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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 October 2015 (08.10.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/150465 A2

(51) International Patent Classification: *C12N 15/82* (2006.01)

(21) International Application Number: PCT/EP2015/057198

(22) International Filing Date: 1 April 2015 (01.04.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/974,466 3 April 2014 (03.04.2014) US

(71) Applicant: **BASF SE** [DE/DE]; 67056 Ludwigshafen (DE).

(72) Inventors: **APONTE, Raphael**; G7 18, 68159 Mannheim (DE). **TRESCH, Stefan**; Rieslingweg 18, 67281 Kirchheim (DE). **HUTZLER, Johannes**; Carl-Bosch-Str. 9, 67165 Waldsee (DE). **MIETZNER, Thomas**; Rehbergstr.62, 76855 Annweiler (DE). **PAULIK, Jill Marie**; 215 King George Loop, Cary, North Carolina 27511 (US). **BROMMER, Chad**; 1903 Ridley St., Raleigh, North Carolina 27608 (US). **SIMON, Anja**; Klingenweg 10, 69469 Weinheim (DE). **VOGT, Florian**; Lange Rötterstr. 64, 68167 Mannheim (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
- *with sequence listing part of description (Rule 5.2(a))*

(54) Title: PLANTS HAVING INCREASED TOLERANCE TO HERBICIDES

(57) Abstract: The present invention refers to aplant or plant part comprising a polynucleotide encoding a wildtype or mutant transketolase polypeptide, the expression of said polynucleotide confers to the plant or plant part tolerance to transketolase-inhibiting herbicides.

PLANTS HAVING INCREASED TOLERANCE TO HERBICIDES

FIELD OF THE INVENTION

This application claims priority to US provisional applications number US 61/974466, filed on April 3, 2014, the content of which is incorporated herein by reference in its entirety.

The present invention relates in general to methods for conferring on plants agricultural level tolerance to herbicides. Particularly, the invention refers to plants having an increased tolerance to herbicides, more specifically to herbicides which inhibit the enzyme transketolase (TK). More specifically, the present invention relates to methods and plants obtained by mutagenesis and cross-breeding and transformation that have an increased tolerance to herbicides, particularly TK-inhibiting herbicides.

BACKGROUND OF THE INVENTION

Transketolase (EC 2.2.1.1) is integral to both the Calvin cycle and the oxidative pentose phosphate pathway of higher-plant chloroplasts. In the Calvin cycle, it catalyses the transfer of a two-carbon ketol group from either D-fructose- 6-phosphate or D-sedoheptulose-7-phosphate to D-glyceraldehyde- 3-phosphate to yield D-xylulose-5- phosphate and either D-erythrose-4-phosphate or D-ribose-5-phosphate, respectively [Flechner et al., *Plant Molecular Biology* 32: 475-484, 1996]. In the pentose phosphate pathway these reversible reactions proceed in the opposite direction. Transketolase (TK) is thiamine pyrophosphate-dependent and is a dimer of 74 kDa subunits [de la Haba et al., J Biol Chem 214 409-426 (1955); Murphy and Walker, Planta 155, 316-320 (1982)]. Yeast transketolase is one of several thiamin diphosphate dependent enzymes whose three-dimensional structures have been determined [Schenk et al., The International Journal of Biochemistry & Cell Biology 30 (1998) 1297-1318]. Together with mutational analysis these structural data have led to detailed understanding of thiamin diphosphate catalyzed reactions. In the homodimer transketolase the two catalytic sites, where dihydroxyethyl groups are transferred from ketose donors to aldose acceptors, are formed at the interface between the two subunits, where the thiazole and pyrimidine rings of thiamin diphosphate are bound. Transketolase is ubiquitous and more than 30 full-length sequences are known. The encoded protein sequences contain two motifs of high homology; one common to all thiamin diphosphate-dependent enzymes and the other a unique transketolase motif [Schenk et al., The International Journal of Biochemistry & Cell Biology 30 (1998) 1297-1318]. Higher plants in general might possess chloroplast and cytosolic TK isoenzymes, and, additionally, TK activity was observed to be localized to the chloroplast, at least in spinach leaves [Flechner et al., *Plant Molecular Biology* 32: 475-484, 1996 and references contained therein]. Since products of the transketolase-catalysed reaction serve as precursors for a number of synthetic compounds this enzyme has been exploited for industrial applications.

The inventors of the present invention have now surprisingly found that over-expression of wildtype and mutant transketolase forms confers in plants tolerance/resistance to particular classes of TK-inhibiting herbicides as compared to the non-transformed and/or non-mutagenized plants or plant cells, respectively. More specifically, the inventors of the present invention have found that TK expression confers tolerance/resistance to cornexistin and/or hydrocornexistin.

Cornexistin and hydroxycornexistin are natural products derived from the fungus *Paecilomyces divaricatus*. Isolation of cornexistin from the cultures of Paecilomyces species was published as early as 1989 by the Sankyo research group. The Sankyo Corporation discovered cornexistin during the screening of biological extracts for herbicidal use (JP2256602). Later work from the DOW Elanco group described identification of hydroxycornexistin also produced in *Paecilomyces variotii* SANK 21086 (US00542478). Both, cornexistin and hydroxycornexistin are highly potent herbicides that have the unique quality of being harmless to corn plants. Because of this quality, both molecules have attracted research interest. Cornexistin showed good activity as a herbicide as well as relative inactivity towards corn plants. Cornexistin and hydroxycornexistin has been synthesized by chemical synthesis only as diastereomeres (Org. Biomol. Chem., 2008, 6, 4012–4025). Nine-membered carbocyclic structures in general are rare in nature and their synthesis as well as the genes involved in the synthesis is still unknown and not described.

The problem of the present invention can be seen as to the provision of novel traits by identifying target polypeptides, the manipulation of which makes plants tolerant to herbicides.

Three main strategies are available for making plants tolerant to herbicides, i.e. (1) detoxifying the herbicide with an enzyme which transforms the herbicide, or its active metabolite, into non-toxic products, such as, for example, the enzymes for tolerance to bromoxynil or to basta (EP242236, EP337899); (2) mutating the target enzyme into a functional enzyme which is less sensitive to the herbicide, or to its active metabolite, such as, for example, the enzymes for tolerance to glyphosate (EP293356, Padgette S. R. et al., J.Biol. Chem., 266, 33, 1991); or (3) overexpressing the sensitive enzyme so as to produce quantities of the target enzyme in the plant which are sufficient in relation to the herbicide, in view of the kinetic constants of this enzyme, so as to have enough of the functional enzyme available despite the presence of its inhibitor.

The problem is solved by the subject-matter of the present invention.

SUMMARY OF THE INVENTION

Accordingly, in one aspect, the present invention provides a plant or plant part comprising a polynucleotide encoding a wildtype or mutated TK polypeptide, the expression of said polynucleotide confers to the plant or plant part tolerance to TK-inhibiting herbicides.

In some aspects, the present invention provides a seed capable of germination into a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In one aspect, the present invention provides a plant cell of or capable of regenerating a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides, wherein the plant cell comprises the polynucleotide operably linked to a promoter.

In another aspect, the present invention provides a plant cell comprising a polynucleotide operably linked to a promoter operable in a cell, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In other aspects, the present invention provides a plant product prepared from a plant or plant part comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In some aspects, the present invention provides a progeny or descendant plant derived from a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, wherein the progeny or descendant plant comprises in at least some of its cells the recombinant polynucleotide operably linked to the promoter, the expression of the wildtype or mutated TK polypeptide conferring to the progeny or descendant plant tolerance to the TK-inhibiting herbicides.

In other aspects, the present invention provides a method for controlling weeds at a locus for growth of a plant, the method comprising: (a) applying an herbicide composition comprising TK-inhibiting herbicides to the locus; and (b) planting a seed at the locus, wherein the seed is capable of producing a plant that comprises in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to

TK-inhibiting herbicides.

In some aspects, the present invention provides a method for controlling weeds at a locus for growth of a plant, the method comprising: applying an herbicidal composition comprising TK-inhibiting herbicides to the locus; wherein said locus is: (a) a locus that contains: a plant or a seed capable of producing said plant; or (b) a locus that is to be after said applying is made to contain the plant or the seed; wherein the plant or the seed comprises in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In one aspect, step (a) occurs before, after, or concurrently with step (b).

In other aspects, the present invention provides a method of producing a plant having tolerance to TK-inhibiting herbicides, the method comprising regenerating a plant from a plant cell transformed with a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In one aspect, the present invention provides a method of producing a progeny plant having tolerance to TK-inhibiting herbicides, the method comprising: crossing a first TK-inhibiting herbicides- tolerant plant with a second plant to produce a TK-inhibiting herbicides- tolerant progeny plant, wherein the first plant and the progeny plant comprise in at least some of their cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In addition, the present invention refers to a method for identifying a TK-inhibiting herbicide by using a wild-type or mutated TK of the present invention encoded by a nucleic acid which comprises the nucleotide sequence of SEQ ID NO: 182 or 183, or a variant, homologue, paralogue or orthologue thereof.

Said method comprises the steps of:

a) generating a transgenic cell or plant comprising a nucleic acid encoding a mutated TK of the present invention, wherein the mutated TK of the present invention is expressed;
b) applying a TK-inhibiting herbicide to the transgenic cell or plant of a) and to a control cell or plant of the same variety;
c) determining the growth or the viability of the transgenic cell or plant and the control cell or plant after application of said test compound, and

d) selecting test compounds which confer reduced growth to the control cell or plant as compared to the growth of the transgenic cell or plant.

Another object refers to a method of identifying a nucleotide sequence encoding a mutated TK which is resistant or tolerant to a TK-inhibiting herbicide, the method comprising:

a) generating a library of mutated TK-encoding nucleic acids,
b) screening a population of the resulting mutated TK-encoding nucleic acids by expressing each of said nucleic acids in a cell or plant and treating said cell or plant with a TK-inhibiting herbicide,
c) comparing the TK-inhibiting herbicide-tolerance levels provided by said population of mutated TK encoding nucleic acids with the TK-inhibiting herbicide-tolerance level provided by a control TK-encoding nucleic acid,
d) selecting at least one mutated TK-encoding nucleic acid that provides a significantly increased level of tolerance to a TK-inhibiting herbicide as compared to that provided by the control TK-encoding nucleic acid.

In a preferred embodiment, the mutated TK-encoding nucleic acid selected in step d) provides at least 2-fold as much tolerance to a TK-inhibiting herbicide as compared to that provided by the control TK-encoding nucleic acid.

The resistance or tolerance can be determined by generating a transgenic plant comprising a nucleic acid sequence of the library of step a) and comparing said transgenic plant with a control plant.

Another object refers to a method of identifying a plant or algae containing a nucleic acid encoding a mutated TK which is resistant or tolerant to a TK-inhibiting herbicide, the method comprising:

a) identifying an effective amount of a TK-inhibiting herbicide in a culture of plant cells or green algae.
b) treating said plant cells or green algae with a mutagenizing agent,
c) contacting said mutagenized cells population with an effective amount of TK-inhibiting herbicide, identified in a),
d) selecting at least one cell surviving these test conditions,
e) PCR-amplification and sequencing of TK genes from cells selected in d) and comparing such sequences to wild-type TK gene sequences, respectively.

In a preferred embodiment, the mutagenizing agent is ethylmethanesulfonate.

Another object refers to an isolated nucleic acid encoding a mutated TK, the nucleic acid comprising the sequence of SEQ ID NO: 182, or 183, or a variant thereof, as defined hereinafter.

In a preferred embodiment, the nucleic acid being identifiable by a method as defined above.

Another object refers to an isolated mutated TK polypeptide, the polypeptide comprising the sequence set forth in SEQ ID NO: 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, 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, or 181, a variant, derivative, orthologue, paralogue or homologue thereof, as defined hereinafter.

In still further aspects, the present invention provides a plant or plant part comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides, wherein the plant or plant part further exhibits a second or third herbicide-tolerant trait.

In another embodiment, the invention refers to a plant cell transformed by and expressing a wild-type or a mutated TK nucleic acid according to the present invention or a plant which has been mutated to obtain a plant expressing, preferably over-expressing a wild-type or a mutated TK nucleic acid according to the present invention, wherein expression of said nucleic acid in the plant cell results in increased resistance or tolerance to a TK -inhibiting herbicide as compared to a wild type variety of the plant cell

In another embodiment, the invention refers to a plant comprising a plant cell according to the present invention, wherein expression of the nucleic acid in the plant results in the plant's increased resistance to TK-inhibiting herbicide as compared to a wild type variety of the plant.

The plants of the present invention can be transgenic or non-transgenic.

Preferably, the expression of the nucleic acid of the invention in the plant results in the plant's increased resistance to TK-inhibiting herbicides as compared to a wild type variety of the plantIn another embodiment, the invention refers to a seed produced by a transgenic plant comprising a plant cell of the present invention, wherein the seed is true breeding for an increased resistance to a TK-inhibiting herbicide as compared to a wild type variety of the seed.

In another embodiment, the invention refers to a method of producing a transgenic plant cell with an increased resistance to a TK-inhibiting herbicide as compared to a wild type variety of the plant cell comprising, transforming the plant cell with an expression cassette comprising a a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide.

In another embodiment, the invention refers to a method of producing a transgenic plant comprising, (a) transforming a plant cell with an expression cassette comprising a a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, and (b) generating a plant with an increased resistance to TK-inhibiting herbicide from the plant cell.

Preferably, the expression cassette further comprises a transcription initiation regulatory region and a translation initiation regulatory region that are functional in the plant.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1

A, B, C, D, E, F, G, and H show a Multiple alignment using clustalw of sequences 1 to 181 depicted below in Table 1.

KEY TO SEQUENCE LISTING

Table 1

SEQ ID	Organism / Origin / Accession number
1	1 [transketolase Spinacia oleracea]
2	transketolase [Spinacia oleracea]
3	gi|2529342|gb|AAD10219.1| transketolase [Spinacia oleracea]
4	gi|223548870|gb|EEF50359.1| transketolase, putative [Ricinus communis]
5	Transketolase_1_AMATU
6	Transketolase_2_AMATU
7	Transketolase_3_AMATU
8	Transketolase_1_KOSCS
9	Transketolase_2_KOSCS
10	Transketolase_LEMPA
11	Partial_Sequence_Transketolase_ARBTH
12	gi|514760041|ref|XP_004964445.1| PREDICTED: transketolase, chloroplastic-like [Setaria italica]
13	gi|568875658|ref|XP_006490909.1| PREDICTED: transketolase-2, chloroplastic-like [Citrus sinensis]
14	gi|351735634|gb|AEQ59483.1| chloroplast transketolase [Cucumis sativus]
15	gi|118487947|gb|ABK95795.1| [Populus trichocarpa]
16	gi|470142472|ref|XP_004306931.1| PREDICTED: [Fragaria vesca subsp. vesca]
17	gi|561015858|gb|ESW14662.1| hypothetical protein PHAVU_007G006600g [Phaseolus vulgaris]
18	gi|449498919|ref|XP_004160671.1| PREDICTED: transketolase, chloroplastic-

	like [Cucumis sativus]
19	gi|356576867|ref|XP_003556551.1| transketolase, chloroplastic [Glycine max]
20	gi|469474163|gb|AGH33875.1| transketolase [Camellia sinensis]
21	gi|462416702|gb|EMJ21439.1| [Prunus persica]
22	gi|222843004|gb|EEE80551.1| transketolase family protein [Populus trichocarpa]
23	gi|460388792|ref|XP_004240048.1| transketolase, chloroplastic-like [Solanum lycopersicum]
24	gi|356536526|ref|XP_003536788.1| transketolase, chloroplastic-like [Glycine max]
25	gi|355516131|gb|AES97754.1| Transketolase [Medicago truncatula]
26	gi|330255441|gb|AEC10535.1| transketolase [Arabidopsis thaliana]
27	gi|118481093|gb|ABK92500.1| [Populus trichocarpa]
28	gi|449474170|ref|XP_004154093.1| transketolase, chloroplastic-like [Cucumis sativus]
29	gi|565357366|ref|XP_006345515.1| transketolase, chloroplastic-like [Solanum tuberosum]
30	gi|502121526|ref|XP_004497357.1| transketolase, chloroplastic-like [Cicer arietinum]
31	Nicotianatabacum
32	Arabidopsisthaliana
33	Theobromacacao
34	Capsellarubella
35	Arabidopsisthaliana
36	Vitisvinifera
37	Solanumlycopersicum
38	Zeamays
39	Arabidopsisthaliana
40	gi|3559814|emb|CAA75777.1| transketolase [Capsicum annuum]
41	gi|568214657|ref|NP_001275202.1| transketolase [Solanum tuberosum]
42	gi|356506190|ref|XP_003521870.1| transketolase, chloroplastic [Glycine max]
43	gi|7329685|emb|CAB82679.1| transketolase-like protein [Arabidopsis thaliana]
44	gi|565465348|ref|XP_006290638.1| [Capsella rubella]
45	gi|28190676|gb|AAO33154.1| transketolase [Oryza sativa]
46	gi|561026815|gb|ESW25455.1| [Phaseolus vulgaris]
47	gi|115466224|ref|NP_001056711.1| Os06g0133800 [Oryza sativa]
48	gi|312282187|dbj|BAJ33959.1| [Thellungiella halophila]
49	gi|573948603|ref|XP_006656605.1| partial Transketolase [Oryza brachyantha]
50	gi|548832450|gb|ERM95246.1 [Amborella trichopoda]
51	gi|548847307|gb|ERN06491.1| [Amborella trichopoda]
52	gi|502146626|ref|XP_004506536.1| [Cicer arietinum]
53	gi|357110873|ref|XP_003557240.1| [Brachypodium distachyon]

54	gi|355506195|gb|AES87337.1| Transketolase [Medicago truncatula]
55	gi|147835837|emb|CAN72939.1| hypothetical protein [Vitis vinifera]
56	gi|225454009|ref|XP_002280760.1| transketolase [Vitis vinifera]
57	gi|297318807|gb|EFH49229.1| transketolase [Arabidopsis lyrata subsp. lyrata]
58	gi|241915985|gb|EER89129.1| [Sorghum bicolor]
59	gi|413953334|gb|AFW85983.1| transketolase isoform 1 [Zea mays]
60	gi|224127366|ref|XP_002320056.1| [Populus trichocarpa]
61	Hordeum vulgare subsp. vulgare
62	Picea sitchensis
63	Arabidopsis thaliana
64	Sorghum bicolor
65	Zea mays
66	Eutrema salsugineum
67	Polygonum tinctorium
68	Oryza sativa Japonica Group
69	Oryza sativa Indica Group
70	Oryza brachyantha
71	Setaria italica
72	Physcomitrella patens
73	Setaria italica
74	Physcomitrella patens
75	Oryza sativa Indica Group
76	gi|115457470|ref|NP_001052335.1| Os04g0266900 [Oryza sativa Japonica]
77	gi|475481099|gb|EMT02862.1| Transketolase [Aegilops tauschii]
78	gi|357167367|ref|XP_003581128.1| transketolase [Brachypodium distachyon]
79	gi|125547501|gb|EAY93323.1| [Oryza sativa Indica Group]
80	gi|514760053|ref|XP_004964448.1| transketolase [Setaria italica]
81	gi|242072546|ref|XP_002446209.1| [Sorghum bicolor]
82	gi|514760061|ref|XP_004964450.1| transketolase [Setaria italica]
83	gi|326515912|dbj|BAJ87979.1| [Hordeum vulgare subsp. vulgare]
84	gi|474352176|gb|EMS63024.1| Transketolase [Triticum urartu]
85	gi|664903|emb|CAA86609.1| transketolase [Craterostigma plantagineum]
86	gi|147788852|emb|CAN60522.1| [Vitis vinifera]
87	gi|300140959|gb|EFJ07676.1| [Selaginella moellendorffii]
88	gi|413953333|gb|AFW85982.1| [Zea mays]
89	gi|414587814|tpg|DAA38385.1| [Zea mays]
90	gi|162683807|gb|EDQ70214.1| [Physcomitrella patens]
91	gi|125595949|gb|EAZ35729.1| [Oryza sativa Japonica Group]
92	gi|222423758|dbj|BAH19845.1| AT3G60750 [Arabidopsis thaliana]
93	Medicago truncatula
94	Sorghum bicolor
95	Volvox carteri f. nagariensis

