

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 March 2006 (30.03.2006)

PCT

(10) International Publication Number
WO 2006/034501 A2

- (51) International Patent Classification: Not classified
- (21) International Application Number:
PCT/US2005/034447
- (22) International Filing Date:
23 September 2005 (23.09.2005)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/612,803 23 September 2004 (23.09.2004) US
- (71) Applicant (for all designated States except US): **UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INC.** [US/US]; 223 Grinter Hall, Gainesville, FL 33611 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **HANSON, Andrew, D.** [US/US]; 3962 SW 5th Place, Gainesville, FL 32607 (US). **GREGORY, Jesse, F., III** [US/US]; 38 SW 5th Place, Gainesville, FL 32607 (US). **DIAZ DE LA GARZA, Rocio** [MX/US]; 2901 SW 13th Street, Apt. 241, Gainesville, FL 32607 (US). **SCHUBERT, Karel, R.** [US/US]; 817 Berry Hill Drive, St. Louis, MO 63132 (US). **HOSSAIN, Tahzeeba** [US/US]; 12800 Town and Four Drive, Apt. A, St. Louis, MO 63141 (US).
- (74) Agents: **PACE, Doran, R.** et al.; Saliwanchik, Lloyd & Saliwanchik, A Professional Association, P.O. Box 142950, Gainesville, FL 32614-2950 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: MATERIALS AND METHODS FOR FOLATE BIOFORTIFICATION IN PLANTS

(57) Abstract: The subject invention concerns materials and methods for enhancing folate content of plants. In one embodiment of a method of the invention, a plant is engineered to increase levels of pteridines in the plant. In an exemplified embodiment, a plant is transformed with a polynucleotide encoding a mammalian-type GCHI polypeptide or a bacterial GCHI polypeptide that is free of feedback control when expressed in plants. The subject invention also concerns plants having enhanced folate content. In one embodiment, a plant of the invention comprises and expresses a polynucleotide encoding a mammalian-type GCHI polypeptide or a bacterial GCHI polypeptide that is free of feedback control when expressed in the plant.



WO 2006/034501 A2

DESCRIPTION

MATERIALS AND METHODS FOR FOLATE BIOFORTIFICATION IN PLANTS

5 This invention was made with government support under National Aeronautical and Space Administration grant number NAG2-1525 and National Science Foundation grant number MCB-0129944. The government has certain rights in the invention.

CROSS-REFERENCE TO A RELATED APPLICATION

10 This application claims the benefit of U.S. Provisional Application Serial No. 60/612,803, filed September 23, 2004, which is hereby incorporated by reference herein in its entirety, including any figures, tables, nucleic acid sequences, amino acid sequences, and
15 drawings.

BACKGROUND OF THE INVENTION

20 Tetrahydrofolate (THF) and its derivatives (collectively termed folates) are essential cofactors for the one-carbon transfer reactions needed to form methionine, serine, purines, and thymidylate (Cossins *et al.*, 1997; Scott *et al.*, 2000). Folates are tripartite molecules that consist of pteridine, *p*-aminobenzoate (PABA), and one or more glutamate moieties (Figure 1A). In plants, the pteridine moiety is produced in the cytosol, PABA is formed in plastids, and the two are coupled together and glutamylated in mitochondria (Figure 1B) (Hanson *et al.*, 2002; Ravanel *et al.*, 2001; Basset *et al.*, 2002; Basset *et al.*, 2004; Goyer *et al.*, 2004). In
25 tomato fruit, the first enzyme of pteridine synthesis, GTP cyclohydrolase I (GCHI), disappears at the onset of ripening (Basset *et al.*, 2002), cutting off the pteridine supply. Ripening tomatoes contain large pools of glutamate (Boggio *et al.*, 2000) and moderate pools of PABA (Basset *et al.*, 2004; Quinlivan *et al.*, 2003).

30 Whereas plants and microorganisms can synthesize folates, humans and other animals lack a complete folate synthesis pathway and so require a dietary supply. For humans, this supply comes mostly from plant sources (Scott *et al.*, 2000; Konings *et al.*, 2001). Green leafy vegetables and legume seeds are folate-rich, but staple foods such as cereals, tubers, and many fruits are relatively poor in folate; related to this, folate intakes are seriously inadequate

in poor countries and suboptimal even in rich ones (Konings *et al.*, 2001; de Bree *et al.*, 1997; Krishnaswamy *et al.*, 2001). The consequences of low folate intake include birth defects, anemia, and increased risk of vascular disease and some cancers (Krishnaswamy *et al.*, 2001; Lucock 2000; Molloy *et al.*, 2001). The March of Dimes has stated that inadequate intake of
5 folate before pregnancy is the most common cause of birth defects, including neural tube defects (NTDs) such as spina bifida and anencephaly. NTDs affect one of every 1,000 live births in the United States (Centers for Disease Control and Prevention, 1996), while the incidence of NTDs in other countries is 10-20 times greater (Lemire, 1998; Moore *et al.*, 1997; Verma, 1978).

10 Dietary folate deficiency can be rectified by adding chemically synthesized folic acid to staple foods (fortification), or by taking pills containing folic acid (supplementation), but these solutions have significant recurrent costs and are hard to implement in developing countries. Biofortification is a promising alternative strategy to enhance the folate content of food organisms by metabolic engineering (Scott *et al.*, 2000; Bouis 2002; DellaPenna 1999).
15 The validity of this approach has been demonstrated in food grade lactic acid bacteria, in which engineering tripled folate production (Sybesma *et al.*, 2003a).

Tomatoes have worldwide importance as a food crop but have relatively low folate content (Konings *et al.*, 2001; Basset *et al.*, 2002). As the plant folate synthesis pathway is now known (Scott *et al.*, 2000; Hanson *et al.*, 2002; Ravelle *et al.*, 2001; Basset *et al.*, 2002;
20 Basset *et al.*, 2004; Goyer *et al.*, 2004), folate engineering can be attempted with plants, and several groups have begun doing this (Hosain *et al.*, 2004; Rafalski *et al.*, 2003, U.S. Patent No. 6,642,435; Zhang *et al.*, 2003; Van der Meer *et al.*, 2001).

As can be understood from the above, there remains a need in the art for means to increase folate content in plants.

BRIEF SUMMARY OF THE INVENTION

The subject invention concerns materials and methods for enhancing folate content of plants. In one embodiment of a method of the invention, a plant is engineered to express increased levels of pteridines in the plant. In an exemplified embodiment, a plant is
30 transformed with a polynucleotide encoding a mammalian-type GCHI that is free of feedback control when expressed in a plant. In another exemplified embodiment, a plant is transformed with a polynucleotide encoding a bacterial GCHI that is not subject to metabolic regulation within a plant cell.

The subject invention also concerns plants having enhanced folate content. In one embodiment, a plant of the invention comprises and expresses a polynucleotide encoding a mammalian-type GCHI polypeptide or a bacterial GCHI polypeptide that is free of feedback control when expressed in the plant.

5

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show structure and biosynthesis of folates. **Figure 1A** shows a chemical structure of tetrahydrofolate, monoglutamyl form. Most plant folates have γ -linked polyglutamyl tails of up to ~6 residues attached to the first glutamate. One-carbon units at various levels of oxidation (from formate to methyl) are attached to N-5 and/or N-10. The pteridine ring of folates and free pteridines can exist in tetrahydro-, dihydro-, and fully oxidized forms. **Figure 1B** shows an outline of the plant folate synthesis pathway. Pteridines are in blue; note that one enzyme, DHN aldolase, interconverts DHN and DHM, and cleaves both to HMDHP (Goyer *et al.*, 2004). The red arrow is the engineered GCHI reaction. Abbreviations: ADC, aminodeoxychorismate; DHN, dihydroneopterin; DHM, dihydromonapterin; HMDHP, 6-hydroxymethyldihydropterin; -P, monophosphate; -PP, pyrophosphate, -PPP, triphosphate.

Figures 2A-2D show GCHI overexpression increases pteridine content in tomato fruit. **Figure 2A** shows the 5' Region of the synthetic GCHI cDNA ORF, with the recoded nucleotides italicized and bolded. **Figure 2B** is a photograph that shows the Western analysis of recombinant GCHI protein in fruit of four representative GCHI⁺ transformants and a vector-alone control (V2), at breaker (Br), red (R), and red-ripe (RR) stages. Tracks contained 80 μ g of protein. **Figure 2C** shows fluorometric HPLC analysis (Ultremex C₁₈ IP column) of oxidized pteridines in red-ripe fruit from a representative GCHI⁺ transformant (M102) and a vector control (V6). CPt, 6-carboxypterin; NPt, neopterin; MPt, monapterin, U1, unknown 1; U2, unknown 2; HMPT, 6-hydroxymethylpterin. **Figure 2D** shows changes in total pteridine content of control (V2) and GCHI⁺ (M9) fruit during ripening. Data are means of three replicates and SE. MG, mature green stage; other abbreviations as above.

Figures 3A and 3B show pteridine species accumulated by GCHI⁺ fruit. **Figure 3A** shows HPLC-fluorometric analysis of unknown peaks 1 and 2 before treatment (control) and after treatment with HCl, α -glucosidase, or β -glucosidase. Glucosidase treatments were for 5 min for peak 1, and 2 h for peak 2. **Figure 3B** shows quantitation of the major pteridines in

30

red-ripe fruit of three representative GCHI⁺ transformants. NPt-G, neopterin glycoside; MPt-G, monapterin glycoside; other abbreviations as in Figure 2C.

Figures 4A-4C show analysis of folates in GCHI⁺ and vector-control fruits by HPLC with EC detection. Figure 4A shows total folate contents of red-ripe fruit of 12 independent GCHI⁺ and 10 independent control transformants. Values for each transformant are averages of two fruit. Figure 4B shows the analysis of the one-carbon substituents of folates from GCHI⁺ and control fruits. Data are mean values for all 12 GCHI⁺ and all 10 control transformants. Note that the analytical procedures used converted 10-formyl-THF to 5,10-methenylTHF. Figure 4C shows the polyglutamyl tail length of 5-methyl-THF from control and GCHI⁺ fruit. Values are means for three GCHI⁺ fruit (M73, M86, M102) and three controls (V6, V13, V15).

Figures 5A-5C show the relationships between the levels of folate, pteridines, and PABA. Figure 5A shows the scatter plot of folate content vs. total pteridine content. Data are for two different fruit from each of the 10 control and 12 GCHI⁺ transgenics. Note that above a threshold of 25 nmol pteridines g⁻¹ FW (white area of frame), folate content in GCHI⁺ fruit is negatively correlated with pteridine level (dashed line; $r^2 = 0.43$, $P = 0.015$). Figure 5B shows the scatter plot of PABA content vs. folate content for the same fruit as in Figure 5A. Note that PABA content is negatively correlated with folate content (dashed line; $r^2 = 0.20$, $P = 0.003$). Figure 5C shows the effect of exogenously supplied PABA on folate levels in fruit of seven GCHI⁺ transgenics. Two matched, breaker-stage fruit from each transgenic were supplied via the stalk with 2 μ mol of PABA in 100 μ l of water, or water alone, for 1 d. Stalks were then removed and the fruit were left to ripen for another 6 d. The half of the fruit nearest the stalk was taken for analysis, since a pilot experiment with [¹⁴C]PABA showed that 96% of the PABA taken up by the fruit was located in this half. The mean folate contents of the water-treated and PABA-treated fruit tissue (1.20 and 5.65 nmol g⁻¹ FW, respectively) are significantly different at $P < 0.0001$.

Figures 6A-6E show expression and characterization of EcGCH. Figure 6A shows analysis of purified EcGCH fusion protein by SDS-PAGE. The molecular mass (MW) of EcGCH determined by SDS-PAGE (Yim *et al.*, 1976) and EcGCH fusion protein with N- and C-terminal extensions were 25,500 and 31,461 Da, respectively. The major protein band detected (*) had an estimated MW of 31,000 Da. Lane 1: supernatant from induced cells transformed with plasmid pTK101; lane 2: Ni-column purified protein fraction; lane 3: supernatant from cells with empty vector. Figures 6B and 6C show Western blot of EcGCH

expressed in *E. coli* detected using either mouse anti-His-tag monoclonal antibody (Scholl *et al.*, 2000) or anti-EcGCH (Scholl *et al.*, 2000). Lane 1: supernatant from induced cells transformed with plasmid pTK101; lane 2: affinity purified EcGCH fusion protein; lane 3: supernatant from cells with empty vector. 2.5 µg protein was loaded onto each lane. Western blot of transgenic and non-transgenic plants. **Figure 6D** shows extract from wild-type *Arabidopsis* (lane 1); extracts from T3 lines 5-21, 6-12, 2-19, respectively (lanes 2-4). 15 µg of protein was loaded in each lane. **Figure 6E** shows extracts from 6-10, 6-13 and non-transgenic control concentrated ~16-fold. Lane 1-2: ~210 µg protein 6-10 and 6-13, respectively; lane 3: ~15 µg 5-21; lane 4: ~210 µg non-transgenic control; lane 5: BENCHMARK prestained MW standards. EcGCH (26.2 kDa) marked with (*).

Figures 7A-7F are photographs that show an analysis and characterization of pterins from transgenic and non-transgenic *Arabidopsis*. **Figure 7A** shows pterins from undialyzed extract from T3 line 2-19 separated by RP chromatography and detected by fluorescence. Pterins were identified by comparison to elution times (**Figure 7B**) and spectral characteristics of Peak 1 (**Figure 7C**) and neopterin (neo) (**Figure 7D**), xanthopterin (xan) and isoxanthopterin (isoxan) (data not shown) standards measured by diode array. Individual HPLC peaks were collected and identification of neopterin was confirmed by MS analysis of Peak 1 (**Figure 7E**) and neopterin standard (**Figure 7F**) ($[M+H]^+$; $m/z=254$).

Figure 8A shows levels of pterins in T2 and T3 transgenic lines and non-transgenic *Arabidopsis*. **Figure 8B** shows values for transgenic T3 line 5-21 (value x 0.02) and non-transgenic (NT) plants (solid bars) were compared to values from the literature (Kohashi *et al.*, 1980) for other plant sources.

Figure 9A shows total folate levels of T3 transgenic lines and non-transgenic (NT) *Arabidopsis*. Total folate (nmol per gfw tissue +/- the standard error) for 3 replicates. Folate values for all transgenic plant lines were significantly different from the non-transgenic plants at the 95% confidence level. **Figure 9B** shows a comparison of folate levels for transgenic and non-transgenic *Arabidopsis* plants (solid bars) and values reported in the literature (48, USDA ARS Nutrient Data Laboratory (World Wide Website: <http://nal.usda.gov>) for other plant sources. **Figure 9C** shows plot of folate values as a function of total pterins in transgenic and non-transgenic lines.

Figure 10 shows a proposed pathway and localization of key enzymes for synthesis of pterins and folates in plants. Glucose ester of PABA; PABA-glc.

Figure 11 shows the analysis and characterization of pterins in leaves of transgenic *Arabidopsis thaliana* expressing the bacterial EcGCH gene and comparison with pterins present in transgenic tomato fruits expressing a synthetic GCH. *Arabidopsis* leaf tissue from transgenic plants expressing EcGCH was extracted and the extract was analyzed by HPLC on an Ultramex reverse phase column. Pterins were detected using a fluorescence detector. Specific pterins were identified using standards for neopterin (N), monapterin (M), 6-carboxypterin (C), hydroxymethylpterin (H), neopterin glycoside (NG) and monapterin glycoside (MG). Additional peaks (1, 2, 3 and 4) were detected but not identified. For comparison purposes, results are shown for transgenic *Arabidopsis* leaves (red line) and ripe fruits of transgenic tomato (black line) expressing GCH. Although 6-carboxypterin, neopterin, monapterin and hydroxymethylpterin were present in extracts of transgenic *Arabidopsis* and tomato fruits, the relative concentration of these pterins were different depending on the plant species. Glycosides of neopterin and monapterin were detected in tomato fruits but not *Arabidopsis* leaves while addition unknown peaks were present in leaves of transgenic *Arabidopsis*.

Figure 12 shows the analysis of pterins in leaves of transgenic *Arabidopsis thaliana* expressing the bacterial EcGCH gene during plant development. *Arabidopsis* leaf tissue from transgenic plants expressing EcGCH was harvested at four different times during plant development (5.5, 8, 9 and 11 weeks after planting) and extracted. The extract was analyzed by HPLC on an Ultrasphere ODS reverse phase column. The total pterins were measured using a fluorescence detector and reported in relative fluorescence units (RFU). Pterins apparently accumulated in leaf tissue during development.

Figure 13 shows the analysis of pterins profiles from seeds of transgenic *Arabidopsis thaliana* expressing the bacterial GCH gene and non-transgenic control plants. *Arabidopsis* seeds from transgenic plants (Line T3-219) expressing EcGCH and non-transgenic wild-type (WT) plants were collected at maturity and extracted. The extract was analyzed by HPLC on an Ultrasphere ODS reverse phase column. Extract from transgenic plants was diluted 20-fold in extraction buffer before HPLC analysis whereas extract from wild-type plants was not diluted. Pterins were detected using a fluorescence detector and reported in relative fluorescence units (RFU). Pterin levels in seeds of transgenic plants were much higher than in leaves of wild-type plants. The peak profiles were also different with additional peaks detected in extracts of seeds of transgenic plants.

Figure 14 shows the analysis of total pterins from seeds of transgenic *Arabidopsis thaliana* expressing the bacterial GCH gene and non-transgenic wild-type control plants. *Arabidopsis* seeds from transgenic plants expressing EcGCH (Lines 219, 515 and 613) and non-transgenic wild-type (WT) plants were collected at maturity and extracted. The extract
5 was analyzed by HPLC on an Ultrasphere ODS reverse phase column. Total pterins were measured using a fluorescence detector and reported in relative fluorescence units (RFU). The total pterins were 2- to 3-times higher for Line 219 than for Lines 515 and 613.

Figure 15 shows the analysis of total pterins from roots, leaves and seeds of transgenic *Arabidopsis thaliana* expressing the bacterial GCH gene and non-transgenic wild-type control plants. *Arabidopsis* roots, leaves and seeds from transgenic plants expressing EcGCH (Line 219) and non-transgenic wild-type (WT) plants were collected, extracted and analyzed by HPLC on an Ultrasphere ODS reverse phase column. Pterins were measured using a fluorescence detector and reported in relative fluorescence units (RFU). Pterin levels of roots were substantially lower than in leaves while levels in seeds of transgenic plants 7- to
15 10-fold higher.

Figure 16 shows the analysis of specific pterins from roots, leaves and seeds of transgenic *Arabidopsis thaliana* expressing the bacterial GCH gene. *Arabidopsis* roots, leaves and seeds from transgenic plants expressing EcGCH (Lines 219) were collected, extracted and analyzed by HPLC on an Ultrasphere ODS reverse phase column. Pterins were
20 detected using a fluorescence detector and reported in relative fluorescence units (RFU). Individual peaks correspond to the following pterin standards: Peak 1, 6-carboxypterin; Peak 2, neopterin; Peak 3, monapterin; Peak 6, hydroxymethylpterin. Peaks 4, 5 and 7 were not identified. The relative distribution of specific pterins differed between roots, leaves and seeds of transgenic plants. Several pterin peaks were not detected in roots. Neopterin,
25 monapterin and 6-carboxypterin were the major pterins detected in all tissues.

Figure 17 shows total folate levels in seeds of transgenic and non-transgenic *Arabidopsis* seeds. Total folates in seeds of transgenic *Arabidopsis thaliana* expressing the bacterial GCH gene Lines (219, 515, 613) and non-transgenic (NT) wild-type control plants were measured using the microbiological assay and expressed in μg per g. Folate levels were
30 increased 2- to 3-fold in transgenic lines when compared to non-transgenic controls.

Figure 18 shows total PABA levels in leaves of transgenic and non-transgenic *Arabidopsis*. Leaves of transgenic lines 219 (blue trace) and 514 (red trace) and non-transgenic Control plants (black trace) were extracted and analyzed for total PABA by HPLC

using fluorescence detection. The peak corresponding to PABA is marked. The total PABA levels were depleted more than 20-fold in the transgenic lines compared to the control plants. These results suggest that the supply of PABA may limit folate biosynthesis in the transgenic plants expressing EcGCH.

5 **Figure 19** shows total folate levels in seeds and leaves of transgenic *Arabidopsis* plants expressing EcGCH when grown with nutrient solution with and without supplemental (exogenous) PABA. Total folates were measured using the microbiological assay and expressed in μg per g. Folate levels, which are elevated in leaves and seeds of transgenic plants relative to control plants, were further elevated when plants were treated with a
10 nutrient solution containing PABA. Total folates were increased 2-fold and 7-fold in seeds and leaves of plants supplemented with PABA, respectively. These results support the proposal that increasing PABA supply could further enhance folate levels in transgenic plants expressing EcGCH.

15 **Figures 20A and 20B** show total PABA levels in leaves of transgenic and non-transgenic *Arabidopsis* treated with different concentrations of PABA. Leaves of transgenic line 219 (right side of Figure 20A) and non-transgenic control plants (left side of Figure 20A) were extracted and analyzed for total PABA by HPLC using fluorescence detection. PABA levels ($\pm$ SE) were substantially lower in the transgenic plants compared to control plants prior to treatment with PABA (Figure 20B). Although total PABA levels increased in both
20 the transgenic and control plants upon treatment with increasing concentrations of PABA, the levels of total PABA were still lower in transgenic plants with increased capacity for pterin synthesis. These results are consistent with the supply of PABA limiting folate biosynthesis in the transgenic plants expressing EcGCH.

25 **Figure 21** shows the analysis of total pterin levels in leaves of transgenic *Arabidopsis thaliana* expressing the bacterial GCH and treated with different concentrations of PABA. *Arabidopsis* leaves from transgenic plants expressing EcGCH and treated with different levels of PABA (0.5 mM, 1.0 mM and 2.0 mM PABA) and without added PABA (WO) were collected and extracted. Pterins were analyzed by HPLC on an Ultrasphere ODS reverse phase column using a fluorescence detector and reported in relative fluorescence units (RFU).
30 The levels of total pterins changed little upon treatment with PABA.

Figure 22 shows total folate levels in leaves of transgenic *Arabidopsis* plants expressing EcGCH when grown with nutrient solution with and without supplemental PABA. Total folates in *Arabidopsis* leaves from transgenic plants expressing EcGCH and treated

with different levels of PABA (0 mM, 0.5 mM, 1.0 mM and 2.0 mM PABA) were measured using the microbiological assay and expressed in μg per g. Folate levels increased upon treatment with increasing concentrations of PABA. Results are presented after the third treatment with PABA. At the highest concentration of PABA, total folate levels were enhanced by 50% over plants not treated with PABA. These results support the conclusion that increasing PABA supply will further enhance folate levels in transgenic plants expressing EcGCH and that engineering plants expressing higher levels of pterins to produce more PABA would further enhance folate levels.

Figure 23 shows total folate levels in leaves of transgenic *Arabidopsis* plants expressing EcGCH when grown with nutrient solution with and without supplemental PABA. Total folates in *Arabidopsis* leaves from transgenic plants expressing EcGCH and treated with different levels of PABA (0 mM, 0.5 mM, 1.0 mM and 2.0 mM PABA) were measured using the microbiological assay and expressed in μg per g. Folate levels increased upon treatment with increasing concentrations of PABA. Results are presented after the sixth treatment with PABA. At the highest concentration of PABA, total folate levels were enhanced 2-fold over plants not treated with PABA. These results support the conclusion that (1) increasing PABA supply leads to substantial increases in folate levels in transgenic plants expressing EcGCH and (2) engineering plants expressing EcGCH to synthesize more PABA leads to further enhancement of folate levels. Total folates were also increased in non-transgenic plants treated with PABA (data not shown) indicating that PABA supply limits folate biosynthesis and providing PABA to plants can increase total folates.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO: 1 is the nucleotide sequence of a synthetic mammalian GCHI cDNA (with recoded nucleotides to provide codons of more common usage in plants) that can be used according to the present invention.

SEQ ID NO: 2 is an amino acid sequence of a mammalian GCHI polypeptide encoded by SEQ ID NO: 1 of the present invention.

SEQ ID NO: 3 is the nucleotide sequence of a mammalian GCHI cDNA that can be used according to the present invention.

SEQ ID NO: 4 is an amino acid sequence of a mammalian GCHI polypeptide encoded by SEQ ID NO: 3 of the present invention.

SEQ ID NO: 5 is a polylinker within an expression vector of the subject invention.

SEQ ID NO: 6 is a forward PCR primer.

SEQ ID NO: 7 is a reverse PCR primer.

SEQ ID NO: 8 is the nucleotide sequence of a GCHI polypeptide from *E. coli* that can be used according to the present invention.

5 SEQ ID NO: 9 is an amino acid sequence of an *E. coli* GCHI polypeptide encoded by SEQ ID NO: 8 of the present invention.

SEQ ID NO: 10 is a sense primer.

SEQ ID NO: 11 is an antisense primer.

10 DETAILED DISCLOSURE OF THE INVENTION

The subject invention concerns materials and methods for increasing folate content of plants by incorporating one or more polynucleotide of the present invention in the genome of plant cells and expressing the polypeptide encoded by the polynucleotide. In one embodiment of a method of the invention, a plant is engineered to increase levels of pteridines in the plant. The subject invention contemplates the use of a polynucleotide that encodes any enzyme that catalyzes the synthesis of dihydroneopterin-PPP from GTP and that is not under feedback control of a plant. One embodiment of the invention concerns the use of a synthetic gene derived from a mammalian GCHI gene. The use of a synthetic gene in plants is unlikely to be subject to retroinhibition *in planta* because 1) retroinhibition of mammalian GCHI is mediated by a feedback regulatory protein (Yoneyama *et al.*, 1998), which plants appear not to have (Basset *et al.*, 2002), and 2) the inhibitory end-product of mammalian GCHI is tetrahydrobiopterin (Yoneyama *et al.*, 1998), which plants apparently lack (Kohashi, 1980). Another embodiment concerns the use of a polynucleotide encoding a bacterial GCHI polypeptide that is not subject to metabolic regulation. A further embodiment concerns the use of one or more polynucleotides encoding one or more enzymes that catalyze synthesis of PABA

In one embodiment, a plant is transformed with a polynucleotide encoding a mammalian-type GCHI polypeptide that is not subject to feedback control in a plant system. In an exemplified embodiment, the polynucleotide encodes a synthetic mammalian-type GCHI polypeptide having the amino acid sequence shown in SEQ ID NO: 2, or an enzymatically functional fragment or variant thereof. In a specific embodiment, the polynucleotide encoding the amino acid sequence shown in SEQ ID NO: 2 comprises the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3, or a sequence encoding an

enzymatically functional fragment or variant of SEQ ID NO: 2 or SEQ ID NO: 4.

In another embodiment, a plant is transformed with a polynucleotide encoding a bacterial GCHI polypeptide that is not subject to feedback control in a plant system. In an exemplified embodiment, the polynucleotide encodes a bacterial GCHI polypeptide having the amino acid sequence shown in SEQ ID NO: 9, or an enzymatically functional fragment or variant thereof. In a specific embodiment, the polynucleotide encoding the amino acid sequence shown in SEQ ID NO: 9 comprises the nucleotide sequence shown in SEQ ID NO: 8, or a sequence encoding an enzymatically functional fragment or variant of SEQ ID NO: 9.

Preferably, a polynucleotide(s) of the invention is expressed in the fruit, seed, leaf, root, or other edible part of a plant. In one embodiment, the plant is tomato. Transformants that overexpress a GCHI polypeptide can be selected using standard methods known in the art. The level of expression of a polynucleotide of the invention can be manipulated using standard methods known in the art, including the use of promoters that provide for low, intermediate or high levels of expression.

In a further embodiment, a plant transformed with a polynucleotide of the invention is a plant that is genetically engineered to express increased levels of PABA relative to PABA levels expressed in a wild-type plant. In one embodiment, the plant is transformed with a polynucleotide encoding one or more enzymes involved in PABA synthesis and that is not under feedback or other regulation by the plant. In a specific embodiment, the enzyme involved in PABA synthesis is a 4-amino-4-deoxychorismate synthase (ADCS) (Viswanathan *et al.* 1995; Basset *et al.* 2004; Genbank Accession No. NC 000913). In another specific embodiment, the enzyme is a 4-amino-4-deoxychorismate lyase (ADCL) (Green *et al.* 1992; Genbank Accession No. M93135). It is contemplated that a plant can be transformed with nucleic acid encoding one or both of ADCS and ADCL. The ADCS and/or ADCL can be from any organism, including a microorganism, such as *E. coli*.

