1, wherein the polypeptide is endogenous or heterologous to the parental microorganism.

3. The recombinant micooorganism of claim 1, wherein the microorganism is derived from a genera selected from Brevibacterium, Arthrobacter, Microbacterium, and Corynebacterium.

4. The recombinant microorganism of claim 3, wherein the microorganism is derived from *Corynebacterium glutamicum.*

5. The recombinant microorganism of claim 1, wherein the polypeptide having glutamate decarboxylase activity is gadA or gadB.

6. The recombinant microorganism of claim 1, wherein the polypeptide having glutamate decarboxylase activity comprises an amino acid sequence of SEQ ID NO:2, or a conservative-substituted variant of or homologous polypeptide to SEQ ID NO:2 having at least 60% identity to SEQ ID NO:2 and having glutamate decarboxylase activity.

7. The recombinant microorganism of claim 6, wherein the polypeptide having glutamate decarboxylase activity is obtained from a species selected from *E. coli, Shigella dysenteriae, Shigella flexneri, Escherichia albertii, Yersinia ruckeri, Edwardsiella tarda, Brucella suis, Brucella ovis, Akkermansia muciniphila, Desulfovibrio desulfuricans, Blastopirellula marina,* and *Photobacterium damselae.*

8. The recombinant microorganism of claim 1, wherein the polypeptide having glutamate decarboxylase activity is encoded by a gene having at least 70-100% identity to polynucleotide SEQ ID NO:1.

9. The recombinant microorganism of claim 10, wherein the polynucleotide having glutamate decarboxylase activity comprises a gadA polynucleotide selected from: (a) a deoxyribonucleic acid (DNA) comprising the nucleotide sequence of SEQ ID NO:1; and (b) a DNA that hybridizes with the nucleotide sequence of SEQ ID NO:1 or a complement thereof and which encodes a polypeptide having glutamate decarboxylase activity.

10. The recombinant microorganism of claim 1, wherein the microorganism is further modified to enhance glutamic acid/GABA antiporter activity compared to a parental microorganism.

11. The recombinant microorganism of claim 12, wherein glutamic acid/GABA antiporter activity is produced by a gadC polypeptide.

12. A method of producing gamma-aminobutyric acid, comprising:

(a) culturing a recombinant cell or an extract of a cell of claim 1 with a suitable carbon source, wherein the cell or cell extract converts the carbon source to gamma-aminobutyric acid; and

(b) isolating the gamma-aminobutyric acid.

13. The method of claim 12, wherein the suitable carbon source is a sugar.

14. The method of claim 13, wherein the sugar is glucose, sucrose, or a combination thereof.

15. A method of producing a recombinant microorganism that produces gamma-aminobutyric acid, comprising transforming a microorganism with a nucleic acid sequence encoding a polypeptide having glutamate decarboxylase activity.