96	Chlamydomonas reinhardtii
97	Geitlerinema sp. PCC 7407
98	Craterostigma plantagineum
99	Pleurocapsa sp. PCC 7327
100	Moorea producens 3L
101	Fischerella muscicola
102	Leptolyngbya sp. Heron Island J
103	Cyanothece sp. PCC 7425
104	Paecilomyces divaricatus (Transketolase, scaffold001g03940)
105	SPTREMBL|A5DSN8_LODEL RecName: Full=Transketolase; EC=2.2.1.1;
106	SPTREMBL|A8PWG5_MALGO RecName: Full=Transketolase; EC=2.2.1.1;
107	SPTREMBL|B6K5F9_SCHJY RecName: Full=Transketolase; EC=2.2.1.1;
108	SPTREMBL|C0LQF6_PICPA RecName: Full=Transketolase; EC=2.2.1.1; Flags: Fragment;
109	SPTREMBL|C7ZAM9_NECH7 SubName: Full=Putative uncharacterized protein;
110	SPTREMBL|D7UPI2_OGAME SubName: Full=Dihydroxyacetone synthase; EC=2.2.1.3;
111	SPTREMBL|D7UPI3_OGAME SubName: Full=DAS-like protein;
112	SPTREMBL|D8Q788_SCHCM SubName: Full=Putative uncharacterized protein;
113	SPTREMBL|F4P3U6_BATDJ RecName: Full=Transketolase; EC=2.2.1.1;
114	SPTREMBL|F5HCT5_CRYNB SubName: Full=Putative uncharacterized protein;
115	SPTREMBL|G3BBY9_CANTC SubName: Full=Putative uncharacterized protein;
116	SPTREMBL|G7DV49_MIXOS SubName: Full=Uncharacterized protein;
117	SPTREMBL|G8C0M6_TETPH RecName: Full=Transketolase; EC=2.2.1.1;
118	SPTREMBL|G9MFC9_HYPVG SubName: Full=Uncharacterized protein;
119	SPTREMBL|H0GCH3_9SACH RecName: Full=Transketolase; EC=2.2.1.1;
120	SPTREMBL|I1BXP2_RHIO9 RecName: Full=Transketolase; EC=2.2.1.1;
121	SPTREMBL|I1RJN7_GIBZE SubName: Full=Dihydroxyacetone synthase; SubName: Full=Uncharacterized protein;
122	SPTREMBL|I2K2H5_DEKBR RecName: Full=Transketolase; EC=2.2.1.1;
123	SPTREMBL|J9N9S0_FUSO4 SubName: Full=Uncharacterized protein;
124	SPTREMBL|J9VL23_CRYNH RecName: Full=Transketolase; EC=2.2.1.1;
125	SPTREMBL|L0PF66_PNEJ8 SubName: Full=I WGS project CAKM00000000 data, strain SE8, contig 242;
126	SPTREMBL|L2FB38_COLGN SubName: Full=Dihydroxyacetone synthase;
127	SPTREMBL|Q5KHG5_CRYNJ SubName: Full=Transketolase, putative;
128	SPTREMBL|Q6BN17_DEBHA SubName: Full=DEHA2F00968p;
129	SPTREMBL|Q6BXS5_DEBHA SubName: Full=DEHA2B00616p;

130	Paecilomyces divaricatus (Transketolase, scaffold020g00760)
131	SPTREMBL|A1DJZ3_NEOFI SubName: Full=Transketolase;
132	SPTREMBL|A5DD66_PICGU SubName: Full=Putative uncharacterized protein;
133	SPTREMBL|A5DSN8_LODEL RecName: Full=Transketolase; EC=2.2.1.1;
134	SPTREMBL|A8PWG5_MALGO RecName: Full=Transketolase; EC=2.2.1.1;
135	SPTREMBL|B6K5F9_SCHJY RecName: Full=Transketolase; EC=2.2.1.1;
136	SPTREMBL|B8M240_TALSN RecName: Full=Transketolase; EC=2.2.1.1;
137	SPTREMBL|B8M3Z5_TALSN SubName: Full=Transketolase, putative; EC=2.2.1.3;
138	SPTREMBL|C0LQF6_PICPA RecName: Full=Transketolase; EC=2.2.1.1; Flags: Fragment;
139	SPTREMBL|C7ZAM9_NECH7 SubName: Full=Putative uncharacterized protein;
140	SPTREMBL|D7UPI2_OGAME SubName: Full=Dihydroxyacetone synthase; EC=2.2.1.3;
141	SPTREMBL|D7UPI3_OGAME SubName: Full=DAS-like protein;
142	SPTREMBL|D8Q788_SCHCM SubName: Full=Putative uncharacterized protein;
143	SPTREMBL|E6ZQ75_SPORE SubName: Full=Probable Transketolase;
144	SPTREMBL|F2QV92_PICP7 SubName: Full=Dihydroxyacetone synthase variant 1; EC=2.2.1.3;
145	SPTREMBL|F4P3U6_BATDJ RecName: Full=Transketolase; EC=2.2.1.1;
146	SPTREMBL|F5HCT5_CRYNB SubName: Full=Putative uncharacterized protein;
147	SPTREMBL|F9GER6_FUSOF SubName: Full=Uncharacterized protein;
148	SPTREMBL|F9XCD0_MYCGM SubName: Full=Uncharacterized protein;
149	SPTREMBL|G2YJE4_BOTF4 SubName: Full=Similar to transketolase;
150	SPTREMBL|G3B8B6_CANTC SubName: Full=Dihydroxyacetone synthase;
151	SPTREMBL|G3BBY9_CANTC SubName: Full=Putative uncharacterized protein;
152	SPTREMBL|G7DV49_MIXOS SubName: Full=Uncharacterized protein;
153	SPTREMBL|G8BNE8_TETPH RecName: Full=Transketolase; EC=2.2.1.1;
154	SPTREMBL|G8C0M6_TETPH RecName: Full=Transketolase; EC=2.2.1.1;
155	SPTREMBL|H0EPU8_GLAL7 SubName: Full=Putative Dihydroxyacetone synthase;
156	Paecilomyces divaricatus (Transketolase, scaffold025g01090)
157	SPTREMBL|A8NV31_COPC7 SubName: Full=Transketolase;
158	SPTREMBL|D8Q4R0_SCHCM SubName: Full=Putative uncharacterized protein;
159	SPTREMBL|E3JQR8_PUCGT SubName: Full=Putative uncharacterized protein;
160	SPTREMBL|F4RJP4_MELLP SubName: Full=Putative uncharacterized protein;

161	SPTREMBL|G7DVE0_MIXOS SubName: Full=Uncharacterized protein;
162	SPTREMBL|H0EEH2_GLAL7 SubName: Full=Putative Transketolase;
163	SPTREMBL|I4YAU9_WALSC SubName: Full=Thiamin diphosphate-binding protein;
164	SPTREMBL|I4YE20_WALSC SubName: Full=Thiamin diphosphate-binding protein;
165	SPTREMBL|K1WGM0_MARBU SubName: Full=Transketolase;
166	SPTREMBL|K5WES2_PHACS SubName: Full=Uncharacterized protein;
167	SPTREMBL|Q55T38_CRYNB SubName: Full=Putative uncharacterized protein;
168	SPTREMBL|L1J2M8_GUITH SubName: Full=Uncharacterized protein;
169	SPTREMBL|A0YNC4_LYNSP SubName: Full=Uncharacterized protein;
170	SPTREMBL|A0ZLZ6_NODSP SubName: Full=Transketolase; EC=2.2.1.1;
171	SPTREMBL|B4WLI3_9SYNE SubName: Full=Transketolase, C-terminal domain protein;
172	SPTREMBL|C9M6X8_9BACT SubName: Full=Transketolase, C-terminal domain protein;
173	SPTREMBL|D1Y3C7_9BACT SubName: Full=Transketolase; EC=2.2.1.1;
174	SPTREMBL|D2Z2G4_9BACT SubName: Full=Transketolase central region;
175	SPTREMBL|D3L2W3_9BACT SubName: Full=Transketolase, thiamine diphosphate binding domain protein;
176	SPTREMBL|D6THL7_9CHLR SubName: Full=Transketolase domain protein;
177	SPTREMBL|D8F7T0_9DELT SubName: Full=Putative 1-deoxy-D-xylulose-5-phosphate synthase;
178	SPTREMBL|F4XPA7_9CYAN SubName: Full=Transketolase;
179	SPTREMBL|G2H8B8_9DELT SubName: Full=Transketolase, C-terminal domain protein;
180	SPTREMBL|G3J1Y8_9GAMM SubName: Full=Transketolase domain-containing protein;
181	SPTREMBL|G9PZY8_9BACT SubName: Full=Putative uncharacterized protein;

DETAILED DESCRIPTION

The articles "a" and "an" are used herein to refer to one or more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "an element" means one or more elements.

As used herein, the word "comprising," or variations such as "comprises" or "comprising," will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The term “control of undesired vegetation or weeds” is to be understood as meaning the killing of weeds and/or otherwise retarding or inhibiting the normal growth of the weeds. Weeds, in the broadest sense, are understood as meaning all those plants which grow in locations where they are undesired. The weeds of the present invention include, for example, dicotyledonous and monocotyledonous weeds. Dicotyledonous weeds include, but are not limited to, weeds of the genera: Sinapis, Lepidium, Galium, Stellaria, Matricaria, Anthemis, Galinsoga, Chenopodium, Urtica, Senecio, Amaranthus, Portulaca, Xanthium, Convolvulus, Ipomoea, Polygonum, Sesbania, Ambrosia, Cirsium, Carduus, Sonchus, Solanum, Rorippa, Rotala, Lindernia, Lamium, Veronica, Abutilon, Emex, Datura, Viola, Galeopsis, Papaver, Centaurea, Trifolium, Ranunculus, and Taraxacum. Monocotyledonous weeds include, but are not limited to, weeds of of the genera: Echinochloa, Setaria, Panicum, Digitaria, Phleum, Poa, Festuca, Eleusine, Brachiaria, Lolium, Bromus, Avena, Cyperus, Sorghum, Agropyron, Cynodon, Monochoria, Fimbristyslis, Sagittaria, Eleocharis, Scirpus, Paspalum, Ischaemum, Sphenoclea, Dactyloctenium, Agrostis, Alopecurus, and Apera. In addition, the weeds of the present invention can include, for example, crop plants that are growing in an undesired location. For example, a volunteer maize plant that is in a field that predominantly comprises soybean plants can be considered a weed, if the maize plant is undesired in the field of soybean plants.

The term “plant” is used in its broadest sense as it pertains to organic material and is intended to encompass eukaryotic organisms that are members of the Kingdom Plantae, examples of which include but are not limited to vascular plants, vegetables, grains, flowers, trees, herbs, bushes, grasses, vines, ferns, mosses, fungi and algae, etc, as well as clones, offsets, and parts of plants used for asexual propagation (e.g. cuttings, pipings, shoots, rhizomes, underground stems, clumps, crowns, bulbs, corms, tubers, rhizomes, plants/tissues produced in tissue culture, etc.). The term “plant” further encompasses whole plants, ancestors and progeny of the plants and plant parts, including seeds, shoots, stems, leaves, roots (including tubers), flowers, florets, fruits, pedicles, peduncles, stamen, anther, stigma, style, ovary, petal, sepal, carpel, root tip, root cap, root hair, leaf hair, seed hair, pollen grain, microspore, cotyledon, hypocotyl, epicotyl, xylem, phloem, parenchyma, endosperm, a companion cell, a guard cell, and any other known organs, tissues, and cells of a plant, and tissues and organs, wherein each of the aforementioned comprise the gene/nucleic acid of interest. The term “plant” also encompasses plant cells, suspension cultures, callus tissue, embryos, meristematic regions, gametophytes, sporophytes, pollen and microspores, again wherein each of the aforementioned comprises the gene/nucleic acid of interest.

Plants that are particularly useful in the methods of the invention include all plants which belong to the superfamily Viridiplantae, in particular monocotyledonous and dicotyledonous plants including fodder or forage legumes, ornamental plants, food crops, trees or shrubs selected from the list comprising Acer spp., Actinidia spp., Abelmoschus spp., Agave sisalana, Agropyron spp., Agrostis stolonifera, Allium spp., Amaranthus spp., Ammophila arenaria, Ananas comosus, Annona spp., Apium graveolens, Arachis spp, Artocarpus spp., Asparagus officinalis, Avena spp. (e.g. Avena sativa, Avena fatua, Avena byzantina, Avena fatua var. sativa, Avena hybrida), Averrhoa carambola, Bambusa sp., Benincasa hispida, Bertholletia excelsea, Beta vulgaris, Brassica spp. (e.g. Brassica napus, Brassica rapa ssp. [canola, oilseed rape, turnip rape]), Cadaba farinosa, Camellia sinensis, Canna indica, Cannabis sativa, Capsicum spp., Carex elata, Carica papaya, Carissa macrocarpa, Carya spp., Carthamus tinctorius, Castanea spp., Ceiba pentandra, Cichorium endivia, Cinnamomum spp., Citrullus lanatus, Citrus spp., Cocos spp., Coffea spp., Colocasia esculenta, Cola spp., Corchorus sp., Coriandrum sativum, Corylus spp., Crataegus spp., Crocus sativus, Cucurbita spp., Cucumis spp., Cynara spp., Daucus carota, Desmodium spp., Dimocarpus longan, Dioscorea spp., Diospyros spp., Echinochloa spp., Elaeis (e.g. Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Eragrostis tef, Erianthus sp., Eriobotrya japonica, Eucalyptus sp., Eugenia uniflora, Fagopyrum spp., Fagus spp., Festuca arundinacea, Ficus carica, Fortunella spp., Fragaria spp., Ginkgo biloba, Glycine spp. (e.g. Glycine max, Soja hispida or Soja max), Gossypium hirsutum, Helianthus spp. (e.g. Helianthus annuus), Hemerocallis fulva, Hibiscus spp., Hordeum spp. (e.g. Hordeum vulgare), Ipomoea batatas, Juglans spp., Lactuca sativa, Lathyrus spp., Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus spp., Luffa acutangula, Lupinus spp., Luzula sylvatica, Lycopersicon spp. (e.g. Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma spp., Malus spp., Malpighia emarginata, Mammea americana, Mangifera indica, Manihot spp., Manilkara zapota, Medicago sativa, Melilotus spp., Mentha spp., Miscanthus sinensis, Momordica spp., Morus nigra, Musa spp., Nicotiana spp., Olea spp., Opuntia spp., Ornithopus spp., Oryza spp. (e.g. Oryza sativa, Oryza latifolia), Panicum miliaceum, Panicum virgatum, Passiflora edulis, Pastinaca sativa, Pennisetum sp., Persea spp., Petroselinum crispum, Phalaris arundinacea, Phaseolus spp., Phleum pratense, Phoenix spp., Phragmites australis, Physalis spp., Pinus spp., Pistacia vera, Pisum spp., Poa spp., Populus spp., Prosopis spp., Prunus spp., Psidium spp., Punica granatum, Pyrus communis, Quercus spp., Raphanus sativus, Rheum rhabarbarum, Ribes spp., Ricinus communis, Rubus spp., Saccharum spp., Salix sp., Sambucus spp., Secale cereale, Sesamum spp., Sinapis sp., Solanum spp. (e.g. Solanum tuberosum, Solanum integrifolium or Solanum lycopersicum), Sorghum bicolor, Spinacia spp., Syzygium spp., Tagetes spp., Tamarindus indica, Theobroma cacao, Trifolium spp., Tripsacum dactyloides, Triticosecale rimpaui, Triticum spp. (e.g. Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum, Triticum monococcum or Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium spp., Vicia spp., Vigna spp., Viola odorata, Vitis spp., Zea mays, Zizania palustris, Ziziphus spp., amaranth, artichoke, asparagus, broccoli, Brussels sprouts, cabbage, canola, carrot, cauliflower, celery, collard greens, flax, kale, lentil, oilseed rape, okra, onion, potato, rice, soybean, strawberry, sugar beet, sugar cane, sunflower, tomato, squash, tea and algae, amongst others. According to a preferred embodiment of the present invention, the plant is a crop plant. Examples of crop plants include inter alia soybean, sunflower, canola, alfalfa, rapeseed, cotton, tomato, potato or tobacco. Further preferebly, the plant is a monocotyledonous plant, such as sugarcane. Further preferably, the plant is a cereal, such as rice, maize, wheat, barley, millet, rye, sorghum or oats.

Generally, the term “herbicide” is used herein to mean an active ingredient that kills, controls or otherwise adversely modifies the growth of plants. The preferred amount or concentration of the herbicide is an "effective amount" or "effective concentration." By "effective amount" and "effective concentration" is intended an amount and concentration, respectively, that is sufficient to kill or inhibit the growth of a similar, wild-type, plant, plant tissue, plant cell, or host cell, but that said amount does not kill or inhibit as severely the growth of the herbicide-resistant plants, plant tissues, plant cells, and host cells of the present invention. Typically, the effective amount of a herbicide is an amount that is routinely used in agricultural production systems to kill weeds of interest. Such an amount is known to those of ordinary skill in the art. Herbicidal activity is exhibited by herbicides useful for the the present invention when they are applied directly to the plant or to the locus of the plant at any stage of growth or before planting or emergence. The effect observed depends upon the plant species to be controlled, the stage of growth of the plant, the application parameters of dilution and spray drop size, the particle size of solid components, the environmental conditions at the time of use, the specific compound employed, the specific adjuvants and carriers employed, the soil type, and the like, as well as the amount of chemical applied. These and other factors can be adjusted as is known in the art to promote non-selective or selective herbicidal action. Generally, it is preferred to apply the herbicide postemergence to relatively immature undesirable vegetation to achieve the maximum control of weeds.

By a "herbicide-tolerant" or "herbicide-resistant" plant, it is intended that a plant that is tolerant or resistant to at least one herbicide at a level that would normally kill, or inhibit the growth of, a normal or wild-type plant. By "herbicide-tolerant wildtype or mutated TK protein" or "herbicide -resistant wildtype or mutated TK protein", it is intended that such a TK protein displays higher TK activity, relative to the TK activity of a wild-type TK protein, when in the presence of at least one herbicide that is known to interfere with TK activity and at a concentration or level of the herbicide that is known to inhibit the TK activity of the wild-type mutated TK protein. Furthermore, the TK activity of such a herbicide-tolerant or herbicide-resistant mutated TK protein may be referred to herein as "herbicide-tolerant" or "herbicide-resistant" TK activity.

Levels of herbicide that normally inhibit growth of a non-tolerant plant are known and readily determined by those skilled in the art. Examples include the amounts recommended by manufacturers for application. The maximum rate is an example of an amount of herbicide that would normally inhibit growth of a non-tolerant plant. For the present invention, the terms "herbicide-tolerant" and "herbicide-resistant" are used interchangeably and are intended to have an equivalent meaning and an equivalent scope. Similarly, the terms "herbicide-tolerance" and "herbicide-resistance" are used interchangeably and are intended to have an equivalent meaning and an equivalent scope. Similarly, the terms "tolerant" and "resistant" are used interchangeably and are intended to have an equivalent meaning and an equivalent scope. As used herein, in regard to an herbicidal composition useful in various embodiments hereof, terms such as TK-inhibiting herbicides, and the like, refer to those agronomically acceptable herbicide active ingredients (A.I.) recognized in the art. Similarly, terms such as fungicide, nematicide, pesticide, and the like, refer to other agronomically acceptable active ingredients recognized in the art.

When used in reference to a particular mutant enzyme or polypeptide, terms such as herbicide-tolerant and herbicide-tolerance refer to the ability of such enzyme or polypeptide to perform its physiological activity in the presence of an amount of an herbicide A.I. that would normally inactivate or inhibit the activity of the wild-type (non-mutant) version of said enzyme or polypeptide. For example, when used specifically in regard to a TK enzyme, it refers specifically to the ability to tolerate a TK-inhibitor. By "herbicide-tolerant wildtype or mutated TK protein" or "herbicide -resistant wildtype or mutated TK protein", it is intended that such a TK protein displays higher TK activity, relative to the TK activity of a wild-type TK protein, when in the presence of at least one herbicide that is known to interfere with TK activity and at a concentration or level of the herbicide that is known to inhibit the TK activity of the wild-type or mutated TK protein. Furthermore, the TK activity of such a herbicide-tolerant or herbicide-resistant wildtype or mutated TK protein may be referred to herein as "herbicide-tolerant" or "herbicide-resistant" TK activity.

As used herein, "recombinant," when referring to nucleic acid or polypeptide, indicates that such material has been altered as a result of human application of a recombinant technique, such as by polynucleotide restriction and ligation, by polynucleotide overlap-extension, or by genomic insertion or transformation. A gene sequence open reading frame is recombinant if that nucleotide sequence has been removed from it natural text and cloned into any type of artificial nucleic acid vector. The term recombinant also can refer to an organism having a recombinant material, e.g., a plant that comprises a recombinant nucleic acid can be considered a recombinant plant.

The term "transgenic plant" refers to a plant that comprises a heterologous polynucleotide. Preferably, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to refer to any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been so altered by the presence of heterologous nucleic acid including those transgenic organisms or cells initially so altered, as well as those created by crosses or asexual propagation from the initial transgenic organism or cell. In some embodiments, a "recombinant" organism is a "transgenic" organism. The term "transgenic" as used herein is not intended to encompass the alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods (e.g., crosses) or by naturally occurring events such as, e.g., self-fertilization, random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non- recombinant transposition, or spontaneous mutation.

As used herein, "mutagenized" refers to an organism or DNA thereof having alteration(s) in the biomolecular sequence of its native genetic material as compared to the sequence of the genetic material of a corresponding wild-type organism or DNA, wherein the alteration(s) in genetic material were induced and/or selected by human action. Examples of human action that can be used to produce a mutagenized organism or DNA include, but are not limited to, as illustrated in regard to herbicide tolerance: tissue culture of plant cells (e.g., calli) and selection thereof with herbicides (e.g., TK-inhibiting herbicides), treatment of plant cells with a chemical mutagen such as EMS and subsequent selection with herbicide(s); or by treatment of plant cells with x-rays and subsequent selection with herbicide(s). Any method known in the art can be used to induce mutations. Methods of inducing mutations can induce mutations in random positions in the genetic material or can induce mutations in specific locations in the genetic material (i.e., can be directed mutagenesis techniques), such as by use of a genoplasty technique.

As used herein, a "genetically modified organism" (GMO) is an organism whose genetic characteristics contain alteration(s) that were produced by human effort causing transfection that results in transformation of a target organism with genetic material from another or "source" organism, or with synthetic or modified-native genetic material, or an organism that is a descendant thereof that retains the inserted genetic material. The source organism can be of a different type of organism (e.g., a GMO plant can contain bacterial genetic material) or from the same type of organism (e.g., a GMO plant can contain genetic material from another plant). As used herein in regard to plants and other organisms, "recombinant," "transgenic," and "GMO" are considered synonyms and indicate the presence of genetic material from a different source; in contrast, "mutagenized" is used to refer to a plant or other organism, or the DNA thereof, in which no such transgenic material is present, but in which the native genetic material has become mutated so as to differ from a corresponding wild-type organism or DNA.