The subject invention also concerns polynucleotide expression constructs comprising a polynucleotide sequence of the present invention encoding a mammalian-type GCHI polypeptide. In one embodiment, an expression construct of the invention comprises a polynucleotide sequence encoding a GCHI polypeptide comprising an amino acid sequence shown in SEQ ID NO: 2 or SEQ ID NO: 4, or an enzymatically functional fragment or variant thereof. In a specific embodiment, the polynucleotide sequence comprises a polynucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3, or a sequence encoding an enzymatically functional fragment or variant of SEQ ID NO: 2 or SEQ ID NO: 4.

The subject invention also concerns polynucleotide expression constructs comprising a polynucleotide sequence of the present invention encoding a bacterial GCHI polypeptide. In one embodiment, an expression construct of the invention comprises a polynucleotide sequence encoding a GCHI polypeptide comprising an amino acid sequence shown in SEQ ID NO: 9, or an enzymatically functional fragment or variant thereof. In a specific embodiment, the polynucleotide sequence comprises a polynucleotide sequence shown in SEQ ID NO: 8, or a sequence encoding an enzymatically functional fragment or variant of SEQ ID NO: 9.

Expression constructs of the invention generally include regulatory elements that are functional in the intended host cell in which the expression construct is to be expressed. Thus, a person of ordinary skill in the art can select regulatory elements for use in bacterial host cells, yeast host cells, plant host cells, insect host cells, mammalian host cells, and human host cells. Regulatory elements include promoters, transcription termination sequences, translation termination sequences, enhancers, and polyadenylation elements. As used herein, the term "expression construct" refers to a combination of nucleic acid sequences that provides for transcription of an operably linked nucleic acid sequence. As used herein, the term "operably linked" refers to a juxtaposition of the components described wherein the components are in a relationship that permits them to function in their intended manner. In general, operably linked components are in contiguous relation.

An expression construct of the invention can comprise a promoter sequence operably linked to a polynucleotide sequence encoding a mammalian-type or bacterial GCHI polypeptide of the invention. Promoters can be incorporated into a polynucleotide using standard techniques known in the art. Multiple copies of promoters or multiple promoters can be used in an expression construct of the invention. In a preferred embodiment, a promoter can be positioned about the same distance from the transcription start site in the expression construct as it is from the transcription start site in its natural genetic environment. Some variation in this distance is permitted without substantial decrease in promoter activity. A transcription start site is typically included in the expression construct.

If the expression construct is to be provided in or introduced into a plant cell, then plant viral promoters, such as, for example, a cauliflower mosaic virus (CaMV) 35S (including the enhanced CaMV 35S promoter (see, for example U.S. Patent No. 5,106,739)) or a CaMV 19S promoter or a cassava vein mosaic can be used. Other promoters that can be used for expression constructs in plants include, for example, prolifera promoter, Ap3

promoter, heat shock promoters, T-DNA 1'- or 2'-promoter of *A. tumefaciens*, polygalacturonase promoter, chalcone synthase A (CHS-A) promoter from petunia, tobacco PR-1a promoter, ubiquitin promoter, actin promoter, alcA gene promoter, pin2 promoter (Xu *et al.*, 1993), maize WipI promoter, maize trpA gene promoter (U.S. Patent No. 5,625,136),
5 maize CDPK gene promoter, and RUBISCO SSU promoter (U.S. Patent No. 5,034,322) can also be used. Tissue-specific promoters, for example fruit-specific promoters, such as the E8 promoter of tomato (accession number: AF515784; Good *et al.* (1994)) can be used. Fruit-specific promoters such as flower organ-specific promoters can be used with an expression construct of the present invention for expressing a polynucleotide of the invention in the
10 flower organ of a plant. Examples of flower organ-specific promoters include any of the promoter sequences described in U.S. Patent Nos. 6,462,185; 5,639,948; and 5,589,610. Seed-specific promoters such as the promoter from a β -phaseolin gene (for example, of kidney bean) or a glycinin gene (for example, of soybean), and others, can also be used. Root-specific promoters, such as any of the promoter sequences described in U.S. Patent No.
15 6,455,760 or U.S. Patent No. 6,696,623, or in published U.S. patent application Nos. 20040078841; 20040067506; 20040019934; 20030177536; 20030084486; or 20040123349, can be used with an expression construct of the invention. Constitutive promoters (such as the CaMV, ubiquitin, actin, or NOS promoter), developmentally-regulated promoters, and inducible promoters (such as those promoters than can be induced by heat, light, hormones,
20 or chemicals) are also contemplated for use with polynucleotide expression constructs of the invention.

For expression in prokaryotic systems, an expression construct of the invention can comprise promoters such as, for example, alkaline phosphatase promoter, tryptophan (trp) promoter, lambda P_L promoter, β -lactamase promoter, lactose promoter, phoA promoter, T3
25 promoter, T7 promoter, or tac promoter (de Boer *et al.*, 1983). Promoters suitable for use with an expression construct of the invention in yeast cells include, but are not limited to, 3-phosphoglycerate kinase promoter, glyceraldehyde-3-phosphate dehydrogenase promoter, metallothionein promoter, alcohol dehydrogenase-2 promoter, and hexokinase promoter.

Expression constructs of the invention may optionally contain a transcription
30 termination sequence, a translation termination sequence, a sequence encoding a signal peptide, and/or enhancer elements. Transcription termination regions can typically be obtained from the 3' untranslated region of a eukaryotic or viral gene sequence. Transcription termination sequences can be positioned downstream of a coding sequence to provide for

efficient termination. A signal peptide sequence is a short amino acid sequence typically present at the amino terminus of a protein that is responsible for the relocation of an operably linked mature polypeptide to a wide range of post-translational cellular destinations, ranging from a specific organelle compartment to sites of protein action and the extracellular environment. Targeting gene products to an intended cellular and/or extracellular destination through the use of an operably linked signal peptide sequence is contemplated for use with the polypeptides of the invention. Classical enhancers are cis-acting elements that increase gene transcription and can also be included in the expression construct. Classical enhancer elements are known in the art, and include, but are not limited to, the CaMV 35S enhancer element, cytomegalovirus (CMV) early promoter enhancer element, and the SV40 enhancer element. Intron-mediated enhancer elements that enhance gene expression are also known in the art. These elements must be present within the transcribed region and are orientation dependent. Examples include the maize shrunken-1 enhancer element (Clancy and Hannah, 2002).

DNA sequences which direct polyadenylation of mRNA transcribed from the expression construct can also be included in the expression construct, and include, but are not limited to, an octopine synthase or nopaline synthase signal. The expression constructs of the invention can also include a polynucleotide sequence that directs transposition of other genes, *i.e.*, a transposon.

Expression constructs can also include one or more dominant selectable marker genes, including, for example, genes encoding antibiotic resistance and/or herbicide-resistance for selecting transformed cells. Antibiotic-resistance genes can provide for resistance to one or more of the following antibiotics: hygromycin, kanamycin, bleomycin, G418, streptomycin, paromomycin, neomycin, and spectinomycin. Kanamycin resistance can be provided by neomycin phosphotransferase (NPT II). Herbicide-resistance genes can provide for resistance to phosphinothricin acetyltransferase or glyphosate. Other markers used for cell transformation screening include genes encoding β -glucuronidase (GUS), β -galactosidase, luciferase, nopaline synthase, chloramphenicol acetyltransferase (CAT), green fluorescence protein (GFP), or enhanced GFP (Yang *et al.*, 1996).

The subject invention also concerns polynucleotide vectors comprising a polynucleotide sequence of the invention that encodes a mammalian-type or bacterial GCHI polypeptide of the invention. Unique restriction enzyme sites can be included at the 5' and 3' ends of an expression construct or polynucleotide of the invention to allow for insertion into a

polynucleotide vector. As used herein, the term "vector" refers to any genetic element, including for example, plasmids, cosmids, chromosomes, phage, virus, and the like, which is capable of replication when associated with proper control elements and which can transfer polynucleotide sequences between cells. Vectors contain a nucleotide sequence that permits the vector to replicate in a selected host cell. A number of vectors are available for expression and/or cloning, and include, but are not limited to, pBR322, pUC series, M13 series, and pBLUESCRIPT vectors (Stratagene, La Jolla, CA).

The subject invention also concerns oligonucleotide probes and primers, such as polymerase chain reaction (PCR) primers, that can hybridize to a coding or non-coding sequence of a polynucleotide of the present invention. Oligonucleotide probes of the invention can be used in methods for detecting nucleic acid sequences related to polynucleotides of the present invention. Oligonucleotide primers of the invention can be used in PCR methods and other methods involving nucleic acid amplification. In a preferred embodiment, a probe or primer of the invention can hybridize to a polynucleotide of the invention under stringent conditions. Probes and primers of the invention can optionally comprise a detectable label or reporter molecule, such as fluorescent molecules, enzymes, radioactive moiety, and the like. Probes and primers of the invention can be of any suitable length for the method or assay in which they are being employed. Typically, probes and primers of the invention will be 10 to 500 or more nucleotides in length. Probes and primers that are 10 to 20, 21 to 30, 31 to 40, 41 to 50, 51 to 60, 61 to 70, 71 to 80, 81 to 90, 91 to 100 or more nucleotides in length are contemplated within the scope of the invention. Probes and primers of the invention can have complete (100%) nucleotide sequence identity with the polynucleotide sequence, or the sequence identity can be less than 100%. For example, sequence identity between a probe or primer and a sequence can be 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75%, 70% or any other percentage sequence identity so long as the probe or primer can hybridize under stringent conditions to a nucleotide sequence of a polynucleotide of the invention. Exemplified probes and primers of the invention include those having the nucleotide sequence of SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 10, and SEQ ID NO: 11, or a functional fragment or variant of any of the SEQ ID NOs: 6, 7, 10, and 11.

Polynucleotides of the present invention can be composed of either RNA or DNA. Preferably, the polynucleotides are composed of DNA. The subject invention also encompasses those polynucleotides that are complementary in sequence to the

polynucleotides disclosed herein. Polynucleotides and polypeptides of the invention can be provided in purified or isolated form.

Because of the degeneracy of the genetic code, a variety of different polynucleotide sequences can encode mammalian-type or bacterial GCHI enzymes of the present invention.

5 A table showing all possible triplet codons (and where U also stands for T) and the amino acid encoded by each codon is described in Lewin (1985). In addition, it is well within the skill of a person trained in the art to create alternative polynucleotide sequences encoding the same, or essentially the same, GCHI enzymes of the subject invention. These variant or alternative polynucleotide sequences are within the scope of the subject invention. As used
10 herein, references to "essentially the same" sequence refers to sequences which encode amino acid substitutions, deletions, additions, or insertions which do not materially alter the functional activity of the polypeptide encoded by the polynucleotides of the present invention. Allelic variants of the nucleotide sequences encoding a mammalian-type GCHI of the invention are also encompassed within the scope of the invention.

15 Polypeptide fragments according to the subject invention typically comprise a contiguous span of about or at least 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110,
20 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, or 151 amino acids of SEQ ID NO: 2, or a contiguous span of about or at least 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72,
25 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170,
30 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, or 221 amino acids of SEQ ID NO: 9.

Polypeptide fragments of the subject invention can be any integer in length from at least about 25 consecutive amino acids to 1 amino acid less than the sequence shown in SEQ ID NO: 2 or SEQ ID NO: 9. Thus, for SEQ ID NO: 2, a polypeptide fragment can be any integer of consecutive amino acids from about 25 to 151 amino acids. For SEQ ID NO: 9, a polypeptide fragment can be any integer of consecutive amino acids from about 25 to 221 amino acids. The term "integer" is used herein in its mathematical sense and thus representative integers include: 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, and/or 221.

Each polypeptide fragment of the subject invention can also be described in terms of its N-terminal and C-terminal positions. For example, combinations of N-terminal to C-terminal fragments of about 25 contiguous amino acids to 1 amino acid less than the full length polypeptide of SEQ ID NO: 2, SEQ ID NO: 4, or SEQ ID NO: 9 are included in the present invention. Thus, using SEQ ID NO: 2 as an example, a 25 consecutive amino acid fragment could correspond to amino acids of SEQ ID NO: 2 selected from the group consisting of 1-25, 2-26, 3-27, 4-28, 5-29, 6-30, 7-31, 8-32, 9-33, 10-34, 11-35, 12-36, 13-37, 14-38, 15-39, 16-40, 17-41, 18-42, 19-43, 20-44, 21-45, 22-46, 23-47, 24-48, 25-49, 26-50, 27-51, 28-52, 29-53, 30-54, 31-55, 32-56, 33-57, 34-58, 35-59, 36-60, 37-61, 38-62, 39-63, 40-64, 41-65, 42-66, 43-67, 44-68, 45-69, 46-70, 47-71, 48-72, 49-73, 50-74, 51-75, 52-76, 53-77, 54-78, 55-79, 56-80, 57-81, 58-82, 59-83, 60-84, 61-85, 62-86, 63-87, 64-88, 65-89, 66-90, 67-91, 68-92, 69-93, 70-94, 71-95, 72-96, 73-97, 74-98, 75-99, 76-100, 77-101, 78-102, 79-103, 80-104, 81-105, 82-106, 83-107, 84-108, 85-109, 86-110, 87-111, 88-112, 89-113, 90-114, 91-115, 92-116, 93-117, 94-118, 95-119, 96-120, 97-121, 98-122, 99-123, 100-124, 101-125, 102-126, 103-127, 104-128, 105-129, 106-130, 107-131, 108-132, 109-133, 110-134, 111-135, 112-136, 113-137, 114-138, 115-139, 116-140, 117-141, 118-142, 119-

143, 120-144, 121-145, 122-146, 123-147, 124-148, 125-149, 126-150, 127-151, and 128-152. Similarly, the amino acids corresponding to all other fragments of sizes between 26 consecutive amino acids and 151 consecutive amino acids of SEQ ID NO: 2 are included in the present invention and can also be immediately envisaged based on these examples.

5 Therefore, additional examples, illustrating various fragments of the polypeptides of SEQ ID NO: 2, SEQ ID NO: 4, or SEQ ID NO: 9 are not individually listed herein in order to avoid unnecessarily lengthening the specification.

Polypeptide fragments comprising:

a) 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45,
 10 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70,
 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95,
 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114,
 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132,
 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, or
 15 151 consecutive amino acids of SEQ ID NO: 2 may alternatively be described by the formula
 “n to c” (inclusive), where “n” equals the N-terminal amino acid position and “c” equals the
 C-terminal amino acid position of the polypeptide. In this embodiment of the invention, “n”
 is an integer having a lower limit of 1 and an upper limit of the total number of amino acids
 of the full length polypeptide minus 24 (e.g., $152-24=128$ for SEQ ID NO: 2). “c” is an
 20 integer between 25 and the number of amino acids of the full length polypeptide sequence
 (152 for SEQ ID NO: 2) and “n” is an integer smaller than “c” by at least 24. Therefore, for
 SEQ ID NO: 2, “n” is any integer selected from the list consisting of: 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, 26, 27, 28, 29, 30, 31, 32, 33, 34,
 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59,
 25 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84,
 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106,
 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124,
 125, 126, 127, 128; and “c” is any integer selected from the group consisting of: 25, 26, 27,
 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52,
 30 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77,
 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101,
 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119,
 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137,

138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, and 152 provided that “n” is a value less than “c” by at least 24. Every combination of “n” and “c” positions are included as specific embodiments of polypeptide fragments of the invention. Examples illustrating the various fragments of a polypeptide contemplated under this formula are not individually listed in order to avoid unnecessarily lengthening the specification. However, all
5 embodiments of a particular polypeptide can be immediately envisaged from the description provided herein. All ranges used to describe any polypeptide fragment embodiment of the present invention are inclusive unless specifically set forth otherwise. Polynucleotides encoding any of the polypeptide fragments of the invention are also contemplated within the
10 scope of the invention.

Substitution of amino acids other than those specifically exemplified or naturally present in a mammalian-type or bacterial GCHI enzyme of the invention are also contemplated within the scope of the present invention. For example, non-natural amino acids can be substituted for the amino acids of a GCHI enzyme, so long as the GCHI enzyme
15 having the substituted amino acids retains substantially the same functional activity as the GCHI in which amino acids have not been substituted. Examples of non-natural amino acids include, but are not limited to, ornithine, citrulline, hydroxyproline, homoserine, phenylglycine, taurine, iodotyrosine, 2,4-diaminobutyric acid, α -amino isobutyric acid, 4-aminobutyric acid, 2-amino butyric acid, γ -amino butyric acid, ϵ -amino hexanoic acid, 6-
20 amino hexanoic acid, 2-amino isobutyric acid, 3-amino propionic acid, norleucine, norvaline, sarcosine, homocitrulline, cysteic acid, τ -butylglycine, τ -butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoro-amino acids, designer amino acids such as β -methyl amino acids, C-methyl amino acids, N-methyl amino acids, and amino acid analogues in general. Non-natural amino acids also include amino acids having derivatized side groups.
25 Furthermore, any of the amino acids in the protein can be of the D (dextrorotary) form or L (levorotary) form. Allelic variants of a protein sequence of mammalian-type GCHI enzyme of the present invention are also encompassed within the scope of the invention.

Amino acids can be generally categorized in the following classes: non-polar, uncharged polar, basic, and acidic. Conservative substitutions whereby a mammalian-type or
30 bacterial GCHI enzyme of the present invention having an amino acid of one class is replaced with another amino acid of the same class fall within the scope of the subject invention so long as the mammalian-type or bacterial GCHI enzyme having the substitution still retains substantially the same functional activity as the mammalian-type or bacterial GCHI enzyme

that does not have the substitution. Polynucleotides encoding a mammalian-type or bacterial GCHI enzyme having one or more amino acid substitutions in the sequence are contemplated within the scope of the present invention. Table 1 below provides a listing of examples of amino acids belonging to each class.

5

Table 1.	
Class of Amino Acid	Examples of Amino Acids
Nonpolar	Ala, Val, Leu, Ile, Pro, Met, Phe, Trp
Uncharged Polar	Gly, Ser, Thr, Cys, Tyr, Asn, Gln
Acidic	Asp, Glu
Basic	Lys, Arg, His

The subject invention also concerns variants of the polynucleotides of the present invention that encode enzymatically active mammalian-type and bacterial GCHI enzymes of the invention. Variant sequences include those sequences wherein one or more nucleotides of the sequence have been substituted, deleted, and/or inserted. The nucleotides that can be substituted for natural nucleotides of DNA have a base moiety that can include, but is not limited to, inosine, 5-fluorouracil, 5-bromouracil, hypoxanthine, 1-methylguanine, 5-methylcytosine, and tritylated bases. The sugar moiety of the nucleotide in a sequence can also be modified and includes, but is not limited to, arabinose, xylulose, and hexose. In addition, the adenine, cytosine, guanine, thymine, and uracil bases of the nucleotides can be modified with acetyl, methyl, and/or thio groups. Sequences containing nucleotide substitutions, deletions, and/or insertions can be prepared and tested using standard techniques known in the art.

Fragments and variants of mammalian-type and bacterial GCHI of the present invention can be generated as described herein and tested for the presence of enzymatic function using standard techniques known in the art. For example, for testing fragments and/or variants of a mammalian-type or bacterial GCHI, the conversion of guanosine triphosphate (GTP) to dihydroneopterin triphosphate (DHN-PPP) can be assayed according to the present invention. Thus, an ordinarily skilled artisan can readily prepare and test

20

fragments and variants of a GCHI enzyme of the invention and determine whether the fragment or variant retains functional enzymatic activity relative to full-length or a non-variant GCHI enzyme.

Polynucleotides and polypeptides contemplated within the scope of the subject invention can also be defined in terms of more particular identity and/or similarity ranges with those sequences of the invention specifically exemplified herein. The sequence identity will typically be greater than 60%, preferably greater than 75%, more preferably greater than 80%, even more preferably greater than 90%, and can be greater than 95%. The identity and/or similarity of a sequence can be 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% as compared to a sequence exemplified herein. Unless otherwise specified, as used herein percent sequence identity and/or similarity of two sequences can be determined using the algorithm of Karlin and Altschul (1990), modified as in Karlin and Altschul (1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul *et al.* (1990). BLAST searches can be performed with the NBLAST program, score = 100, wordlength = 12, to obtain sequences with the desired percent sequence identity. To obtain gapped alignments for comparison purposes, Gapped BLAST can be used as described in Altschul *et al.* (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (NBLAST and XBLAST) can be used. See NCBI/NIH website.

The subject invention also contemplates those polynucleotide molecules having sequences which are sufficiently homologous with the polynucleotide sequences exemplified herein so as to permit hybridization with that sequence under standard stringent conditions and standard methods (Maniatis *et al.*, 1982). As used herein, "stringent" conditions for hybridization refers to conditions wherein hybridization is typically carried out overnight at 20-25 C below the melting temperature (T_m) of the DNA hybrid in 6x SSPE, 5x Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. The melting temperature, T_m , is described by the following formula (Beltz *et al.*, 1983):

$$T_m = 81.5 C + 16.6 \log[Na^+] + 0.41(\%G+C) - 0.61(\% \text{ formamide}) - 600/\text{length of duplex}$$

in base pairs.

Washes are typically carried out as follows:

(1) Twice at room temperature for 15 minutes in 1x SSPE, 0.1% SDS (low stringency wash).

(2) Once at $T_m - 20^\circ\text{C}$ for 15 minutes in 0.2x SSPE, 0.1% SDS (moderate stringency wash).

As used herein, the terms "nucleic acid" and "polynucleotide" refer to a deoxyribonucleotide, ribonucleotide, or a mixed deoxyribonucleotide and ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, would encompass known analogs of natural nucleotides that can function in a similar manner as naturally-occurring nucleotides. The polynucleotide sequences include the DNA strand sequence that is transcribed into RNA and the strand sequence that is complementary to the DNA strand that is transcribed. The polynucleotide sequences also include both full-length sequences as well as shorter sequences derived from the full-length sequences. Allelic variations of the exemplified sequences also fall within the scope of the subject invention. The polynucleotide sequence includes both the sense and antisense strands either as individual strands or in the duplex.

The subject invention also concerns cells transformed with a polynucleotide of the present invention encoding a mammalian-type or bacterial GCHI polypeptide of the invention. In one embodiment, the cell is transformed with a polynucleotide sequence comprising a sequence encoding the amino acid sequence shown in SEQ ID NO: 2, or an enzymatically functional fragment or variant thereof. In a specific embodiment, the cell is transformed with a polynucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3, or a sequence encoding an enzymatically functional fragment or variant of SEQ ID NO: 2. In another embodiment, the cell is transformed with a polynucleotide sequence comprising a sequence encoding the amino acid sequence shown in SEQ ID NO: 9, or an enzymatically functional fragment or variant thereof. In a specific embodiment, the cell is transformed with a polynucleotide sequence shown in SEQ ID NO: 8, or a sequence encoding an enzymatically functional fragment or variant of SEQ ID NO: 9. The cell can also comprise a polynucleotide encoding one or more enzymes that catalyze PABA synthesis and that are not under feedback or other regulation by the cell. In one embodiment, the enzyme is an ADCS or an ADCL enzyme.

Preferably, the polynucleotide sequence is provided in an expression construct of the invention. The transformed cell can be a prokaryotic cell, for example, a bacterial cell such as *E. coli*, *Synechocystis* sp., or *B. subtilis*, or the transformed cell can be a eukaryotic cell, for example, a plant cell, including protoplasts, or an animal cell. Plant cells include, but are not limited to, dicotyledonous, monocotyledonous, algal, and fungal, and conifer cells. In one

embodiment, the plant cell is a cell from tomato. In another embodiment, the plant cell is a cell from *Arabidopsis*. Animal cells include human cells, mammalian cells, avian cells, and insect cells.

The subject invention also concerns plants, including transgenic plants, having enhanced folate content. Plants, plant tissues, and plant cells transformed with or bred to contain a polynucleotide of the invention are contemplated by the present invention. Plant tissue includes, but is not limited to, seed, scion, and rootstock. In one embodiment, a plant of the invention comprises and expresses a polynucleotide encoding a mammalian-type or bacterial GCHI polypeptide free of feedback control. In one embodiment, the plant, plant tissue, or plant cell comprises a polynucleotide sequence comprising a sequence encoding the amino acid sequence shown in SEQ ID NO: 2, SEQ ID NO: 4, or SEQ ID NO: 9. In a specific embodiment, the plant comprises a polynucleotide sequence comprising the nucleotide sequence shown in SEQ ID NO: 1, SEQ ID NO: 3, or SEQ ID NO: 8, or a sequence encoding an enzymatically functional fragment or variant of SEQ ID NO: 2, SEQ ID NO: 4, or SEQ ID NO: 9, respectively. The plant can also comprise a polynucleotide encoding one or more enzymes that catalyze PABA synthesis and that are not under feedback or other regulation by the plant. In one embodiment, the enzyme is an ADCS or an ADCL enzyme. In one embodiment, the plant is transgenic for a polynucleotide that encodes a mammalian type or a bacterial GCHI and transgenic for one or more polynucleotides that encode one or more enzymes that catalyze synthesis of PABA.

Plants within the scope of the present invention include monocotyledonous plants, such as, for example, rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, turfgrasses, and millet. Plants within the scope of the present invention also include dicotyledonous plants, such as, for example, tomato, cucumber, squash, peas, alfalfa, melon, chickpea, chicory, clover, kale, lentil, soybean, beans, tobacco, potato, sweet potato, yams, cassava, radish, broccoli, spinach, cabbage, rape, apple trees, citrus (including oranges, mandarins, grapefruit, lemons, limes and the like), grape, cotton, sunflower, strawberry, and lettuce. In one embodiment, the plant, plant tissue, or plant cell is tomato. In another embodiment, the plant, plant tissue, or plant cell is thale cress (*Arabidopsis*) Herb plants containing a polynucleotide of the invention are also contemplated within the scope of the invention. Herb plants include parsley, sage, rosemary, thyme, and the like. Also contemplated within the scope of the present invention are algae and fungi that comprise a polynucleotide of the present invention. Algae contemplated include those in the

Cyanophyta, Chlorophyta, Pyrrophyta, Chrysophyta, Rhodophyta, and Phaeophyta divisions. Fungi contemplated include those of Ascomycetes and Basidiomycetes.

Techniques for transforming plant cells with a gene are known in the art and include, for example, *Agrobacterium* infection, biolistic methods, electroporation, calcium chloride treatment, PEG-mediated transformation, *etc.* U. S. Patent No. 5,661,017 teaches methods and materials for transforming an algal cell with a heterologous polynucleotide. Transformed cells can be selected, redifferentiated, and grown into plants that contain and express a polynucleotide of the invention using standard methods known in the art. The seeds and other plant tissue and progeny of any transformed or transgenic plant cells or plants of the invention are also included within the scope of the present invention.

The subject invention also concerns methods for producing a plant that exhibits increased folate content relative to a wild type plant, wherein a polynucleotide encoding a mammalian-type GCHI enzyme or a bacterial GCHI enzyme of the present invention and/or one or more polynucleotide molecules encoding one or more enzymes that catalyze synthesis of PABA is introduced into a plant cell and the polypeptide(s) encoded by the polynucleotide(s) is expressed. In one embodiment, the polynucleotide or polynucleotides is incorporated into the genome of the plant cell and a plant is grown from the plant cell. In a preferred embodiment, the plant grown from the plant cell stably expresses the incorporated polynucleotide or polynucleotides. In one embodiment, the plant is grown in the presence of exogenously added PABA. In another embodiment, the plant is genetically engineered to produce increased levels of PABA relative to a wild type plant. In one embodiment, the plant cell transformed with a polynucleotide encoding a mammalian-type GCHI or a bacterial GCHI is transgenic for a polynucleotide encoding one or more enzymes that catalyze synthesis of PABA. In a preferred embodiment, a plant is transformed or engineered to express increased levels of both pteridines and PABA in the plant.

MATERIALS AND METHODS FOR EXAMPLES 1-6

Expression Vector Construction.