1/1

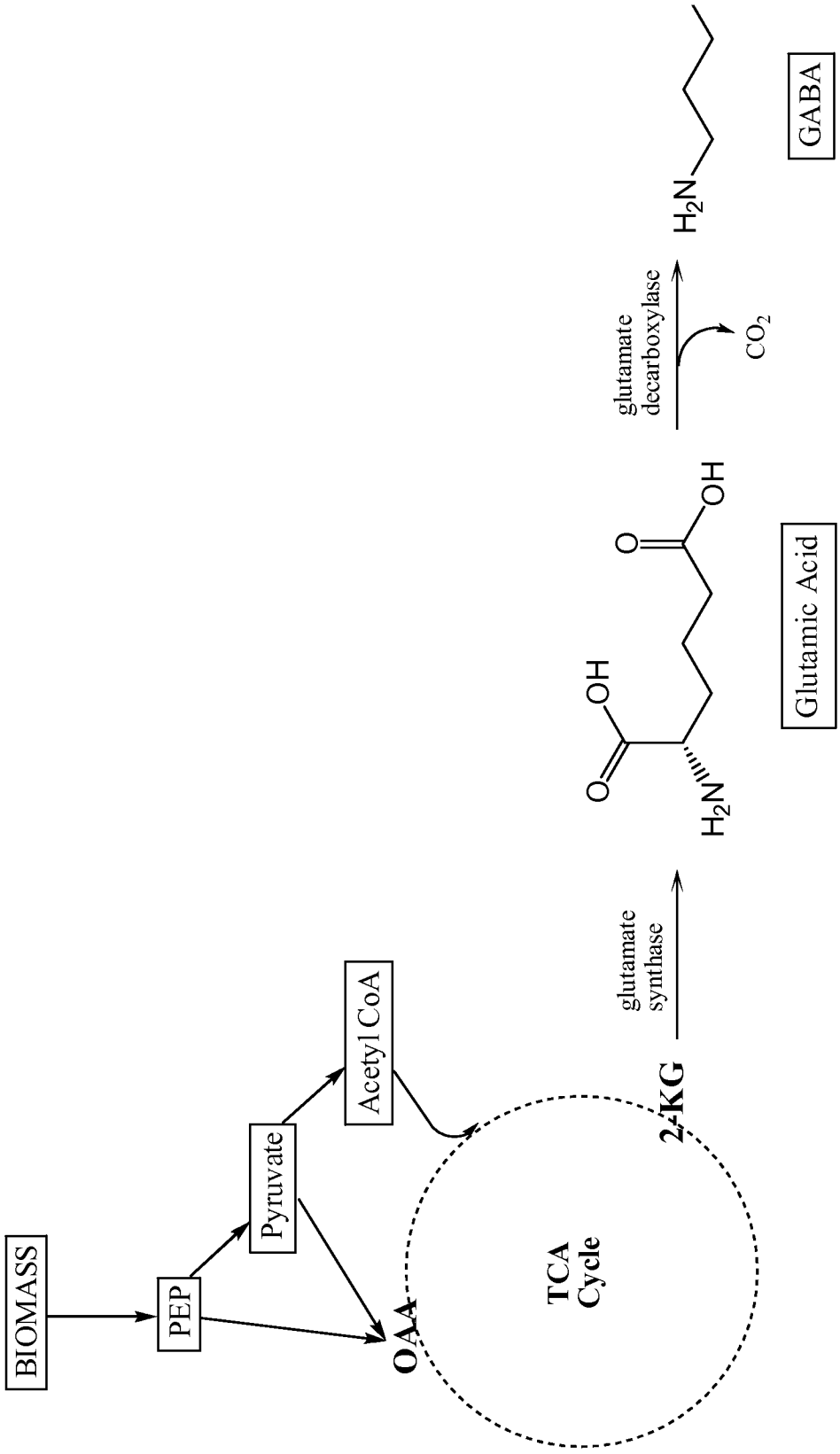
BIOMASS
PEP
Pyruvate
Acetyl CoA
OAA
TCA Cycle
2-KG
glutamate synthase
Glutamic Acid
glutamate decarboxylase
CO_2
GABA

FIGURE 1

International application No
PCT/US2011/034612

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N1/21 C12N9/88 C12P13/00
ADD.

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)
C12N C12P

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, BIOSIS, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/103547 A1 (BASF SE [DE]; ZELDER OSKAR [DE]; JEONG WEOL KYU [KR]; KLOPPROGGE CORIN) 27 August 2009 (2009-08-27) page 1 - page 7 page 21 - page 28; claims 1-27 -----	1-15

☐ 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 document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 July 2011	27/07/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Devijver, Kristof

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2011/034612

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

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

   a. (means)

   [ ] on paper

   [X] in electronic form

   b. (time)

   [X] in the international application as filed

   [ ] together with the international application in electronic form

   [ ] subsequently to this Authority for the purpose of search

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2011/034612

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2009103547 A1	27-08-2009	CN 101945997 A	12-01-2011
		EP 2247724 A1	10-11-2010
		JP 2011512144 A	21-04-2011
		KR 20110004368 A	13-01-2011
		US 2010324258 A1	23-12-2010