As used herein, "wild-type" or "corresponding wild-type plant" means the typical form of an organism or its genetic material, as it normally occurs, as distinguished from, e.g., mutagenized and/or recombinant forms. Similarly, by “control cell” or "similar, wild-type, plant, plant tissue, plant cell or host cell" is intended a plant, plant tissue, plant cell, or host cell, respectively, that lacks the herbicide-resistance characteristics and/or particular polynucleotide of the invention that are disclosed herein. The use of the term "wild-type" is not, therefore, intended to imply that a plant, plant tissue, plant cell, or other host cell lacks recombinant DNA in its genome, and/or does not possess herbicide-resistant characteristics that are different from those disclosed herein.

As used herein, "descendant" refers to any generation plant. In some embodiments, a descendant is a first, second, third, fourth, fifth, sixth, seventh, eight, ninth, or tenth generation plant.

As used herein, "progeny" refers to a first generation plant.

The term "seed" comprises seeds of all types, such as, for example, true seeds, caryopses, achenes, fruits, tubers, seedlings and similar forms. In the context of Brassica and Sinapis species, "seed" refers to true seed(s) unless otherwise specified. For example, the seed can be seed of transgenic plants or plants obtained by traditional breeding methods. Examples of traditional breeding methods can include cross-breeding, selfing, back-crossing, embryo rescue, in-crossing, out-crossing, inbreeding, selection, asexual propagation, and other traditional techniques as are known in the art.

Although exemplified with reference to specific plants or plant varieties and their hybrids, in various embodiments, the presently described methods using TK-inhibiting herbicides can be employed with a variety of commercially valuable plants. TK-inhibiting herbicides-tolerant plant lines described as useful herein can be employed in weed control methods either directly or indirectly, i. e. either as crops for herbicide treatment or as TK-inhibiting herbicides-tolerance trait donor lines for development, as by traditional plant breeding, to produce other varietal and/or hybrid crops containing such trait or traits. All such resulting variety or hybrids crops, containing the ancestral TK-inhibiting herbicides-tolerance trait or traits can be referred to herein as progeny or descendant of the ancestral, TK-inhibiting herbicides-tolerant line(s). Such resulting plants can be said to retain the "herbicide tolerance characteristic(s)" of the ancestral plant, i.e. meaning that they possess and express the ancestral genetic molecular components responsible for the trait.

In one aspect, the present invention provides a plant or plant part comprising a polynucleotide encoding a wildtype or mutated TK polypeptide, the expression of said polynucleotide confers to the plant or plant part tolerance to TK-inhibiting herbicides.

In a preferred embodiment, the plant has been previously produced by a process comprising recombinantly preparing a plant by introducing and over-expressing a wild-type or mutated TK transgene according to the present invention, as described in greater detail hereinfter.

In another preferred embodiment, the plant has been previously produced by a process comprising in situ mutagenizing plant cells, to obtain plant cells which express a mutated TK.

In another embodiment, the polynucleotide encoding the wildtype or mutated TK polypeptide polypeptide comprises the nucleic acid sequence set forth in SEQ ID NO: 182 or 183 a variant or derivative thereof.

In other embodiments, the wildtype or mutated TK polypeptide for use according to the present invention is a functional variant having, over the full-length of the variant, at least about 80%, illustratively, at least about 80%, 90%, 95%, 98%, 99% or more amino acid sequence identity to SEQ ID NO: 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, 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, or 181.

In another embodiment, the wildtype or mutated TK polypeptide for use according to the present invention is a functional fragment of a polypeptide having the amino acid sequence set forth in SEQ ID NO: 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, 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, or 181.

It is recognized that the TK polynucleotide molecules and TK polypeptides of the invention encompass polynucleotide molecules and polypeptides comprising a nucleotide or an amino acid sequence that is sufficiently identical to nucleotide sequences set forth in SEQ ID Nos: 182 or 183, or to the amino acid sequences set forth in SEQ ID Nos: 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, 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, or 181. The term "sufficiently identical" is used herein to refer to a first amino acid or nucleotide sequence that contains a sufficient or minimum number of identical or equivalent (e.g., with a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences have a common structural domain and/or common functional activity.

Generally, "sequence identity" refers to the extent to which two optimally aligned DNA or amino acid sequences are invariant throughout a window of alignment of components, e.g., nucleotides or amino acids. An "identity fraction" for aligned segments of a test sequence and a reference sequence is the number of identical components that are shared by the two aligned sequences divided by the total number of components in reference sequence segment, i.e., the entire reference sequence or a smaller defined part of the reference sequence. "Percent identity" is the identity fraction times 100. Optimal alignment of sequences for aligning a comparison window are well known to those skilled in the art and may be conducted by tools such as the local homology algorithm of Smith and Waterman, the homology alignment algorithm of Needleman and Wunsch, the search for similarity method of Pearson and Lipman, and preferably by computerized implementations of these algorithms such as GAP, BESTFIT, FASTA, and TFASTA available as part of the GCG. Wisconsin Package. (Accelrys Inc. Burlington, Mass.)

Polynucleotides and Oligonucleotides

By an "isolated polynucleotide", including DNA, RNA, or a combination of these, single or double stranded, in the sense or antisense orientation or a combination of both, dsRNA or otherwise, we mean a polynucleotide which is at least partially separated from the polynucleotide sequences with which it is associated or linked in its native state. Preferably, the isolated polynucleotide is at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated. As the skilled addressee would be aware, an isolated polynucleotide can be an exogenous polynucleotide present in, for example, a transgenic organism which does not naturally comprise the polynucleotide. Furthermore, the terms "polynucleotide(s)", "nucleic acid sequence(s)", "nucleotide sequence(s)", “nucleic acid(s)”, “nucleic acid molecule” are used interchangeably herein and refer to nucleotides, either ribonucleotides or deoxyribonucleotides or a combination of both, in a polymeric unbranched form of any length.

The term “mutated TK nucleic acid” refers to a TK nucleic acid having a sequence that is mutated from a wild-type TK nucleic acid and that confers increased TK-inhibiting herbicide tolerance to a plant in which it is expressed. Furthermore, the term “mutated transketolase (mutated TK)” refers to the replacement of an amino acid of the wild-type primary sequences of SEQ ID NO: 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, 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, or 181, or a variant, a derivative, a homologue, an orthologue, or paralogue thereof, with another amino acid. The expression "mutated amino acid" will be used below to designate the amino acid which is replaced by another amino acid, thereby designating the site of the mutation in the primary sequence of the protein.

In a preferred embodiment, the TK nucleotide sequence encoding a mutated TK comprises the sequence of SEQ ID NO: 182, or 183, or a variant or derivative thereof

Furthermore, it will be understood by the person skilled in the art that the TK nucleotide sequences encompasse homologues, paralogues and and orthologues of SEQ ID NO: 182 or 183, as defined hereinafter.

The term "variant" with respect to a sequence (e.g., a polypeptide or nucleic acid sequence such as – for example – a transcription regulating nucleotide sequence of the invention) is intended to mean substantially similar sequences. For nucleotide sequences comprising an open reading frame, variants include those sequences that, because of the degeneracy of the genetic code, encode the identical amino acid sequence of the native protein. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis and for open reading frames, encode the native protein comprising the sequence of SEQ ID NO: 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, 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, or 181, as well as those that encode a polypeptide having amino acid substitutions relative to the native protein, e.g. the mutated TK according to the present invention as disclosed herein. Generally, nucleotide sequence variants of the invention will have at least 30, 40, 50, 60, to 70%, e.g., preferably 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, to 79%, generally at least 80%, e.g., 81%-84%, at least 85%, e.g., 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, to 98% and 99% nucleotide “sequence identity” to the nucleotide sequence of SEQ ID NO: 182 or 183. The % identity of a polynucleotide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. Unless stated otherwise, the query sequence is at least 45 nucleotides in length, and the GAP analysis aligns the two sequences over a region of at least 45 nucleotides. Preferably, the query sequence is at least 150 nucleotides in length, and the GAP analysis aligns the two sequences over a region of at least 150 nucleotides. More preferably, the query sequence is at least 300 nucleotides in length and the GAP analysis aligns the two sequences over a region of at least 300 nucleotides. Even more preferably, the GAP analysis aligns the two sequences over their entire length.

Polypeptides

By "substantially purified polypeptide" or "purified" a polypeptide is meant that has been separated from one or more lipids, nucleic acids, other polypeptides, or other contaminating molecules with which it is associated in its native state. It is preferred that the substantially purified polypeptide is at least 60% free, more preferably at least 75% free, and more preferably at least 90% free from other components with which it is naturally associated. As the skilled addressee will appreciate, the purified polypeptide can be a recombinantly produced polypeptide. The terms "polypeptide" and "protein" are generally used interchangeably and refer to a single polypeptide chain which may or may not be modified by addition of non-amino acid groups. It would be understood that such polypeptide chains may associate with other polypeptides or proteins or other molecules such as co-factors. The terms "proteins" and "polypeptides" as used herein also include variants, mutants, modifications, analogous and/or derivatives of the polypeptides of the invention as described herein.

The % identity of a polypeptide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. The query sequence is at least 25 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 25 amino acids. More preferably, the query sequence is at least 50 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 50 amino acids. More preferably, the query sequence is at least 100 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 100 amino acids. Even more preferably, the query sequence is at least 250 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 250 amino acids. Even more preferably, the GAP analysis aligns the two sequences over their entire length.

With regard to a defined polypeptide, it will be appreciated that % identity figures higher than those provided above will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the TK polypeptide of the invention comprises an amino acid sequence which is at least 40%, more preferably at least 45%, more preferably at least 50%, more preferably at least 55%, more preferably at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least

93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to SEQ ID NO: 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, 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, or 181.

By "variant" polypeptide is intended a polypeptide derived from the protein of SEQ ID NO: 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, 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, or 181, by deletion (so-called truncation) or addition of one or more amino acids to the N-terminal and/or C-terminal end of the native protein; deletion or addition of one or more amino acids at one or more sites in the native protein; or substitution of one or more amino acids at one or more sites in the native protein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Methods for such manipulations are generally known in the art.

“Derivatives” of a protein encompass peptides, oligopeptides, polypeptides, proteins and enzymes having amino acid substitutions, deletions and/or insertions relative to the unmodified protein in question and having similar biological and functional activity as the unmodified protein from which they are derived. Thus, functional variants and fragments of the TK polypeptides, and nucleic acid molecules encoding them, also are within the scope of the present invention, and unless specifically described otherwise, irrespective of the origin of said polypeptide and irrespective of whether it occurs naturally. Various assays for functionality of a TK polypeptide can be employed. For example, a functional variant or fragment of the TK polypeptide can be assayed to determine its ability to confer TK-inhibiting herbicides detoxification. By way of illustration, a TK-inhibiting herbicides detoxification rate can be defined as a catalytic rate sufficient to provide a determinable increase in tolerance to TK-inhibiting herbicides in a plant or plant part comprising a recombinant polynucleotide encoding the variant or fragment of the TK polypeptide, wherein the plant or plant part expresses the variant or fragment at up to about 0.5%, illustratively, about 0.05 to about 0.5%, about 0.1 to about 0.4%, and about 0.2 to about 0.3%, of the total cellular protein relative to a similarly treated control plant that does not express the variant or fragment.

In a preferred embodiment, the wildtype or mutated TK polypeptide is a functional variant or fragment of a transketolase having the amino acid sequence set forth in SEQ ID NO: 1 or 2, wherein the functional variant or fragment has at least about 80% amino acid sequence identity to SEQ ID NO:1 or 2. For the avoidance of doubts, SEQ ID NO: 2 is identical to SEQ ID NO: 1 except that SEQ ID NO: 2 lacks the N-terminal transit peptide comprising amino acids 1-73 of SEQ ID NO: 1 [MAASSSLSTL SHHQTLLSHP KTHLPTTPAS SLLVPTTSSK VNGVLLKSTS SSRRLRVGSA SAVVRAAAVE ALE] (see Table 2a and 2b herein below for corresponding amino acid residues)

In other embodiments, the functional variant or fragment further has a TK-inhibiting herbicides detoxification rate defined as a catalytic rate sufficient to provide a determinable increase in tolerance to TK-inhibiting herbicides in a plant or plant part comprising a recombinant polynucleotide encoding the variant or fragment, wherein the plant or plant part expresses the variant or fragment at up to about 0.5% of the total cellular protein to a similarly treated control plant that does not express the variant or fragment.

“Homologues” of a protein encompass peptides, oligopeptides, polypeptides, proteins and enzymes having amino acid substitutions, deletions and/or insertions relative to the unmodified protein in question and having similar biological and functional activity as the unmodified protein from which they are derived.

In addition, one of ordinary skill in the art will further appreciate that changes can be introduced by mutation into the nucleotide sequences of the invention thereby leading to changes in the amino acid sequence of the encoded proteins without altering the biological activity of the proteins. Thus, for example, an isolated polynucleotide molecule encoding a mutated TK polypeptide having an amino acid sequence that differs from that of SEQ ID NO: 1 or 2 can be created by introducing one or more nucleotide substitutions, additions, or deletions into the corresponding nucleotide sequence, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Such variant nucleotide sequences are also encompassed by the present invention. For example, preferably, conservative amino acid substitutions may be made at one or more predicted preferably nonessential amino acid residues. A "nonessential" amino acid residue is a residue that can be altered from the wild-type sequence of a protein without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity.

A deletion refers to removal of one or more amino acids from a protein.

An insertion refers to one or more amino acid residues being introduced into a predetermined site in a protein. Insertions may comprise N-terminal and/or C-terminal fusions as well as intra-sequence insertions of single or multiple amino acids. Generally, insertions within the amino acid sequence will be smaller than N- or C-terminal fusions, of the order of about 1 to 10 residues. Examples of N- or C-terminal fusion proteins or peptides include the binding domain or activation domain of a transcriptional activator as used in the yeast two-hybrid system, phage coat proteins, (histidine)-6-tag, glutathione S-transferase-tag, protein A, maltose-binding protein, dihydrofolate reductase, Tag•100 epitope, c-myc epitope, FLAG®-epitope, lacZ, CMP (calmodulin-binding peptide), HA epitope, protein C epitope and VSV epitope.

A substitution refers to replacement of amino acids of the protein with other amino acids having similar properties (such as similar hydrophobicity, hydrophilicity, antigenicity, propensity to form or break α-helical structures or β-sheet structures). Amino acid substitutions are typically of single residues, but may be clustered depending upon functional constraints placed upon the polypeptide and may range from 1 to 10 amino acids; insertions will usually be of the order of about 1 to 10 amino acid residues. A conservative amino acid substitution is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Such substitutions would not be made for conserved amino acid residues, or for amino acid residues residing within a conserved motif. Conservative substitution tables are well known in the art (see for example Creighton (1984) Proteins. W.H. Freeman and Company (Eds).

Amino acid substitutions, deletions and/or insertions may readily be made using peptide synthetic techniques well known in the art, such as solid phase peptide synthesis and the like, or by recombinant DNA manipulation. Methods for the manipulation of DNA sequences to produce substitution, insertion or deletion variants of a protein are well known in the art. For example, techniques for making substitution mutations at predetermined sites in DNA are well known to those skilled in the art and include M13 mutagenesis, T7-Gen in vitro mutagenesis (USB, Cleveland, OH), QuickChange Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-directed mutagenesis or other site-directed mutagenesis protocols.

“Derivatives” further include peptides, oligopeptides, polypeptides which may, compared to the amino acid sequence of the naturally-occurring form of the protein, such as the protein of interest, comprise substitutions of amino acids with non-naturally occurring amino acid residues, or additions of non-naturally occurring amino acid residues. “Derivatives” of a protein also encompass peptides, oligopeptides, polypeptides which comprise naturally occurring altered (glycosylated, acylated, prenylated, phosphorylated, myristoylated, sulphated etc.) or non-naturally altered amino acid residues compared to the amino acid sequence of a naturally-occurring form of the polypeptide. A derivative may also comprise one or more non-amino acid substituents or additions compared to the amino acid sequence from which it is derived, for example a reporter molecule or other ligand, covalently or non-covalently bound to the amino acid sequence, such as a reporter molecule which is bound to facilitate its detection, and non-naturally occurring amino acid residues relative to the amino acid sequence of a naturally-occurring protein. Furthermore, “derivatives” also include fusions of the naturally-occurring form of the protein with tagging peptides such as FLAG, HIS6 or thioredoxin (for a review of tagging peptides, see Terpe, Appl. Microbiol. Biotechnol. 60, 523-533, 2003).

“Orthologues” and “paralogues” encompass evolutionary concepts used to describe the ancestral relationships of genes. Paralogues are genes within the same species that have originated through duplication of an ancestral gene; orthologues are genes from different organisms that have originated through speciation, and are also derived from a common ancestral gene. A non-limiting list of examples of such orthologues is shown in Table 1. It will be understood by the person skilled in the art that the sequences of SEQ ID NOs:2-181 as listed in Table 1 represent orthologues and paralogues to SEQ ID NO:1.

It is well-known in the art that paralogues and orthologues may share distinct domains harboring suitable amino acid residues at given sites, such as binding pockets for particular substrates or binding motifs for interaction with other proteins.

The term "domain" refers to a set of amino acids conserved at specific positions along an alignment of sequences of evolutionarily related proteins. While amino acids at other positions can vary between homologues, amino acids that are highly conserved at specific positions indicate amino acids that are likely essential in the structure, stability or function of a protein. Identified by their high degree of conservation in aligned sequences of a family of protein homologues, they can be used as identifiers to determine if any polypeptide in question belongs to a previously identified polypeptide family.

The term “motif” or “consensus sequence” refers to a short conserved region in the sequence of evolutionarily related proteins. Motifs are frequently highly conserved parts of domains, but may also include only part of the domain, or be located outside of conserved domain (if all of the amino acids of the motif fall outside of a defined domain).

Specialist databases exist for the identification of domains, for example, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)), or Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). A set of tools for in silico analysis of protein sequences is available on the ExPASy proteomics server (Swiss Institute of Bioinformatics (Gasteiger et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31:3784-3788(2003)). Domains or motifs may also be identified using routine techniques, such as by sequence alignment.

Methods for the alignment of sequences for comparison are well known in the art, such methods include GAP, BESTFIT, BLAST, FASTA and TFASTA. GAP uses the algorithm of Needleman and Wunsch ((1970) J Mol Biol 48: 443-453) to find the global (i.e. spanning the complete sequences) alignment of two sequences that maximizes the number of matches and minimizes the number of gaps. The BLAST algorithm (Altschul et al. (1990) J Mol Biol 215: 403-10) calculates percent sequence identity and performs a statistical analysis of the similarity between the two sequences. The software for performing BLAST analysis is publicly available through the National Centre for Biotechnology Information (NCBI). Homologues may readily be identified using, for example, the ClustalW multiple sequence alignment algorithm (version 1.83), with the default pairwise alignment parameters, and a scoring method in percentage (See Figure 1). Global percentages of similarity and identity may also be determined using one of the methods available in the MatGAT software package (Campanella et al., BMC Bioinformatics. 2003 Jul 10;4:29. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences.). Minor manual editing may be performed to optimise alignment between conserved motifs, as would be apparent to a person skilled in the art. Furthermore, instead of using full-length sequences for the identification of homologues, specific domains may also be used. The sequence identity values may be determined over the entire nucleic acid or amino acid sequence or over selected domains or conserved motif(s), using the programs mentioned above using the default parameters. For local alignments, the Smith-Waterman algorithm is particularly useful (Smith TF, Waterman MS (1981) J. Mol. Biol 147(1);195-7).

The proteins of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel (1985) PNAS, 82:488-492; Kunkel et al. (1987) Methods in Enzymol. 154:367-382; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) Techniques in Molecular Biology (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff et al. (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D. C), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be preferable.

Alternatively, variant nucleotide sequences can be made by introducing mutations randomly along all or part of a coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened to identify mutants that encode proteins that retain activity. For example, following mutagenesis, the encoded protein can be expressed recombinantly, and the activity of the protein can be determined using standard assay techniques.

The inventors of the present invention have found that by substituting one or more of the key amino acid residues of the TK enzyme of SEQ ID NO: 1 or 2, e.g. by employing one of the above described methods to mutate the TK encoding nucleic acids, the tolerance or resistance to particular TK-inhibiting herbicides could be remarkably increased Preferred substitutions of mutated TK are those that increase the herbicide tolerance of the plant, but leave the biological activitiy of the oxidase activity substantially unaffected.

Accordingly, in another object of the present invention refers to a TK polypeptide, comprising the sequence of SEQ ID NO: 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, 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, or 181, a variant, derivative, orthologue, paralogue or homologue thereof, the key amino acid residues of which is substituted by any other amino acid.

It will be understood by the person skilled in the art that amino acids located in a close proximity to the positions of amino acids mentioned below may also be substituted. Thus, in another embodiment the variant of SEQ ID NO: 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, 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, or 181, a variant, derivative, orthologue, paralogue or homologue thereof comprises a mutated TK, wherein an amino acid ±3, ±2 or ±1 amino acid positions from a key amino acid is substituted by any other amino acid.

Based on techniques well-known in the art, a highly characteristic sequence pattern can be developed, by means of which further of mutated TK candidates with the desired activity may be searched.