The pMON10086 vector (Klee *et al.*, U.S. Patent No. 5,512,466) containing the tomato E8 promoter (Deikman *et al.*, 1992), pea RuBisCo ssu terminator, and the *nptII* kanamycin resistance gene was modified by ablating the *NotI* site and ligating a polylinker (5'-GGATCCGCGGCCGCGAATTCAGGCCTGGTACCGGCGCGCCAGATCT-3') (SEQ ID NO: 5) into the unique *BamHI* site between the E8 promoter and the terminator. A

mammalian GCHI cDNA (GenBank accession no. BE136861) was modified by using PCR primers to change the sequence context of the start codon to the plant consensus TAAACA-ATG (Koziel *et al.*, 1996) and to replace 17 rare codons (Figure 2A). This synthetic cDNA was inserted between the *NotI* and *AscI* sites of the modified vector, and introduced into
5 *Agrobacterium tumefaciens* strain ABI by electroporation.

Transgenic Plants.

The sequence-verified GCHI construct and the modified vector were used to transform tomato (*Lycopersicon esculentum* Mill., cv. Micro-Tom) essentially as described
10 (Tieman *et al.*, 2001). Transformants were selected and regenerated on media containing 100 $\mu\text{g ml}^{-1}$ kanamycin. Kanamycin-resistant plantlets were screened for the GCHI construct by PCR using a forward primer located in the E8 promoter (5'-CTTTCTTGTTCCCATTTCTC-3') (SEQ ID NO: 6) and a reverse primer from the GCHI coding region (5'-ATGCACATGTGTGTCGCTTC-3') (SEQ ID NO: 7); vector-alone transformants were
15 verified using primers for the *nptII* gene. Positives were transplanted to soil and grown to maturity in a growth chamber (16-h day, 23°C, photosynthetic flux density 200 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$; 8-h night, 20°C), irrigating with nutrient solution. Fruit were harvested at the stages indicated in the text. For uniformity, red and red-ripe stages were respectively defined as 3 and 7 d after breaker stage. In PABA-feeding experiments, breaker-stage fruit were cut at the
20 base of the stalk, supplied via the stalk with 2 μmol of PABA in 100 μl of water, or water only, for 1 d, then destalked and allowed to ripen for a further 6 d at 23°C.

Antibodies and Western Analysis.

To produce recombinant antigen, the synthetic GCHI cDNA was cloned into the
25 EcoRI and XhoI sites of pET28b (Novagen, Madison, WI), which adds hexahistidine tags to both termini. The construct was electroporated into *Escherichia coli* BL21 (DE3) CodonPlus-RIL cells (Stratagene), and the recombinant protein was isolated from IPTG-induced cells by Ni^{2+} affinity chromatography under denaturing conditions followed by preparative SDS-polyacrylamide gel electrophoresis. Rabbit antibodies were prepared by
30 Cocalico Inc. (Reamstown, PA). To screen for GCHI expression, proteins were extracted from pericarp tissue by grinding in 0.1 M Tris-HCl, pH 8.0 containing 15 mM ascorbate, 2 mM DTT, and 3% (w/v) PVPP. After centrifuging to clear, samples (80 μg protein) were separated on SDS-polyacrylamide gels, blotted to nitrocellulose, and probed as described

(Nuccio *et al.*, 1999) with antiserum diluted 1:2000. No cross-reaction with tomato GCHI was detectable.

Pteridine and PABA Analysis.

Representative fruit segments (0.70 g) were pulverized in liquid N₂ and homogenized with 7 ml of methanol. For pteridine analysis, a 600- μ l portion of the homogenate was mixed with 250 μ l of CHCl₃ and 50 μ l of water, and shaken for 40 min before adding another 225 μ l of CHCl₃ and 340 μ l of water. After shaking for 20 min and centrifuging to break the emulsion, the aqueous phase was removed, dried *in vacuo*, and redissolved in 200 μ l of water. Samples were oxidized by adding 0.1 volume of a solution of 1% I₂ and 2% KI (w/v) in 1 M HCl, and incubating in darkness for 1 h; excess I₂ was then removed by adding 10 μ l of 5% (w/v) Na-ascorbate (Fukushima *et al.*, 1980). The oxidized samples were separated by HPLC, using an Ultremex C₁₈ RP column (5 μ m, 250 X 4.6 mm, Phenomenex, Belmont, CA) or a Synergi Fusion-RP 80 column (4 μ m, 250 X 4.6 mm, Phenomenex) eluted isocratically with 10 mM Na-phosphate (pH 6.0) at 1.5 ml min⁻¹. Peaks were detected by a Waters 2475 fluorescence detector (350 nm excitation, 450 nm emission) and identified by reference to standards and by spectral properties. To investigate pteridine oxidation state, 10 mM β -mercaptoethanol and 2% (w/v) Na-ascorbate were added to the methanol extraction medium, and the oxidation step was omitted. Pteridine conjugate peaks were recovered from the mobile phase by an ion exchange procedure (Stea *et al.*, 1980) and treated with HCl (1 N, 100°C, 1 h) or with yeast α -glucosidase or almond β -glucosidase (2 units in 40 μ l of 10 mM Na-phosphate, pH 6.0, 37°C, 5-120 min). The 6-carboxypterin, neopterin, monapterin, and 6-hydroxymethylpterin peaks were quantified relative to standard pteridines; conjugate peaks were quantified relative to the corresponding free pteridine since hydrolysis did not change fluorescence yield. Pteridine standards were from Schircks Laboratories (Jona, Switzerland). A 5-ml portion of the methanol extract above was used for analysis of total PABA (*i.e.*, free PABA + PABA glucose ester) by acid hydrolysis, cation exchange chromatography, ethyl acetate partitioning, and fluorometric HPLC as described (Quinlivan *et al.*, 2003).

Folate Analysis.

Folates were extracted from representative fruit segments (0.5-1.0 g) by Polytron homogenization in 10 ml of 50 mM HEPES/50 mM CHES, adjusted to pH 7.9 with HCl, containing 1 mM CaCl₂, 2% (w/v) Na-ascorbate, and 10 mM β -mercaptoethanol, followed by

boiling for 10 min, then centrifuging (13,000 × g, 10 min). The pellet was re-extracted the same way. The combined extracts were treated with 1 ml of dialyzed rat plasma at 37°C for 2 h to deglutamylate folates. Samples were then boiled for 15 min, centrifuged as above, filtered through glass wool and applied to folate affinity columns prepared as described (Gregory *et al.*, 1988). After washing the columns with 5 ml of 25 mM K-phosphate, pH 7.0, plus 1% Na-ascorbate (buffer 1) containing 1 M NaCl, then with 5 ml of buffer 1 alone, they were eluted with 5 ml of HPLC mobile phase A (see below) containing 1% ascorbic acid. Samples of the eluate (400 µl) were taken for HPLC analysis with electrochemical detection (Bagley *et al.*, 2000) using a Prodigy 5 µm ODS2 column (150 × 3.2 mm, Phenomenex) and a four-channel detector (CoulArray Model 5600A, ESA, Chelmsford, MA) with potentials set at 0, 300, 500, and 600 mV. The mobile phase was a binary mixture of (A) 28 mM K₂HPO₄ and 0.59 mM H₃PO₄, pH 2.5 and (B) a mixture of 75% (v/v) A and 25% CH₃CN with a 55-min nonlinear elution program from 90% A to 100% B at 1 ml min⁻¹. Detector response was calibrated using tetrahydrofolate (THF), 5-methyl-, 5,10-methenyl-, 5 formyl-THF, and folic acid standards from Schircks. For analysis of polyglutamyl tail length (Bagley *et al.*, 2000) the rat plasma treatment was omitted. Folate extracts from human erythrocytes were used to identify the retention times of 5-methyl-THF polyglutamates (Pfeiffer *et al.*, 1996).

MATERIALS AND METHODS FOR EXAMPLES 7-10

20 Materials.

E. coli K12 MG1655 was a gift from the Laboratory of Genetics, University of Wisconsin-Madison (*E. coli* Genome Project; Laboratory of Frederick Blattner). Restriction enzymes used were from New England Biolabs (Beverly, MA). DNA markers were from Gene Choice (PGC Scientific Corporation, Frederick, MD) and the BENCHMARK protein ladder was from Bio-Rad Laboratories (Hercules, CA). Neopterin, GTP, isoxanthopterin, and xanthopterin were obtained from Sigma Chemical Company (St. Louis, MO). FINALE (glufosonate-ammonium: butanoic acid, 2-amino-4 [hydroxymethylphosphinyl] monoammonium salt; AgrEvo USA Company) was purchased from a local nursery.

30 PCR Amplification of the *folE* Gene.

The bacterial gene encoding GCH1 (*folE*, GenBank accession number AE000304) was amplified by polymerase chain reaction (PCR) from *E. coli* K12 MG1655. *E. coli* DNA was purified according to the DNEASY Tissue Kit (Qiagen, Valencia, CA) and used as

template for PCR amplification. The sense primer 5'-CCCATCACTCAGTAAAGAAGCGGC-3' (SEQ ID NO: 10) and antisense primer 5'-CCGTTGTGATGACGCACAGCG-3' (SEQ ID NO: 11) were synthesized based on the nucleotide sequence of the *folE* gene (Blattner *et al.*, 1997). Primers were designed to insert the DNA fragment in the correct reading frame in a PET-BLUE 2 BLUNT vector purchased from Novagen (Madison, WI). PCR was performed according to Sambrook *et al.* (1989) and the product was gel purified using the Qiagen QIAQUICK GEL EXTRACTION KIT. To confirm the identity of the amplified *folE* gene, DNA sequencing was performed at the Protein and Nucleic Acid Core Laboratory at the Washington University School of Medicine (St. Louis, MO).

Construction of the Bacterial Expression Vector.

The PCR insert was converted to a blunt phosphorylated form and ligated with blunt dephosphorylated PET-BLUE 2 VECTOR from Novagen. The resulting plasmid, pTK101, was transformed into NOVA BLUE SINGLES bacterial cells. Recombinant plasmids were isolated and transformed into pET-Blue compatible expression host strain Tuner DE3 pLac1.

Bacterial Expression and Purification of His/HSV-Tagged Protein.

Protein expression was induced with isopropyl β -D-thiogalactopyranoside. The expressed protein was purified by Ni-column chromatography using Novagen Ni-NTA-His-binding resin.

Protein Detection and Identification.

Protein samples were analyzed by SDS-PAGE (12.5%) to determine purity and to estimate molecular mass of the protein (Laemmli, 1970). Protein was transferred electrophoretically to a Protran Pure Nitrocellulose membrane (Schleicher & Schuell Bioscience, Keene, NH). The histidine-tagged EcGCH fusion protein was detected using mouse anti-His-tag monoclonal antibody as primary antibody and alkaline-phosphatase conjugated goat anti-mouse IgG as secondary antibody. Gel-purified fusion protein was used as an antigen for production of polyclonal antibodies in rabbits by Bethyl Laboratories (Montgomery, TX) using standard protocols. Alkaline-phosphatase conjugated goat anti-rabbit IgG was used to detect rabbit anti-EcGCH.

Vector Construction and Transformation of *Arabidopsis thaliana*.

The plant transformation vector PC Gus-Bar was provided by J. Koo (Danforth Plant Science Center). Basta-resistance gene (Bar) encodes phosphinothricin acetyltransferase (PAT) and provides resistance to phosphinothricin (Basta). DNA of PC Gus-Bar and pTK101 were digested with NcoI and EcoRI, DNA fragments were separated on a 1% agarose gel and purified from the gel. Fragments containing PC-Gus-BAR (without the coding sequence for GUS) and the folE DNA insert (including short N-terminal and C-terminal extensions) were ligated using the ROCHE RAPID LIGATION KIT to obtain the plant transformation vector pTK202. Products of ligation were introduced into DH5 α cells and selected on LB media containing kanamycin. Colonies containing the folE coding sequence in the desired orientation relative to the CaMV 35S promoter were identified by restriction with EcoRI and NcoI and electrophoresis in a 1% agarose gel. Competent *Agrobacterium tumefaciens* GV3010 cells were transformed with pTK202 (An, 1987). Transformation of *Arabidopsis* with *A. tumefaciens* GV3010 carrying the plasmid pTK202 was performed using a standard *Arabidopsis* transformation protocol (Bechtold *et al.*, 1993; Clough *et al.*, 1998). Plants were grown in a Conviron growth chamber (22 °C, 50% RH, 200 μ mol light, 10 hr photoperiod).

Selection of Transformed Plants.

To select for transformed plants, plants at the four-leaf stage were sprayed with Finale™ (BASTA; phosphinothricin; 1:400 v/v). BASTA selection was repeated every fifth day for a total of five applications. Leaf discs were collected from the BASTA-selected plants for DNA extraction using the REDEXTRACT-N-AMP Plant PCR kit (Sigma). Extracted DNA was used as template for PCR detection using the folE primers listed above and plants containing the folE gene were transplanted and maintained. Seeds from the primary (T1) generation were planted and resultant T2 plants were subjected to another round of BASTA selection and characterization via PCR and GCH1 assays. The process was repeated to obtain non-segregating T3 transgenic lines.

Extraction of Leaf Tissue for GCH and Pterin Assays.

Leaf tissue from transgenic lines and wild-type Columbia was harvested before the onset of flowering, frozen in liquid nitrogen and stored at -80 °C. Tissue was extracted in 10 mM Tris-HCl buffer (pH 8.0) with 0.2 g of insoluble polyvinyl-pyrrolidone (Sigma Cat.

No. P6755) per g tissue. Extract was filtered through MIRACLOTH and centrifuged at 10,000 x g for 15 min. The supernatant (crude extract) was used for analysis of pterins, protein and GCH activity. Crude extract was extensively dialyzed before GCH1 assay.

5 GCH1 Assay.

The activity of purified EcGCH was determined by measuring the production of neopterin from GTP (Milstien *et al.*, 1996; Duch *et al.*, 1984). The enzymatic product, dihydroneopterin triphosphate, was first oxidized and then dephosphorylated to form neopterin. The latter was analyzed by reverse phase (RP) HPLC on a Beckman
10 ULTRASPHERE C18 RP column. Neopterin was eluted isocratically with 0.5% acetonitrile and 0.1% tetrahydrofuran in water at a flow rate of 1 ml/min. Neopterin and other pterins were detected by fluorescence (365 nm excitation; 446 nm emission) using a Jasco FP 1520 fluorescence detector and by absorbance using a Beckman 168 Diode Array detector. Protein was measured using the Bradford assay (Bradford, 1976).

15

Analysis of pterins.

Unconjugated pterins in leaves were analyzed by RP chromatography using the same conditions as used to detect neopterin in the GCH1 assay. Pterins were detected by fluorescence and absorbance at 280 and 330 nm. Diode array and fluorescence data were
20 compared to authentic standards and literature values (Win, 2000; Kohashi, 1980; Kohashi *et al.*, 1980). Pterin concentrations were estimated by integration of the major fluorescent pterin peaks using neopterin as a standard. The elution of pterins was compared to the elution of selected pterin standards and the identity of these pterins was further analyzed by MS analysis at the Protein and Nucleic Acids Core Facility at Washington University. Acidified
25 samples in 0.5% acetonitrile and 1% THF collected after HPLC separation were introduced into a Finnegan LC-DECA ion trap MS and analyzed in positive-ion mode.

Folate Analysis.

Total folate was analyzed in the Vitamin Metabolism Laboratory, Jean Mayer Human
30 Nutrition Research Center on Aging at Tufts University. Leaf tissue from transgenic *A. thaliana* (T2 and T3 lines) was extracted in extraction buffer (20 g Bis-Tris, 20 g Na-ascorbate and 500 µl mercaptoethanol per liter water) and extracts were treated with conjugase enzymes to free folate for bacterial growth. The microbial assay with

Lactobacillus casei was used to measure total folate content (Grossowicz *et al.*, 1980; Wilson *et al.*, 1982; Horne *et al.*, 1998). Triplicate samples from each T3 line were analyzed in replicated analyses (N=8). Results were expressed as nmol folate per g fresh weight (gfw) tissue. Controls for folate recovery and effectiveness of conjugase treatment were included.

5

MATERIALS AND METHODS FOR EXAMPLES 11-15

Materials.

Arabidopsis thaliana ecotype Columbia was transformed with plant transformation vector (pTK202) containing an *E. coli* gene insert encoding GTP cyclohydrolase 1 (EcGCH) as described in Hossain *et al.* 2004. Neopterin, GTP, isoxanthopterin, and xanthopterin were obtained from Sigma Chemical Company (St. Louis, MO). Non-segregating (T3) transgenic plants expressing the EcGCH were selected on the basis of BASTA (glufosonate-ammonium: butanoic acid, 2-amino-4 [hydroxymethylphosphinyl] monoammonium salt) resistance, PCR, and protein detection on Western blots. Plants were grown in a Conviron growth chamber (22°C, 50% RH, 200 µmol light, 10 hr photoperiod).

10
15

Extraction of Leaf Tissue for Total Pterin Assays.

Leaf tissue from transgenic lines and wild-type Columbia was harvested before the onset of flowering, frozen in liquid nitrogen and stored at -80°C. Seeds were collected from dried siliques and stored under the same conditions. Tissue was extracted by grinding in 10 mM Tris-HCl buffer (pH 8.0) with 0.2 g of insoluble polyvinyl-polypyrrolidone (Sigma Cat. No. P6755) per g tissue. Tissue extracts were filtered through two-layers of MIRACLOTH (Calbiochem) and centrifuged at 10,000 x g for 15 min. The supernatant (crude extract) was used for HPLC analysis of pterins.

20
25

Analysis of total pterins.

Unconjugated pterins in leaves were analyzed by reverse phase (RP) HPLC on a Beckman ULTRASPHERE C18 RP column. Pterins and standards were eluted isocratically with 0.5% acetonitrile and 0.1% tetrahydrofuran in water at a flow rate of 1 ml/min and detected by fluorescence (365 nm excitation; 446 nm emission) using a Jasco FP 1520 fluorescence detector and by absorbance using a Beckman 168 Diode Array detector. Diode array and fluorescence data were compared to authentic standards and literature values (Win

30

2000; Kohashi 1980; Kohashi *et al.*, 1980). Total pterin concentrations were estimated by integration of the major fluorescent pterin peaks using neopterin as a standard.

Total PABA Analysis.

5 *Arabidopsis* samples (400-500 mg) were triturated in liquid N₂ and resuspended in 5 ml of methanol. The suspension was centrifuged and the supernatant collected. The pellet was washed with 2 ml of methanol, centrifuged and the second supernatant was combined with the initial 5 ml fraction. Duplicate 3.5-ml portions of the methanol extract were used for analysis of total PABA (i.e., free PABA plus PABA glucose ester) by acid hydrolysis, cation
10 exchange chromatography, ethyl acetate partitioning, reverse phase HPLC with fluorometric detection as previously described (Quinlivan *et al.*, 2003).

Characterization of Plant Pterins.

Arabidopsis samples (250 mg) were triturated in liquid N₂ and resuspended in 2.5 ml
15 of methanol. A 600- μ l portion of the homogenate was mixed with 250 μ l of CHCl₃ and 50 μ l of water and shaken for 40 min before adding another 225 μ l of CHCl₃ and 340 μ l of water. After shaking for 10 min, the sample was centrifuged to break the emulsion. The aqueous phase was removed, dried *in vacuo*, and redissolved in 200 μ l of water. Samples were oxidized by adding 0.1 vol of a solution of 1% I₂ and 2% KI (wt/vol) in 1 M HCl and
20 incubating in darkness for 1 h; excess I₂ was then removed by adding 10 μ l of 5% (wt/vol) Na-ascorbate (Fukushima and Nixon, 1980). The oxidized samples were separated by HPLC with a 4- μ m, 250 x 4.6-mm Synergi Fusion-RP 80 column (Phenomenex, 411 Madrid Avenue, Torrance, CA 90501-1430) by isocratic elution with 10 mM Na-phosphate (pH 6.0) with a flowrate of 1.5 ml·min⁻¹ (Diaz de la Garza *et al.*, 2004). Peaks were detected using a
25 Waters 2475 fluorescence detector (350-nm excitation and 450-nm emission) and identified by comparison with reference standards and by spectral properties.

Folate Analysis.

 Total folate was analyzed in the Vitamin Metabolism Laboratory, Jean Mayer Human
30 Nutrition Research Center on Aging at Tufts University. Leaf and seed tissue from transgenic *A. thaliana* was extracted in extraction buffer (20 g Bis-Tris, 20 g Na-ascorbate and 500 μ l mercaptoethanol per liter water) and extracts were treated with conjugase to free folate for bacterial growth. The microbial assay with *Lactobacillus casei* was used to

measure total folate content (Grossowicz *et al.*, 1981; Wilson & Horne 1982; Horne and Patterson 1988). Triplicate samples were analyzed in replicated analyses (N=8). Results were expressed as μg folate per g fresh weight (gfw) tissue. Controls for folate recovery and effectiveness of conjugase treatment were included.

5

All patents, patent applications, provisional applications, and publications referred to or cited herein are incorporated by reference in their entirety, including all figures and tables, to the extent they are not inconsistent with the explicit teachings of this specification.

10

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

15 EXAMPLE 1—OVEREXPRESSING GCHI IN RIPENING FRUIT INCREASES PTERIDINE LEVELS

A synthetic GCHI gene was constructed by partially recoding a mammalian GCHI cDNA (Figure 2A) and by adjusting the sequence context around the start codon to fit the plant consensus (Koziel *et al.*, 1996). This synthetic gene was placed behind the ripening-specific E8 promoter (Deikman *et al.*, 1992) in an *Agrobacterium* binary vector, and introduced into tomato plants. Screening fruit by Western analysis identified transformants that expressed a protein of the correct size (27 kDa) in a ripening-specific fashion (Figure 2B). Twelve such GCHI⁺ transformants, spanning a range of protein expression levels, were chosen for further analysis along with ten empty vector controls. GCHI⁺ plants and their fruit were not visibly different from control plants.

Pteridines were analyzed by fluorometric HPLC after conversion to their fully oxidized, fluorescent forms. The resulting pteridine profiles of GCHI⁺ fruits differed greatly from those of vector controls (Figure 2C). Measured at the red-ripe stage, the total pteridine content of the GCHI⁺ fruit was from 3- to 140-fold higher than the average for controls. Monitoring pteridines during ripening showed that GCHI⁺ fruit began to diverge from controls after the mature green stage and that their pteridine contents were highest at the red-ripe stage (Figure 2D). This pattern of pteridine accumulation fits with the expected activity of the E8-promoter that was used to drive GCHI expression (Deikman *et al.*, 1992).

30

EXAMPLE 2—PTERIDINE SPECIATION

The major pteridines detected in GCHI⁺ fruit included neopterin, monapterin, and 6-hydroxymethylpterin, which are oxidized forms of folate synthesis intermediates, as well as 6-carboxypterin, a catabolite of pteridines and folates (Figure 2C). The identity of 6-hydroxymethylpterin was confirmed by using HPLC-mass spectrometry to check its mass ($[M+H]^+ = m/z$ 194) and fragmentation pattern relative to an authentic standard. GCHI⁺ fruit also showed peaks at ~7.5 and ~9.5 min that did not match any of our standards (Figure 2C). Acid hydrolysis converted these unknowns to neopterin and monapterin, respectively, as did treatment with almond β -glucosidase but not yeast α -glucosidase (Figure 3A). This result indicates that both unknown peaks are β -D-glycosides. The relative proportions of the various pteridines were similar in all GCHI⁺ fruit, regardless of absolute levels (Figure 3B). Control fruit had small peaks with the same retention times as most of the pteridines in GCHI⁺ fruit (not shown) but containing too little material to identify spectrally.

The *in-vivo* oxidation states of the pteridines in representative GCHI⁺ fruit were studied by comparing samples extracted in the presence of β -mercaptoethanol and ascorbate (to maintain native oxidation state), and given no oxidation treatment, with sister samples that were extracted and oxidized as usual. Since reduced pteridines do not fluoresce, the difference in peak size between oxidized and non-oxidized treatments measures reduced pteridines. This method indicated that pteridine pools were typically ~50-90% reduced at breaker and red stages, but $\leq$ 20% reduced at the red-ripe stage. The reduced pteridines appeared to be dihydro, not tetrahydro, forms since alkaline I₂ treatment did not yield any pterin, a diagnostic cleavage product of tetrahydro pteridines in these conditions (Fukushima *et al.*, 1980).

EXAMPLE 3—PTERIDINE ACCUMULATION ENHANCES FOLATE CONTENT

The total folate contents of control fruit fell in the range 0.8-2.3 nmol g⁻¹ fresh weight (FW) whereas almost all the GCHI⁺ fruit exceeded this range (Figure 4A) and the mean for the GCHI⁺ population was double that for the controls (2.99 vs. 1.47 nmol g⁻¹ FW; difference significant at $P < 0.001$). The increase in folate was contributed mainly by 5-methyl-THF and 5,10-methenyl-THF, both of which were major forms in vector-alone control fruit (Figure 4B). (The acidic mobile phase used in the HPLC analysis causes quantitative conversion of 10-formyl-THF to 5,10-methenyl-THF, so that the 5,10-methenyl-THF peak included both

10-formyl-THF and any preexisting 5,10-methenyl-THF.) Analysis of the polyglutamyl forms of 5-methyl-THF showed a similar distribution in GCHI⁺ and control fruit, with hexaglutamate dominant in both (Figure 4C). Other folates also existed mainly as polyglutamates in both GCHI⁺ and control fruit (not shown). Taken together, these data show that the extra folate in GCHI⁺ fruit was all in normal forms.

EXAMPLE 4—EXCESS PTERIDINE ACCUMULATION MAY DEPRESS FOLATE PRODUCTION

A scatter plot of total pteridine level vs. folate content (Figure 5A) shows that maximal folate was attained with a total pteridine level of ~25 nmol g⁻¹ FW, and that further increase in pteridine was of no benefit. In fact, there was a significant negative correlation between folate content and pteridine level for fruit with >25 nmol pteridine g⁻¹ FW ($r^2 = 0.43$, $P = 0.015$) (Figure 5A). Significant negative correlations were also found between folate and each of the individual pteridines except hydroxymethylpterin. Of these correlations, that with neopterin ($r^2 = 0.60$, $P = 0.002$) was the strongest. Such negative trends suggest that excess pteridines, beyond being unnecessary for high folate production, may actually antagonize it.

EXAMPLE 5—FOLATE ACCUMULATION IS LINKED TO PABA DEPLETION

Because the total PABA pool in non-engineered tomatoes is about the same size (1 to 2 nmol g⁻¹ FW; Basset *et al.*, 2004; Quinlivan *et al.*, 2003) as the folate increase in GCHI⁺ fruit, the relationship between the levels of folate and PABA was examined. A significant negative correlation was evident, and the GCHI⁺ fruit with the highest folate contents had very little PABA (Figure 5B). Furthermore, it can be calculated from the data of Figure 5A that in the five fruits richest in folate, 90-97% of the PABA moieties present in the fruit had been incorporated into folate.

EXAMPLE 6—EXOGENOUS PABA INCREASES FOLATE ACCUMULATION

To test whether PABA depletion limits folate accumulation, GCHI⁺ fruit were harvested at the breaker stage, supplied via the cut stalk with PABA (or water as a control), and allowed to ripen. Fruit tissue was then analyzed for PABA and folates. Among fruit from seven different transformants, those given PABA had much higher PABA levels than controls (0.0-0.0 vs. 0-00 nmol g⁻¹ FW) and were 2.5- to 10-fold richer in folate (Figure 5C).

The folate levels in the water-control GCHI⁺ fruit (open bars, Figure 5C), which ripened off the plant, were below those in matching fruit that ripened on the plant (Figure 4A).

EXAMPLE 7—CLONING AND EXPRESSION OF *E. COLI* GCHI GENE

5 The EcGCH (*folE*) gene was PCR amplified, cloned and expressed in *E. coli*. The expressed His-tagged fusion protein was purified by nickel-affinity chromatography, analyzed by SDS-PAGE (Figure 6A), and detected on western blots using anti-His antibody (Figure 6B). Antiserum raised against affinity and gel-purified fusion protein detected the same protein band (Figure 6C), while preimmune serum did not react with the purified fusion
10 protein. GCH1 activity of the purified fusion protein was 20-fold higher than the activity of the fusion protein isolated from uninduced transformed cells (data not shown). The identity of the expressed protein was confirmed by peptide sequence analysis of the affinity-purified protein after trypsin digestion and *de novo* sequencing on an ABI Q-ToF mass spectrometer equipped with a Hewlett Packard capillary LC system (data not shown). These results
15 confirm the identity of the expressed protein and indicate that the EcGCH fusion protein was enzymatically active.

EXAMPLE 8—TRANSFORMATION OF PLANTS WITH *E. COLI* GCHI GENE

 The *folE* gene sequence was inserted into a modified pCAMBIA plant transformation
20 vector and the resulting vector was used to transform *A. thaliana* cv. Columbia. The primary (T1) transformants were selected for BASTA resistance and screened for the presence of the *folE* gene sequence by PCR. Extracts from BASTA-resistant, PCR-positive transgenic plants and non-transgenic Columbia were tested for GCH1 activity. Only lines with GCH1 activity were retained. Activity was not detected in wild-type *Arabidopsis* while GCH1 activity was
25 detected in extracts of transgenic T1 lines. Transgenic plants from primary T1 lines 2, 5 and 6 expressing EcGCH (N=26) were subjected to a second round of BASTA selection and PCR analysis. Three individual T2 plants (2-14; 2-15; 2-19; 5-14; 5-15; 5-21; 6-10; 6-12; 6-13) were selected from each of three primary transgenic lines for further analysis. Plants derived from these nine lines were subjected to another round of BASTA selection and testing to
30 obtain non-segregating, PCR-positive T3 transgenic lines. Leaf tissue from each of the nine transgenic lines was collected for further analysis.

 To confirm that the bacterial gene was expressed, leaf tissue from the T3 plants was extracted and analyzed by SDS-PAGE and immunoblotting. An immunoreactive band

corresponding in size to the predicted molecular weight (26.2 kDa) of the *folE* gene product EcGCH was detected in these extracts but not in wild-type control (Figure 6D). While the quantity of EcGCH protein detected in extracts from plants derived from any one of the three primary transgenic lines was quite similar; the differences between lines 2, 5 and 6 were more significant. Plants derived from primary lines 2 and 5 had substantially more immuno-detectable protein than plants derived from line 6 (Figure 6D). When ~15 µg of total leaf protein was loaded onto a mini-gel, lines 2 and 5 exhibited a distinct immunoreactive band of the correct MW (Note: Several smaller bands were detected when the blot was over exposed to detect EcGCH in line 6 suggesting that proteolytic cleavage may have occurred), whereas there was very little, if any, band visible in extracts from line 6 under the same conditions. However, when extract from derivatives of primary line 6 was concentrated and ~210 µg protein was loaded onto a larger gel, an immuno-detectable band was observed in line 6 (Figure 6E) with additional smaller bands (not shown). The results of this analysis suggest that the levels of EcGCH in line 6 are 10- to 20-fold lower than the levels in lines 2 and 5.

To determine if EcGCH expressed in *Arabidopsis* was active, leaf extracts from T3 plants of each of the three primary lines were tested for GCH1 activity using the standard fluorometric assay. GCH1 activity was detected in transgenic plants but not in non-transgenic controls under equivalent conditions (data not shown). Crude extracts from different T3 plants derived from the same primary line exhibited similar GCH1 activity, whereas there were substantial differences in GCH1 activity between plants derived from different primary lines. Plants from primary lines 2 and 5 exhibited measurable activity, while activity was not detectable above background in extracts of plants derived from line 6 under standard assay conditions. These differences in GCH1 activity were positively correlated with the differences in the amount of EcGCH detected via immunoblot reactions.

EXAMPLE 9—PTERIN SYNTHESIS IN TRANSGENIC PLANTS

To test whether expression of EcGCH in transgenic plants increases pterin synthesis, undialyzed crude leaf extracts of transgenic lines were subjected to HPLC analysis using the same chromatographic conditions as for the analysis of neopterin in the GCH1 assay. This analysis confirmed the accumulation of high amounts of fluorescent pterin-like metabolites in leaf extracts of T2 (data not shown) and T3 plants (Figures 7A-7B) compared with non-transgenic plants. The fluorescence properties and UV/visible spectra as well as preliminary

analysis by MS indicate that the fluorescent compounds detected in the crude extracts were unconjugated pterins, referred to hereafter as pterins. The compound eluting at 4.5 min (Peak 1) was tentatively identified as neopterin based on comparison of elution time (Figures 7A-7B), UV/visible spectrum (Figures 7C-7D) and mass spectrum with neopterin standard (Figures 7E-7F). Two other fluorescent metabolites present in extracts of transgenic plants exhibited similar retention times, spectral properties and fluorescence characteristics (Win, 2000; Kohashi, 1980; Kohashi *et al.*, 1980) to xanthopterin and isoxanthopterin (Figure 7B), products of pterin degradation.

Examination of fluorescence profiles indicated that the most abundant pterins can be detected in crude extracts of non-transgenic plants but the levels are very low. The concentration of pterins increased for each line from the T2 to the T3 generation with levels in T3 plants from primary line 2 and 5 being 10- to 20-fold higher than the corresponding levels in derivatives of line 6, as summarized graphically (Figure 8A). The levels of neopterin observed in transgenic plants were up to 1100-fold higher than the corresponding levels in non-transgenic plants and the quantities of pterins present in transgenic lines were highly correlated with the amount of EcGCH detected (see Figures 6D and 6E), supporting the conclusion that these products accumulated in response to the expression of the transgene. The concentration of pterins in non-transgenic *Arabidopsis* were comparable to those reported for other plant species (Figure 8B), whereas the levels in transgenic plants were 750- to 1250-fold higher than controls (Figures 8A and 8B).

EXAMPLE 10—SYNTHESIS AND ACCUMULATION OF FOLATES IN TRANSGENIC PLANTS

To determine whether the expression of EcGCH increased synthesis and accumulation of folates, the total folate levels of transgenic and non-transgenic plants were measured by microbial bioassay and statistically analyzed using a one-way ANOVA followed by pairwise comparison of means using both the Student's t-test and the Tukey-Kramer method (Figure 9A). The mean value for total folate for non-transgenic *Arabidopsis* was 1.41 nmol folate/gfw leaf tissue. This is within the range of values reported for leaves and seeds from other plant species (Figure 9B) (Cossins, 2000). The folate values for primary transgenic lines were up to 3.3-fold higher than non-transgenic *Arabidopsis*, with values for replicate samples ranging from 2.27 to 4.70 nmol/gfw with an average of 3.40, 3.40 and 2.55 nmol/gfw for plants

derived from line 2, 5 and 6, respectively. Based on the statistical analysis, the folate contents of all transgenic lines were significantly higher ($p = 0.05$) than control. The folate content of line 6, which had 10- to 20-fold lower levels of EcGCH and total pterins, was ~50% lower than the folate values for lines 2 and 5. The average value for the ratio of pterin to folate levels for non-transgenic *Arabidopsis* (this study) and other plant species (values reported in the literature) was 1.40, whereas the ratio for transgenic T3 lines ranged between 55 and 644 with an average of 350. The observed increase in total folate levels as a function of pterin levels (Figure 9C) appeared to saturate. The maximum values for transgenic lines were approximately 2-fold higher than those reported for spinach, a plant considered to be rich in folates (Figure 9B). These results demonstrate the enhancement of folate levels in plants through bioengineering.

EXAMPLE 11—PTERIN ANALYSIS

In results reported earlier (Hossain *et al.* 2004), total pterin levels in extracts of leaves of non-segregating transgenic *Arabidopsis* expressing EcGCH accumulated to levels up to 1200-fold higher than the levels present in leaves of non-transgenic control plants. In these initial studies, at least 8-10 different pterins were shown to accumulate in transgenic plants by HPLC analysis of leaf extracts. The identity of the major pterin peak as neopterin was confirmed by comparison of the UV-visible spectrum and mass spectrum with those of authentic neopterin.

To test whether some of the unidentified pterins in extracts of transgenic *Arabidopsis* leaves could simply be conjugated forms of pterins, samples were heated at 100 C in 1 N HCl for 30 to 60 minutes to break potential glycosidic or phosphoryl bonds. Several peaks disappeared upon acid treatment suggesting that these unidentified pterins might be conjugated derivatives (data not shown). To further explore whether these acid-labile forms might be glycosylated or phosphorylated conjugates of the pterins, the *Arabidopsis* leaf extracts were treated enzymatically with either beta-glucosidase or alkaline phosphatase, respectively. Beta-glucosidase treatment had little effect with a slight increase in peak height in an early eluting form. Likewise, treatment with alkaline phosphatase had limited effects.

To further identify and characterize the other major pterins accumulating in the transgenic *Arabidopsis* plants, leaf extracts were analyzed by HPLC using two different chromatography systems and the elution times of pterins in leaf extracts were compared to those of authentic neopterin, monapterin, hydroxymethylpterin, carboxypterin, neopterin

glycoside and monapterin glycoside standards. Results upon chromatographic separation with an Ultremex column are shown in Figure 11. Similar results were obtained using a Synergi column (data not shown). These results (red trace in Figure 11) were compared to those obtained under the same conditions using extract of tomato fruit (black trace) engineered to express a synthetic GCH1 gene derived from an animal source and optimized for plant expression (Diaz de la Garza *et al.*, (2004). Diaz de la Garza *et al.* (2004) demonstrated the presence of neopterin glycoside and monapterin glycoside in transgenic tomato fruit expressing GCH1. The major unconjugated pterins identified in the *Arabidopsis* extract included neopterin, monapterin, 6-carboxypterin and hydroxymethylpterin. These pterins accounted for ~60-70% of the total pterins detected in the transgenic *Arabidopsis* leaf extract. The same four pterins were detected in the transgenic tomato fruit extract although the total concentrations and the relative distribution between and among these four pterins differed in transgenic *Arabidopsis* and tomato. Diaz de la Garza *et al.* (2004) demonstrated the presence of neopterin glycoside and monapterin glycoside in transgenic tomato fruit expressing GCH1. In contrast with results with tomato, neopterin glycoside and monapterin glycoside were not detected in *Arabidopsis* leaf extracts. At least four unidentified pterins (Unknown 1, 2, 3 and 4) were present in *Arabidopsis* leaf extracts but not detected in tomato extracts. At least 4 additional unidentified peaks were present and we hypothesized that these might be conjugated forms, e.g. glycosylated or phosphorylated forms of pterins.

EXAMPLE 12—DEVELOPMENTAL AND TISSUE SPECIFIC ACCUMULATION OF PTERINS

In Hossain *et al.* 2004, the presence of high levels of pterins in leaf tissue of mature transgenic *Arabidopsis* expressing a bacterial GCH1 were demonstrated. In the more recent studies, leaf tissue was collected periodically from transgenic and non-transgenic plants grown under controlled environmental conditions through normal vegetative growth. Leaf tissue was extracted and analyzed by HPLC for changes in pterin profiles and for total pterin levels. The levels of total pterins in transgenic plants increased with plant age (Figure 12) while the relative distribution between different pterin forms remained relatively constant (data not shown). Levels of total pterins in non-transgenic lines remained extremely low and did not change during development (data not shown).

To determine whether pterins were present and accumulated in different plant tissues, roots, leaves and seeds were collected from transgenic and non-transgenic *Arabidopsis* plants

and analyzed for total pterins and pterin profiles by HPLC. Results of the pterin analysis for extracts of seeds of transgenic (after 20-fold dilution) and non-transgenic (undiluted) *Arabidopsis* are shown in Figure 13. Pterin profiles of transgenic seeds were similar to those obtained from leaves of transgenic plants. Analysis of pterin accumulation in seeds of different transgenic lines followed the same pattern as in the leaves of the same lines with the highest levels in Line 219 and the lowest in line 613 (Figure 14). The levels of total pterins in seeds from different transgenic lines corresponded with the relative levels of pterins observed in leaves of the same lines.

Pterins accumulated in roots, leaves and seeds of transgenic plants over those observed in control plants (Figure 15). In contrast to leaves, the levels of pterins in seeds of non-transgenic controls were easily detectable. Pterins accumulated to much higher levels in seeds than in leaves or roots (Figure 15). The profile of different forms of pterins accumulating differed with tissue. The levels of specific pterins differed between roots, leaves and seeds (Figure 16). Roots contained primarily 6-carboxypterin, neopterin, monapterin and hydroxymethylpterin whereas seeds and leaves contained additional forms of pterins.

EXAMPLE 13—ACCUMULATION OF FOLATES IN SEEDS OF TRANSGENIC PLANTS

Seeds of transgenic and non-transgenic controls were analyzed for total folates using the microbiological assay. Values ranged from 11 to 13.8 μg per g seed weight (Figure 17) compared to 1.2 $\mu\text{g/g}$ fresh weight (gfw) for the leaves of transgenic plants. Folate levels in seeds of transgenic lines 219, 515 and 613 were enhanced 2.5 to 3.5 fold over those in non-transgenic seeds. These increases in total folates of seeds paralleled those observed in leaves. This is the first demonstration of the engineering of increased folate levels in the seeds of plants.

EXAMPLE 14—ANALYSIS OF TOTAL PABA LEVELS IN LEAVES OF TRANSGENIC AND NON-TRANSGENIC PLANTS

Results of the analysis of total PABA levels in leaves of transgenic and non-transgenic *Arabidopsis* are presented in Figure 18. Total PABA levels were reduced by more than 20-fold in the transgenic lines 514 and 219 when compared to the non-transgenic control

plants. The depletion of total PABA suggests that the supply of PABA is a limiting factor in folate biosynthesis.

EXAMPLE 15—ANALYSIS OF TOTAL PABA, PTERINS AND FOLATES IN LEAVES
AND SEEDS OF TRANSGENIC AND NON-TRANSGENIC PLANTS TREATED WITH
AN EXOGENOUS SUPPLY OF PABA

Transgenic plants were grown with and without exogenous supply of PABA added to the nutrient solution. In the initial experiments, plants were grown without added PABA or with a single concentration of PABA and leaves and seeds were harvested for total folate analysis. In these initial studies, the levels of total folates, which are already elevated by a factor of 2- to 4-fold in transgenic plants over those in non-transgenic plants were further elevated by a factor of 7-fold and 2-fold respectively in seeds and leaves of the transgenic plants treated with PABA when compared to transgenic plants not treated with PABA (Figure 19). These results clearly support the conclusion that PABA supply limits folate biosynthesis in plants.

To further support this, transgenic and non-transgenic plants were treated and grown with different levels of PABA (0, 0.5, 1.0 and 2.0 mM). Results of the analysis of total PABA levels in the transgenic and non-transgenic lines treated with different levels of PABA are shown in Figure 20A. Results at the beginning of the study prior to supplying exogenous PABA are presented in Figure 20B. Total PABA levels were depleted by a factor of 3 in the transgenic plants at this early stage of plant growth and development. Total PABA levels increased in both non-transgenic and transgenic plants with increasing levels of exogenous PABA. The levels however were still lower in the transgenic lines at all concentrations of PABA which indicates that the PABA taken up by the plants was consumed to produce folates due to the increased synthesis of pterins in these plants expressing EcGCH. The levels of total pterins in transgenic plants treated with different levels of PABA were fairly similar (Figure 21).

Results of the analysis of total folates in the leaves of the transgenic plants exposed to different levels of exogenous PABA at two different sampling dates are shown in Figures 22 and 23. The levels of total folates increased with increasing PABA concentration in the nutrient solution. When compared to plants grown without added PABA, the levels of total folates observed at the highest PABA levels at the earlier and later sampling dates were 50% and 100% higher, respectively. The levels of total folates also increased in non-transgenic

plants with increasing concentrations of PABA. Thus, increasing PABA supply in plants can independently result in increased folate levels.

5 It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application. In addition, any elements or limitations of any invention or embodiment thereof disclosed herein can be combined with any and/or all other elements or limitations
10 (individually or in any combination) or any other invention or embodiment thereof disclosed herein, and all such combinations are contemplated with the scope of the invention without limitation thereto.

REFERENCES

- U.S. Patent No. 6,642,435
- U.S. Patent No. 5,512,466
- 5 U.S. Patent No. 5,625,136
- U.S. Patent No. 5,034,322
- 10 U.S. Patent No. 5,589,610
- U.S. Patent No. 5,639,948
- U.S. Patent No. 6,462,185
- 15 U.S. Patent No. 5,106,739
- U.S. Patent No. 6,455,760
- 20 U.S. Patent No. 6,696,623
- U.S. Patent No. 5,661,017
- U.S. Published Application No. 20040078841
- 25 U.S. Published Application No. 20040067506
- U.S. Published Application No. 20040019934
- 30 U.S. Published Application No. 20030177536
- U.S. Published Application No. 20030084486
- U.S. Published Application No. 20040123349
- 35 Altschul, S. F. *et al.* (1990) "Basic Local Alignment Search Tool" *J. Mol. Biol.* 215:402-410.
- Altschul, S. F. *et al.* (1997) "Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs" *Nucl. Acids Res.* 25:3389-3402.
- 40 An, G. (1987) In: *Methods in Enzymology*, 153, eds. Wu, R. and Grossman, L. (Academic Press, New York), pp. 292-305.
- Bagley, P. J. & Selhub, J. (2000) "Analysis of folate form distribution by affinity followed by reversed- phase chromatography with electrical detection" *Clin. Chem.* 46(3):404-111.
- 45 Basset, G., Quinlivan, E. P., Ziemak, M.J., Díaz de la Garza, R., Fischer, M., Schiffmann, S., Bacher, A., Gregory, J. F. III & Hanson, A. D. (2002) "Folate synthesis in plants: the

first step of the pterin branch is mediated by a unique bimodular GTP cyclohydrolase I” *Proc. Natl. Acad. Sci. USA* 99(19):12489-12494.

- 5 Basset, G. J., Quinlivan, E. P., Ravel, S., Rébeillé, F., Nichols, B. P., Shinozaki, K., Seki, M., Adams-Phillips, L. C., Giovannoni, J. J., Gregory, J. F. III, Hanson A. D. (2004) “Folate synthesis in plants: the p-aminobenzoate branch is initiated by a bifunctional PabA-PabB protein that is targeted to plastids” *Proc. Natl. Acad. Sci. USA* 101(6):1496-1501.
- 10 Bechtold, N., Ellis, J. & Pelletier, G. (1993) *C. R. Acad. Sci. Paris Life Sci.* 316:1194-1199.
- Beltz, G. A., Jacobs, K. A., Eickbush, T. H., Cherbas, P. T., Kafatos, F. C. (1983) “Isolation of multigene families and determination of homologies by filter hybridization methods” *Methods of Enzymology*, R. Wu, L. Grossman and K. Moldave [eds.]
15 Academic Press, New York 100:266-285.
- Blattner, F. R., Plunkett, G. III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F. *et al.* (1997) *Science* 277:1453-1474.
- 20 Bouis, H. E. (2002) “Plant breeding: a new tool for fighting micronutrient malnutrition” *J. Nutr.* 132(3):491S-494S.
- Bradford, M. M. (1976) *Anal. Biochem.* 72:248-254.
- 25 Centers for Disease Control and Prevention. (1996) *Morbidity and Mortality Weekly Report* 44 (SS-4):1-13.
- Clancy, M. and Hannah, L.C. (2002) “Splicing of the maize Sh1 first intron is essential for enhancement of gene expression, and a T-rich motif increases expression without affecting splicing” *Plant Physiol.* 130(2):918-29.
- 30 Clough, S. J. & Bent, A. F. (1998) *Plant J.* 16:735-743.
- 35 Cossins, E. A. & Chen, L. (1997) “Folates and one-carbon metabolism in plants and fungi” *Phytochemistry* 45(3):437-452.
- Cossins, E. A. (2000) *Can. J. Bot.* 78:691-708.
- 40 de Boer, H. A., Comstock, L. J., Vasser, M. (1983) “The tac promoter: a functional hybrid derived from the trp and lac promoters” *Proc. Natl. Acad. Sci. USA* 80(1):21-25.
- de Bree, A., van Dusseldorp, M., Brouwer, I. A., van het Hof, K. H. & Steegers-Theunissen, R.P. (1997) “Folate intake in Europe: recommended, actual and desired intake” *Eur. J. Clin. Nutr.* 51(10):643-660.
- 45 Boggio, S. B., Palatnik, J. F., Heldt, H. W. & Valle, E. M. (2000) “Changes in amino acid composition and nitrogen metabolizing enzymes in ripening fruits of *Lycopersicon esculentum* Mill” *Plant Sci.* 159(1):125-133.

- Deikman, J., Kline, R. & Fischer, R. L. (1992) *Plant Physiol.* 100:2013-2017.
- 5 DellaPenna, D. (1999) "Nutritional genomics: manipulating plant micronutrients to improve human health" *Science* 285(5426):375-379.
- Diaz de la Garza, R. D. Quinlivan, E. P., Klaus, S. M. J., Basset, G. J. C., Gregory, J. F., Hansen, A. D. (2004) *Proc. Natl. Acad. Sci. USA* 101: 13720-13725.
- 10 Duch, D. S., Browns, S. W., Wolf, J. H. & Nichol, C. A. (1984) "Biopterin cofactor biosynthesis: GTP cyclohydrolase, neopterin and biopterin in tissues and body fluids of mammalian species" *Life Sci.* 35:1895-1901.
- 15 Fell, D. A. (1998) "Increasing the flux in metabolic pathways: A metabolic control analysis perspective" *Biotechnol. Bioeng.* 58(2-3):121-124.
- Forrest, H. S. & Van Baalen, C. (1970) "Microbiology of unconjugated pteridines" *Annu. Rev. Microbiol.* 24:91-108.
- 20 Fukushima, T. & Nixon, J. C. (1980) "Analysis of reduced forms of biopterin in biological tissues and fluids" *Anal. Biochem.* 102(1):176-188.
- Good, X. *et al.* (1994) "Reduced ethylene synthesis by transgenic tomatoes expressing S-adenosylmethionine hydrolase" *Plant Molec. Biol.* 26:781-790.
- 25 Goyer, A., Illarionova, V., Roje, S., Fischer, M., Bacher, A. & Hanson, A. D. (2004) "Folate biosynthesis in higher plants. cDNA cloning, heterologous expression, and characterization of dihydroneopterin aldolases" *Plant Physiol.* 135(1):103-111.
- 30 Green, J. M., Merkel, W. K., Nichols, B. P. (1992) "Characterization and Sequence of *Escherichia coli* pabC, the gene encoding aminodeoxychorismate lyase, a pyridoxal phosphate-containing enzyme" *Journal of Bacteriology* 174:5317-5323.
- 35 Gregory, J. F. III & Toth, J. P. (1988) "Chemical synthesis of deuterated folate monoglutamate and in vivo assessment of urinary excretion of deuterated folates in man" *Anal. Biochem.* 170(1):94-104.
- Grossowicz, N., Waxman, S. & Schreiber, C. (1981) "Cryoprotected *Lactobacillus casei*: an approach to standardization of microbiological assay of folic acid in serum" *Clin. Chem.* 27:745-747.
- 40 Hanson, A. D. & Gregory, J. F. III (2002) "Synthesis and turnover of folates in plants" *Curr. Opin. Plant Biol.* 5(3):244-249.
- 45 Horne, D. W. & Patterson, D. (1988) "Lactobacillus Casei Microbiological Assay of Folic Acid Derivatives in 96-well Microtiter Plates" *Clin. Chem.* 34:2357-2359.

- Hossain, T., Rosenberg, I., Selhub, J., Kishore, G., Beachy, R. & Schubert, K. (2004) "Enhancement of folates in plants through metabolic engineering" *Proc. Natl. Acad. Sci. USA* 101(14):5158-5163.
- 5 Karlin S. and Altschul, S. F. (1990) "Methods for Assessing the Statistical Significance of Molecular Sequence Features by Using General Scoring Schemes" *Proc. Natl. Acad. Sci. USA* 87:2264-2268.
- 10 Karlin S. and Altschul, S. F. (1993) "Applications and Statistics for Multiple High-Scoring Segments in Molecular Sequences" *Proc. Natl. Acad. Sci. USA* 90:5873-5877.
- Kohashi, M. (1980) "Isolation of Six Unconjugated Pteridines From Soybean (*Glycin max* L. Tsurunoko) Seeds" *J. Biochem.* 87:1581-1586.
- 15 Kohashi, M., Tomita, K. & Iwai, K. (1980) *Agric Biol Chem.* 44:2089-2094.
- Konings, E. J., Roomans, H. H., Dorant, E., Goldbohm, R. A., Saris, W. H. & van den Brandt, P. A. (2001) "Folate intake of the Dutch population according to newly established liquid chromatography data for foods" *Am. J. Clin. Nutr.* 73(4):765-776.
- 20 Koziel, M. G., Carozzi, N. B. & Desai, N. (1996) "Optimizing expression of transgenes with an emphasis on post-transcriptional events" *Plant Mol. Biol.* 32(1-2):393-405.
- 25 Krishnaswamy, K. & Madhavan Nair, K. (2001) "Importance of folate in human nutrition" *Br. J. Nutr.* 85:S115-S124.
- Laemmli, U. K. (1970) *Nature* 227:680-685.
- 30 Lemire, R. J. (1998) *JAMA* 259:558-562.
- Lewin, B. (1985) *Genes II*, John Wiley & Sons, Inc., p. 96.
- 35 Lin, X. L. & White, R. H. (1988) "Structure of Solfapterin (*erythro*-Neopterin-3'-D-2-Deoxy-2-Aminoglucopyranoside) Isolated from the Thermophilic Archaeobacterium *Sulfolobus solfataricus*" *J. Bacteriol.* 170:1396-1398.
- Lucock, M. (2000) "Folic acid: nutritional biochemistry, molecular biology, and role in disease processes" *Mol. Genet. Metab.* 71(1-2):121-138.
- 40 Maniatis, T., E.F. Fritsch, J. Sambrook (1982) "Nuclease *Bal31*" *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- 45 Milstien, S., Jaffe, H., Kowlessur, D. & Bonner, T. I. (1996) "Purification and Cloning of the GTP Cyclohydrolase I Feedback Regulatory Protein, GFRP" *J. Biol. Chem.* 271:19743-19751.
- Molloy, A. M. & Scott, J. M. (2001) "Folates and prevention of disease" *Public Health Nutr.* 4(2B):601-609.

Moore, C. A., Li, S., Li, Z. Hong, S., Gu, H., Berry, R. J., Mulinare, J. & Ericson, J. D. (1997) *Am. J. Med. Gen.* 73:113-118

5 Nuccio, M. L. & Thomas, T. L. (1999) "ATS1 and ATS3: two novel embryo-specific genes in *Arabidopsis thaliana*" *Plant Mol. Biol.* 39(6):1153-1163.

Pfeiffer, C. M. & Gregory, J. F. III (1996) "Enzymatic deconjugation of erythrocyte polyglutamyl folates during preparation for folate assay: investigation with reversed-phase liquid chromatography" *Clin. Chem.* 42(11):1847-1854.

10 Quinlivan, E. P., Roje, S., Basset, G., Shachar-Hill, Y., Gregory, J. F. III & Hanson, A. D. (2003) "The folate precursor p-aminobenzoate is reversibly converted to its glucose ester in the plant cytosol" *J. Biol. Chem.* 278(23):20731-20737.

15 Ravanel, S., Cherest, H., Jabrin, S., Grunwald, D., Surdin-Kerjan, Y., Douce, R. & Rébeillé, F. (2001) "Tetrahydrofolate biosynthesis in plants: molecular and functional characterization of dihydrofolate synthetase and three isoforms of folypolyglutamate synthetase in *Arabidopsis thaliana*" *Proc. Natl. Acad. Sci. USA* 98(26):15360-15365.

20 Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning. A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York), pp 14.18-14.19.

Scholl, T. O. *et al.* (2000) "Folic Acid: Influence on the Outcome of Pregnancy" *Am. J. Clin. Nutr.* 71S:1295S-1303S.

25 Scott, J., Rebeille, F. & Fletcher, J. (2000) "Folic acid and folates: the feasibility for nutritional enhancement in plant foods" *J. Sci. Food Agric.* 80:795-824.

30 Stea, B., Halpern, R. M., Halpern, B. C. & Smith, R. A. (1980) "Quantitative determination of pterins in biological fluids by high-performance liquid chromatography" *J. Chromatogr.* 188(2):363-375.