Searching for further mutated TK candidates by applying a suitable sequence pattern would also be encompassed by the present invention. It will be understood by a skilled reader that the present sequence pattern is not limited by the exact distances between two adjacent amino acid residues of said pattern. Each of the distances between two neighbours in the above patterns may, for example, vary independently of each other by up to ±10, ± 5, ±3, ±2 or ±1 amino acid positions without substantially affecting the desired activity.

Furthermore, by applying the method of site directed mutagenesis, in particular saturation mutagenes (see e.g. Schenk et al., Biospektrum 03/2006, pages 277-279), the inventors of the present invention have identified and generated specific amino acid subsitutions and combinations thereof, which - when introduced into a plant by transforming and expressing the respective mutated TK encoding nucleic acid - confer increased herbicide resistance or tolerance to a TK inhibiting herbicide to said plant.

Thus, in a particularly preferred embodiment, the variant or derivative of the mutated TK refers to a TK polypeptide comprising SEQ ID NO: 1, a orthologue, paralogue, or homologue thereof, wherein the amino acid sequence differs from the wildtype amino acid sequence of a TK polypeptide at one or more positions corresponding to the following positions of SEQ ID NO:1: 265, 267, 337, 342, 343, 458, 459, 460, 461, 463, 511, 512, 513, 514, 515, 544.

Examples of differences at these amino acid positions include, but are not limited to, one or more of the following:
the amino acid at or corresponding to position 265 is other than isoleucine;
the amino acid at or corresponding to position 267 is other than isoleucine;
the amino acid at or corresponding to position 337 is other than tyrosine;
the amino acid at or corresponding to position 342 is other than serine;
the amino acid at or corresponding to position 343 is other than alanine;
the amino acid at or corresponding to position 458 is other than leucine;
the amino acid at or corresponding to position 459 is other than alanine;

the amino acid at or corresponding to position 460 is other than serine;
the amino acid at or corresponding to position 461 is other than serine;
the amino acid at or corresponding to position 463 is other than methionine;
the amino acid at or corresponding to position 511 is other than Threonine;
the amino acid at or corresponding to position 512 is other than phenylalanine;
the amino acid at or corresponding to position 513 is other than phenylalanine;
the amino acid at or corresponding to position 514 is other than valine;
the amino acid at or corresponding to position 515 is other than phenylalanine
the amino acid at or corresponding to position 544 is other than leucine;

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 265 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 267 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 337 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Phe, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Arg, His, Lys, Asp, Glu, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 343 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Val, Leu, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 458 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of

SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 459 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Val, Leu, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 460 is Arg, His, Lys, Asp, Glu, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 461 is Arg, His, Lys, Asp, Glu, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 511 is Arg, His, Lys, Asp, Glu, Ser, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Phe, Met, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 512 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 513 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 514 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Leu, Ile, Phe, Met, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 515 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 544 is Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Ile, Met, Phe, Tyr, or Trp.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Arg, His, Lys, Asp, Glu, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Phe, Tyr, or Trp, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Arg, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is His, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Lys, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Asp, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Glu, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Thr, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Asn, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Gln, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Cys, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Gly, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Pro, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Ala, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Val, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Leu, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which:

the amino acid at or corresponding to position 342 is Ile, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Met, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Phe, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Tyr, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Trp, and the amino acid at or corresponding to position 343 is Pro.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Arg, His, Lys, Asp, Glu, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Val, Leu, Ile, Met, Phe, Tyr, or Trp, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Arg, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is His, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Lys, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Asp, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Glu, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Thr, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Asn, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Gln, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Cys, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Gly, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Pro, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Ala, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Val, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Leu, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Ile, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Met, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Phe, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Tyr, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 342 is Trp, and the amino acid at or corresponding to position 343 is Gly.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which:

the amino acid at or corresponding to position 463 is Val, Cys, Ala, Ser, Gly, and the amino acid at or corresponding to position 544 is Thr, Ala, Ser, Cys, Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Val, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Val, and the amino acid at or corresponding to position 544 is Ala.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Val, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Val, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Val, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 544 is Ala.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ala, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ala, and the amino acid at or corresponding to position 544 is Ala.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ala, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ala, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ala, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ser, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which:

the amino acid at or corresponding to position 463 is Ser, and the amino acid at or corresponding to position 544 is Ala.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ser, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ser, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Ser, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 544 is Ala.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, Cys, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Thr, Cys, Ser, Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which:

the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, Cys, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Thr, Cys, Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which:

the amino acid at or corresponding to position 463 is Gly, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Thr.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Cys.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 463 is Cys, and the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 544 is Ser.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Val.

In another preferred embodiment, the mutated TK polypeptide comprises a sequence of SEQ ID NO: 1 a variant, derivative, orthologue, paralogue of homologue thereof, in which: the amino acid at or corresponding to position 511 is Ser, and the amino acid at or corresponding to position 514 is Asn, and the amino acid at or corresponding to position 544 is Val.

It will be within the knowledge of the skilled artisan to identify conserved regions and motifs shared between the homologues, orthologues and paralogues encoded by SEQ ID NO: 182 or 183, such as those depicted in Table 1. Having identified such conserved regions that may represent suitable binding motifs, amino acids corresponding to the amino acids listed below in Table 2a and 2b, can be chosen to be subsituted by any other amino acid, for example by conserved amino acids, preferably by the amino acid substitutions described SUPRA using SEQ ID NO:1 as reference.

Table 2a and 2b provides an overview of positions in the orthologues and homologues to SEQ ID NO:1, i.e. the corresponding positions in SEQ ID NOs: 1 to 181.

Table2a

ID	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5	Pos 6	Pos 7	Pos 8
1	I265	I267	Y337	S342	A343	L458	A459	S460
2	I195	I197	Y267	S272	A273	L388	A389	S390
3	I265	I267	Y337	S342	A343	L458	A459	S460
4	I276	I278	Y347	S352	A353	L468	A469	S470
5	I264	I266	Y336	S341	A342	L457	A458	S459
6	I263	I265	Y335	S340	A341	L456	A457	S458
7	I235	I237	H307	A312	P313	L428	A429	P430
8	I264	I266	Y336	A341	A342	L457	A458	S459
9	I266	I268	Y338	S343	A344	L459	A460	S461
10	I273	I275	Y345	S350	A351	L466	A467	S468
11	I247	I249	Y319	A324	A325	L440	A441	S442
12	I268	I270	Y340	S345	A346	L461	A462	S463
13	I272	I274	Y344	S349	A350	L465	A466	S467
14	I266	I268	Y338	S343	A344	L459	A460	S461
15	I268	I270	Y340	S345	A346	L461	A462	S463
16	I267	I269	Y339	A344	A345	L460	A461	S462
17	I264	I266	Y336	S341	A342	L457	A458	S459
18	I269	I271	Y341	S346	A347	L462	A463	S464
19	I263	I265	Y335	S340	A341	L456	A457	S458
20	I271	I273	Y343	S348	A349	L464	A465	S466
21	I276	I278	Y348	A353	A354	L469	A470	S471
22	I268	I270	Y340	S345	A346	L461	A462	S463
23	I270	I272	Y342	S347	A348	L463	A464	S465
24	I266	I268	Y338	S343	A344	L459	A460	S461
25	I259	I261	Y331	A336	A337	L452	A453	S454
26	I265	I267	Y337	A342	A343	L458	A459	S460
27	I268	I270	Y340	S345	A346	L461	A462	S463
28	I269	I271	Y341	S346	A347	L462	A463	S464
29	I270	I272	Y342	S347	A348	L463	A464	S465
30	I263	I265	Y335	A340	A341	L456	A457	S458
31	I268	I270	Y340	S345	A346	L461	A462	S463
32	I264	I267	Y337	A342	A343	L458	A459	S460
33	I269	I271	Y341	S346	A347	L462	A463	S464
34	I263	I265	Y335	A340	A341	L456	A457	S458
35	I264	I267	Y337	A342	A343	L458	A459	S460

36	I253	I255	Y325	S330	A331	L446	A447	S448
37	I265	I267	Y337	S342	A343	L458	A459	S460
38	I200	I202	Y272	S277	A278	L393	A394	S395
39	I264	I266	Y336	A341	A342	L457	A458	S459
40	I268	I270	Y340	S345	A346	L461	A462	S463
41	I265	I267	Y337	S342	G343	L458	A459	S460
42	I255	I257	Y327	S332	A333	L448	A449	S450
43	I265	I267	Y337	A342	A343	L471	A472	S473
44	I265	I267	Y337	A342	A343	L458	A459	S460
45	I268	I270	Y340	S345	A346	L461	A462	S463
46	I254	I256	Y326	S331	A332	L447	A448	S449
47	I203	I205	Y275	S280	A281	L396	A397	S398
48	I260	I262	Y332	A337	A338	L453	A454	S455
49	I199	I201	Y271	S276	A277	L392	A393	S394
50	I276	I278	Y348	S353	A354	L469	A470	S471
51	I225	I227	Y297	S302	A303	L418	A419	S420
52	I259	I261	Y331	S336	A337	L452	A453	S454
53	I266	I268	Y338	A343	A344	L459	A460	S461
54	I259	I261	Y331	S336	A337	L452	A453	S454
55	I269	I271	Y341	S346	A347	L462	A463	S464
56	I269	I271	Y341	S346	A347	L462	A463	S464
57	I258	I260	Y330	S335	A336	L451	A452	S453
58	I158	I160	Y230	S235	A236	L351	A352	S353
59	I158	I160	Y230	S235	A236	L351	A352	S353
60	I158	I160	Y230	S235	A236	L351	A352	S353
61	I204	I206	Y276	A281	A282	L397	A398	S399
62	I274	I276	Y346	S351	A352	L467	A468	S469
63	I158	I160	Y230	A235	A236	L351	A352	S353
64	I158	I160	Y230	S235	A236	L351	A352	S353
65	I153	I155	Y225	S230	A231	L346	A347	S348
66	I260	I262	Y332	A337	A338	L453	A454	S455
67	I144	I146	Y216	A221	A222	L337	A338	S339
68	I158	I160	Y230	S235	A236	L351	A352	S353
69	I239	I155	Y225	S230	A231	L346	A347	S348
70	I236	I238	Y308	S313	A314	L429	A430	T431
71	I244	V311	Y381	G386	P387	L502	A503	P504
72	I215	I240	Y310	A315	A316	L431	A432	S433
73	I244	I246	Y316	S321	A322	L437	A438	T439
74	I215	I217	Y287	A292	A293	L408	A409	S410
75	I239	I241	Y311	S316	A317	L432	A433	T434

76	I239	I241	Y311	S316	A317	L432	A433	T434
77	I158	I160	Y230	A235	A236	L351	A352	S353
78	I246	I248	H318	S323	A324	L439	A440	S441
79	I239	I241	Y311	S316	A317	L432	A433	T434
80	I343	V345	Y415	G420	P421	L536	A537	P538
81	I251	I253	Y323	T328	A329	L444	A445	T446
82	I299	V301	Y371	G376	P377	L492	A493	P494
83	I232	I234	H304	S309	A310	L425	A426	S427
84	I153	I155	Y225	A230	A231	L346	A347	S348
85	I198	I200	Y270	N275	A276	L391	T392	L393
86	I253	I255	Y325	S330	A331	L446	A447	S448
87	I158	I160	Y230	S235	A236	L351	A352	S353
88	I146	I148	Y218	S223	A224	L339	A340	S341
89	I158	I160	Y230	N235	A236	L351	A352	T353
90	I158	I160	Y230	A235	A236	L351	A352	S353
91	I114	I116	Y186	S191	A192	L307	A308	S309
92	I94	I96	Y166	A171	A172	L287	A288	S289
93	I105	I107	Y177	S182	A183	L298	A299	S300
94	I123	V125	Y195	G200	A201	V316	A317	S318
95	I240	I242	H312	A317	P318	L433	A434	P435
96	I240	I242	H312	A317	P318	L433	A434	P435
97	I193	I195	A265	A270	A271	L386	T387	H388
98	I43	I45	Y115	S120	A121	L236	A237	S238
99	I193	I195	A265	A270	A271	L386	T387	H388
100	I193	I195	H265	A270	A271	L386	A387	P388
101	I193	I195	A265	A270	A271	L386	T387	H388
102	I193	I195	A265	A270	A271	L386	A387	K388
103	I194	I196	A266	A271	A272	L387	T388	H389
104	I190	I192	G261	N266	P267	L383	T384	G385
105	I188	I190	G259	A264	P265	L381	T382	G383
106	V192	I194	A263	S268	P269	L385	T386	G387
107	I193	I195	H264	S269	P270	L386	T387	G388
108	I185	I187	H255	S260	P261	L377	T378	G379
109	-	-	A264	G269	T270	L381	W382	N383
110	V199	A201	H270	N275	S276	L391	S392	V393
111	I205	C207	A275	A280	D281	L396	S397	P398
112	V214	V216	G285	A290	A291	L408	M409	E410
113	I190	I192	E261	S266	P267	L383	T384	G385
114	V219	V221	G290	A295	A296	L413	C414	E415
115	I188	I190	A259	A264	P265	L381	T382	P383

116	I227	I229	H298	N303	P304	L420	T421	G422
117	I190	I192	Y261	S266	P267	L384	T385	P386
118	I206	C208	A277	A282	A283	L400	S401	P402
119	I189	I191	A260	S265	A266	L383	T384	P385
120	I189	I191	E260	A265	P266	L382	T383	G384
121	I217	C219	A288	V293	A294	L411	S412	P413
122	V187	I189	A258	S263	P264	L380	T381	P382
123	I225	C227	A295	A300	A301	L418	T419	P420
124	I191	I193	H262	A267	P268	L384	T385	G386
125	I190	I192	C260	A265	P266	L382	T383	A384
126	V247	C249	F318	G323	A324	L441	V442	N443
127	V215	V217	G286	Q291	A292	L410	L411	E412
128	I189	C191	A260	A265	A266	L385	S386	P387
129	I201	C203	A272	A277	A278	L393	S394	P395
130	I209	C211	A280	V285	A286	L401	S402	P403
131	V199	C201	A270	G275	A276	L394	V395	T396
132	I191	C193	A261	A266	A267	L378	S379	P380
133	I188	I190	G259	A264	P265	L381	T382	G383
134	V192	I194	A263	S268	P269	L385	T386	G387
135	I193	I195	H264	S269	P270	L386	T387	G388
136	I190	I192	G261	N266	P267	L383	T384	S385
137	I225	C227	A296	S301	L302	L419	V420	N421
138	I185	I187	H255	S260	P261	L377	T378	G379
139	-	-	A264	G269	T270	L381	W382	N383
140	V199	A201	H270	N275	S276	L391	S392	V393
141	I205	C207	A275	A280	D281	L396	S397	P398
142	V214	V216	G285	A290	A291	L408	M409	E410
143	I193	I195	A264	A269	P270	L386	T387	G388
144	V200	C202	H271	S276	A277	L392	S393	V394
145	I190	I192	E261	S266	P267	L383	T384	G385
146	V219	V221	G290	A295	A296	L413	C414	E415
147	I226	I228	H297	N302	P303	L419	T420	P421
148	I192	C194	A263	A268	D269	L386	T387	P388
149	V224	C226	A295	A300	A301	L418	S419	P420
150	I199	C201	A270	A275	A276	L391	S392	P393
151	I188	I190	A259	A264	P265	L381	T382	P383
152	I227	I229	H298	N303	P304	L420	T421	G422
153	I190	I192	H260	S265	P266	L383	T384	P385
154	I190	I192	Y261	S266	P267	L384	T385	P386
155	V238	C240	A309	G314	G315	-	-	-

156	V199	I201	E263	E268	T269	L350	A351	G352
157	V201	I203	P265	S270	T271	L351	E352	G353
158	V201	I203	D265	T270	P271	L351	E352	G353
159	V202	I204	Q266	S271	N272	L353	E354	G355
160	V221	I223	E253	C258	N259	L272	K273	G274
161	V203	I205	D267	S272	T273	L354	E355	G356
162	V202	I204	K266	T271	T272	L352	E353	G354
163	V201	I203	E265	E270	S271	L352	E353	G354
164	V204	I206	E268	E273	S274	L355	E356	G357
165	V165	I167	K229	S234	T235	L315	A316	G317
166	V200	I202	A264	T269	P270	L352	E353	G354
167	V237	I239	K301	S306	P307	L386	E387	G388
168	V199	I201	E263	T268	C269	L354	E355	G356
169	V201	I203	E265	S270	E271	L351	E352	G353
170	L182	Q184	G253	K258	A259	V338	S339	N340
171	V205	I207	E269	S274	T275	L355	E356	G357
172	I197	I199	P266	K271	A272	L359	A360	G361
173	I196	I198	P265	K270	A271	L361	A362	G363
174	I185	I187	P254	K259	A260	L350	A351	G352
175	I195	I197	P264	K269	A270	L362	A363	S364
176	L185	Q187	D256	K261	P262	V341	S342	N343
177	V197	S199	S262	P267	L268	L375	A376	A377
178	I195	S197	S260	T265	A266	L353	M354	L355
179	I198	S200	A263	A268	A269	L358	K359	K360
180	V197	I199	I261	S266	T267	L347	E348	G349
181	I194	I196	P263	K268	P269	L360	A361	G362

Table 2b

ID	Pos 9	Pos 10	Pos 11	Pos 12	Pos 13	Pos 14	Pos 15	Pos 16
1	S461	M463	T511	F512	F513	V514	F515	L544
2	S391	M393	T441	F442	F443	V444	F445	L474
3	S461	M463	T511	F512	F513	V514	F515	L544
4	S471	M473	T521	F522	F523	V524	F525	L554
5	S460	M462	T510	F511	F512	V513	F514	L543
6	S459	M461	T509	F510	F511	V512	F513	L542
7	S431	M433	T481	F482	F483	I484	F485	L514
8	S460	M462	T510	F511	F512	V513	F514	L543
9	S462	M464	T512	F513	F514	V515	F516	L545
10	S469	M471	T519	F520	F521	V522	F523	L552
11	S443	M445	T493	F494	F495	V496	F497	L526
12	S464	M466	T514	F515	F516	V517	F518	L547

13	S468	M470	T518	F519	F520	V521	F522	L551
14	S462	M464	T512	F513	F514	V515	F516	L545
15	S464	M466	T514	F515	F516	V517	F518	L547
16	S463	M465	T513	F514	F515	V516	F517	L546
17	S460	M462	T510	F511	F512	V513	F514	L543
18	S465	M467	T515	F516	F517	V518	F519	L548
19	S459	M461	T509	F510	F511	V512	F513	L542
20	S467	M469	T517	F518	F519	V520	F521	L550
21	S472	M474	T522	F523	F524	V525	F526	L555
22	S464	M466	T514	F515	F516	V517	F518	L547
23	S466	M468	T516	F517	F518	V519	F520	L549
24	S462	M464	T512	F513	F514	V515	F516	L545
25	S455	M457	T505	F506	F507	V508	F509	L538
26	S461	M463	T511	F512	F513	V514	F515	L544
27	S464	M466	T514	F515	F516	V517	F518	L547
28	S465	M467	T515	F516	F517	V518	F519	L548
29	S466	M468	T516	F517	F518	V519	F520	L549
30	S459	M461	T509	F510	F511	V512	F513	L542
31	S464	M466	T514	F515	F516	V517	F518	L547
32	S461	M463	T511	F512	F513	V514	F515	L544
33	S465	M467	T515	F516	F517	V518	F519	L548
34	S459	M461	T509	F510	F511	V512	F513	L542
35	S461	M463	T511	F512	F513	V514	F515	L544
36	S449	M451	T499	F500	F501	V502	F503	L532
37	S461	M463	T511	F512	F513	V514	F515	L544
38	S396	M398	T446	F447	F448	V449	F450	L479
39	S460	M462	T510	F511	F512	V513	F514	L543
40	S464	M466	T514	F515	F516	V517	F518	L547
41	S461	M463	T511	F512	F513	V514	F515	L544
42	S451	M453	T501	F502	F503	V504	F505	L534
43	S474	M476	T524	F525	F526	V527	F528	L557
44	S461	M463	T511	F512	F513	V514	F515	L544
45	S464	M466	T514	F515	F516	V517	F518	L547
46	S450	M452	T500	F501	F502	V503	F504	L533
47	S399	M401	T449	F450	F451	V452	F453	L482
48	S456	M458	T506	F507	F508	V509	F510	L539
49	S395	M397	T445	F446	F447	V448	F449	L478
50	S472	M474	T521	F522	F523	V524	F525	L554
51	S421	M423	T471	F472	F473	V474	F475	L504
52	S455	M457	T505	F506	F507	V508	F509	L538