35 Sybesma, W., Starrenburg, M., Kleerebezem, M., Mierau, I., de Vos, W. M. & Hugenholtz, J. (2003a) "Increased production of folate by metabolic engineering of *Lactococcus lactis*" *Appl. Environ. Microbiol.* 69(6):3069-3076.

40 Sybesma, W., Starrenburg, M., Tijsseling, L., Hoefnagel, M. H. & Hugenholtz, J. (2003b) "Effects of cultivation conditions on folate production by lactic acid bacteria" *Appl. Environ. Microbiol.* 69(8):4542-4548.

Tieman, D. M., Ciardi, J. A., Taylor, M. G. & Klee, H. J. (2001) "Members of the tomato LeEIL (EIN3-like) gene family are functionally redundant and regulate ethylene responses throughout plant development" *Plant J.* 26(1):47-58.

45 Van der Meer, I. M., Bovy, A. G. & Bosch, D. (2001) "Plant-based raw material: improved food quality for better nutrition via plant genomics" *Curr. Opin. Biotechnol.* 12(5):488-492.

Verma, I. C. (1978) *Lancet* 1:879-80.

- Viswanathan, V. K., Green, J. M., Nichols, B. P. (1995) "Kinetic Characterization of 4-Amino-4-Deoxychorismate Synthase from *Escherichia coli*" *Journal of Bacteriology* 177(20):5918-5923.
- 5 Wilson, S. D. & Horne, D. W. (1982) "Use of Glycerol-Cryoprotected *Lactobacillus Casei* for Microbiological Assay of Folic Acid" *Clin. Chem.* 28:1198-1200.
- 10 Win, T. (2000) *Dissertation, Institute of Chemistry and Biology of the Marine Environment* (Universität of Oldenburg, 26111 Oldenburg Germany), pp. 63-66.
- Xu, D., McElroy, D., Thornburg, R. W., Wu, R. *et al.* (1993) "Systemic induction of a potato pin2 promoter by wounding, methyl jasmonate, and abscisic acid in transgenic rice plants" *Plant Molecular Biology* 22:573-588.
- 15 Yang, T. T. *et al.* (1996) "Optimized Codon Usage and Chromophore Mutations Provide Enhanced Sensitivity with the Green Fluorescent Protein" *Nucleic Acid Research* 24(22):4592-4593.
- 20 Yim, J. J. *et al.* (1976) "Characteristics of Guanosine Triphosphate Cyclohydrolase I Purified From *Escherichia coli*" *J. Biol. Chem.* 251:5087-5094.
- 25 Yoneyama, T. & Hatakeyama, K. (1998) "Decameric GTP cyclohydrolase I forms complexes with two pentameric GTP cyclohydrolase I feedback regulatory proteins in the presence of phenylalanine or of a combination of tetrahydrobiopterin and GTP" *J. Biol. Chem.* 273(32):20102-20108.
- 30 Zhang, G. F., Maudens, K. E., Storozhenko, S., Mortier, K. A., Van Der Straeten, D. & Lambert, W. E. (2003) "Determination of total folate in plant material by chemical conversion into para-aminobenzoic acid followed by high performance liquid chromatography combined with on-line postcolumn derivatization and fluorescence detection" *J. Agric. Food Chem.* 51(27):7872-7878.

CLAIMS

We claim:

- 5 1. A method for increasing levels of a folate in a plant, said method comprising incorporating in a plant a polynucleotide that encodes an enzyme that catalyzes the synthesis of dihydroneopterin-PPP from GTP, wherein said enzyme is not under feedback control of the plant.
- 10 2. The method according to claim 1, wherein said enzyme is a GTP cyclohydrolase I (GCHI) enzyme.
3. The method according to claim 2, wherein said GCHI enzyme is a mammalian or a bacterial GCHI enzyme.
- 15 4. The method according to claim 3, wherein said bacterial GCHI enzyme is from *E. coli*.
5. The method according to claim 3, wherein said mammalian GCHI enzyme comprises
20 the amino acid sequence shown in SEQ ID NO: 2 or SEQ ID NO: 4, or an enzymatically functional fragment of SEQ ID NO: 2 or SEQ ID NO: 4.
6. The method according to claim 3, wherein said mammalian GCHI enzyme is encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID
25 NO: 3, or a nucleotide sequence that hybridizes under stringent conditions with the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3.
7. The method according to claim 3, wherein said bacterial GCHI enzyme comprises the amino acid sequence shown in SEQ ID NO: 9, or an enzymatically functional fragment of
30 SEQ ID NO: 9.

8. The method according to claim 3, wherein said bacterial GCHI enzyme is encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 8, or a nucleotide sequence that hybridizes under stringent conditions with the nucleotide sequence shown in SEQ ID NO: 8.

5

9. The method according to claim 1, wherein the plant is a monocotyledonous plant.

10. The method according to claim 9, wherein the monocotyledonous plant is selected from the group consisting of rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, turfgrasses, and millet.

10

11. The method according to claim 1, wherein the plant is a dicotyledonous plant.

12. The method according to claim 11, wherein the dicotyledonous plant is selected from the group consisting of tomato, cucumber, squash, peas, beans, peanuts, spinach, broccoli, alfalfa, melon, chickpea, chicory, clover, kale, lentil, soybean, tobacco, potato, sweet potato, yams, cassava, radish, cabbage, rape, apple trees, citrus (including oranges, mandarins, grapefruit, lemons, and limes), grape, cotton, sunflower, strawberry, and lettuce.

15

20 13. The method according to claim 1, wherein the plant is a fungus.

14. The method according to claim 1, wherein the plant is an alga.

15. The method according to claim 1, wherein the folate is tetrahydrofolate.

25

16. The method according to claim 1, wherein the plant is grown in the presence of p-aminobenzoate (PABA).

17. The method according to claim 1, wherein the plant is engineered to express increased levels of PABA.

30

18. The method according to claim 17, wherein the plant is transgenic for one or more polynucleotide molecules encoding one or more enzymes that catalyze synthesis of PABA.

19. The method according to claim 17, wherein the plant is transformed with one or more polynucleotide molecules encoding one or more enzymes that catalyze synthesis of PABA.

20. The method according to claim 18 or 19, wherein said one or more enzymes is 4-amino-4-deoxychorismate synthase or 4-amino-4-deoxychorismate lyase.

21. A plant, plant tissue, or plant cell that synthesizes, or is capable of synthesizing, increased levels of a folate, said plant, plant tissue, or plant cell comprising a polynucleotide that encodes an enzyme that catalyzes the synthesis of dihydroneopterin-PPP from GTP, wherein said enzyme is not under feedback control of the plant.

22. The plant, plant tissue, or plant cell according to claim 21, wherein said enzyme is a GTP cyclohydrolase I (GCHI) enzyme.

23. The plant, plant tissue, or plant cell according to claim 22, wherein said GCHI enzyme is a mammalian or a bacterial GCHI enzyme.

24. The plant, plant tissue, or plant cell according to claim 23, wherein said bacterial GCHI enzyme is from *E. coli*.

25. The plant, plant tissue, or plant cell according to claim 23, wherein said mammalian GCHI enzyme comprises the amino acid sequence shown in SEQ ID NO: 2 or SEQ ID NO: 4, or an enzymatically functional fragment of SEQ ID NO: 2 or SEQ ID NO: 4.

26. The plant, plant tissue, or plant cell according to claim 23, wherein said mammalian GCHI enzyme is encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3, or a nucleotide sequence that hybridizes under stringent conditions with the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3.

27. The plant, plant tissue, or plant cell according to claim 23, wherein said bacterial GCHI enzyme comprises the amino acid sequence shown in SEQ ID NO: 9, or an enzymatically functional fragment of SEQ ID NO: 9.

28. The plant, plant tissue, or plant cell according to claim 23, wherein said bacterial GCHI enzyme is encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 8, or a nucleotide sequence that hybridizes under stringent conditions with the
5 nucleotide sequence shown in SEQ ID NO: 8.

29. The plant, plant tissue, or plant cell according to claim 21, wherein said plant is a monocotyledonous plant or said plant tissue or plant cell is from a monocotyledonous plant.

10 30. The plant, plant tissue, or plant cell according to claim 29, wherein said monocotyledonous plant is selected from the group consisting of rice, wheat, barley, oats, rye, sorghum, maize, sugarcane, pineapple, onion, bananas, coconut, lilies, turfgrasses, and millet.

15 31. The plant, plant tissue, or plant cell according to claim 21, wherein said plant is a dicotyledonous plant or said plant tissue or plant cell is from a dicotyledonous plant.

32. The plant, plant tissue, or plant cell according to claim 31, wherein said dicotyledonous plant is selected from the group consisting of tomato, cucumber, squash, peas,
20 beans, peanuts, spinach, broccoli, alfalfa, melon, chickpea, chicory, clover, kale, lentil, soybean, tobacco, potato, sweet potato, yams, cassava, radish, cabbage, rape, apple trees, citrus (including oranges, mandarins, grapefruit, lemons, and limes), grape, cotton, sunflower, strawberry, and lettuce.

25 33. The plant, plant tissue, or plant cell according to claim 21, wherein the plant is a fungus.

34. The plant, plant tissue, or plant cell according to claim 21, wherein the plant is an alga.

30 35. The plant, plant tissue, or plant cell according to claim 21, wherein the folate is tetrahydrofolate.

36. The plant, plant tissue, or plant cell according to claim 21, wherein the plant is grown in the presence of PABA.

37. The plant, plant tissue, or plant cell according to claim 21, wherein the plant, plant
5 tissue, or plant cell is engineered to express increased levels of PABA.

38. The plant, plant tissue, or plant cell according to claim 21, wherein the plant is transgenic for one or more polynucleotide molecules encoding one or more enzymes that catalyze synthesis of PABA.

10

39. The plant, plant tissue, or plant cell according to claim 36, wherein the plant, plant tissue, or plant cell is transformed with one or more polynucleotide molecules encoding one or more enzymes that catalyze synthesis of PABA.

15

40. The plant, plant tissue, or plant cell according to claim 38 or 39, wherein said one or more enzymes is 4-amino-4-deoxychorismate synthase or 4-amino-4-deoxychorismate lyase.

20

41. A polynucleotide comprising a nucleotide sequence encoding a GCHI enzyme comprising the amino acid sequence shown in SEQ ID No: 2 or SEQ ID NO: 4, or an enzymatically functional fragment of SEQ ID NO: 2 or SEQ ID NO: 4.

25

42. The polynucleotide according to claim 41, wherein said polynucleotide comprises the nucleotide sequence shown in SEQ ID No: 1 or SEQ ID NO: 3, or a nucleotide sequence that hybridizes under stringent conditions with the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3.

43. An expression construct comprising a polynucleotide of claim 41.

30

44. The expression construct according to claim 43, wherein said polynucleotide comprises the nucleotide sequence shown in SEQ ID No: 1 or SEQ ID NO: 3, or a nucleotide sequence that hybridizes under stringent conditions with the nucleotide sequence shown in SEQ ID NO: 1 or SEQ ID NO: 3.

1/26

FIG. 1A

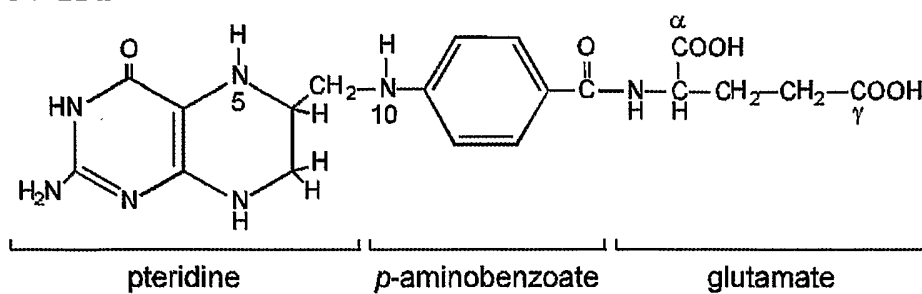


FIG. 1B

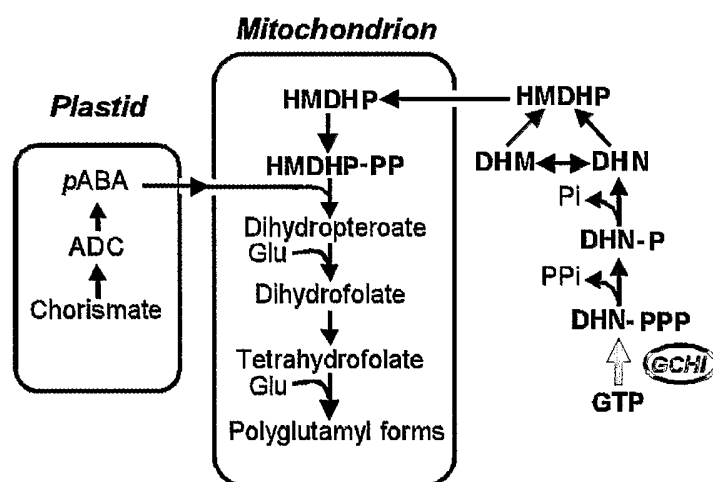


FIG. 2A

atggagaagcc**aag**aggagtcaggtgcaccaatgggttctctccgag**ag**agagct**t**
M E K P R G V R C T N G F S E R E L
cc**aag**acc**agg**agc**ttct**ccacctgctgagaagtc**tag**accacctgaggc**t**
P R P G A S P P A E K S R P P E A

FIG. 2B

V2			M9			M18			M67			M75		
Br	R	RR	Br	R	RR	Br	R	RR	Br	R	RR	Br	R	RR

FIG. 2C

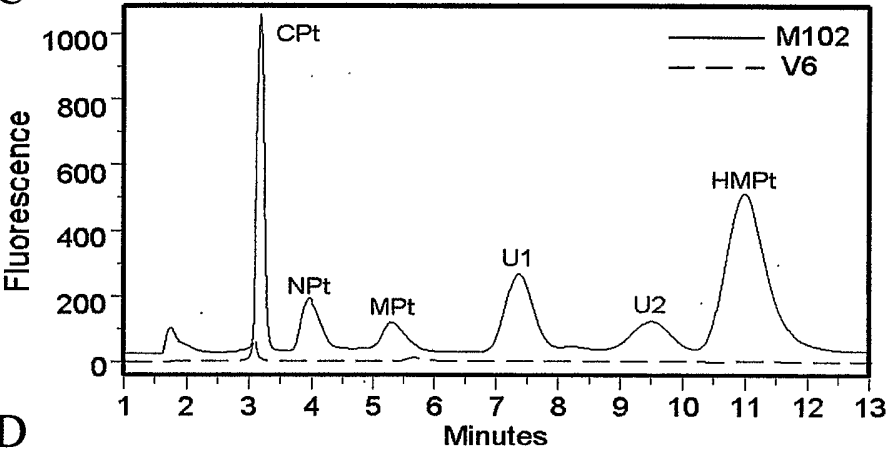
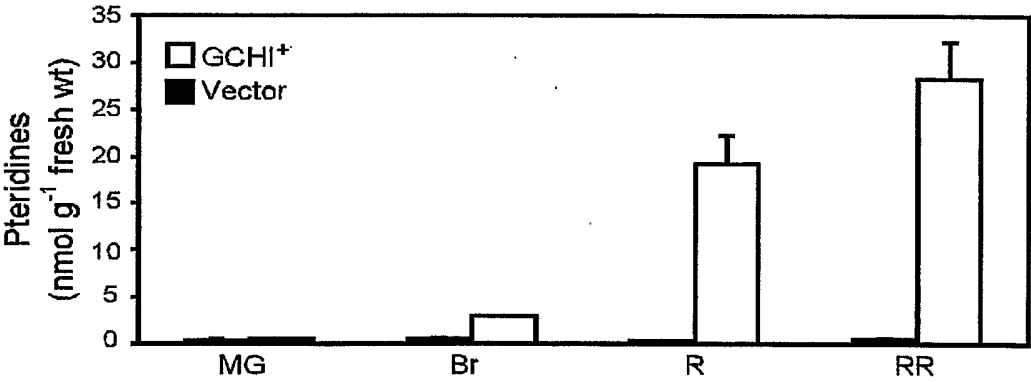


FIG. 2D



3/26

FIG. 3A

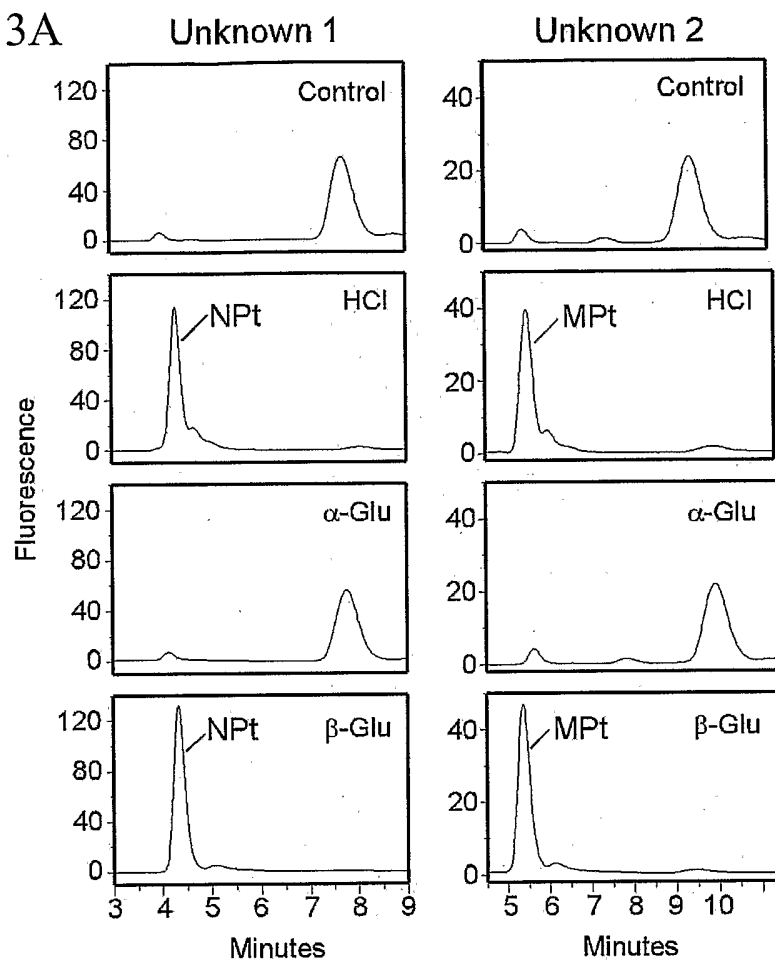
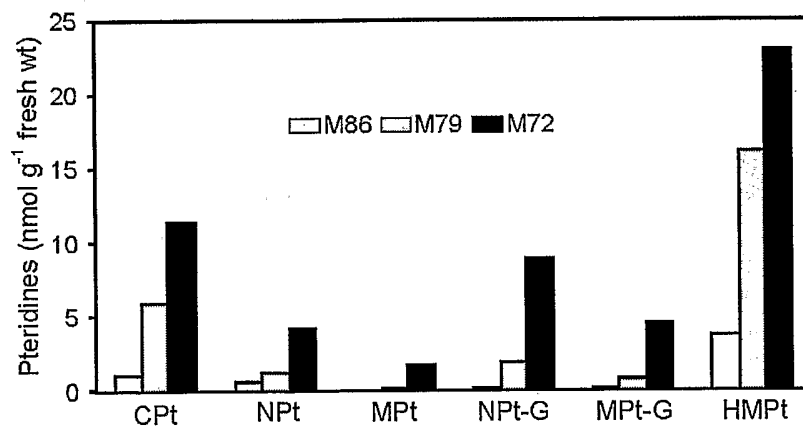


FIG. 3B



4/26

FIG. 4A

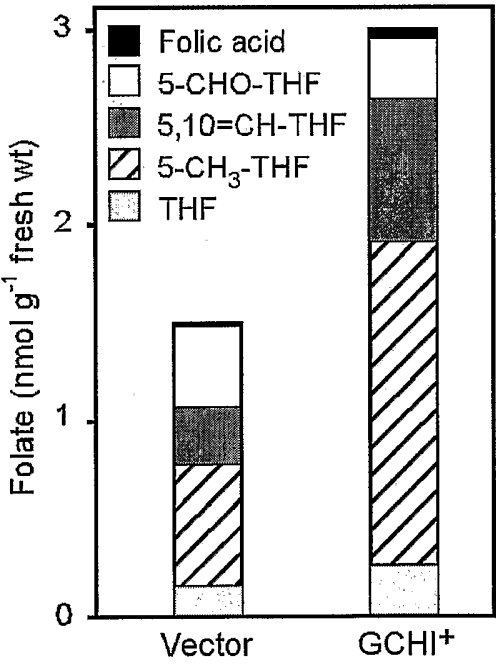
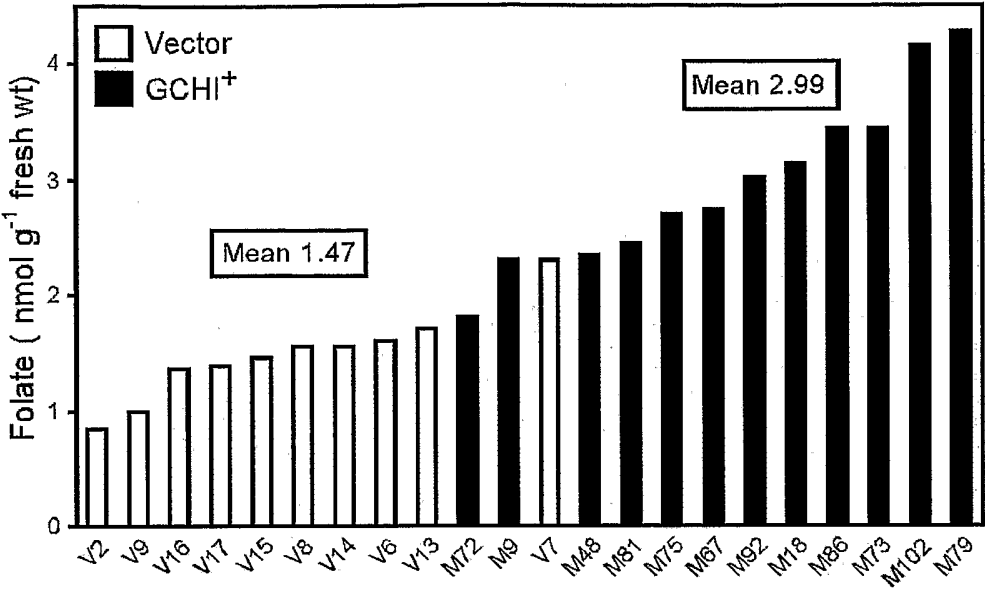