53	S462	M464	T512	F513	F514	V515	F516	L545
54	S455	M457	T505	F506	F507	V508	F509	L538
55	S465	L467	T515	F516	F517	V518	F519	L548
56	S465	L467	T515	F516	F517	V518	F519	L548
57	S454	M456	T504	F505	F506	V507	F508	L537
58	S354	M356	T404	F405	F406	V407	F408	L437
59	S354	M356	T404	F405	F406	V407	F408	L437
60	S354	M356	T404	F405	F406	V407	F408	L437
61	S400	M402	T450	F451	F452	V453	F454	L483
62	S470	M472	T520	F521	F522	V523	F524	L553
63	S354	M356	T404	F405	F406	V407	F408	L437
64	S354	M356	T404	F405	F406	V407	F408	L437
65	S349	M351	T399	F400	F401	V402	F403	L432
66	S456	M458	T506	F507	F508	V509	F510	L539
67	S340	M342	T390	F391	F392	V393	F394	L423
68	S354	M356	T404	F405	F406	V407	F408	L437
69	S349	M351	T399	F400	F401	V402	F403	L432
70	S432	M434	T482	F483	F484	V485	F486	L515
71	S505	M507	T555	Y556	F557	A558	F559	V588
72	S434	M436	T484	F485	F486	I487	F488	L517
73	S440	M442	T490	F491	F492	V493	F494	L523
74	S411	M413	T461	F462	F463	I464	F465	L494
75	S435	M437	T485	F486	F487	V488	F489	L518
76	S435	M437	T485	F486	F487	V488	F489	L518
77	S354	M356	T404	F405	F406	V407	F408	L437
78	S442	M444	T492	F493	F494	V495	F496	L525
79	S435	M437	T485	F486	F487	V488	F489	L518
80	S539	M541	T589	Y590	F591	A592	F593	V622
81	S447	M449	T497	F498	F499	V500	F501	L530
82	S495	M497	T545	Y546	F547	A548	F549	V578
83	S428	M430	T478	F479	F480	V481	F482	L511
84	S349	M351	T399	F400	F401	V402	F403	L432
85	S394	M396	T444	Y445	F446	V447	F448	L477
86	S449	M451	T499	F500	F501	V502	F503	-
87	S354	M356	T404	F405	F406	V407	F408	L437
88	S342	M344	T392	F393	F394	V395	F396	L425
89	S354	M356	T404	F405	F406	V407	F408	L437
90	S354	M356	T404	F405	F406	I407	F408	L437
91	S310	M312	T360	F361	F362	V363	F364	L393
92	S290	M292	T340	F341	F342	V343	F344	L373

93	S301	M303	T351	F352	F353	V354	F355	L384
94	S319	M321	T369	Y370	L371	A372	F373	V402
95	S436	L438	T486	F487	Y488	I489	F490	L519
96	S436	L438	T486	F487	Y488	I489	F490	L519
97	S389	L391	T439	F440	L441	V442	F443	L472
98	S239	M241	T289	F290	F291	V292	F293	L322
99	S389	N391	T439	F440	L441	V442	F443	L472
100	S389	L391	T439	F440	L441	I442	F443	L472
101	S389	L391	T439	F440	L441	V442	F443	L472
102	S389	N391	T439	F440	L441	V442	F443	L472
103	S390	Y392	T440	F441	L442	V443	F444	L473
104	S386	N388	T440	F441	L442	N443	F444	L473
105	S384	L386	T439	F440	L441	N442	F443	L472
106	S388	L390	T442	F443	L444	N445	F446	L475
107	S389	L391	T443	F444	L445	N446	F447	L476
108	S380	L382	T435	F436	L437	N438	F439	L468
109	S384	Q386	T435	F436	F437	M438	F439	E468
110	S394	N396	T455	F456	F457	M458	F459	E488
111	S399	N401	T460	F461	F462	M463	F464	T493
112	S411	F413	S467	Y468	L469	A470	F471	V500
113	S386	L388	T440	F441	F442	N443	F444	L473
114	S416	F418	S472	Y473	F474	I475	F476	V505
115	S384	L386	T439	F440	L441	N442	F443	L472
116	S423	L425	T479	F480	A481	N482	F483	L512
117	S387	L389	T442	F443	L444	N445	F446	L475
118	S403	H405	S464	F465	F466	M467	F468	M497
119	S386	L388	T441	F442	L443	N444	F445	L474
120	S385	L387	T439	F440	L441	N442	F443	L472
121	S414	N416	S475	F476	F477	M478	F479	T508
122	S383	N385	T438	F439	L440	N441	F442	L471
123	S421	H423	S482	F483	F484	M485	F486	T515
124	S387	L389	T441	F442	L443	N444	F445	L474
125	S385	L387	T439	F440	L441	N442	F443	L472
126	S444	K446	T496	F497	F498	M499	F500	E529
127	S413	F415	T469	F470	F471	M472	F473	I502
128	S388	N390	S449	F450	F451	M452	F453	T482
129	S396	N398	S457	F458	F459	M460	F461	T490
130	S404	N406	S465	F466	F467	M468	F469	M498
131	S397	K399	T451	F452	L453	M454	F455	E484
132	S381	N383	S442	F443	F444	M445	F446	L475

133	S384	L386	T439	F440	L441	N442	F443	L472
134	S388	L390	T442	F443	L444	N445	F446	L475
135	S389	L391	T443	F444	L445	N446	F447	L476
136	S386	N388	T440	F441	F442	N443	F444	L473
137	S422	K424	T474	F475	F476	M477	F478	E507
138	S380	L382	T435	F436	L437	N438	F439	L468
139	S384	Q386	T435	F436	F437	M438	F439	E468
140	S394	N396	T455	F456	F457	M458	F459	E488
141	S399	N401	T460	F461	F462	M463	F464	T493
142	S411	F413	S467	Y468	L469	A470	F471	V500
143	S389	L391	T443	F444	L445	N446	F447	L476
144	S395	L397	T456	F457	Y458	M459	F460	A489
145	S386	L388	T440	F441	F442	N443	F444	L473
146	S416	F418	S472	Y473	F474	I475	F476	V505
147	S422	L424	T478	F479	L480	N481	F482	L511
148	S389	N391	S450	F451	F452	M453	F454	T483
149	S421	N423	S482	F483	F484	M485	F486	T515
150	S394	N396	S455	F456	F457	M458	F459	I488
151	S384	L386	T439	F440	L441	N442	F443	L472
152	S423	L425	T479	F480	A481	N482	F483	L512
153	S386	L388	T441	F442	L443	N444	F445	V474
154	S387	L389	T442	F443	L444	N445	F446	L475
155	-	-	T434	F435	F436	M437	F438	E467
156	S353	G355	T398	F399	S400	A401	F402	E431
157	S354	G356	T397	F398	S399	A400	F401	E430
158	S354	G356	T397	F398	A399	A400	F401	E430
159	S356	G358	T399	F400	S401	A402	F403	E432
160	S275	G277	-	-	-	-	-	-
161	S357	G359	T402	F403	T404	A405	F406	E435
162	S355	G357	T398	F399	S400	A401	F402	E431
163	S355	G357	T400	F401	S402	A403	F404	E433
164	S358	G360	T403	F404	S405	A406	F407	E436
165	S318	G320	T361	F362	S363	A364	F365	E394
166	S355	G357	T398	F399	G400	A401	F402	E431
167	S389	G391	T433	F434	S435	A436	F437	E466
168	S357	G359	T400	F401	A402	A403	F404	D433
169	S354	G356	T397	F398	S399	A400	F401	D430
170	S341	Y343	T383	F384	A385	A386	F387	I416
171	S358	G360	T401	F402	S403	A404	F405	D434
172	S362	K364	G404	F405	G406	M407	F408	V438

173	S364	K366	D406	F407	G408	V409	F410	V440
174	S353	K355	D395	F396	G397	V398	F399	V429
175	S365	K367	D407	F408	G409	V410	F411	V441
176	S344	Y346	T386	F387	A388	A389	F390	I419
177	D378	K380	S420	F421	A422	S423	F424	P454
178	D356	G358	S398	F399	A400	C401	F402	P432
179	D361	G363	S403	F404	A405	C406	F407	P437
180	S350	G352	T393	F394	S395	A396	F397	D426
181	S363	K365	D405	F406	G407	V408	F409	V439

Another object refers to a method of identifying a nucleotide sequence encoding a mutated TK which is resistant or tolerant to a TK-inhibiting herbicide, the method comprising:

a) generating a library of mutated TK-encoding nucleic acids,
b) screening a population of the resulting mutated TK-encoding nucleic acids by expressing each of said nucleic acids in a cell or plant and treating said cell or plant with a TK-inhibiting herbicide,
c) comparing the TK-inhibiting herbicide-tolerance levels provided by said population of mutated TK encoding nucleic acids with the TK-inhibiting herbicide-tolerance level provided by a control TK-encoding nucleic acid,
d) selecting at least one mutated TK-encoding nucleic acid that provides a significantly increased level of tolerance to a TK-inhibiting herbicide as compared to that provided by the control TK-encoding nucleic acid.

In a preferred embodiment, the mutated TK-encoding nucleic acid selected in step d) provides at least 2-fold as much resistance or tolerance of a cell or plant to a TK-inhibiting herbicide as compared to that provided by the control TK-encoding nucleic acid.

In a further preferred embodiment, the mutated TK-encoding nucleic acid selected in step d) provides at least 2-fold, at least 5-fold, at least 10-fold, at least 20-fold, at least 50-fold, at least 100-fold, at least 500-fold, as much resistance or tolerance of a cell or plant to a TK-inhibiting herbicide as compared to that provided by the control TK-encoding nucleic acid.

The resistance or tolerance can be determined by generating a transgenic plant or host cell, preferably a plant cell, comprising a nucleic acid sequence of the library of step a) and comparing said transgenic plant with a control plant or host cell, preferably a plant cell.

Another object refers to a method of identifying a plant or algae containing a nucleic acid comprising a nucleotide sequence encoding a wild-type or mutated TK which is resistant or tolerant to a TK-inhibiting herbicide, the method comprising:

a) identifying an effective amount of a TK-inhibiting herbicide in a culture of plant cells or green algae that leads to death of said cells.
b) treating said plant cells or green algae with a mutagenizing agent,
c) contacting said mutagenized cells population with an effective amount of TK-inhibiting

herbicide, identified in a),

d) selecting at least one cell surviving these test conditions,

e) PCR-amplification and sequencing of TK genes from cells selected in d) and comparing such sequences to wild-type TK gene sequences, respectively.

In a preferred embodiment, said mutagenizing agent is ethylmethanesulfonate (EMS).

Many methods well known to the skilled artisan are available for obtaining suitable candidate nucleic acids for identifying a nucleotide sequence encoding a mutated TK from a variety of different potential source organisms including microbes, plants, fungi, algae, mixed cultures etc. as well as environmental sources of DNA such as soil. These methods include inter alia the preparation of cDNA or genomic DNA libraries, the use of suitably degenerate oligonucleotide primers, the use of probes based upon known sequences or complementation assays (for example, for growth upon tyrosine) as well as the use of mutagenesis and shuffling in order to provide recombined or shuffled mutated TK-encoding sequences.

Nucleic acids comprising candidate and control TK encoding sequences can be expressed in yeast, in a bacterial host strain, in an alga or in a higher plant such as tobacco or Arabidopsis and the relative levels of inherent tolerance of the TK encoding sequences screened according to a visible indicator phenotype of the transformed strain or plant in the presence of different concentrations of the selected TK-inhibiting herbicide. Dose responses and relative shifts in dose responses associated with these indicator phenotypes (formation of brown color, growth inhibition, herbicidal effect etc) are conveniently expressed in terms, for example, of GR50 (concentration for 50% reduction of growth) or MIC (minimum inhibitory concentration) values where increases in values correspond to increases in inherent tolerance of the expressed TK. For example, in a relatively rapid assay system based upon transformation of a bacterium such as E. coli, each mutated TK encoding sequence may be expressed, for example, as a DNA sequence under expression control of a controllable promoter such as the lacZ promoter and taking suitable account, for example by the use of synthetic DNA, of such issues as codon usage in order to obtain as comparable a level of expression as possible of different TK sequences. Such strains expressing nucleic acids comprising alternative candidate TK sequences may be plated out on different concentrations of the selected TK-inhibiting herbicide in, optionally, a tyrosine supplemented medium and the relative levels of inherent tolerance of the expressed TK enzymes estimated on the basis of the extent and MIC for inhibition of the formation of the brown, ochronotic pigment.

In another embodiment, candidate nucleic acids are transformed into plant material to generate a transgenic plant, regenerated into morphologically normal fertile plants which are then measured for differential tolerance to selected TK-inhibiting herbicides as described in the Example section hereinafter. Many suitable methods for transformation using suitable selection markers such as kanamycin, binary vectors such as from Agrobacterium and plant regeneration as, for example, from tobacco leaf discs are well known in the art. Optionally, a control population of plants is likewise transformed with a nucleic acid expressing the control TK.

Alternatively, an untransformed dicot plant such as Arabidopsis or Tobacco can be used as a control since this, in any case, expresses its own endogenous TK. The average, and distribution, of herbicide tolerance levels of a range of primary plant transformation events or their progeny to TK-inhibiting herbicides described supra are evaluated in the normal manner based upon plant damage, meristematic bleaching symptoms etc. at a range of different concentrations of herbicides. These data can be expressed in terms of, for example, GR50 values derived from dose/response curves having "dose" plotted on the x-axis and "percentage kill", "herbicidal effect", "numbers of emerging green plants" etc. plotted on the y-axis where increased GR50 values correspond to increased levels of inherent tolerance of the expressed TK. Herbicides can suitably be applied pre-emergence or post-emergence.

Another object of the present invention refers to an isolated nucleic acid encoding a mutated TK as disclosed SUPRA, wherein the nucleic acid comprises the nucleotide sequence of SEQ ID NO: 182, or 183, or a variant or derivative thereof.

In one embodiment, the nucleic acid is identifiable by a method as defined above.

In a preferred embodiment, the encoded mutated TK is a variant of SEQ ID NO: 1, which includes one or more of the following:
the amino acid corresponding to position 265 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Met,Phe,Tyr, or Trp;
the amino acid corresponding to position 267 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Met,Phe,Tyr, or Trp;
the amino acid corresponding to position 337 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Phe, or Trp;
the amino acid corresponding to position 342 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Phe,Tyr, or Trp;
the amino acid corresponding to position 343 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Val,Leu,Ile,Met,Phe,Tyr, or Trp,
the amino acid corresponding to position 458 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Ile,Met,Phe,Tyr, or Trp,
the amino acid corresponding to position 459 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Val,Leu,Ile,Met,Phe,Tyr, or Trp
the amino acid corresponding to position 460 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Phe,Tyr, or Trp,
the amino acid corresponding to position 461 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Phe,Tyr, or Trp.
the amino acid corresponding to position 463 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Phe,Tyr, or Trp.
the amino acid corresponding to position 511 of SEQ ID NO:1 is Arg,His,Lys,Asp,Glu,Ser,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Phe,Met,Tyr, or Trp.
the amino acid corresponding to position 512 of SEQ ID NO:1 is

Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Tyr, or Trp.
the amino acid corresponding to position 513 of SEQ ID NO:1 is
Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Tyr, or Trp.
the amino acid corresponding to position 514 of SEQ ID NO:1 is
Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Leu,Ile,Phe,Met,Tyr, or Trp.
the amino acid corresponding to position 515 of SEQ ID NO:1 is
Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Leu,Ile,Met,Tyr, or Trp.
the amino acid corresponding to position 544 of SEQ ID NO:1 is
Arg,His,Lys,Asp,Glu,Ser,Thr,Asn,Gln,Cys,Gly,Pro,Ala,Val,Ile,Met,Phe,Tyr, or Trp

In other aspects, the present invention encompasses a progeny or a descendant of a TK-inhibiting herbicides-tolerant plant of the present invention as well as seeds derived from the TK-inhibiting herbicides-tolerant plants of the invention and cells derived from the TK-inhibiting herbicides-tolerant plants of the invention.

In some embodiments, the present invention provides a progeny or descendant plant derived from a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, wherein the progeny or descendant plant comprises in at least some of its cells the recombinant polynucleotide operably linked to the promoter, the expression of the wildtype or mutated TK polypeptide conferring to the progeny or descendant plant tolerance to the TK-inhibiting herbicides.

In one embodiment, seeds of the present invention preferably comprise the TK-inhibiting herbicides-tolerance characteristics of the TK-inhibiting herbicides-tolerant plant. In other embodiments, a seed is capable of germination into a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the progeny or descendant plant tolerance to the TK-inhibiting herbicides.

In some embodiments, plant cells of the present invention are capable of regenerating a plant or plant part. In other embodiments, plant cells are not capable of regenerating a plant or plant part. Examples of cells not capable of regenerating a plant include, but are not limited to, endosperm, seed coat (testa & pericarp), and root cap.

In another embodiment, the present invention provides a plant cell of or capable of regenerating a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to the TK-inhibiting herbicides, wherein the plant cell comprises the recombinant polynucleotide operably linked to a promoter.

In other embodiments, the present invention provides a plant cell comprising a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the cell tolerance to the TK-inhibiting herbicides.

In another embodiment, the invention refers to a plant cell transformed by a nucleic acid encoding a mutated TK polypeptide according to the present invention or to a plant cell which has been mutated to obtain a plant expressing a nucleic acid encoding a mutated TK polypeptide according to the present invention, wherein expression of the nucleic acid in the plant cell results in increased resistance or tolerance to a TK-inhibiting herbicide as compared to a wild type variety of the plant cell. Preferably, the mutated TK polypeptide encoding nucleic acid comprises a polynucleotide sequence selected from the group consisting of: a) a polynucleotide as shown in SEQ ID NO: 182, or 183, or a variant or derivative thereof; b) a polynucleotide encoding a polypeptide as shown in SEQ ID NO: 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, 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, or 181, or a variant or derivative thereof; c) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) or b); and d) a polynucleotide complementary to the polynucleotide of any of a) through c).

In some aspects, the present invention provides a plant product prepared from the TK-inhibiting herbicides-tolerant plants hereof. In some embodiments, examples of plant products include, without limitation, grain, oil, and meal. In one embodiment, a plant product is plant grain (e.g., grain suitable for use as feed or for processing), plant oil (e.g., oil suitable for use as food or biodiesel), or plant meal (e.g., meal suitable for use as feed).

In one embodiment, a plant product prepared from a plant or plant part is provided, wherein the plant or plant part comprises in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the a plant or plant part tolerance to the TK-inhibiting herbicides.

In another embodiment, the invention refers to a method of producing a transgenic plant cell with an increased resistance to a TK-inhibiting herbicide as compared to a wild type variety of the plant cell comprising, transforming the plant cell with an expression cassette comprising a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide.

In another embodiment, the invention refers to a method of producing a transgenic plant comprising, (a) transforming a plant cell with an expression cassette comprising a a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, and (b) generating a plant with an increased resistance to TK-inhibiting herbicide from the plant cell.

In some aspects, the present invention provides a method for producing a TK-inhibiting herbicides-tolerant plant. In one embodiment, the method comprises: regenerating a plant from a plant cell transformed with a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to the TK-inhibiting herbicides.

The term “expression/expressing” or “gene expression” means the transcription of a specific gene or specific genes or specific genetic construct. The term “expression” or “gene expression” in particular means the transcription of a gene or genes or genetic construct into structural RNA (rRNA, tRNA) or mRNA with or without subsequent translation of the latter into a protein. The process includes transcription of DNA and processing of the resulting mRNA product.

To obtain the desired effect, i.e. plants that are tolerant or resistant to the TK-inhibiting herbicide derivative herbicide of the present invention, it will be understood that the at least one nucleic acid is “over-expressed” by methods and means known to the person skilled in the art.

The term “increased expression” or “overexpression” as used herein means any form of expression that is additional to the original wild-type expression level. Methods for increasing expression of genes or gene products are well documented in the art and include, for example, overexpression driven by appropriate promoters, the use of transcription enhancers or translation enhancers. Isolated nucleic acids which serve as promoter or enhancer elements may be introduced in an appropriate position (typically upstream) of a non-heterologous form of a polynucleotide so as to upregulate expression of a nucleic acid encoding the polypeptide of interest. For example, endogenous promoters may be altered in vivo by mutation, deletion, and/or substitution (see, Kmiec, US 5,565,350; Zarling et al., WO9322443), or isolated promoters may be introduced into a plant cell in the proper orientation and distance from a gene of the present invention so as to control the expression of the gene. If polypeptide expression is desired, it is generally desirable to include a polyadenylation region at the 3'-end of a polynucleotide coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The 3' end sequence to be added may be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene. An intron sequence may also be added to the 5' untranslated region (UTR) or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates in the cytosol. Inclusion of a spliceable intron in the transcription unit in both plant and animal expression constructs has been shown to increase gene expression at both the mRNA and protein levels up to 1000-fold (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Such intron enhancement of gene expression is typically greatest when placed near the 5' end of the transcription unit. Use of the maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. For general information see: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

Where appropriate, nucleic acid sequences may be optimized for increased expression in a transformed plant. For example, coding sequences that comprise plant-preferred codons for improved expression in a plant can be provided. See, for example, Campbell and Gowri (1990) Plant Physiol., 92: 1-11 for a discussion of host-preferred codon usage. Methods also are known in the art for preparing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray et al. (1989) Nucleic Acids Res. 17:477-498, herein incorporated by reference.

Consequently, wildtype/mutated TK nucleic acids of the invention are provided in expression cassettes for expression in the plant of interest. The cassette will include regulatory sequences operably linked to a mutated TK nucleic acid sequence of the invention. The term "regulatory element" as used herein refers to a polynucleotide that is capable of regulating the transcription of an operably linked polynucleotide. It includes, but not limited to, promoters, enhancers, introns, 5' UTRs, and 3' UTRs. By "operably linked" is intended a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame. The cassette may additionally contain at least one additional gene to be cotransformed into the organism. Alternatively, the additional gene(s) can be provided on multiple expression cassettes.

Such an expression cassette is provided with a plurality of restriction sites for insertion of the wildtype/mutated TK nucleic acid sequence to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes. The expression cassette of the present invention will include in the 5'-3' direction of transcription, a transcriptional and translational initiation region (i.e., a promoter), a wildtype/mutated TK encoding nucleic acid sequence of the invention, and a transcriptional and translational termination region (i.e., termination region) functional in plants. The promoter may be native or analogous, or foreign or heterologous, to the plant host and/or to the wildtype/mutated TK nucleic acid sequence of the invention. Additionally, the promoter may be the natural sequence or alternatively a synthetic sequence. Where the promoter is "foreign" or "heterologous" to the plant host, it is intended that the promoter is not found in the native plant into which the promoter is introduced. Where the promoter is "foreign" or "heterologous" to the wildtype/mutated TK nucleic acid sequence of the invention, it is intended that the promoter is not the native or naturally occurring promoter for the operably linked wildtype/mutated TK nucleic acid sequence of the invention. As used herein, a chimeric gene comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence. While it may be preferable to express the wildtype/mutated TK nucleic acids of the invention using heterologous promoters, the native promoter sequences may be used. Such constructs would change expression levels of the wildtype/mutated TK protein in the plant or plant cell. Thus, the phenotype of the plant or plant cell is altered.