FIG. 4B

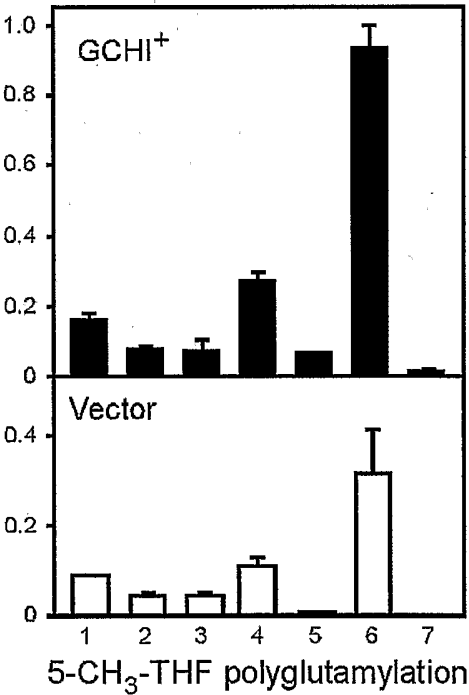


FIG. 4C

5/26

FIG. 5A

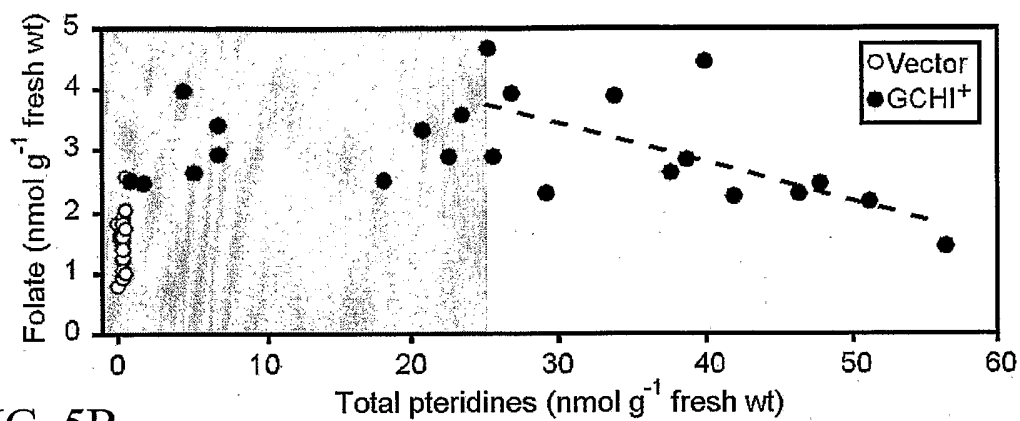


FIG. 6A



FIG. 6B



FIG. 6C

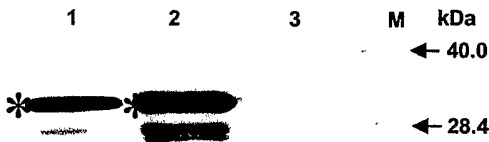


FIG. 6D



FIG. 6E



7/26

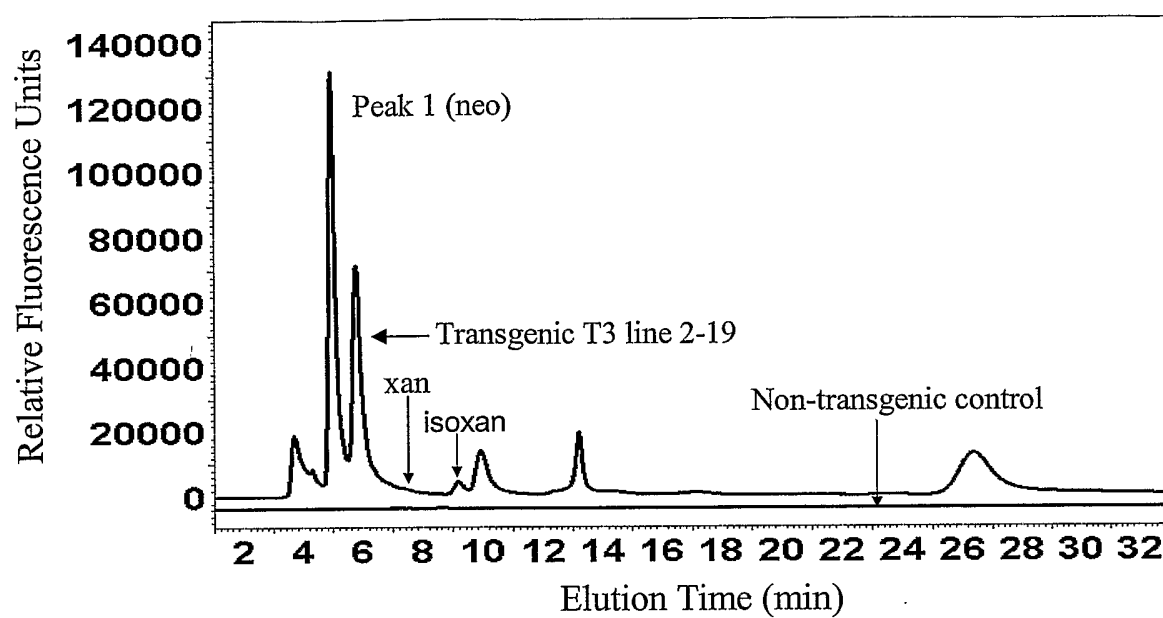


FIG. 7A

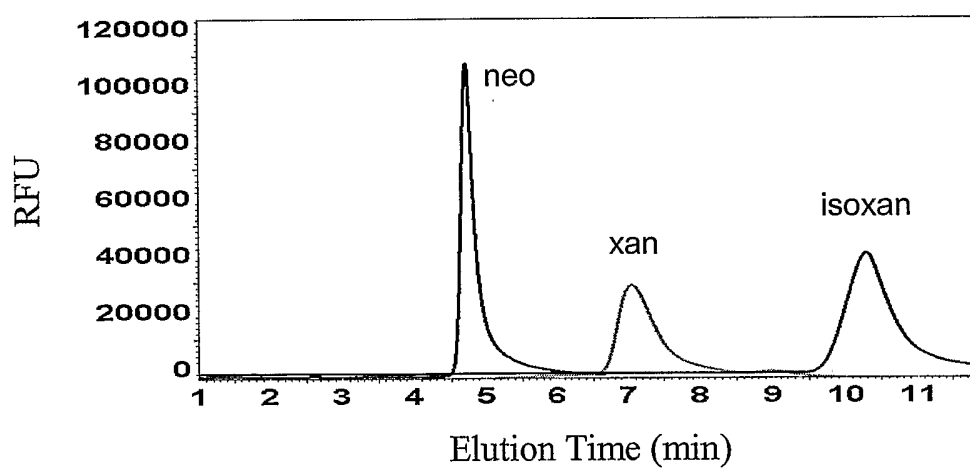


FIG. 7B

8/26

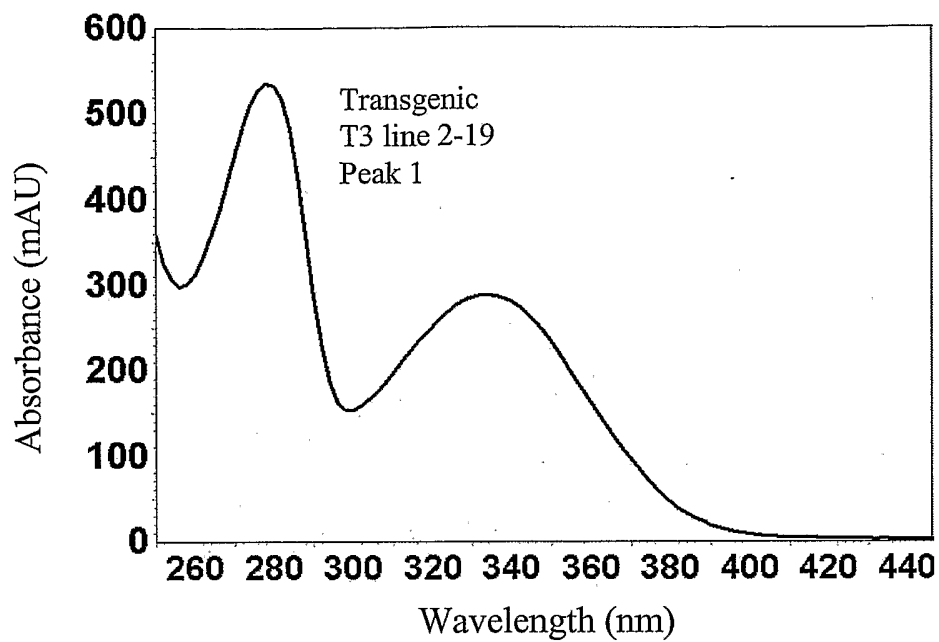


FIG. 7C

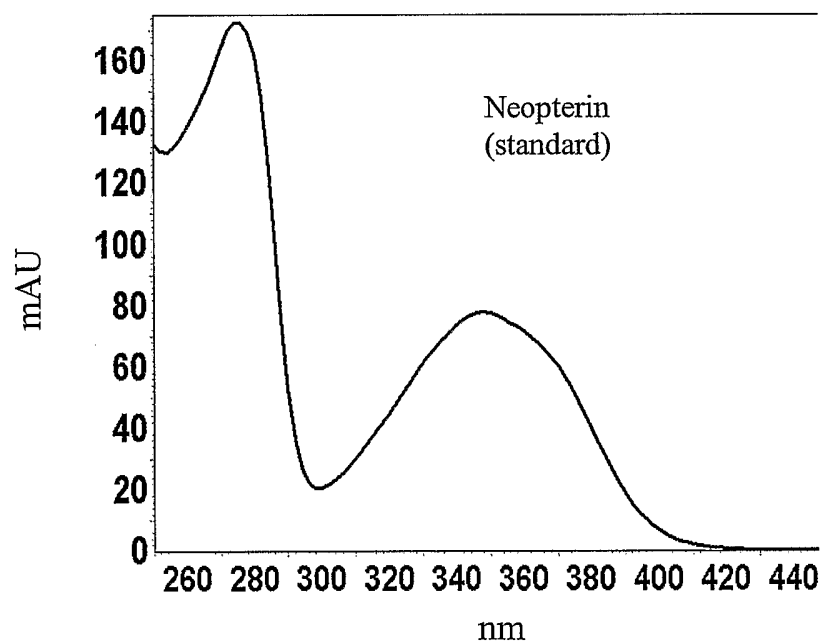


FIG. 7D

9/26

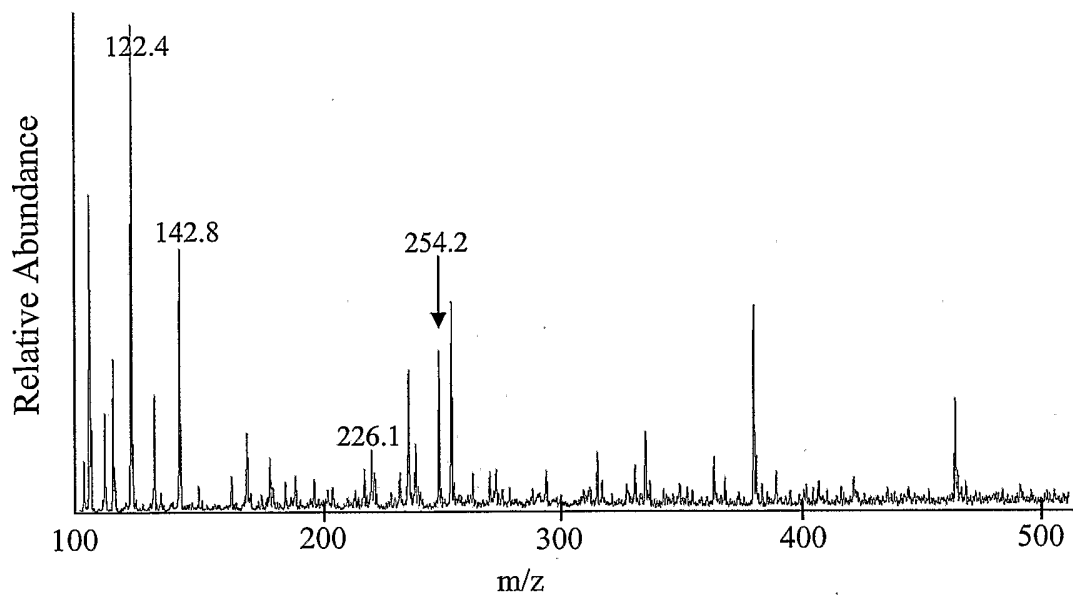


FIG. 7E

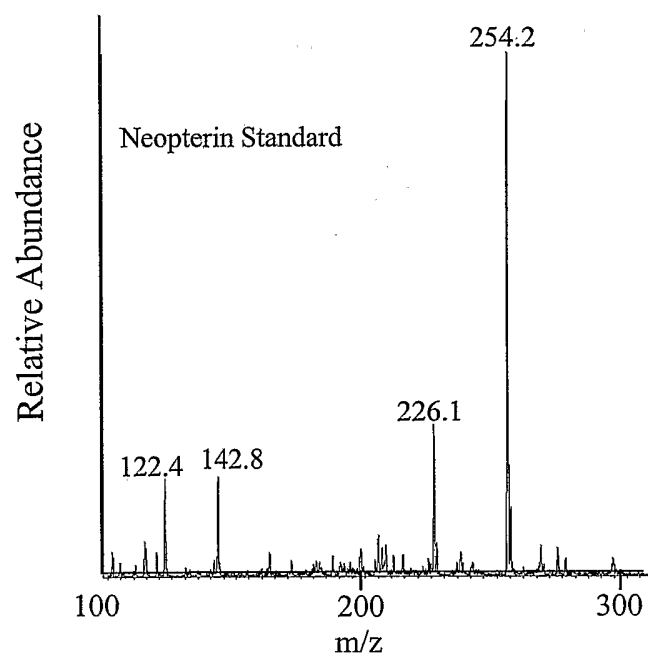


FIG. 7F

10/26

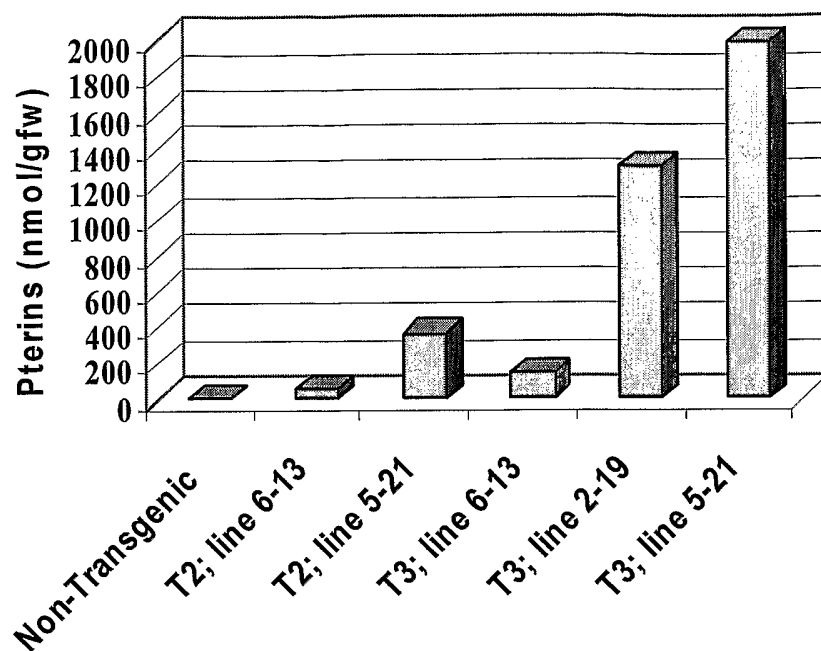


FIG. 8A

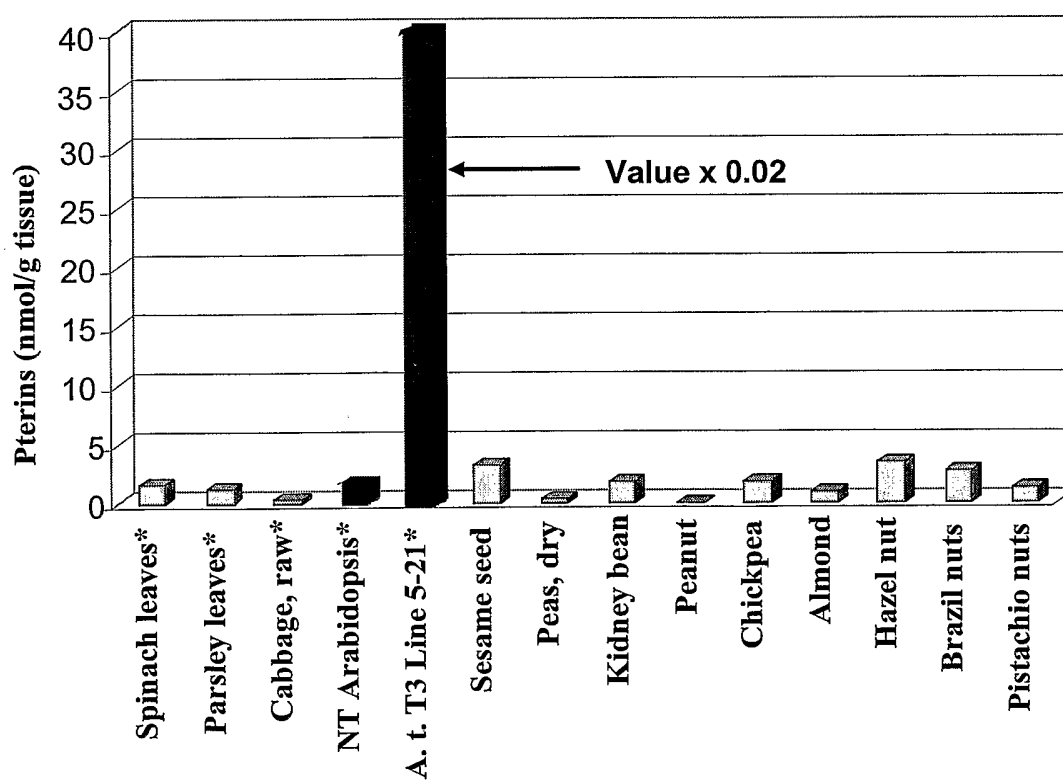


FIG. 8B

11/26

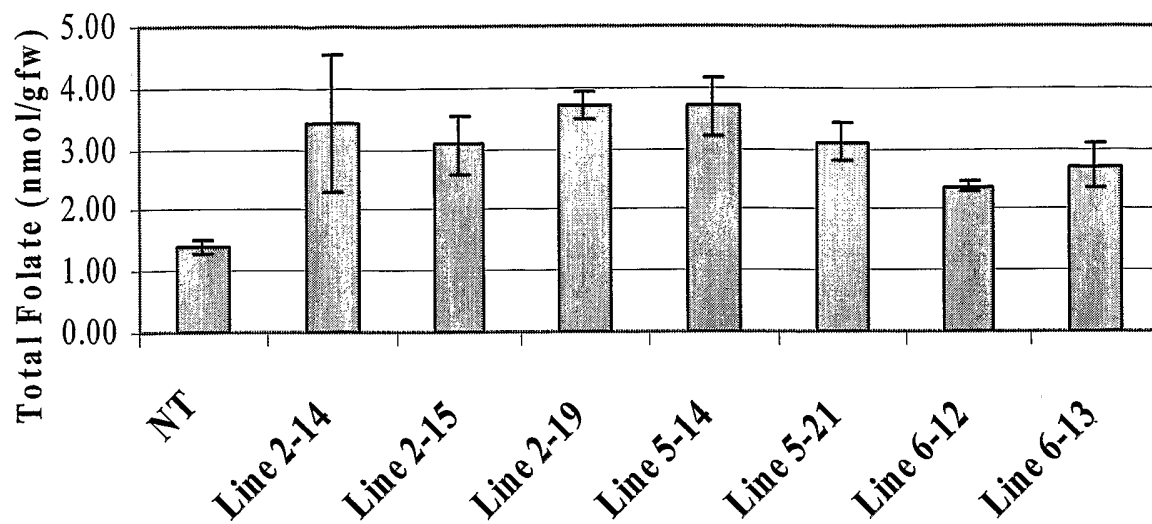


FIG. 9A

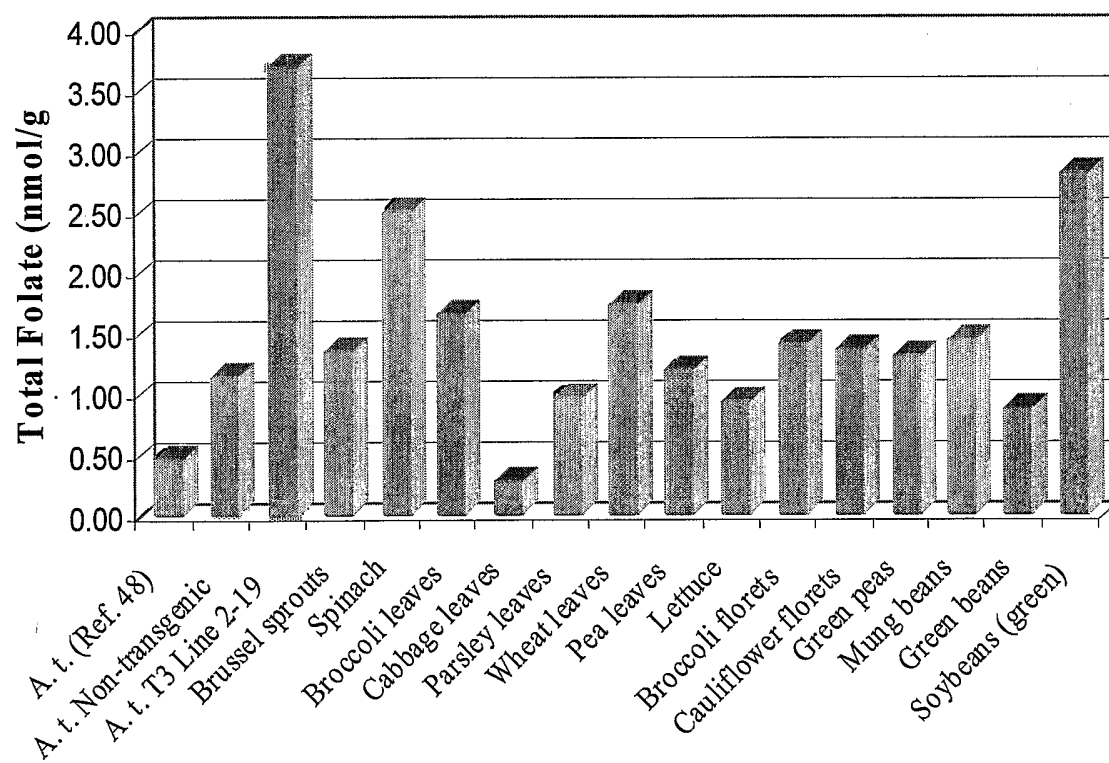


FIG. 9B

12/26

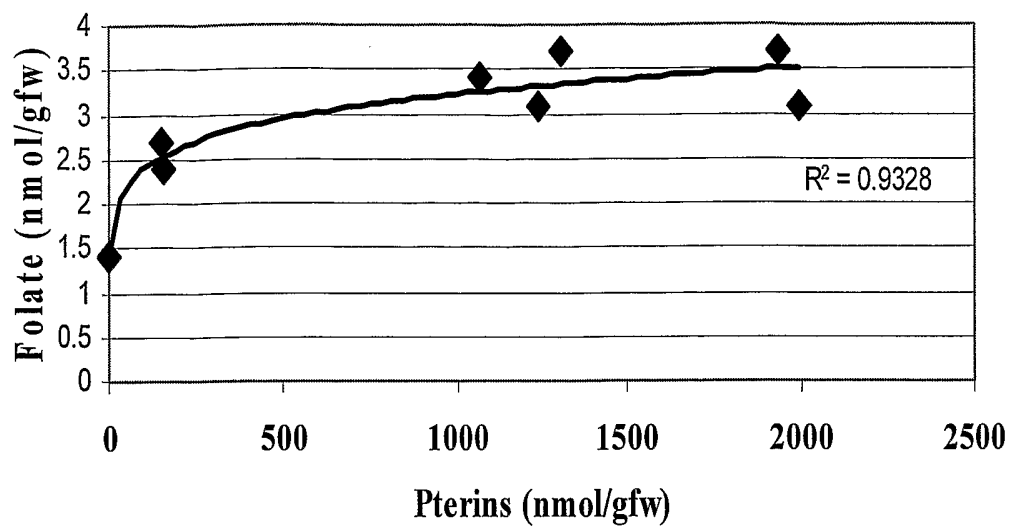


FIG. 9C

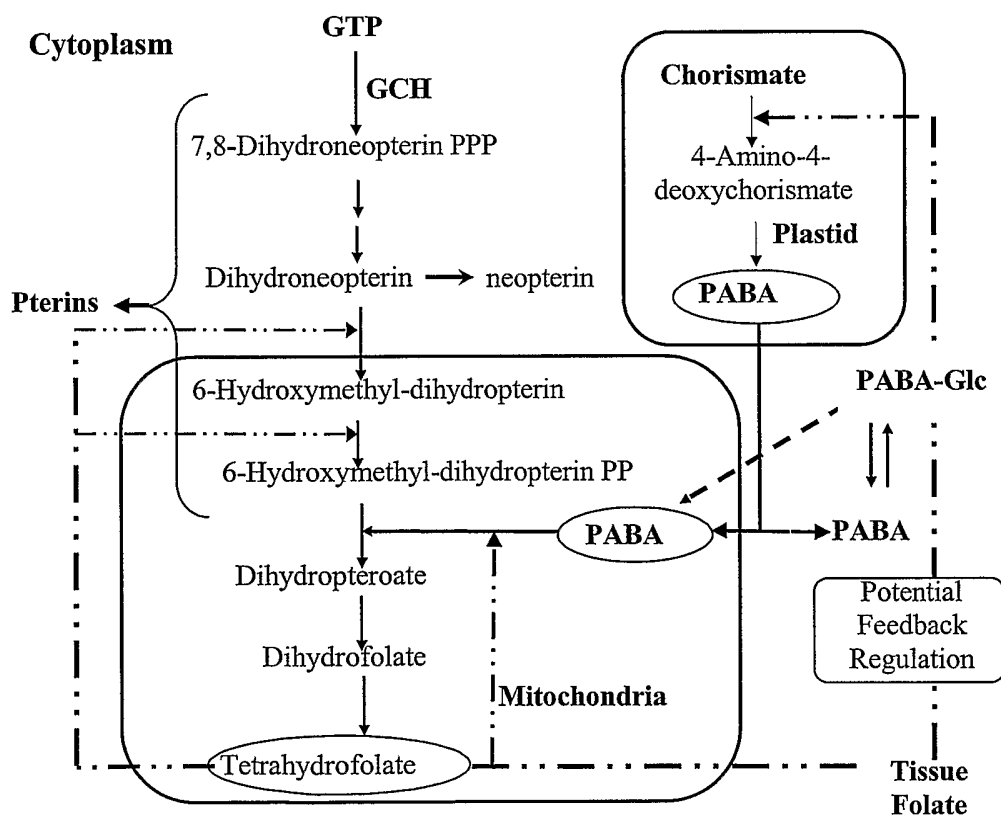


FIG. 10

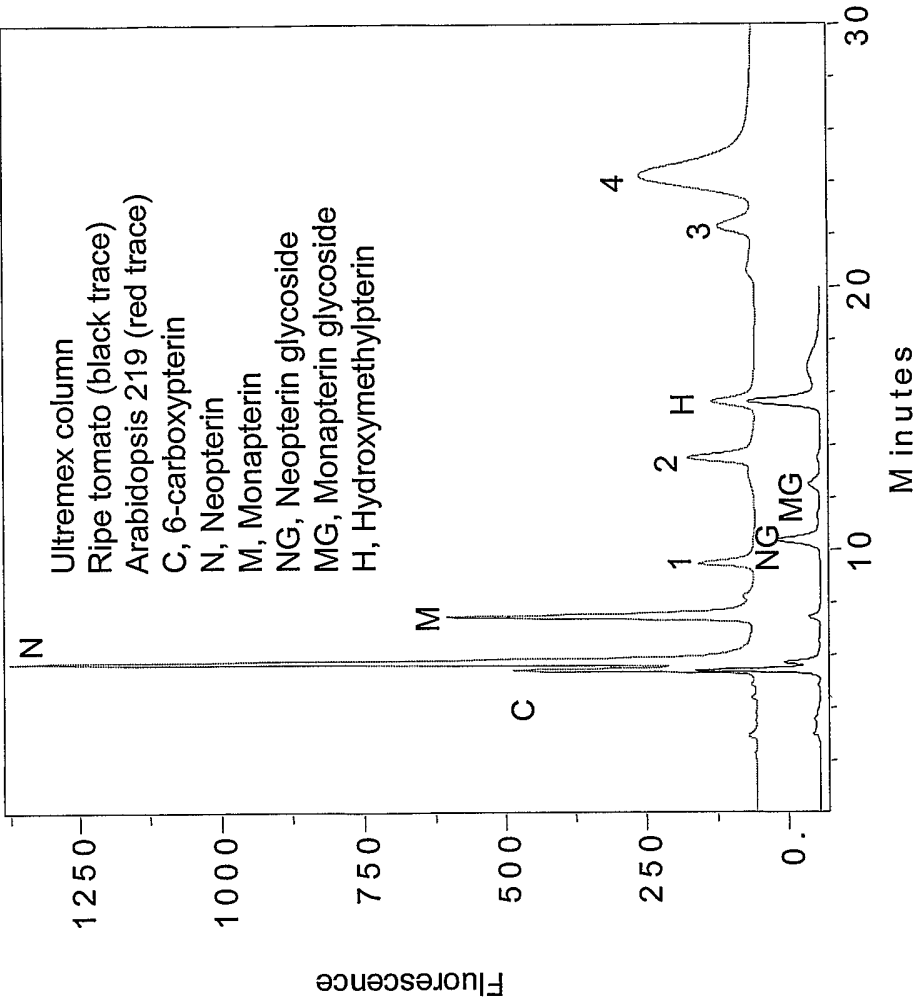


FIG.11

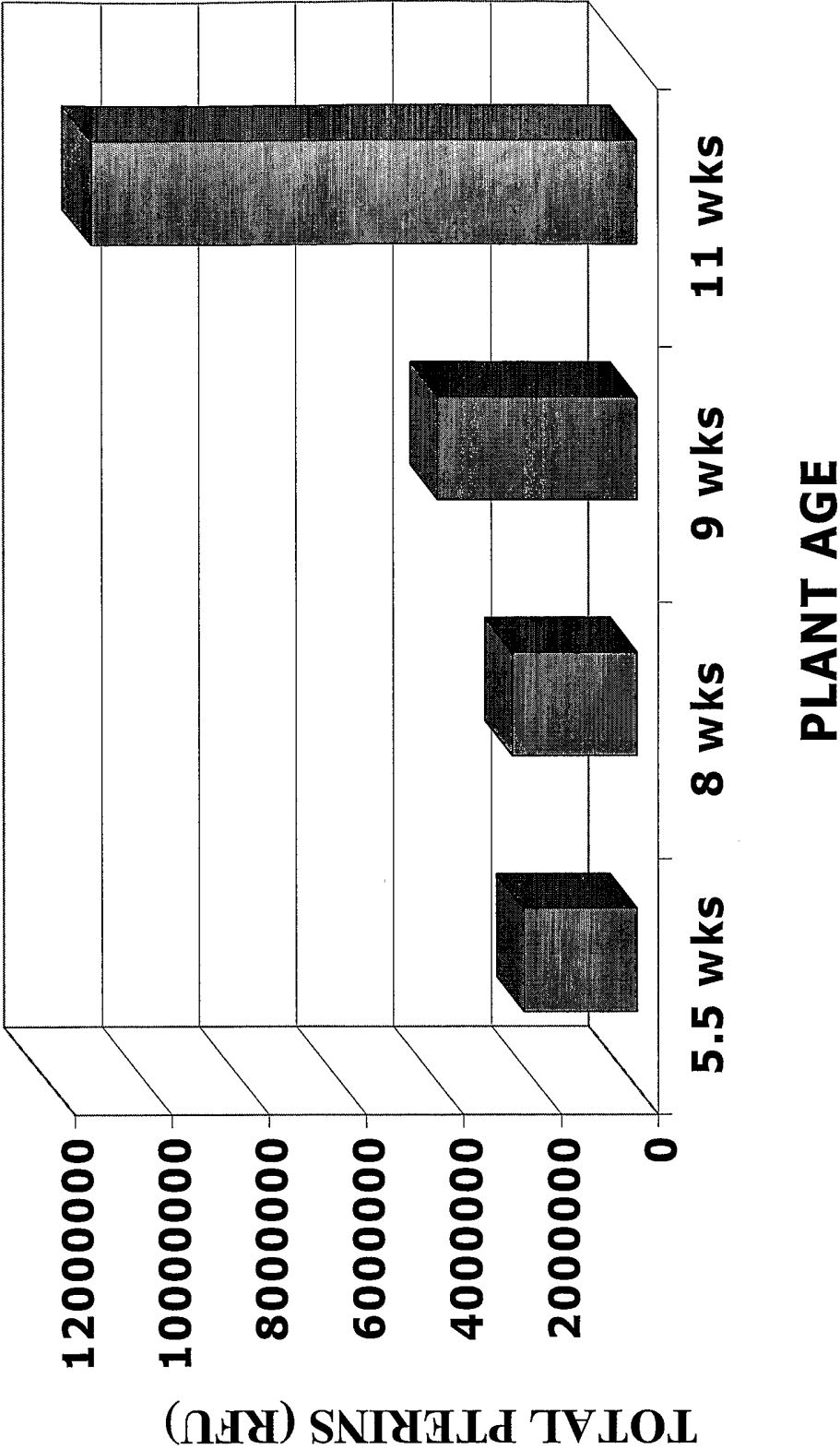


FIG. 12

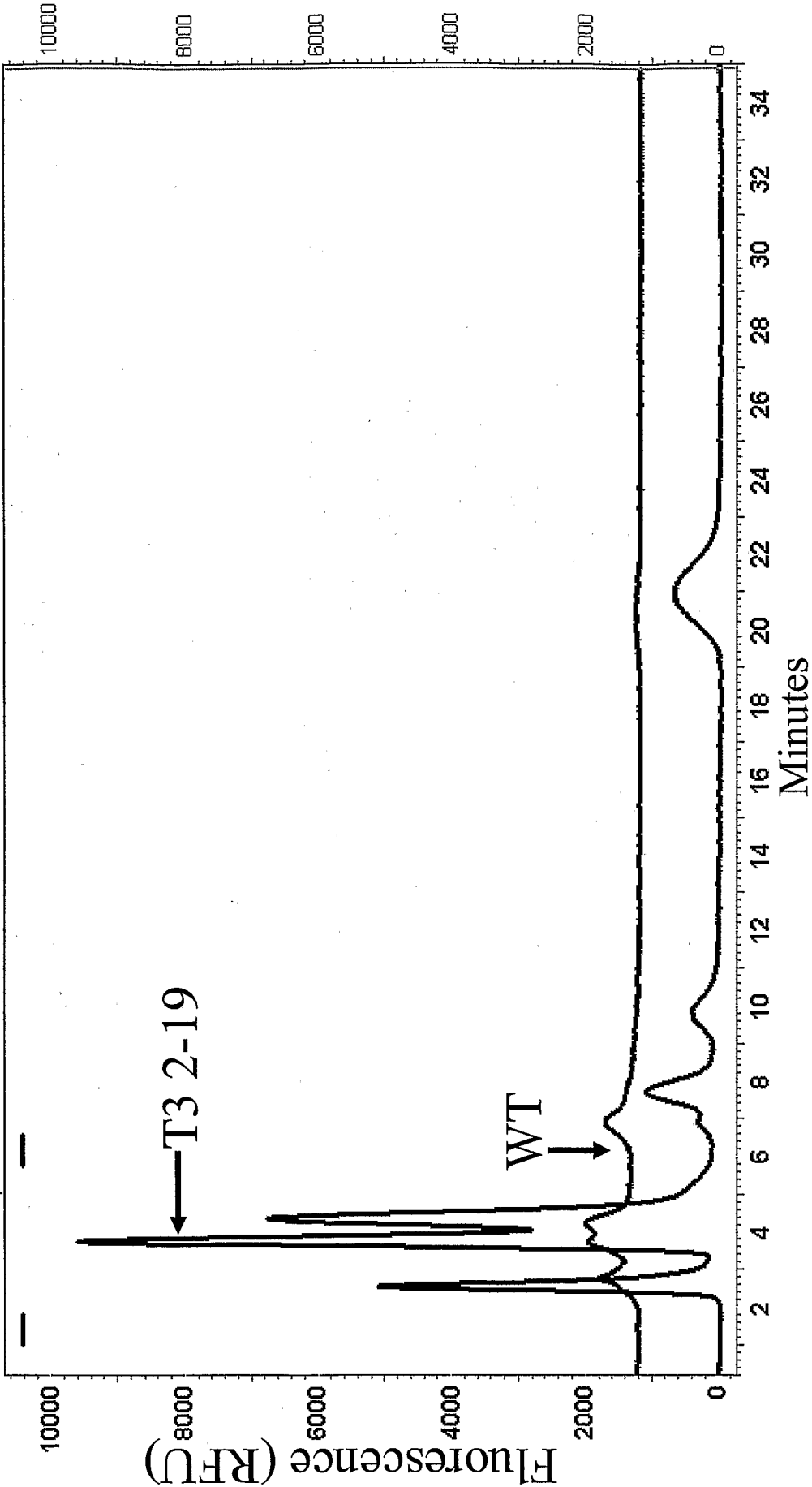


FIG. 13

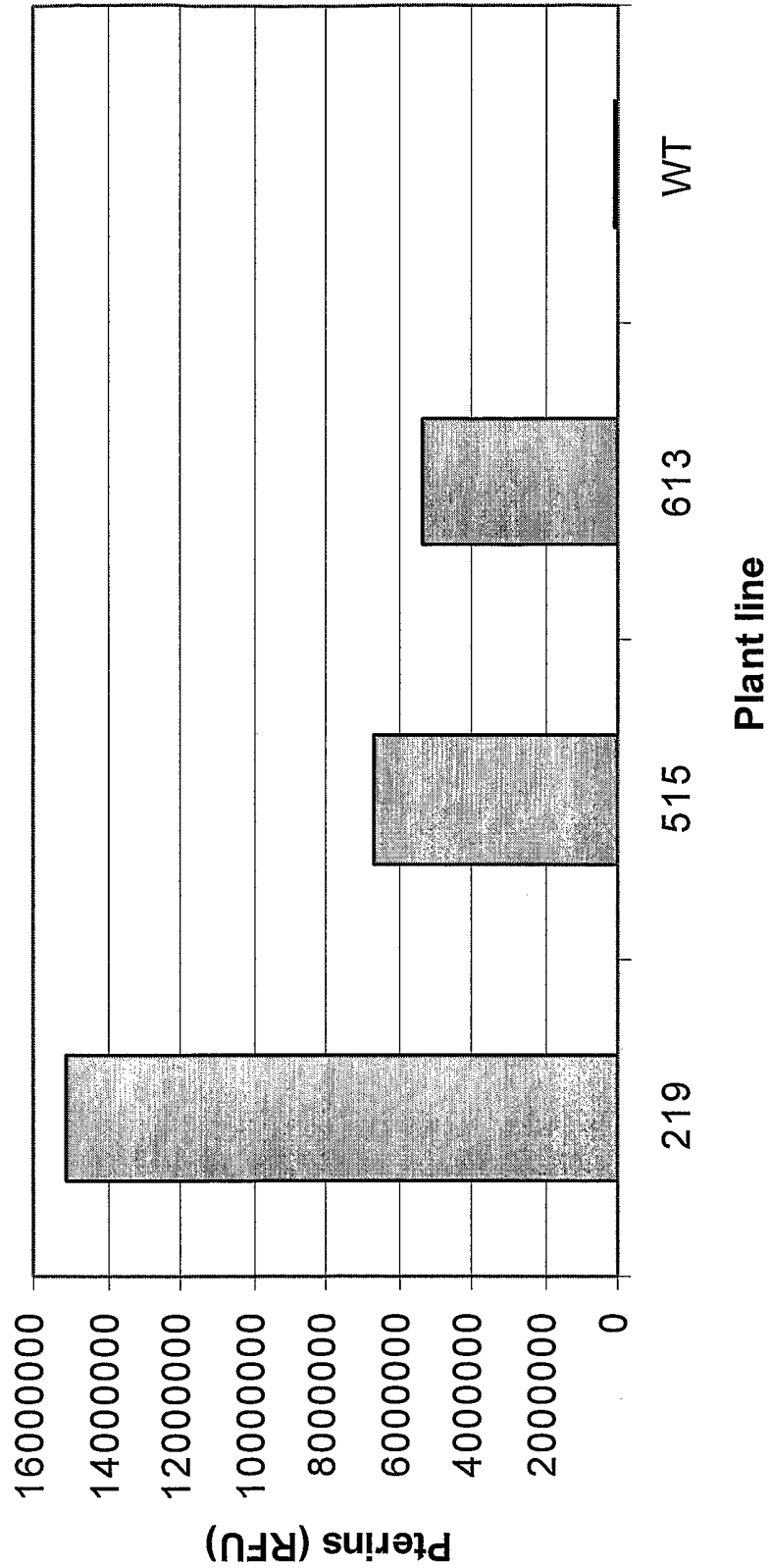


FIG. 14

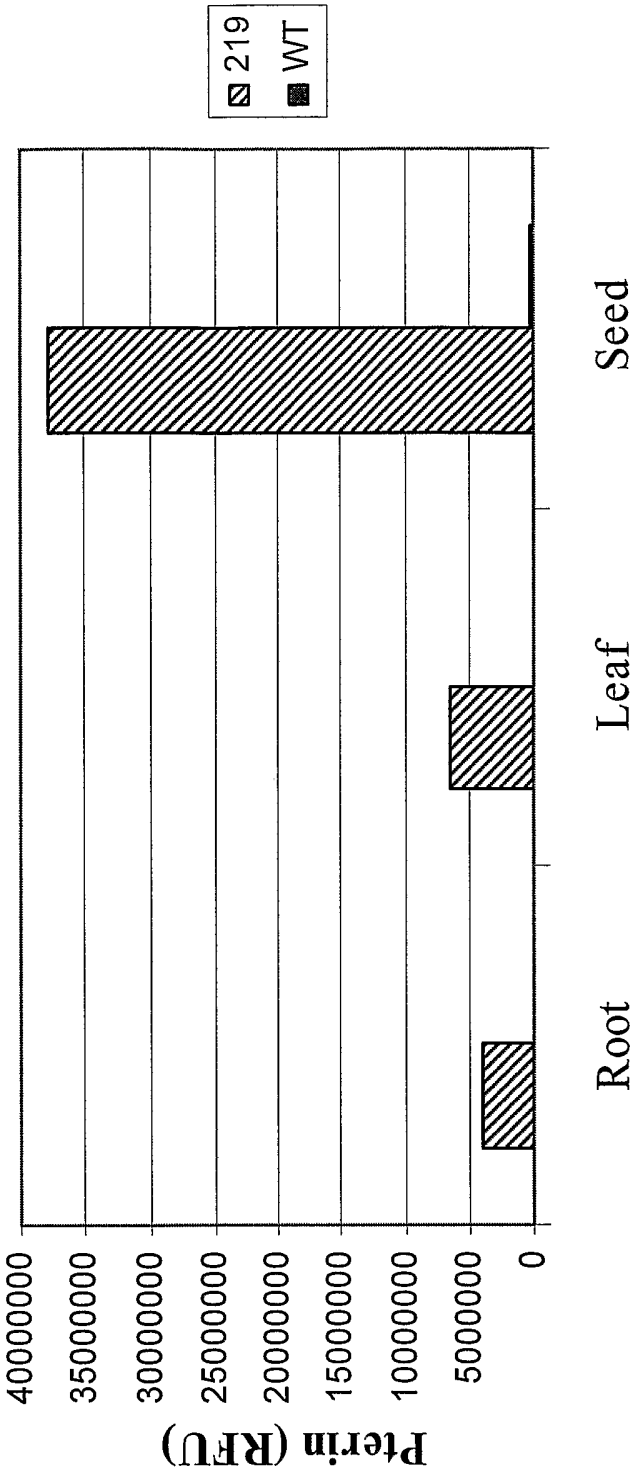
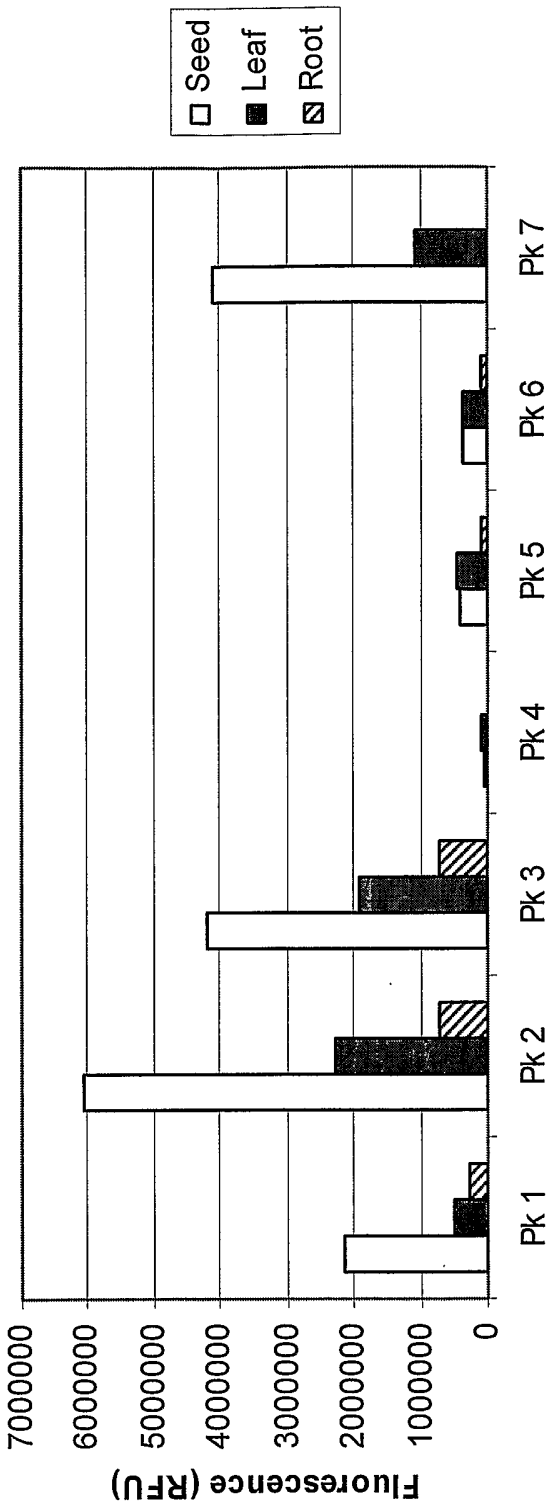


FIG. 15



Peak 1: 6-carboxypterin
Peak 2: neopterin
Peak 3: monapterin
Peak 4: Unknown 1
Peak 5: Unknown 2
Peak 6: Hydroxymethylpterin
Peak 7: Unknown 3

FIG. 16

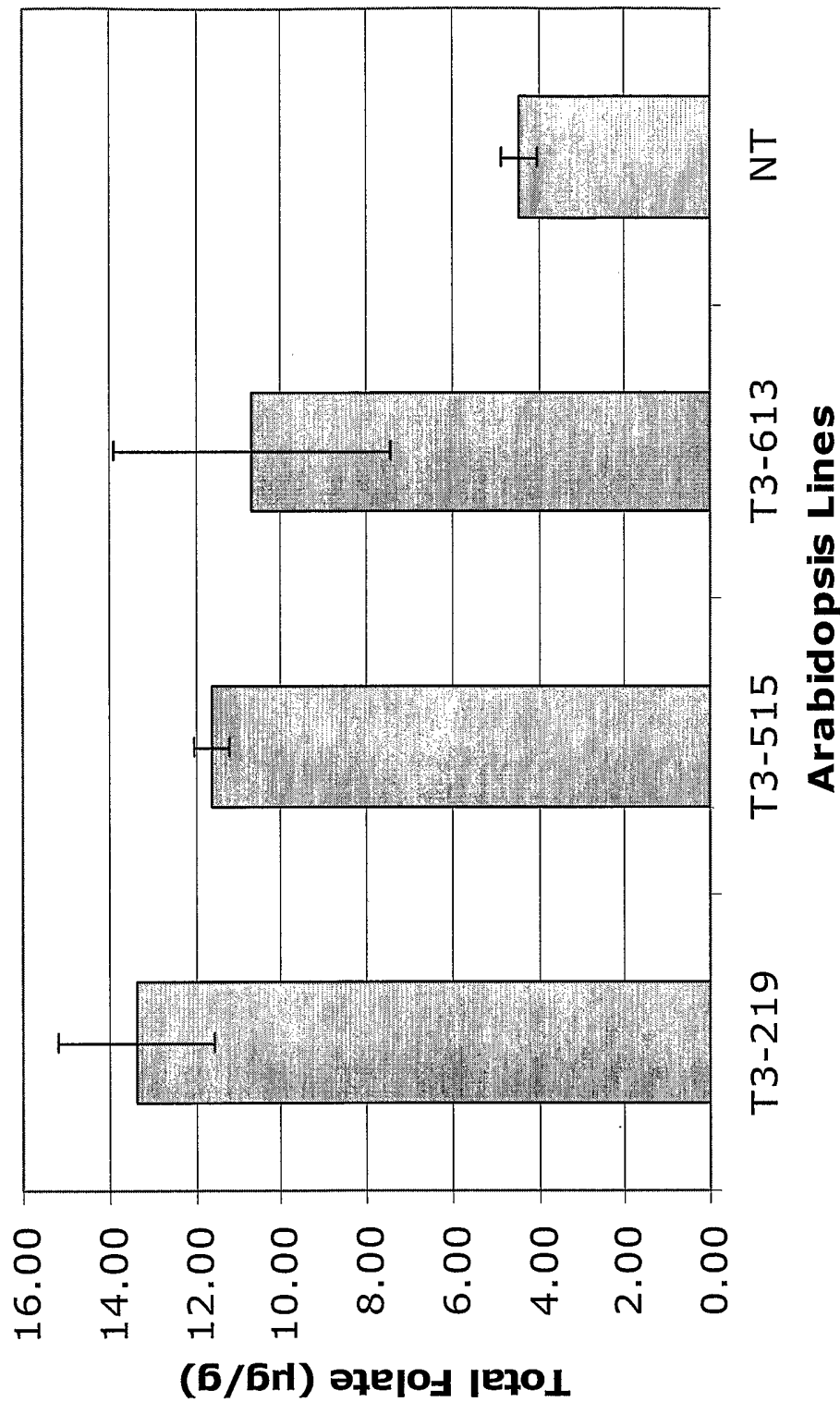
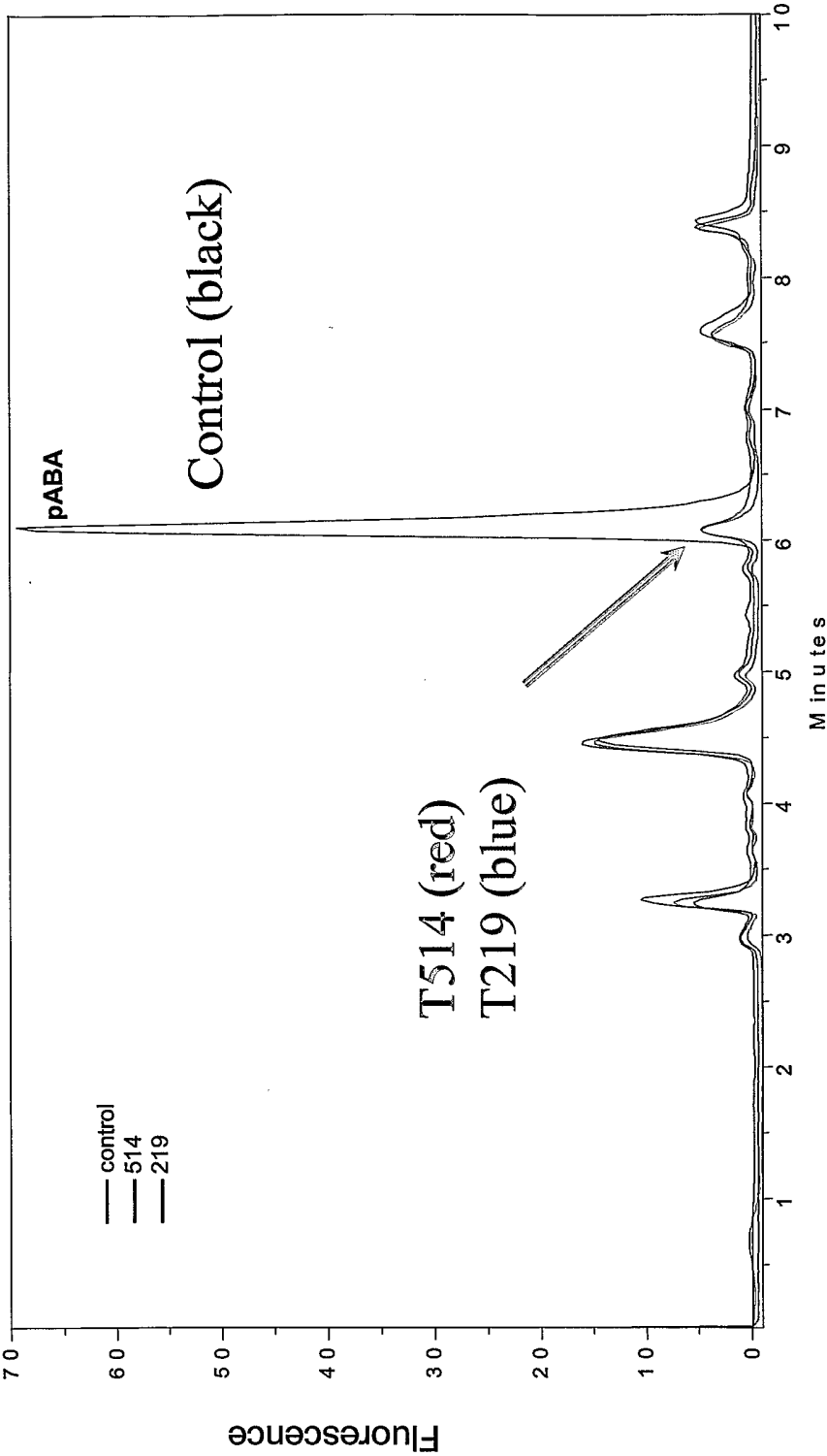


FIG. 17

FIG. 18



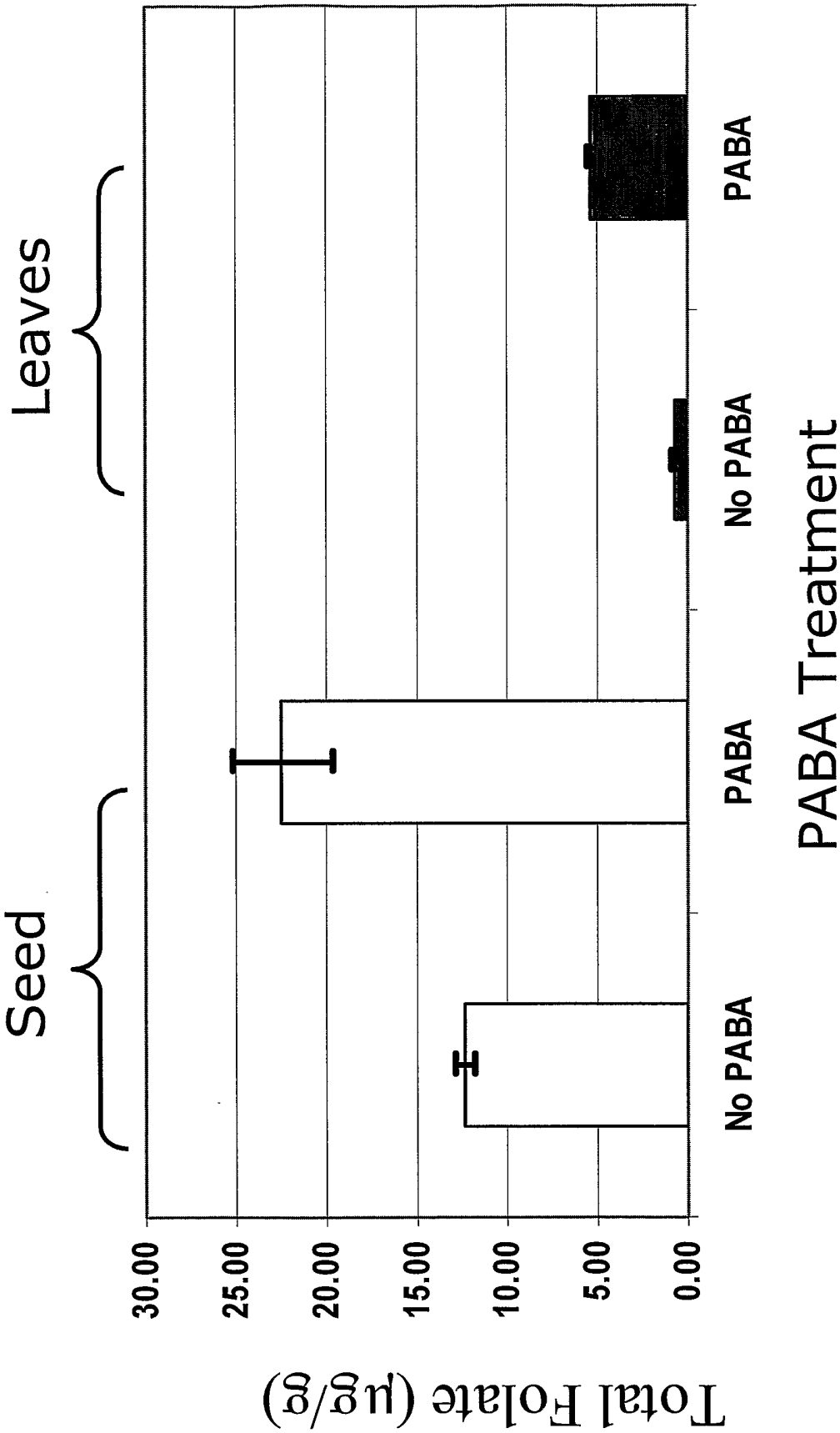
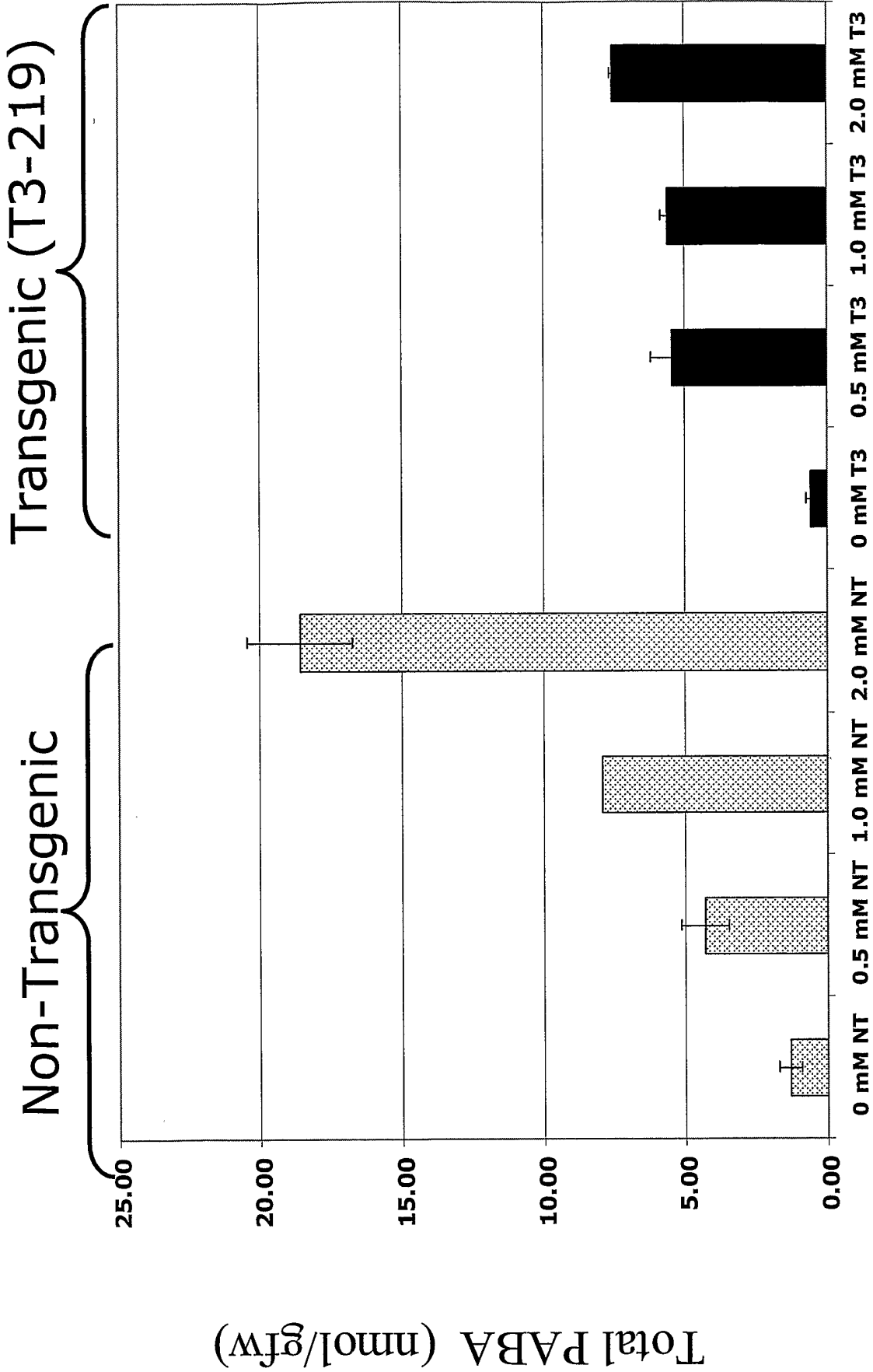