The termination region may be native with the transcriptional initiation region, may be native with the operably linked wildtype/mutated TK sequence of interest, may be native with the plant host, or may be derived from another source (i.e., foreign or heterologous to the promoter, the wildtype/mutated TK nucleic acid sequence of interest, the plant host, or any combination thereof). Convenient termination regions are available from the Ti-plasmid of A. tumefaciens , such as the octopine synthase and nopaline synthase termination regions. See also Guerineau et al. (1991) Mol. Gen. Genet. 262: 141-144; Proudfoot (1991) Cell 64:671-674; Sanfacon et al. (1991) Genes Dev. 5: 141-149; Mogen et al. (1990) Plant Cell 2: 1261-1272; Munroe et al. (1990) Gene 91: 151-158; Ballas t al. (1989) Nucleic Acids Res. 17:7891-7903; and Joshi et al. (1987) Nucleic Acid Res. 15:9627-9639. Where appropriate, the gene(s) may be optimized for increased expression in the transformed plant. That is, the genes can be synthesized using plant-preferred codons for improved expression. See, for example, Campbell and Gowri (1990) Plant Physiol. 92: 1-11 for a discussion of host-preferred codon usage. Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray et al. (1989) Nucleic Acids Res. 17:477-498, herein incorporated by reference.

While the polynucleotides of the invention may find use as selectable marker genes for plant transformation, the expression cassettes of the invention can include another selectable marker gene for the selection of transformed cells. Selectable marker genes, including those of the present invention, are utilized for the selection of transformed cells or tissues. Marker genes include, but are not limited to, genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). See generally, Yarranton (1992) Curr. Opin. Biotech. 3 :506-511 ; Christophers on et al (1992) Proc. Natl. Acad. ScL USA 89:6314-6318; Yao et al. (1992) Cell 71:63-72; Reznikoff (1992) Mol Microbiol 6:2419-2422; Barkley et al (1980) in The Operon, pp. 177-220; Hu et al (1987) Cell 48:555-566; Brown et al (1987) Cell 49:603-612; Figge et al (1988) Cell 52:713-722; Deuschle et al (1989) Proc. Natl Acad. AcL USA 86:5400-5404; Fuerst et al (1989) Proc. Natl Acad. ScL USA 86:2549-2553; Deuschle et al (1990) Science 248:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines et al (1993) Proc. Natl Acad. ScL USA 90: 1917-1921; Labow et al (1990) Mol Cell Biol 10:3343-3356; Zambretti et al (1992) Proc. Natl Acad. ScL USA 89:3952-3956; Bairn et al (1991) Proc. Natl Acad. ScL USA 88:5072-5076; Wyborski et al (1991) Nucleic Acids Res. 19:4647-4653; Hillenand-Wissman (1989) Topics Mol Struc. Biol 10: 143- 162; Degenkolb et al (1991) Antimicrob. Agents Chemother. 35: 1591-1595; Kleinschnidt et al (1988) Biochemistry 27: 1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen et al (1992) Proc. Natl Acad. ScL USA 89:5547- 5551; Oliva et al (1992) Antimicrob. Agents Chemother. 36:913-919; Hlavka et al (1985) Handbook of Experimental Pharmacology, Vol. 78 (Springer-Verlag, Berlin); Gill et al (1988) Nature 334:721-724. Such disclosures are herein incorporated by reference. The above list of selectable marker genes is not meant to be limiting. Any selectable marker gene can be used in the present invention.

Further, additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well - characterized sequences that may be deleterious to gene expression. The G-C content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. Also, if desired, sequences can be readily modified to avoid predicted hairpin secondary mRNA structures. Nucleotide sequences for enhancing gene expression can also be used in the plant expression vectors. These include, for example, introns of the maize Adh gene Adh1-S intron 1, 2, and 6 (Callis et al. Genes and Development 1 : 1183-1200, 1987), and leader sequences, (W-sequence) from the Tobacco Mosaic virus (TMV), Maize Chlorotic Mottle Virus and Alfalfa Mosaic Virus (Gallie et al. Nucleic Acid Res. 15:8693-8711 , 1987 and Skuzeski et al. Plant Mol. Biol. 15:65-79, 1990). The first intron from the shrunken-1 locus of maize has been shown to increase expression of genes in chimeric gene constructs. U.S. Pat. Nos. 5,424,412 and 5,593,874 disclose the use of specific introns in gene expression constructs, and Gallie et al. (Plant Physiol. 106:929-939, 1994) also have shown that introns are useful for regulating gene expression on a tissue specific basis. To further enhance or to optimize gene expression, the plant expression vectors of the invention also may contain DNA sequences containing matrix attachment regions (MARs). Plant cells transformed with such modified expression systems, then, may exhibit overexpression or constitutive expression of a nucleotide sequence of the invention.

The invention further provides an isolated recombinant expression vector comprising the expression cassette containing a wildtype/mutated TK nucleic acid nucleic acid as described above, wherein expression of the vector in a host cell results in increased tolerance to a TK-inhibiting herbicide as compared to a wild type variety of the host cell. As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid,” which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “expression vectors.” In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, “plasmid” and “vector” can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses, and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cells and those that direct expression of the nucleotide sequence only in certain host cells or under certain conditions. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce polypeptides or peptides, including fusion polypeptides or peptides, encoded by nucleic acids as described herein (e.g., mutated TK polypeptides, fusion polypeptides, etc.)

Expression vectors may additionally contain 5' leader sequences in the expression construct. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyo carditis 5' noncoding region) (Elroy-Stein et al. (1989) PNAS, 86:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Gallie et al. (1995) Gene 165(2):233-238), MDMV leader (Maize Dwarf Mosaic Virus) (Virology 154:9-20), and human immunoglobulin heavy-chain binding protein (BiP) (Macejak et al. (1991) Nature 353:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling et al. (1987) Nature 325:622-625); tobacco mosaic virus leader (TMV) (Gallie et al. (1989) in Molecular Biology of RNA, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel et al. (1991) Virology 81 :382-385). See also, Della-Cioppa et al. (1987) Plant Physiol. 84:965-968.

Other methods known to enhance translation also can be utilized, for example, introns, and the like. In preparing an expression vector, the various nucleic acid fragments may be manipulated, so as to provide for the nucleic acid sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the nucleic acid fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous nucleic acid, removal of restriction sites, or the like. For this purpose, in vitro mutagenesis, primer repair, restriction, annealing, resubstitutions, e.g., transitions and transversions, may be involved.

A number of promoters can be used in the practice of the invention. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, tissue-preferred, or other promoters for expression in plants.

Constitutive promoters include, for example, the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell et al. (1985) Nature 313:810-812); rice actin (McElroy et al. (1990) Plant Cell 2: 163-171); ubiquitin (Christensen et al. (1989) Plant Mol. Biol. 12:619-632 and Christensen et al. (1992) Plant Mol. Biol. 18:675-689); pEMU (Last et al. (1991) Theor. Appl. Genet. 81 :581- 588); MAS (Velten et al. (1984) EMBO J. 3:2723- 2730); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, U.S. Patent Nos. 5,608,149; 5,608, 144; 5,604,121 ; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

Tissue-preferred promoters can be utilized to target enhanced expression within a particular plant tissue. Such tissue-preferred promoters include, but are not limited to, leaf- preferred promoters, root-preferred promoters, seed- preferred promoters, and stem-preferred promoters. Some examples of tissue-preferred promoters are described by, e.g., Yamamoto et al. (1997) Plant J. 12(2):255-265; Kawamata et al. (1997) Plant Cell Physiol. 38(7):792-803; Hansen et al. (1997) Mol. Gen Genet. 254(3):337-343; Russell et al. (1997) Transgenic Res. 6(2): 157-168; Rinehart et al. (1996) Plant Physiol. 1 12(3): 1331-1341 ; Van Camp et al. (1996) Plant Physiol. 112(2):525-535; Canevascini et al. (1996) Plant Physiol. 1 12(2):513- 524; Yamamoto et al. (1994) Plant Cell Physiol. 35(5):773-778; Lam (1994) Results Probl. Cell Differ. 20:181- 196; Orozco ef al. (1993) Plant Mol Biol. 23(6): 1 129-1138; Matsuoka et al. (1993) Voc Natl. Acad. ScL USA 90(20):9586-9590; and Guevara-Garcia et al. (1993) Plant J 4(3):495-505. Promoters can be modified, if necessary, for weak expression.

In some embodiments, the nucleic acids of interest can be targeted to the chloroplast for expression. In this manner, where the nucleic acid of interest is not directly inserted into the chloroplast, the expression vector will additionally contain a chloroplast- targeting sequence comprising a nucleotide sequence that encodes a chloroplast transit peptide to direct the gene product of interest to the chloroplasts. Such transit peptides are known in the art. With respect to chloroplast-targeting sequences, "operably linked" means that the nucleic acid sequence encoding a transit peptide (i.e., the chloroplast-targeting sequence) is linked to the desired coding sequence of the invention such that the two sequences are contiguous and in the same reading frame. See, for example, Von Heijne et al. (1991) Plant Mol. Biol. Rep. 9: 104-126; Clark et al. (1989) J Biol. Chem. 264:17544-17550; Della-Cioppa et al. (1987) Plant Physiol. 84:965-968; Romer et al. (1993) Biochem. Biophys. Res. Commun. 196: 1414-1421 ; and Shah et al. (1986) Science 233:478-481. For example, a chloroplast transit peptide known in the art can be fused to the amino acid sequence of a TK polypeptide of the invention by operably linking a choloroplast-targeting sequence to the 5'-end of a nucleotide sequence encoding the TK polypeptide.

Chloroplast targeting sequences are known in the art and include the chloroplast small subunit of ribulose-l,5-bisphosphate carboxylase (Rubisco) (de Castro Silva Filho et al. (1996) Plant Mol. Biol. 30:769-780; Schnell et al. (1991) J Biol. Chem. 266(5):3335-3342); EPSPS (Archer et al. (1990) J Bioenerg. Biomemb. 22(6):789-810); tryptophan synthase (Zhao et al. (1995) J Biol. Chem. 270(11):6081-6087); plastocyanin (Lawrence et al. (1997) J Biol. Chem. 272(33):20357-20363); chorismate synthase (Schmidt et al. (1993) J Biol. Chem. 268(36):27447-27457); and the light harvesting chlorophyll a/b binding protein (LHBP) (Lamppa et al. (1988) J Biol. Chem. 263: 14996-14999). See also Von Heijne et al. (1991) Plant Mol. Biol. Rep. 9: 104- 126; Clark et al. (1989) J Biol. Chem. 264: 17544-17550; Della-Cioppa et al. (1987) Plant Physiol. 84:965-968; Romer et al. (1993) Biochem Biophys. Res. Commun. 196: 1414-1421 ; and Shah et al. (1986) Science 233:478-481.

Methods for transformation of chloroplasts are known in the art. See, for example, Svab et al. (1990) Proc. Natl. Acad. ScL USA 87:8526-8530; Svab and Maliga (1993) Proc. Natl. Acad. Sci. USA 90:913-917; Svab and Maliga (1993) EMBO J. 12:601-606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. Additionally, plastid transformation can be accomplished by transactivation of a silent plastid-borne transgene by tissue-preferred expression of a nuclear-encoded and plastid-directed RNA polymerase. Such a system has been reported in McBride et al. (1994) Proc. Natl. Acad. Sci. USA 91 :7301- 7305.

The nucleic acids of interest to be targeted to the chloroplast may be optimized for expression in the chloroplast to account for differences in codon usage between the plant nucleus and this organelle. In this manner, the nucleic acids of interest may be synthesized using chloroplast-preferred codons. See, for example, U.S. Patent No. 5,380,831, herein incorporated by reference.

Numerous plant transformation vectors and methods for transforming plants are available. See, for example, An, G. et al. (1986) Plant PysioL, 81 :301-305; Fry, J., et al. (1987) Plant Cell Rep. 6:321-325; Block, M. (1988) Theor. Appl. Genet .16: 161 -1 1 A; Hinchee, et al. (1990) Stadler. Genet. Symp.2032\2.203-2\2; Cousins, et al. (1991) Aust. J. Plant Physiol. 18:481-494; Chee, P. P. and Slightom, J. L. (1992) Gene.l I 8:255-260; Christou, et al. (1992) Trends. Biotechnol. 10:239-246; Halluin, et al. (1992) Bio/Technol. 10:309-314; Dhir, et al. (1992) Plant Physiol. 99:81-88; Casas et al. (1993) Proc. Nat. Acad Sd. USA 90: 1 1212-1 1216; Christou, P. (1993) In Vitro Cell. Dev. Biol.-Plant; 29P.119-124; Davies, et al. (1993) Plant Cell Rep. 12: 180-183; Dong, J. A. and Mchughen, A. (1993) Plant ScL 91 : 139-148; Franklin, C. I. and Trieu, T. N. (1993) Plant. Physiol. 102: 167; Golovkin, et al. (1993) Plant ScL 90:41-52; Guo Chin ScL Bull. 38:2072-2078; Asano, et al. (1994) Plant Cell Rep. 13; Ayeres N. M. and Park, W. D. (1994) Crit. Rev. Plant. Sci. 13:219-239; Barcelo, et al. (1994) Plant. J. 5:583-592; Becker, et al. (1994) Plant. J. 5:299-307; Borkowska et al. (1994) Acta. Physiol Plant. 16:225-230; Christou, P. (1994) Agro. Food. Ind. Hi Tech. 5: 17-27; Eapen et al. (1994) Plant Cell Rep. 13:582-586; Hartman, et al. (1994) Bio-Technology 12: 919923; Ritala, et al. (1994) Plant. Mol. Biol. 24:317-325; and Wan, Y. C. and Lemaux, P. G. (1994) Plant Physiol. 104:3748.

In some embodiments, the methods of the invention involve introducing a polynucleotide construct into a plant. By "introducing" is intended presenting to the plant the polynucleotide construct in such a manner that the construct gains access to the interior of a cell of the plant. The methods of the invention do not depend on a particular method for introducing a polynucleotide construct to a plant, only that the polynucleotide construct gains access to the interior of at least one cell of the plant. Methods for introducing polynucleotide constructs into plants are known in the art including, but not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods. The term “introduction” or “transformation” as referred to herein further means the transfer of an exogenous polynucleotide into a host cell, irrespective of the method used for transfer. Plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a genetic construct of the present invention and a whole plant regenerated there from. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include leaf disks, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g., apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g., cotyledon meristem and hypocotyl meristem). The polynucleotide may be transiently or stably introduced into a host cell and may be maintained non-integrated, for example, as a plasmid. Alternatively, it may be integrated into the host genome. The resulting transformed plant cell may then be used to regenerate a transformed plant in a manner known to persons skilled in the art.

By "stable transformation" is intended that the polynucleotide construct introduced into a plant integrates into the genome of the plant and is capable of being inherited by descendent thereof. By "transient transformation" is intended that a polynucleotide construct introduced into a plant does not integrate into the genome of the plant.

For the transformation of plants and plant cells, the nucleotide sequences of the invention are inserted using standard techniques into any vector known in the art that is suitable for expression of the nucleotide sequences in a plant or plant cell. The selection of the vector depends on the preferred transformation technique and the target plant species to be transformed. In an embodiment of the invention, the encoding nucleotide sequence is operably linked to a plant promoter, e.g. a promoter known in the art for high-level expression in a plant cell, and this construct is then introduced into a plant cell that is susceptible to TK-inhibiting herbicides; and a transformed plant is regenerated. In some embodiments, the transformed plant is tolerant to exposure to a level of TK-inhibiting herbicides that would kill or significantly injure a plant regenerated from an untransformed cell. This method can be applied to any plant species or crops.

Methodologies for constructing plant expression vectors and introducing foreign nucleic acids into plants are generally known in the art. For example, foreign DNA can be introduced into plants, using tumor-inducing (Ti) plasmid vectors. Other methods utilized for foreign DNA delivery involve the use of PEG mediated protoplast transformation, electroporation, microinjection whiskers, and biolistics or microprojectile bombardment for direct DNA uptake. Such methods are known in the art. (U.S. Pat. No. 5,405,765 to Vasil et al.; Bilang et a (1991) Gene 100: 247-250; Scheid et al.al, (1991) MoL Gen. Genet., 228: 104- 1 12; Guerche et al., (1987) Plant Science 52: 1 1 1 -1 16; Neuhause et al., (1987) Theor. Appl Genet. 75: 30-36; Klein et al., (1987) Nature 327: 70-73; Howell et al., (1980) Science 208: 1265; Horsch et al., (1985) Science 227: 1229-1231 ; DeBlock et al., (1989) Plant Physiology 91 : 694-701 ; Methods for Plant Molecular Biology (Weissbach and Weissbach, eds.) Academic Press, Inc. (1988) and Methods in Plant Molecular Biology (Schuler and Zielinski, eds.) Academic Press, Inc. (1989).

Other suitable methods of introducing nucleotide sequences into plant cells include microinjection as described by, e.g., Crossway et al. (1986) Biotechniques 4:320-334, electroporation as described by e.g., Riggs et al. (1986) Proc. Natl. Acad. ScL USA 83:5602- 5606, Agrobacterium-mediated transformation as described by e.g., Townsend et al., U.S. Patent No. 5,563,055, Zhao et al., U.S. Patent No. 5,981,840, direct gene transfer as described by, e.g., Paszkowski et al. (1984) EMBO J. 3:2717-2722, and ballistic particle acceleration as described by, e.g., U.S. Patent Nos. 4,945,050; 5,879,918; 5,886,244; and 5,932,782; Tomes et al. (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in Plant Cell, Tissue, and Organ Culture: Fundamental

Methods, ed. Gamborg and Phillips (Springer- Verlag, Berlin); McCabe et al. (1988) Biotechnology 6:923-926); and Led transformation (WO 00/28058). Also see, Weissinger et al., (1988) Ann. Rev. Genet. 22:421-477; Sanford et al, (1987) Particulate Science and Technology 5:27-37 (onion); Christou et al, (1988) Plant Physiol. 87:671-674 (soybean); McCabe et al., (1988) Bio/Technology 6:923-926 (soybean); Finer and McMullen (1991) In Vitro Cell Dev. Biol. 27P: 175-182 (soybean); Singh et al, (1998) Theor. Appl. Genet. 96:319-324 (soybean); Datta et al., (1990) Biotechnology 8:736-740 (rice); Klein et al., (1988) PNAS, 85:4305-4309 (maize); Klein et al., (1988) Biotechnology 6:559-563 (maize); U.S. Patent Nos. 5,240,855; 5,322,783; and 5,324,646; Tomes et al., (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in Plant Cell, Tissue, and Organ Culture: Fundamental Methods, ed. Gamborg (Springer- Verlag, Berlin) (maize); Klein et al., (1988) Plant Physiol. 91 :440-444 (maize); Fromm et al., (1990) Biotechnology 8:833-839 (maize); Hooykaas-Van Slogteren et al., (1984) Nature (London) 31 1 :763-764; Bowen et al, U.S. Patent No. 5,736,369 (cereals); Bytebier et al, (1987) PNAS 84:5345-5349 (Liliaceae); De Wet et al., (1985) in The Experimental Manipulation of Ovule Tissues, ed. Chapman et al, (Longman, New York), pp. 197-209 (pollen); Kaeppler et al., (1990) Plant Cell Reports 9:415-418 and Kaeppler et al., (1992) Theor. Apph Genet. 84:560-566 (whisker-mediated transformation); D'Halluin et al., (1992) Plant Cell 4: 1495-1505 (electroporation); Li et al., (1993) Plant Cell Reports 12:250- 255 and Christou and Ford (1995) Annals of Botany 75:407-413 (rice); Osjoda et al, (1996) Nature Biotechnology 14:745-750 (maize via Agrobacterium tumefaciens); each of which is herein incorporated by reference.

Transgenic plants, including transgenic crop plants, are preferably produced via Agrobacterium-mediated transformation. An advantageous transformation method is the transformation in planta. To this end, it is possible, for example, to allow the agrobacteria to act on plant seeds or to inoculate the plant meristem with agrobacteria. It has proved particularly expedient in accordance with the invention to allow a suspension of transformed agrobacteria to act on the intact plant or at least on the flower primordia. The plant is subsequently grown on until the seeds of the treated plant are obtained (Clough and Bent, Plant J. (1998) 16, 735–743). Methods for Agrobacterium-mediated transformation of rice include well known methods for rice transformation, such as those described in any of the following: European patent application EP 1198985 A1, Aldemita and Hodges (Planta 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), which disclosures are incorporated by reference herein as if fully set forth. In the case of corn transformation, the preferred method is as described in either Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) or Frame et al. (Plant Physiol 129(1): 13-22, 2002), which disclosures are incorporated by reference herein as if fully set forth. Said methods are further described by way of example in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press (1993) 128-143 and in Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). The nucleic acids or the construct to be expressed is preferably cloned into a vector, which is suitable for transforming

Agrobacterium tumefaciens, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Agrobacteria transformed by such a vector can then be used in known manner for the transformation of plants, such as plants used as a model, like Arabidopsis (Arabidopsis thaliana is within the scope of the present invention not considered as a crop plant), or crop plants such as, by way of example, tobacco plants, for example by immersing bruised leaves or chopped leaves in an agrobacterial solution and then culturing them in suitable media. The transformation of plants by means of Agrobacterium tumefaciens is described, for example, by Höfgen and Willmitzer in Nucl. Acid Res. (1988) 16, 9877 or is known inter alia from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38.