FIG. 19



PABA TREATMENT
FIG. 20A

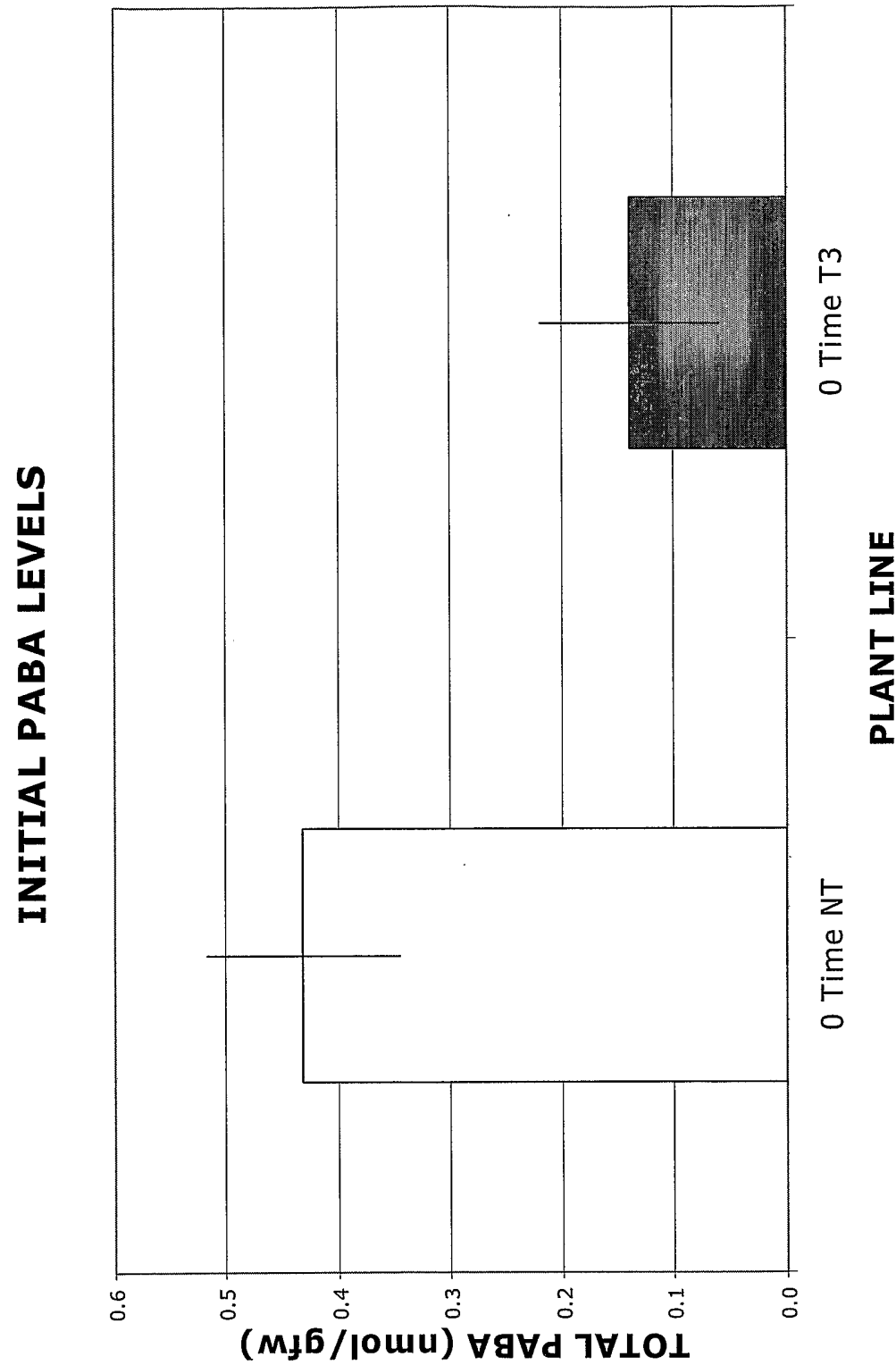


FIG. 20B

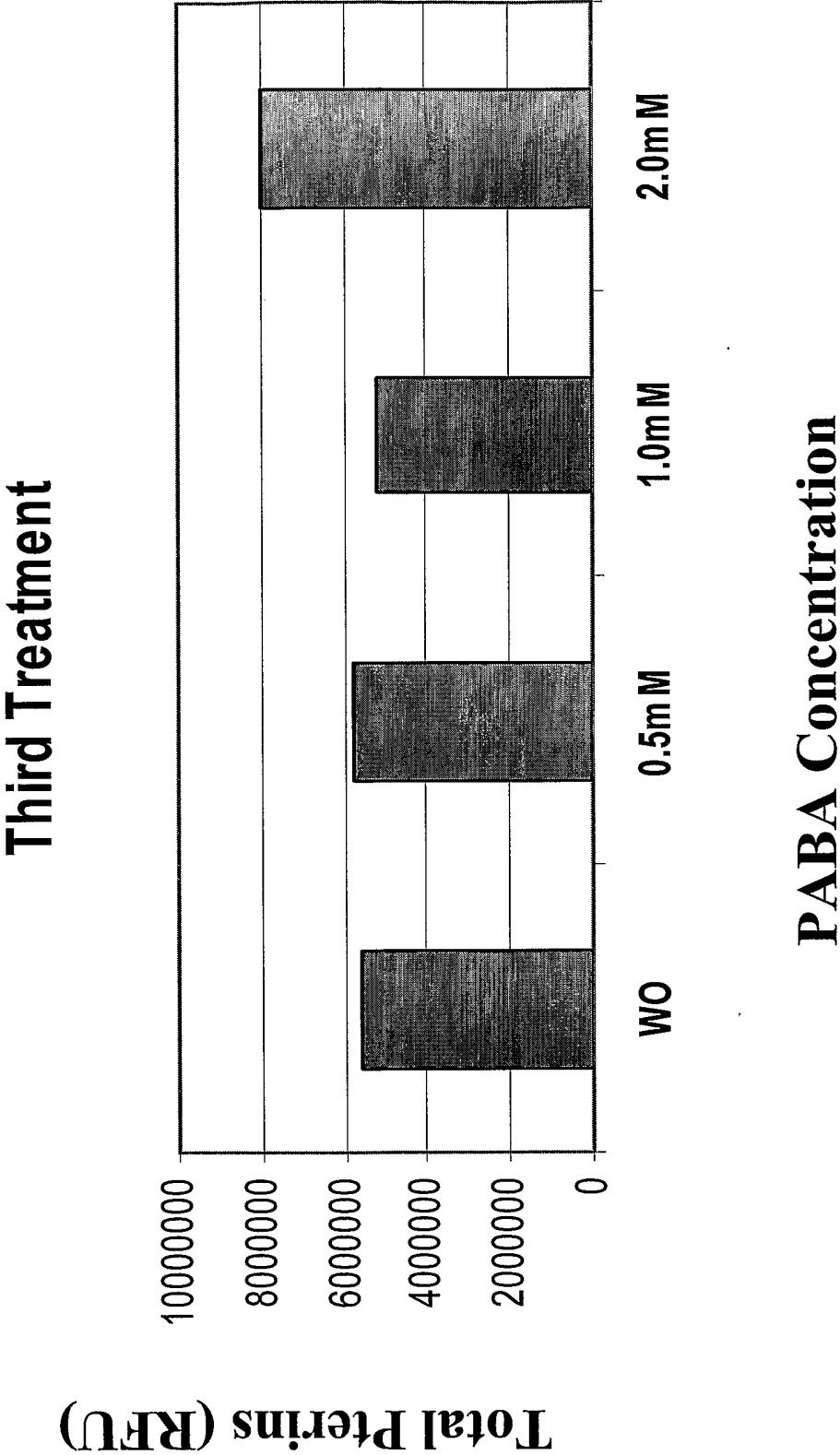


FIG. 21

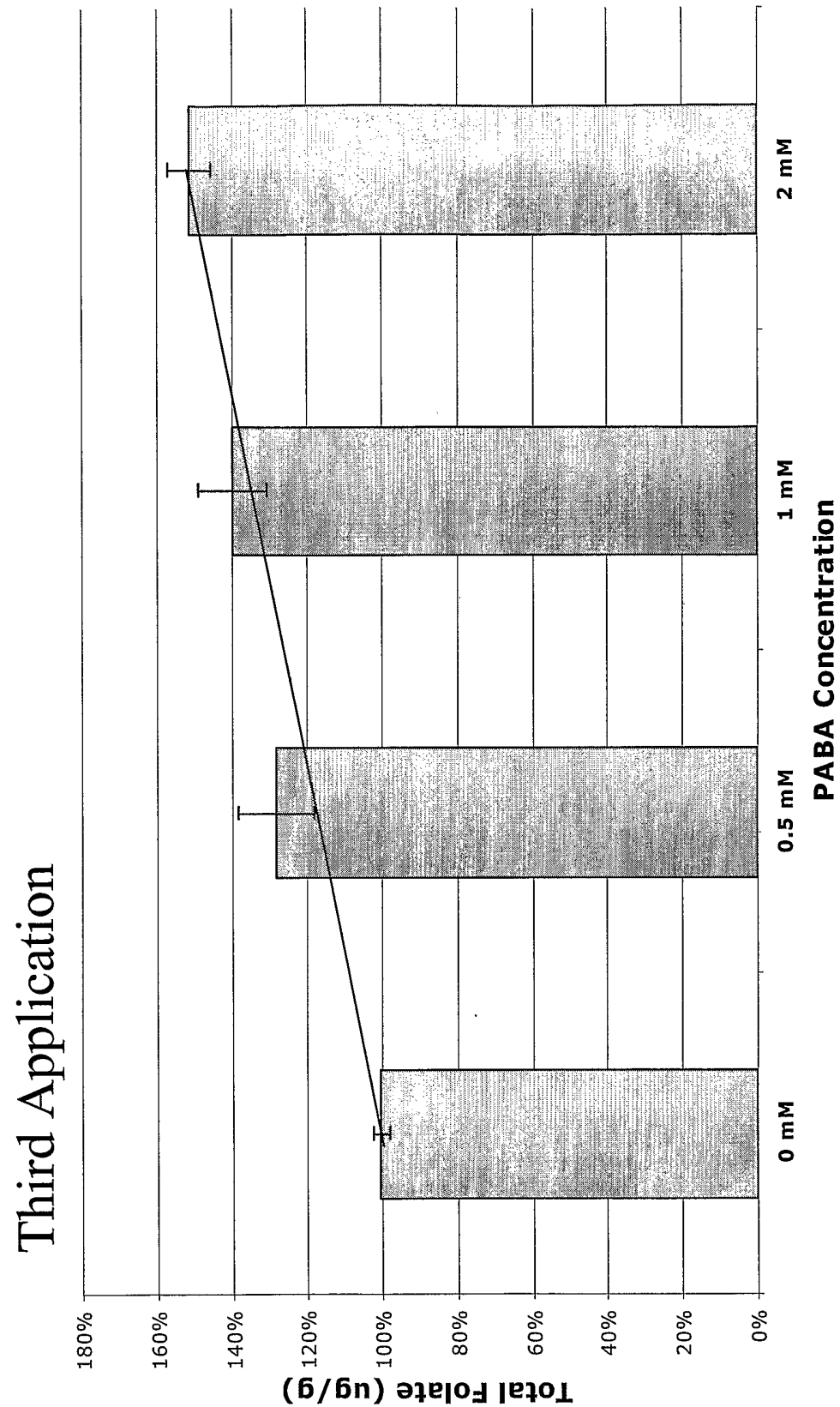
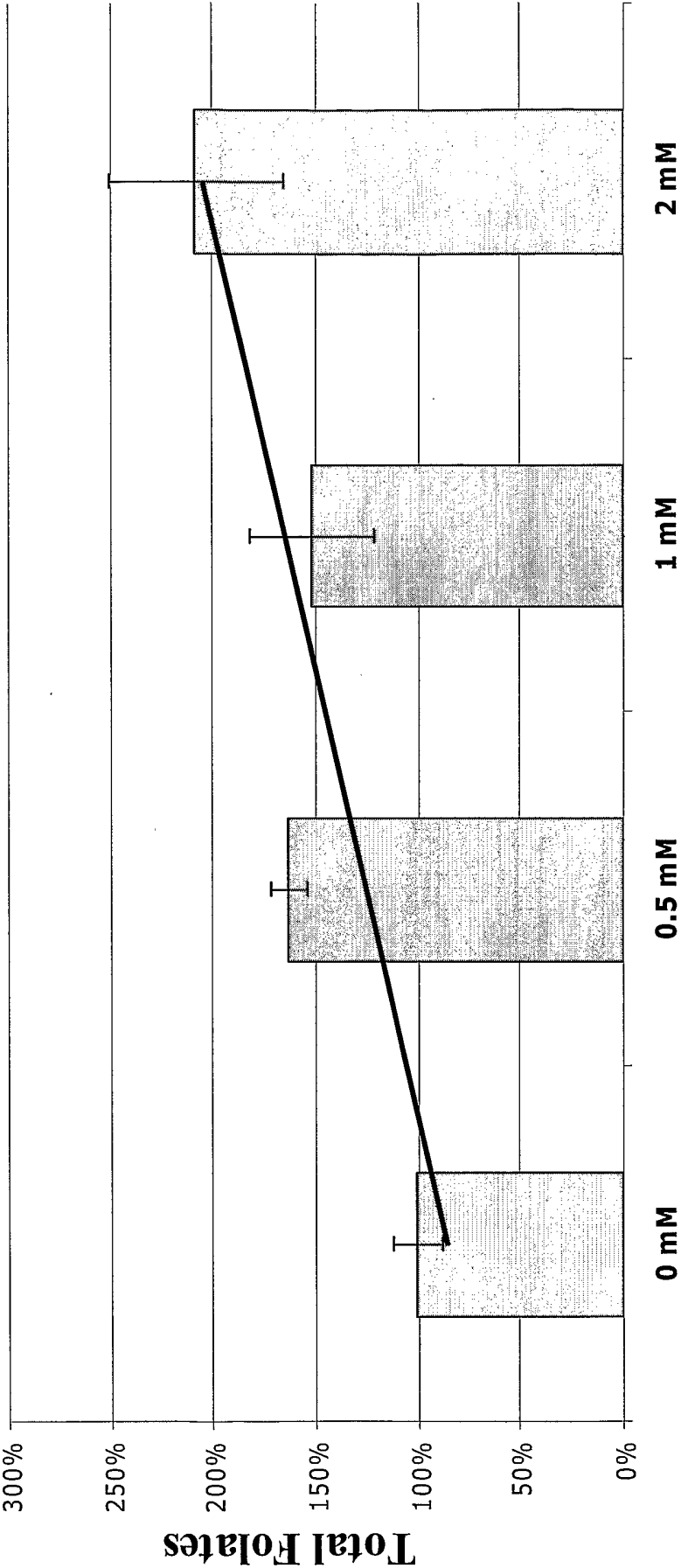


FIG. 22

Sixth Application



PABA Concentration

FIG. 23

SEQUENCE LISTING

<110> University of Florida Research Foundation, Inc.
Hanson, Andrew D.
Gregory, III, Jesse F.
Diaz de la Garza, Rocio
Schubert, Karel
Hossain, Tahzeeba

<120> Materials and Methods for Folate Biofortification in Plants

<130> UF-425XC1

<150> US 60/612,803
<151> 2004-09-23

<160> 11

<170> PatentIn version 3.2

<210> 1
<211> 457
<212> DNA
<213> Unknown

<220>
<223> Synthetic mammalian GCHI gene

<220>
<221> CDS
<222> (1)..(456)

<220>
<221> misc_feature
<222> (421)..(421)
<223> n is a, c, g, or t

<400> 1
atg gag aag cca aga gga gtc agg tgc acc aat ggg ttc tcc gag aga 48
Met Glu Lys Pro Arg Gly Val Arg Cys Thr Asn Gly Phe Ser Glu Arg
1 5 10 15
gag ctt cca aga cca gga gct tct cca cct gct gag aag tct aga cca 96
Glu Leu Pro Arg Pro Gly Ala Ser Pro Pro Ala Glu Lys Ser Arg Pro
20 25 30
cct gag gct aag ggc gca cag ccg gcc gac gcc tgg aag gca ggg cgg 144
Pro Glu Ala Lys Gly Ala Gln Pro Ala Asp Ala Trp Lys Ala Gly Arg
35 40 45
cac cgc agc gag gag gaa aac cag gtg aac ctc ccc aaa ctg gcg gct 192
His Arg Ser Glu Glu Glu Asn Gln Val Asn Leu Pro Lys Leu Ala Ala
50 55 60
gct tac tcg tcc att ctg ctc tcg ctg ggc gag gac ccc cag cgg cag 240
Ala Tyr Ser Ser Ile Leu Leu Ser Leu Gly Glu Asp Pro Gln Arg Gln
65 70 75 80

2

```

ggg ctg ctc aag acg ccc tgg agg gcg gcc acc gcc atg cag tac ttc      288
Gly Leu Leu Lys Thr Pro Trp Arg Ala Ala Thr Ala Met Gln Tyr Phe
      85                      90                      95

acc aag gga tac cag gag acc atc tca gat gtc ctg aat gat gct ata      336
Thr Lys Gly Tyr Gln Glu Thr Ile Ser Asp Val Leu Asn Asp Ala Ile
      100                      105                      110

ttt gat gaa gat cat gac gag atg gtg att gtg aag gac ata gat atg      384
Phe Asp Glu Asp His Asp Glu Met Val Ile Val Lys Asp Ile Asp Met
      115                      120                      125

ttc tcc atg tgt gag cat cac ctt gtt cca ttt gta ngg agg gtc cat      432
Phe Ser Met Cys Glu His His Leu Val Pro Phe Val Xaa Arg Val His
      130                      135                      140

att ggc tat ctt cct aaa caa gca a      457
Ile Gly Tyr Leu Pro Lys Gln Ala
145                      150

```

<210> 2

<211> 152

<212> PRT

<213> Unknown

<220>

<221> misc_feature

<222> (141)..(141)

<223> The 'Xaa' at location 141 stands for Arg, Gly, or Trp.

<220>

<223> Synthetic Construct

<400> 2

```

Met Glu Lys Pro Arg Gly Val Arg Cys Thr Asn Gly Phe Ser Glu Arg
1      5      10      15

Glu Leu Pro Arg Pro Gly Ala Ser Pro Pro Ala Glu Lys Ser Arg Pro
      20      25      30

Pro Glu Ala Lys Gly Ala Gln Pro Ala Asp Ala Trp Lys Ala Gly Arg
      35      40      45

His Arg Ser Glu Glu Glu Asn Gln Val Asn Leu Pro Lys Leu Ala Ala
      50      55      60

Ala Tyr Ser Ser Ile Leu Leu Ser Leu Gly Glu Asp Pro Gln Arg Gln
65      70      75      80

Gly Leu Leu Lys Thr Pro Trp Arg Ala Ala Thr Ala Met Gln Tyr Phe
      85      90      95

Thr Lys Gly Tyr Gln Glu Thr Ile Ser Asp Val Leu Asn Asp Ala Ile
      100      105      110

Phe Asp Glu Asp His Asp Glu Met Val Ile Val Lys Asp Ile Asp Met
      115      120      125

```

Phe Ser Met Cys Glu His His Leu Val Pro Phe Val Xaa Arg Val His
 130 135 140

Ile Gly Tyr Leu Pro Lys Gln Ala
 145 150

<210> 3

<211> 457

<212> DNA

<213> Unknown

<220>

<223> Mammalian GCHI gene

<220>

<221> CDS

<222> (1)..(456)

<220>

<221> misc_feature

<222> (421)..(421)

<223> n is a, c, g, or t

<400> 3

atg gag aag ccg cgg gga gtc agg tgc acc aat ggg ttc tcc gag cgg 48
 Met Glu Lys Pro Arg Gly Val Arg Cys Thr Asn Gly Phe Ser Glu Arg
 1 5 10 15

gag ctg ccg cgg ccc ggg gcc agc ccg cct gcc gag aag tcc cgg ccg 96
 Glu Leu Pro Arg Pro Gly Ala Ser Pro Pro Ala Glu Lys Ser Arg Pro
 20 25 30

ccc gag gcc aag ggc gca cag ccg gcc gac gcc tgg aag gca ggg cgg 144
 Pro Glu Ala Lys Gly Ala Gln Pro Ala Asp Ala Trp Lys Ala Gly Arg
 35 40 45

cac cgc agc gag gag gaa aac cag gtg aac ctc ccc aaa ctg gcg gct 192
 His Arg Ser Glu Glu Glu Asn Gln Val Asn Leu Pro Lys Leu Ala Ala
 50 55 60

gct tac tcg tcc att ctg ctc tcg ctg ggc gag gac ccc cag cgg cag 240
 Ala Tyr Ser Ser Ile Leu Leu Ser Leu Gly Glu Asp Pro Gln Arg Gln
 65 70 75 80

ggg ctg ctc aag acg ccc tgg agg gcg gcc acc gcc atg cag tac ttc 288
 Gly Leu Leu Lys Thr Pro Trp Arg Ala Ala Thr Ala Met Gln Tyr Phe
 85 90 95

acc aag gga tac cag gag acc atc tca gat gtc ctg aat gat gct ata 336
 Thr Lys Gly Tyr Gln Glu Thr Ile Ser Asp Val Leu Asn Asp Ala Ile
 100 105 110

ttt gat gaa gat cat gac gag atg gtg att gtg aag gac ata gat atg 384
 Phe Asp Glu Asp His Asp Glu Met Val Ile Val Lys Asp Ile Asp Met
 115 120 125

ttc tcc atg tgt gag cat cac ctt gtt cca ttt gta ngg agg gtc cat 432

Phe Ser Met Cys Glu His His Leu Val Pro Phe Val Xaa Arg Val His
 130 135 140

att ggc tat ctt cct aaa caa gca a 457
 Ile Gly Tyr Leu Pro Lys Gln Ala
 145 150

<210> 4
 <211> 152
 <212> PRT
 <213> Unknown

<220>
 <221> misc_feature
 <222> (141)..(141)
 <223> The 'Xaa' at location 141 stands for Arg, Gly, or Trp.

<220>
 <223> Synthetic Construct

<400> 4

Met Glu Lys Pro Arg Gly Val Arg Cys Thr Asn Gly Phe Ser Glu Arg
 1 5 10 15

Glu Leu Pro Arg Pro Gly Ala Ser Pro Pro Ala Glu Lys Ser Arg Pro
 20 25 30

Pro Glu Ala Lys Gly Ala Gln Pro Ala Asp Ala Trp Lys Ala Gly Arg
 35 40 45

His Arg Ser Glu Glu Glu Asn Gln Val Asn Leu Pro Lys Leu Ala Ala
 50 55 60

Ala Tyr Ser Ser Ile Leu Leu Ser Leu Gly Glu Asp Pro Gln Arg Gln
 65 70 75 80

Gly Leu Leu Lys Thr Pro Trp Arg Ala Ala Thr Ala Met Gln Tyr Phe
 85 90 95

Thr Lys Gly Tyr Gln Glu Thr Ile Ser Asp Val Leu Asn Asp Ala Ile
 100 105 110

Phe Asp Glu Asp His Asp Glu Met Val Ile Val Lys Asp Ile Asp Met
 115 120 125

Phe Ser Met Cys Glu His His Leu Val Pro Phe Val Xaa Arg Val His
 130 135 140

Ile Gly Tyr Leu Pro Lys Gln Ala
 145 150

<210> 5
 <211> 46
 <212> DNA
 <213> Artificial sequence

<220>
 <223> synthetic polylinker

 <400> 5
 ggatccgcgg ccgcgaattc aggcctggta ccggcgcgcc agatct 46

 <210> 6
 <211> 20
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> PCR primer

 <400> 6
 ctttcttggtt cccatttctc 20

 <210> 7
 <211> 20

 <212> DNA
 <213> Artificial sequence

 <220>
 <223> PCR primer

 <400> 7
 atgcacatgt gtgtcgcttc 20

 <210> 8
 <211> 669
 <212> DNA
 <213> Escherichia coli

 <400> 8
 atgccatcac tcagtaaaga agcggccctg gttcatgaag cgtagattgc gcgaggactg 60
 gaaacaccgc tgcgcccgcc cgtgcatgaa atggataacg aaacgcgcaa aagccttatt 120
 gctggtcata tgaccgaaat catgcagctg ctgaatctcg acctggctga tgacagtttg 180
 atggaaacgc cgcacgcgat cgctaaaatg tatgtcgatg aaattttctc cggctctggat 240
 tacgcccaatt tcccgaaaat caccctcatt gaaaacaaaa tgaaggctga tgaaatggtc 300
 accgtgcgcg atatcactct gaccagcacc tgtgaacacc attttggttac catcgatggc 360
 aaagcgacgg tggcctatat cccgaaagat tcggatgatg gtctgtcaaa aattaaccgc 420
 attgtgcagt tctttgccc a gctccgcag gtgcaggaac gtctgacgca gcaaattctt 480
 attgcgctac aaacgctgct gggcaccaat aacgtggctg tctcgatcga cgcgggtgcat 540
 tactgctga aggcgcgtgg catccgcgat gcaaccagtg ccacgacaac gacctctctt 600
 ggtggattgt tcaaattccag tcagaatacg cgccacgagt ttctgcgcgc tgtgcgtcat 660

cacaactga

669

<210> 9

<211> 222

<212> PRT

<213> Escherichia coli

<400> 9

Met Pro Ser Leu Ser Lys Glu Ala Ala Leu Val His Glu Ala Leu Val
 1 5 10 15

Ala Arg Gly Leu Glu Thr Pro Leu Arg Pro Pro Val His Glu Met Asp
 20 25 30

Asn Glu Thr Arg Lys Ser Leu Ile Ala Gly His Met Thr Glu Ile Met
 35 40 45

Gln Leu Leu Asn Leu Asp Leu Ala Asp Asp Ser Leu Met Glu Thr Pro
 50 55 60

His Arg Ile Ala Lys Met Tyr Val Asp Glu Ile Phe Ser Gly Leu Asp
 65 70 75 80

Tyr Ala Asn Phe Pro Lys Ile Thr Leu Ile Glu Asn Lys Met Lys Val
 85 90 95

Asp Glu Met Val Thr Val Arg Asp Ile Thr Leu Thr Ser Thr Cys Glu
 100 105 110

His His Phe Val Thr Ile Asp Gly Lys Ala Thr Val Ala Tyr Ile Pro
 115 120 125

Lys Asp Ser Val Ile Gly Leu Ser Lys Ile Asn Arg Ile Val Gln Phe
 130 135 140

Phe Ala Gln Arg Pro Gln Val Gln Glu Arg Leu Thr Gln Gln Ile Leu
 145 150 155 160

Ile Ala Leu Gln Thr Leu Leu Gly Thr Asn Asn Val Ala Val Ser Ile
 165 170 175

Asp Ala Val His Tyr Cys Val Lys Ala Arg Gly Ile Arg Asp Ala Thr
 180 185 190

Ser Ala Thr Thr Thr Thr Ser Leu Gly Gly Leu Phe Lys Ser Ser Gln
 195 200 205

Asn Thr Arg His Glu Phe Leu Arg Ala Val Arg His His Asn
 210 215 220

<210> 10

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer

<400> 10

cccatcactc agtaaagaag cggc

24

<210> 11

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer

<400> 11

ccgttgatgacgcacagc g

21