One transformation method known to those of skill in the art is the dipping of a flowering plant into an Agrobacteria solution, wherein the Agrobacteria contains the TK nucleic acid, followed by breeding of the transformed gametes. Agrobacterium mediated plant transformation can be performed using for example the GV3101(pMP90) (Koncz and Schell, 1986, Mol. Gen. Genet. 204:383-396) or LBA4404 (Clontech) Agrobacterium tumefaciens strain. Transformation can be performed by standard transformation and regeneration techniques (Deblaere et al., 1994, Nucl. Acids. Res. 13:4777-4788; Gelvin, Stanton B. and Schilperoort, Robert A, Plant Molecular Biology Manual, 2nd Ed. - Dordrecht : Kluwer Academic Publ., 1995. - in Sect., Ringbuc Zentrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R. and Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton : CRC Press, 1993 360 S., ISBN 0-8493-5164-2). For example, rapeseed can be transformed via cotyledon or hypocotyl transformation (Moloney et al., 1989, Plant Cell Report 8:238-242; De Block et al., 1989, Plant Physiol. 91:694-701). Use of antibiotics for Agrobacterium and plant selection depends on the binary vector and the Agrobacterium strain used for transformation. Rapeseed selection is normally performed using kanamycin as selectable plant marker. Agrobacterium mediated gene transfer to flax can be performed using, for example, a technique described by Mlynarova et al., 1994, Plant Cell Report 13:282-285. Additionally, transformation of soybean can be performed using for example a technique described in European Patent No. 0424 047, U.S. Patent No. 5,322,783, European Patent No. 0397 687, U.S. Patent No. 5,376,543, or U.S. Patent No. 5,169,770. Transformation of maize can be achieved by particle bombardment, polyethylene glycol mediated DNA uptake, or via the silicon carbide fiber technique. (See, for example, Freeling and Walbot “The maize handbook” Springer Verlag: New York (1993) ISBN 3-540-97826-7). A specific example of maize transformation is found in U.S. Patent No. 5,990,387, and a specific example of wheat transformation can be found in PCT Application No. WO 93/07256.

In some embodiments, polynucleotides of the present invention may be introduced into plants by contacting plants with a virus or viral nucleic acids. Generally, such methods involve incorporating a polynucleotide construct of the invention within a viral DNA or RNA molecule. It is recognized that the polypeptides of the invention may be initially synthesized as part of a viral polyprotein, which later may be processed by proteolysis in vivo or in vitro to produce the desired recombinant polypeptide. Further, it is recognized that promoters of the invention also encompass promoters utilized for transcription by viral RNA polymerases. Methods for introducing polynucleotide constructs into plants and expressing a protein encoded therein, involving viral DNA or RNA molecules, are known in the art. See, for example, U.S. Patent Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367 and 5,316,931; herein incorporated by reference. The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick et a (1986) Plant Cell Reports 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved.

The present invention may be used for transformation of any plant species, including, but not limited to, monocots and dicots. Examples of plant species of interest include, but are not limited to, corn or maize (Zea mays), Brassica sp. (e.g., B. napus, B. rapa, B. juncea), particularly those Brassica species useful as sources of seed oil, alfalfa (Medicago sativa), rice (Oryza sativa), rye (Secale cereale), sorghum (Sorghum bicolor, Sorghum vulgare), millet e.g., pearl millet (Pennisetum glaucum), proso millet (Panicum miliaceum), foxtail millet (Setaria italica), finger millet (Eleusine coracana)), sunflower (Helianthus annu), saffiower (Carthamus tinctorius), wheat (Triticum aestivum, T. Turgidum ssp. durum), soybean (Glycine max), tobacco (Nicotiana tabacum), potato (Solarium tuberosum), peanuts (Arachis hypogaea), cotton (Gossypium barbadense, Gossypium hirsutum), sweet potato (Ipomoea batatus), cassava (Manihot esculenta), coffee (Coffea spp.), coconut (Cocos nucifera), pineapple (Ananas comosus), citrus trees (Citrus spp.), cocoa (Theobroma cacao), tea (Camellia sinensis), banana (Musa spp.), avocado (Persea americana), fig (Ficus casica), guava (Psidium guajava), mango (Mangifera indica), olive (Olea europaea), papaya (Carica papaya), cashew (Anacardium occidentale), macadamia (Macadamia integrifolia), almond (Prunus amygdalus), sugar beets (Beta vulgaris), sugarcane (Saccharum spp.), oats, barley, vegetables, ornamentals, and conifers. Preferably, plants of the present invention are crop plants (for example, sunflower, Brassica sp., cotton, sugar, beet, soybean, peanut, alfalfa, safflower, tobacco, corn, rice, wheat, rye, barley triticale, sorghum, millet, etc.).

In addition to the transformation of somatic cells, which then have to be regenerated into intact plants, it is also possible to transform the cells of plant meristems and in particular those cells which develop into gametes. In this case, the transformed gametes follow the natural plant development, giving rise to transgenic plants. Thus, for example, seeds of Arabidopsis are treated with agrobacteria and seeds are obtained from the developing plants of which a certain proportion is transformed and thus transgenic [Feldman, KA and Marks MD (1987). Mol Gen Genet 208:274-289; Feldmann K (1992). In: C Koncz, N-H Chua and J Shell, eds, Methods in Arabidopsis Research. Word Scientific, Singapore, pp.

274-289]. Alternative methods are based on the repeated removal of the inflorescences and incubation of the excision site in the center of the rosette with transformed agrobacteria, whereby transformed seeds can likewise be obtained at a later point in time (Chang (1994). Plant J. 5: 551-558; Katavic (1994). Mol Gen Genet, 245: 363-370). However, an especially effective method is the vacuum infiltration method with its modifications such as the “floral dip” method. In the case of vacuum infiltration of Arabidopsis, intact plants under reduced pressure are treated with an agrobacterial suspension [Bechthold, N (1993). C R Acad Sci Paris Life Sci, 316: 1194-1199], while in the case of the ”floral dip” method the developing floral tissue is incubated briefly with a surfactant-treated agrobacterial suspension [Clough, SJ and Bent AF (1998) The Plant J. 16, 735-743]. A certain proportion of transgenic seeds are harvested in both cases, and these seeds can be distinguished from non-transgenic seeds by growing under the above-described selective conditions. In addition the stable transformation of plastids is of advantages because plastids are inherited maternally is most crops reducing or eliminating the risk of transgene flow through pollen. The transformation of the chloroplast genome is generally achieved by a process which has been schematically displayed in Klaus et al., 2004 [Nature Biotechnology 22 (2), 225-229]. Briefly the sequences to be transformed are cloned together with a selectable marker gene between flanking sequences homologous to the chloroplast genome. These homologous flanking sequences direct site specific integration into the plastome. Plastidal transformation has been described for many different plant species and an overview is given in Bock (2001) Transgenic plastids in basic research and plant biotechnology. J Mol Biol. 2001 Sep 21; 312 (3):425-38 or Maliga, P (2003) Progress towards commercialization of plastid transformation technology. Trends Biotechnol. 21, 20-28. Further biotechnological progress has recently been reported in form of marker free plastid transformants, which can be produced by a transient co-integrated maker gene (Klaus et al., 2004, Nature Biotechnology 22(2), 225-229). The genetically modified plant cells can be regenerated via all methods with which the skilled worker is familiar. Suitable methods can be found in the abovementioned publications by S.D. Kung and R. Wu, Potrykus or Höfgen and Willmitzer.

Generally after transformation, plant cells or cell groupings are selected for the presence of one or more markers which are encoded by plant-expressible genes co-transferred with the gene of interest, following which the transformed material is regenerated into a whole plant. To select transformed plants, the plant material obtained in the transformation is, as a rule, subjected to selective conditions so that transformed plants can be distinguished from untransformed plants. For example, the seeds obtained in the above-described manner can be planted and, after an initial growing period, subjected to a suitable selection by spraying. A further possibility consists in growing the seeds, if appropriate after sterilization, on agar plates using a suitable selection agent so that only the transformed seeds can grow into plants. Alternatively, the transformed plants are screened for the presence of a selectable marker such as the ones described above.

Following DNA transfer and regeneration, putatively transformed plants may also be evaluated, for instance using Southern analysis, for the presence of the gene of interest, copy number and/or genomic organisation. Alternatively or additionally, expression levels of the newly introduced DNA may be monitored using Northern and/or Western analysis, both techniques being well known to persons having ordinary skill in the art.

The generated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed and homozygous second-generation (or T2) transformants selected, and the T2 plants may then further be propagated through classical breeding techniques. The generated transformed organisms may take a variety of forms. For example, they may be chimeras of transformed cells and non-transformed cells; clonal transformants (e.g., all cells transformed to contain the expression cassette); grafts of transformed and untransformed tissues (e.g., in plants, a transformed rootstock grafted to an untransformed scion).

Preferably, the expression of the nucleic acid in the plant results in the plant's increased resistance to TK-inhibiting herbicide as compared to a wild type variety of the plant.

In another embodiment, the invention refers to a plant, comprising a plant cell according to the present invention, wherein expression of the nucleic acid in the plant results in the plant's increased resistance to TK-inhibiting herbicide as compared to a wild type variety of the plant.

The plants described herein can be either transgenic crop plants or non-transgenic plants.

In addition to the general definition, give SUPRA, "transgenic", “transgene” or "recombinant" means with regard to, for example, a nucleic acid sequence, an expression cassette, gene construct or a vector comprising the nucleic acid sequence or an organism transformed with the nucleic acid sequences, expression cassettes or vectors according to the invention, all those constructions brought about by recombinant methods in which either

(a) the nucleic acid sequences encoding proteins useful in the methods of the invention, or
(b) genetic control sequence(s) which is operably linked with the nucleic acid sequence according to the invention, for example a promoter, or
(c) a) and b)

are not located in their natural genetic environment or have been modified by recombinant methods, it being possible for the modification to take the form of, for example, a substitution, addition, deletion, inversion or insertion of one or more nucleotide residues in order to allow for the expression of the mutated TK of the present invention. The natural genetic environment is understood as meaning the natural genomic or chromosomal locus in the original plant or the presence in a genomic library. In the case of a genomic library, the natural genetic environment of the nucleic acid sequence is preferably retained, at least in part. The environment flanks the nucleic acid sequence at least on one side and has a sequence length of at least 50 bp, preferably at least 500 bp, especially preferably at least 1000 bp, most preferably at least 5000 bp. A naturally occurring expression cassette – for example the naturally occurring combination of the natural promoter of the nucleic acid sequences with the corresponding nucleic acid sequence encoding a polypeptide useful in the methods of the present invention, as defined above – becomes a transgenic expression cassette when this expression cassette is modified by non-natural, synthetic ("artificial") methods such as, for example, mutagenic treatment. Suitable methods are described, for example, in US 5,565,350 or WO 00/15815.

A transgenic plant for the purposes of the invention is thus understood as meaning, as above, that the nucleic acids of the invention are not at their natural locus in the genome of said plant, it being possible for the nucleic acids to be expressed homologously or heterologously. However, as mentioned, transgenic also means that, while the nucleic acids according to the invention or used in the inventive method are at their natural position in the genome of a plant, the sequence has been modified with regard to the natural sequence, and/or that the regulatory sequences of the natural sequences have been modified. Transgenic is preferably understood as meaning the expression of the nucleic acids according to the invention at an unnatural locus in the genome, i.e. homologous or, preferably, heterologous expression of the nucleic acids takes place. Preferred transgenic plants are mentioned herein. Furthermore, the term “transgenic” refers to any plant, plant cell, callus, plant tissue, or plant part, that contains all or part of at least one recombinant polynucleotide. In many cases, all or part of the recombinant polynucleotide is stably integrated into a chromosome or stable extra-chromosomal element, so that it is passed on to successive generations. For the purposes of the invention, the term “recombinant polynucleotide” refers to a polynucleotide that has been altered, rearranged, or modified by genetic engineering. Examples include any cloned polynucleotide, or polynucleotides, that are linked or joined to heterologous sequences. The term “recombinant” does not refer to alterations of polynucleotides that result from naturally occurring events, such as spontaneous mutations, or from non-spontaneous mutagenesis followed by selective breeding.

Plants containing mutations arising due to non-spontaneous mutagenesis and selective breeding are referred to herein as non-transgenic plants and are included in the present invention. In embodiments wherein the plant is transgenic and comprises multiple mutated TK nucleic acids, the nucleic acids can be derived from different genomes or from the same genome. Alternatively, in embodiments wherein the plant is non-transgenic and comprises multiple mutated TK nucleic acids, the nucleic acids are located on different genomes or on the same genome.

In certain embodiments, the present invention involves herbidicide-resistant plants that are produced by mutation breeding. Such plants comprise a polynucleotide encoding a mutated TK and are tolerant to one or more TK-inhibiting herbicides. Such methods can involve, for example, exposing the plants or seeds to a mutagen, particularly a chemical mutagen such as, for example, ethyl methanesulfonate (EMS) and selecting for plants that have enhanced tolerance to at least one or more TK-inhibiting herbicide.

However, the present invention is not limited to herbicide-tolerant plants that are produced by a mutagenesis method involving the chemical mutagen EMS. Any mutagenesis method known in the art may be used to produce the herbicide-resistant plants of the present invention. Such mutagenesis methods can involve, for example, the use of any one or more of the following mutagens: radiation, such as X-rays, Gamma rays (e.g., cobalt 60 or cesium 137), neutrons, (e.g., product of nuclear fission by uranium 235 in an atomic reactor), Beta radiation (e.g., emitted from radioisotopes such as phosphorus 32 or carbon 14), and ultraviolet radiation (preferably from 2500 to 2900 nm), and chemical mutagens such as base analogues (e.g., 5-bromo-uracil), related compounds (e.g., 8-ethoxy caffeine), antibiotics (e.g., streptonigrin), alkylating agents (e.g., sulfur mustards, nitrogen mustards, epoxides, ethylenamines, sulfates, sulfonates, sulfones, lactones), azide, hydroxylamine, nitrous acid, or acridines. Herbicide-resistant plants can also be produced by using tissue culture methods to select for plant cells comprising herbicide-resistance mutations and then regenerating herbicide-resistant plants therefrom. See, for example, U.S. Patent Nos. 5,773,702 and 5,859,348, both of which are herein incorporated in their entirety by reference. Further details of mutation breeding can be found in "Principals of Cultivar Development" Fehr, 1993 Macmillan Publishing Company the disclosure of which is incorporated herein by reference

The plant of the present invention comprises at least one mutated TK nucleic acid or over-expressed wild-type TK nucleic acid, and has increased tolerance to a TK-inhibiting herbicide as compared to a wild-type variety of the plant. It is possible for the plants of the present invention to have multiple wild-type or mutated TK nucleic acids from different genomes since these plants can contain more than one genome. For example, a plant contains two genomes, usually referred to as the A and B genomes. Because TK is a required metabolic enzyme, it is assumed that each genome has at least one gene coding for the TK enzyme (i.e. at least one TK gene). As used herein, the term “TK gene locus” refers to the position of an TK gene on a genome, and the terms “TK gene” and “TK nucleic acid” refer to a nucleic acid encoding the TK enzyme. The TK nucleic acid on each genome differs in its nucleotide sequence from an TK nucleic acid on another genome. One of skill in the art can determine the genome of origin of each TK nucleic acid through genetic crossing and/or either sequencing methods or exonuclease digestion methods known to those of skill in the art.

The present invention includes plants comprising one, two, three, or more mutated TK alleles, wherein the plant has increased tolerance to a TK-inhibiting herbicide as compared to a wild-type variety of the plant. The mutated TK alleles can comprise a nucleotide sequence selected from the group consisting of a polynucleotide as defined in SEQ ID NO: 182 or 183, or a variant or derivative thereof, a polynucleotide encoding a polypeptide as defined in SEQ ID NO: 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, 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, or 181, or a variant or derivative, homologue, orthologue, paralogue thereof, a polynucleotide comprising at least 60 consecutive nucleotides of any of the aforementioned polynucleotides; and a polynucleotide complementary to any of the aforementioned polynucleotides.

“Alleles” or “allelic variants” are alternative forms of a given gene, located at the same chromosomal position. Allelic variants encompass Single Nucleotide Polymorphisms (SNPs), as well as Small Insertion/Deletion Polymorphisms (INDELs). The size of INDELs is usually less than 100 bp. SNPs and INDELs form the largest set of sequence variants in naturally occurring polymorphic strains of most organisms

The term “variety” refers to a group of plants within a species defined by the sharing of a common set of characteristics or traits accepted by those skilled in the art as sufficient to distinguish one cultivar or variety from another cultivar or variety. There is no implication in either term that all plants of any given cultivar or variety will be genetically identical at either the whole gene or molecular level or that any given plant will be homozygous at all loci. A cultivar or variety is considered “true breeding” for a particular trait if, when the true-breeding cultivar or variety is self-pollinated, all of the progeny contain the trait. The terms “breeding line” or “line” refer to a group of plants within a cultivar defined by the sharing of a common set of characteristics or traits accepted by those skilled in the art as sufficient to distinguish one breeding line or line from another breeding line or line. There is no implication in either term that all plants of any given breeding line or line will be genetically identical at either the whole gene or molecular level or that any given plant will be homozygous at all loci. A breeding line or line is considered “true breeding” for a particular trait if, when the true-breeding line or breeding line is self-pollinated, all of the progeny contain the trait. In the present invention, the trait arises from a mutation in a TK gene of the plant or seed.

The herbicide-resistant plants of the invention that comprise polynucleotides encoding mutated TK polypeptides also find use in methods for increasing the herbicide-resistance of a plant through conventional plant breeding involving sexual reproduction. The methods comprise crossing a first plant that is a herbicide-resistant plant of the invention to a second plant that may or may not be resistant to the same herbicide or herbicides as the first plant or may be resistant to different herbicide or herbicides than the first plant. The second plant can be any plant that is capable of producing viable progeny plants (i.e., seeds) when crossed with the first plant. Typically, but not necessarily, the first and second plants are of the same species. The methods can optionally involve selecting for progeny plants that comprise the mutated TK polypeptides of the first plant and the herbicide resistance characteristics of the second plant. The progeny plants produced by this method of the present invention have increased resistance to a herbicide when compared to either the first or second plant or both. When the first and second plants are resistant to different herbicides, the progeny plants will have the combined herbicide tolerance characteristics of the first and second plants. The methods of the invention can further involve one or more generations of backcrossing the progeny plants of the first cross to a plant of the same line or genotype as either the first or second plant. Alternatively, the progeny of the first cross or any subsequent cross can be crossed to a third plant that is of a different line or genotype than either the first or second plant. The present invention also provides plants, plant organs, plant tissues, plant cells, seeds, and non-human host cells that are transformed with the at least one polynucleotide molecule, expression cassette, or transformation vector of the invention. Such transformed plants, plant organs, plant tissues, plant cells, seeds, and non-human host cells have enhanced tolerance or resistance to at least one herbicide, at levels of the herbicide that kill or inhibit the growth of an untransformed plant, plant tissue, plant cell, or non-human host cell, respectively. Preferably, the transformed plants, plant tissues, plant cells, and seeds of the invention are Arabidopsis thaliana and crop plants.

It is to be understood that the plant of the present invention can comprise a wild type TK nucleic acid in addition to a mutated TK nucleic acid. It is contemplated that the TK-inhibiting herbicide tolerant lines may contain a mutation in only one of multiple TK isoenzymes. Therefore, the present invention includes a plant comprising one or more mutated TK nucleic acids in addition to one or more wild type TK nucleic acids.

In another embodiment, the invention refers to a seed produced by a transgenic plant comprising a plant cell of the present invention, wherein the seed is true breeding for an increased resistance to a TK-inhibiting herbicide as compared to a wild type variety of the seed.

In other aspects, TK-inhibiting herbicides-tolerant plants of the present invention can be employed as TK-inhibiting herbicides-tolerance trait donor lines for development, as by traditional plant breeding, to produce other varietal and/or hybrid crops containing such trait or traits. All such resulting variety or hybrids crops, containing the ancestral TK-inhibiting herbicides-tolerance trait or traits can be referred to herein as progeny or descendant of the ancestral, TK-inhibiting herbicides-tolerant line(s).

In other embodiments, the present invention provides a method for producing a TK-inhibiting herbicides-tolerant plant. The method comprises: crossing a first TK-inhibiting herbicides-tolerant plant with a second plant to produce a TK-inhibiting herbicides-tolerant progeny plant, wherein the first plant and the progeny plant comprise in at least some of their cells a polynucleotide operably linked to a promoter operable in plant cells, the recombinant polynucleotide being effective in the cells of the first plant to express a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

In some embodiments, traditional plant breeding is employed whereby the TK-inhibiting herbicides-tolerant trait is introduced in the progeny plant resulting therefrom. In one embodiment, the present invention provides a method for producing a TK-inhibiting herbicides-tolerant progeny plant, the method comprising: crossing a parent plant with a TK-inhibiting herbicides-tolerant plant to introduce the TK-inhibiting herbicides-tolerance characteristics of the TK-inhibiting herbicides-tolerant plant into the germplasm of the progeny plant, wherein the progeny plant has increased tolerance to the TK-inhibiting herbicides relative to the parent plant. In other embodiments, the method further comprises the step of introgressing the TK-inhibiting herbicides-tolerance characteristics through traditional plant breeding techniques to obtain a descendent plant having the TK-inhibiting herbicides-tolerance characteristics.

In other aspects, plants of the invention include those plants which, in addition to being TK-inhibiting herbicides-tolerant, have been subjected to further genetic modifications by breeding, mutagenesis or genetic engineering, e.g. have been rendered tolerant to applications of specific other classes of herbicides, such as AHAS inhibitors; auxinic herbicides; bleaching herbicides such as hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors or phytoene desaturase (PDS) inhibitors; EPSPS inhibitors such as glyphosate; glutamine synthetase (GS) inhibitors such as glufosinate; lipid biosynthesis inhibitors such as acetyl CoA carboxylase (ACCase) inhibitors; or oxynil {i.e. bromoxynil or ioxynil) herbicides as a result of conventional methods of breeding or genetic engineering, Thus, TK-inhibiting herbicides-tolerant plants of the invention can be made resistant to multiple classes of herbicides through multiple genetic modifications, such as resistance to both glyphosate and glufosinate or to both glyphosate and a herbicide from another class such as HPPD inhibitors, AHAS inhibitors, or ACCase inhibitors. These herbicide resistance technologies are, for example, described in Pest Management Science (at volume, year, page): 61, 2005, 246; 61, 2005, 258; 61, 2005, 277; 61, 2005, 269; 61, 2005, 286; 64, 2008, 326; 64, 2008, 332; Weed Science 57, 2009, 108; Australian Journal of Agricultural Research 58, 2007, 708; Science 316, 2007, 1185; and references quoted therein. For example, TK-inhibiting herbicides-tolerant plants of the invention, in some embodiments, may be tolerant to ACCase inhibitors, such as "dims" {e.g., cycloxydim, sethoxydim, clethodim, or tepraloxydim), "fops" {e.g. , clodinafop, diclofop, fluazifop, haloxyfop, or quizalofop), and "dens" (such as pinoxaden); to auxinic herbicides, such as dicamba; to EPSPS inhibitors, such as glyphosate; to other TK inhibitors; and to GS inhibitors, such as glufosinate.

In addition to these classes of inhibitors, TK-inhibiting herbicides-tolerant plants of the invention may also be tolerant to herbicides having other modes of action, for example, chlorophyll/carotenoid pigment inhibitors, cell membrane disrupters, photosynthesis inhibitors, cell division inhibitors, root inhibitors, shoot inhibitors, and combinations thereof.

Such tolerance traits may be expressed, e.g. : as mutant or wildtype HPPD proteins, as mutant AHASL proteins, mutant ACCase proteins, mutant EPSPS proteins, or mutant glutamine synthetase proteins; or as mutant native, inbred, or transgenic aryloxyalkanoate dioxygenase (AAD or DHT), haloarylnitrilase (BXN), 2,2-dichloropropionic acid dehalogenase (DEH), glyphosate-N- acetyltransferase (GAT), glyphosate decarboxylase (GDC), glyphosate oxidoreductase (GOX), glutathione-S-transferase (GST), phosphinothricin acetyltransferase (PAT or bar), or CYP450s proteins having an herbicide-degrading activity. TK-inhibiting herbicides- tolerant plants hereof can also be stacked with other traits including, but not limited to, pesticidal traits such as Bt Cry and other proteins having pesticidal activity toward coleopteran, lepidopteran, nematode, or other pests; nutrition or nutraceutical traits such as modified oil content or oil profile traits, high protein or high amino acid concentration traits, and other trait types known in the art.

Furthermore, in other embodiments, TK-inhibiting herbicides-tolerant plants are also covered which are, by the use of recombinant DNA techniques and/or by breeding and/or otherwise selected for such characteristics, rendered able to synthesize one or more insecticidal proteins, especially those known from the bacterial genus Bacillus, particularly from Bacillus thuringiensis, such as [delta]-endotoxins, e.g. CrylA(b), CrylA(c), CrylF, CrylF(a2), CryllA(b), CrylllA, CrylllB(bl) or Cry9c; vegetative insecticidal proteins (VIP), e.g. VIP1, VIP2, VIP3 or VIP3A; insecticidal proteins of bacteria colonizing nematodes, e.g. Photorhabdus spp. or Xenorhabdus spp.; toxins produced by animals, such as scorpion toxins, arachnid toxins, wasp toxins, or other insect-specific neurotoxins; toxins produced by fungi, such streptomycete toxins; plant lectins, such as pea or barley lectins; agglutinins; proteinase inhibitors, such as trypsin inhibitors, serine protease inhibitors, patatin, cystatin or papain inhibitors; ribosome-inactivating proteins (RIP), such as ricin, maize-RIP, abrin, luffin, saporin or bryodin; steroid metabolism enzymes, such as 3-hydroxy-steroid oxidase, ecdysteroid-IDP-glycosyl-transferase, cholesterol oxidases, ecdysone inhibitors or HMG-CoA-reductase; ion channel blockers, such as blockers of sodium or calcium channels; juvenile hormone esterase; diuretic hormone receptors (helicokinin receptors); stilben synthase, bibenzyl synthase, chitinases or glucanases. In the context of the present invention these insecticidal proteins or toxins are to be understood expressly also as pre-toxins, hybrid proteins, truncated or otherwise modified proteins. Hybrid proteins are characterized by a new combination of protein domains, (see, e.g. WO 02/015701). Further examples of such toxins or genetically modified plants capable of synthesizing such toxins are disclosed, e.g., in EP-A 374 753, WO 93/007278, WO 95/34656, EP-A 427 529, EP-A 451 878, WO 03/18810 und WO 03/52073. The methods for producing such genetically modified plants are generally known to the person skilled in the art and are described, e.g. in the publications mentioned above. These insecticidal proteins contained in the genetically modified plants impart to the plants producing these proteins tolerance to harmful pests from all taxonomic groups of arthropods, especially to beetles (Coeloptera), two-winged insects (Diptera), and moths (Lepidoptera) and to nematodes (Nematoda).

In some embodiments, expression of one or more protein toxins (e.g., insecticidal proteins) in the TK-inhibiting herbicides-tolerant plants is effective for controlling organisms that include, for example, members of the classes and orders: Coleoptera such as the American bean weevil Acanthoscelides obtectus; the leaf beetle Agelastica alni; click beetles (Agriotes lineatus, Agriotes obscurus, Agriotes bicolor); the grain beetle Ahasverus advena; the summer schafer Amphimallon solstitialis; the furniture beetle Anobium punctatum; Anthonomus spp. (weevils); the Pygmy mangold beetle Atomaria linearis; carpet beetles

(Anthrenus spp., Attagenus spp.); the cowpea weevil Callosobruchus maculates; the fried fruit beetle Carpophilus hemipterus; the cabbage seedpod weevil Ceutorhynchus assimilis; the rape winter stem weevil Ceutorhynchus picitarsis; the wireworms Conoderus vespertinus and Conoderus falli; the banana weevil Cosmopolites sordidus; the New Zealand grass grub Costelytra zealandica; the June beetle Cotinis nitida; the sunflower stem weevil

Cylindrocopturus adspersus; the larder beetle Dermestes lardarius; the corn rootworms Diabrotica virgifera, Diabrotica virgifera virgifera, and Diabrotica barberi; the Mexican bean beetle Epilachna varivestis; the old house borer Hylotropes bajulus; the lucerne weevil Hypera postica; the shiny spider beetle Gibbium psylloides; the cigarette beetle Lasioderma serricorne; the Colorado potato beetle Leptinotarsa decemlineata; Lyctus beetles {Lyctus spp. , the pollen beetle Meligethes aeneus; the common cockshafer Melolontha melolontha; the American spider beetle Mezium americanum; the golden spider beetle Niptus hololeuc s; the grain beetles Oryzaephilus surinamensis and Oryzaephilus Mercator; the black vine weevil Otiorhynchus sulcatus; the mustard beetle Phaedon cochleariae, the crucifer flea beetle Phyllotreta cruciferae; the striped flea beetle Phyllotreta striolata; the cabbage steam flea beetle Psylliodes chrysocephala; Ptinus spp. (spider beetles); the lesser grain borer Rhizopertha dominica; the pea and been weevil Sitona lineatus; the rice and granary beetles Sitophilus oryzae and Sitophilus granaries; the red sunflower seed weevil Smicronyx fulvus; the drugstore beetle Stegobium paniceum; the yellow mealworm beetle Tenebrio molitor, the flour beetles Tribolium castaneum and Tribolium confusum; warehouse and cabinet beetles {Trogoderma spp.); the sunflower beetle Zygogramma exclamationis; Dermaptera (earwigs) such as the European earwig Forficula auricularia and the striped earwig Labidura riparia; Dictyoptera such as the oriental cockroach Blatta orientalis; the greenhouse millipede Oxidus gracilis; the beet fly Pegomyia betae; the frit fly Oscinella frit; fruitflies (Dacus spp., Drosophila spp.); Isoptera (termites) including species from the familes Hodotermitidae, Kalotermitidae, Mastotermitidae, Rhinotermitidae, Serritermitidae, Termitidae, Termopsidae; the tarnished plant bug Lygus lineolaris; the black bean aphid Aphis fabae; the cotton or melon aphid Aphis gossypii; the green apple aphid Aphis pomi; the citrus spiny whitefly Aleurocanthus spiniferus; the sweet potato whitefly Bemesia tabaci; the cabbage aphid Brevicoryne brassicae; the pear psylla Cacopsylla pyricola; the currant aphid Cryptomyzus ribis; the grape phylloxera Daktulosphaira vitifoliae; the citrus psylla Diaphorina citri; the potato leafhopper Empoasca fabae; the bean leafhopper Empoasca Solana; the vine leafhopper Empoasca vitis; the woolly aphid Eriosoma lanigerum; the European fruit scale Eulecanium corni; the mealy plum aphid Hyalopterus arundinis; the small brown planthopper Laodelphax striatellus; the potato aphid Macrosiphum euphorbiae; the green peach aphid Myzus persicae; the green rice leafhopper Nephotettix cinticeps; the brown planthopper Nilaparvata lugens; the hop aphid Phorodon humuli; the bird-cherry aphid Rhopalosiphum padi; the grain aphid Sitobion avenae; Lepidoptera such as Adoxophyes orana (summer fruit tortrix moth); Archips podana (fruit tree tortrix moth); Bucculatrix pyrivorella (pear leafminer); Bucculatrix thurberiella (cotton leaf perforator); Bupalus piniarius (pine looper); Carpocapsa pomonella (codling moth); Chilo suppressalis (striped rice borer); Choristoneura fumiferana (eastern spruce budworm); Cochylis hospes (banded sunflower moth); Diatraea grandiosella (southwestern corn borer); Eupoecilia ambiguella (European grape berry moth); Helicoverpa armigera (cotton bollworm); Helicoverpa zea (cotton bollworm); Heliothis vires cens (tobacco budworm), Homeosoma electellum (sunflower moth); Homona magnanima (oriental tea tree tortrix moth); Lithocolletis blancardella (spotted tentiform leafminer); Lymantria dispar (gypsy moth); Malacosoma neustria (tent caterpillar); Mamestra brassicae (cabbage armyworm); Mamestra configurata (Bertha armyworm); Operophtera brumata (winter moth); Ostrinia nubilalis (European corn borer), Panolis flammea (pine beauty moth), Phyllocnistis citrella (citrus leafminer); Pieris brassicae (cabbage white butterfly); Rachiplusia ni (soybean looper); Spodoptera exigua (beet armywonn); Spodoptera littoralis (cotton leafworm); Sylepta derogata (cotton leaf roller); Trichoplusia ni (cabbage looper); Orthoptera such as the common cricket Acheta domesticus, tree locusts (Anacridium spp.), the migratory locust Locusta migratoria, the twostriped grasshopper Melanoplus bivittatus, the differential grasshopper Melanoplus differ entialis, the redlegged grasshopper Melanoplus femurrubrum, the migratory grasshopper Melanoplus sanguinipes, the northern mole cricket Neocurtilla hexadectyla, the red locust Nomadacris septemfasciata, the shortwinged mole cricket Scapteriscus abbreviatus, the southern mole cricket Scapteriscus borellii, the tawny mole cricket Scapteriscus vicinus, and the desert locust Schistocerca gregaria; Symphyla such as the garden symphylan Scutigerella immaculata; Thysanoptera such as the tobacco thrips Frankliniella fusca, the flower thrips Frankliniella intonsa, the western flower thrips Frankliniella occidentalism the cotton bud thrips Frankliniella schultzei, the banded greenhouse thrips Hercinothrips femoralis, the soybean thrips Neohydatothrips variabilis, Kelly's citrus thrips Pezothrips kellyanus, the avocado thrips Scirtothrips perseae, the melon thrips Thrips palmi, and the onion thrips Thrips tabaci; and the like, and combinations comprising one or more of the foregoing organisms.

In some embodiments, expression of one or more protein toxins (e.g., insecticidal proteins) in the TK-inhibiting herbicides-tolerant plants is effective for controlling flea beetles, i.e. members of the flea beetle tribe of family Chrysomelidae, preferably against Phyllotreta spp., such as Phyllotreta cruciferae and/or Phyllotreta triolata. In other embodiments, expression of one or more protein toxins {e.g., insecticidal proteins) in the TK-inhibiting herbicides- tolerant plants is effective for controlling cabbage seedpod weevil, the Bertha armyworm, Lygus bugs, or the diamondback moth.

Furthermore, in one embodiment, TK-inhibiting herbicides-tolerant plants are also covered which are, e.g. by the use of recombinant DNA techniques and/or by breeding and/or otherwise selected for such traits, rendered able to synthesize one or more proteins to increase the resistance or tolerance of those plants to bacterial, viral or fungal pathogens. The methods for producing such genetically modified plants are generally known to the person skilled in the art.

Furthermore, in another embodiment, TK-inhibiting herbicides-tolerant plants are also covered which are, e.g. by the use of recombinant DNA techniques and/or by breeding and/or otherwise selected for such traits, rendered able to synthesize one or more proteins to increase the productivity (e.g. oil content), tolerance to drought, salinity or other growth-limiting environmental factors or tolerance to pests and fungal, bacterial or viral pathogens of those plants.

Furthermore, in other embodiments, TK-inhibiting herbicides-tolerant plants are also covered which are, e.g. by the use of recombinant DNA techniques and/or by breeding and/or otherwise selected for such traits, altered to contain a modified amount of one or more substances or new substances, for example, to improve human or animal nutrition, e.g. oil crops that produce health-promoting long-chain omega-3 fatty acids or unsaturated omega-9 fatty acids (e.g. Nexera(R) rape, Dow Agro Sciences, Canada).

Furthermore, in some embodiments, TK-inhibiting herbicides-tolerant plants are also covered which are, e.g. by the use of recombinant DNA techniques and/or by breeding and/or otherwise selected for such traits, altered to contain increased amounts of vitamins and/or minerals, and/or improved profiles of nutraceutical compounds.

In one embodiment, TK-inhibiting herbicides-tolerant plants of the present invention, relative to a wild-type plant, comprise an increased amount of, or an improved profile of, a compound selected from the group consisting of: glucosinolates (e.g., glucoraphanin (4-methylsulfinylbutyl-glucosinolate), sulforaphane, 3-indolylmethyl-glucosinolate(glucobrassicin), l -methoxy-3-indolylmethyl-glucosinolate (neoglucobrassicin)); phenolics (e.g., flavonoids (e.g., quercetin, kaempferol), hydroxycinnamoyl derivatives (e.g., 1 ,2,2'- trisinapoylgentiobiose, 1 ,2-diferuloylgentiobiose, l ,2'-disinapoyl-2-feruloylgentiobiose, 3-0- caffeoyl-quinic (neochlorogenic acid)); and vitamins and minerals (e.g., vitamin C, vitamin E, carotene, folic acid, niacin, riboflavin, thiamine, calcium, iron, magnesium, potassium, selenium, and zinc).

In another embodiment, TK-inhibiting herbicides-tolerant plants of the present invention, relative to a wild-type plant, comprise an increased amount of, or an improved profile of, a compound selected from the group consisting of: progoitrin; isothiocyanates; indoles (products of glucosinolate hydrolysis); glutathione; carotenoids such as beta-carotene, lycopene, and the xanthophyll carotenoids such as lutein and zeaxanthin; phenolics comprising the flavonoids such as the flavonols (e.g. quercetin, rutin), the flavans/tannins (such as the procyanidins comprising coumarin, proanthocyanidins, catechins, and anthocyanins); flavones; phytoestrogens such as coumestans, lignans, resveratrol, isoflavones e.g. genistein, daidzein, and glycitein; resorcyclic acid lactones; organosulphur compounds; phytosterols; terpenoids such as carnosol, rosmarinic acid, glycyrrhizin and saponins; chlorophyll; chlorphyllin, sugars, anthocyanins, and vanilla.
[0129] In other embodiments, TK-inhibiting herbicides-tolerant plants of the present

invention, relative to a wild-type plant, comprise an increased amount of, or an improved profile of, a compound selected from the group consisting of: vincristine, vinblastine, taxanes (e.g., taxol (paclitaxel), baccatin III, 10-desacetylbaccatin III, 10-desacetyl taxol, xylosyl taxol, 7- epitaxol, 7-epibaccatin III, 10-desacetylcephalomannine, 7-epicephalomannine, taxotere, cephalomannine, xylosyl cephalomannine, taxagifine, 8-benxoyloxy taxagifine, 9-acetyloxy taxusin, 9-hydroxy taxusin, taiwanxam, taxane Ia, taxane Ib, taxane Ic, taxane Id, GMP paclitaxel, 9-dihydro 13-acetylbaccatin III, 10-desacetyl-7-epitaxol, tetrahydrocannabinol (THC), cannabidiol (CBD), genistein, diadzein, codeine, morphine, quinine, shikonin, ajmalacine, serpentine, and the like.

In other aspects, a method for treating a plant of the present invention is provided.

In some embodiments, the method comprises contacting the plant with an agronomically acceptable composition. In one embodiment, the agronomically acceptable composition comprises an auxinic herbicide A. I.

In another aspect, the present invention provides a method for preparing a descendent seed. The method comprises planting a seed of or capable of producing a plant of the present invention. In one embodiment, the method further comprises growing a descendent plant from the seed; and harvesting a descendant seed from the descendent plant. In other embodiments, the method further comprises applying a TK-inhibiting herbicides herbicidal composition to the descendent plant.

In another embodiment, the invention refers to harvestable parts of the transgenic plant according to the present invention. Preferably, the harvestable parts comprise the TK nucleic acid or TK protein of the present invention. The harvestable parts may be seeds, roots, leaves and/or flowers comprising the TK nucleic acid or TK protein or parts thereof. Preferred parts of soy plants are soy beans comprising the TK nucleic acid or TK protein.

In another embodiment, the invention refers to products derived from a transgenic plant according to the present invention, parts thereof or harvestable parts thereof. A preferred plant product is fodder, seed meal, oil, or seed-treatment-coated seeds. Preferably, the meal and/or oil comprise the TK nucleic acids or TKproteins.

In another embodiment, the invention refers to a method for the production of a product, which method comprises

a) growing the plants of the invention or obtainable by the methods of invention and
b) producing said product from or by the plants of the invention and/or parts, e.g. seeds, of these plants.

In a further embodiment the method comprises the steps

a) growing the plants of the invention,

b) removing the harvestable parts as defined above from the plants and
c) producing said product from or by the harvestable parts of the invention.

The product may be produced at the site where the plant has been grown, the plants and/or parts thereof may be removed from the site where the plants have been grown to produce the product. Typically, the plant is grown, the desired harvestable parts are removed from the plant, if feasible in repeated cycles, and the product made from the harvestable parts of the plant. The step of growing the plant may be performed only once each time the methods of the invention is performed, while allowing repeated times the steps of product production e.g. by repeated removal of harvestable parts of the plants of the invention and if necessary further processing of these parts to arrive at the product. It is also possible that the step of growing the plants of the invention is repeated and plants or harvestable parts are stored until the production of the product is then performed once for the accumulated plants or plant parts. Also, the steps of growing the plants and producing the product may be performed with an overlap in time, even simultaneously to a large extend or sequentially. Generally the plants are grown for some time before the product is produced.

In one embodiment the products produced by said methods of the invention are plant products such as, but not limited to, a foodstuff, feedstuff, a food supplement, feed supplement, fiber, cosmetic and/or pharmaceutical. Foodstuffs are regarded as compositions used for nutrition and/or for supplementing nutrition. Animal feedstuffs and animal feed supplements, in particular, are regarded as foodstuffs.

In another embodiment the inventive methods for the production are used to make agricultural products such as, but not limited to, plant extracts, proteins, amino acids, carbohydrates, fats, oils, polymers, vitamins, and the like.

It is possible that a plant product consists of one or more agricultural products to a large extent.

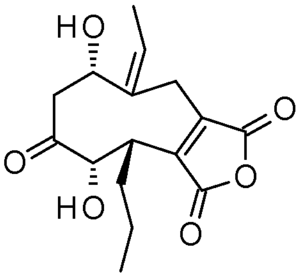
Herbicides

Generally, if the TK-inhibiting herbicides (also referred to as compounds A hereinafter) and/or the herbicidal compounds B as described herein, which can be employed in the context of the present invention, are capable of forming geometrical isomers, for example E/Z isomers, it is possible to use both, the pure isomers and mixtures thereof, in the compositions useful for the present the invention. If the TK-inhibting herbicides A and/or the herbicidal compounds B as described herein have one or more centers of chirality and, as a consequence, are present as enantiomers or diastereomers, it is possible to use both, the pure enantiomers and diastereomers and their mixtures, in the compositions according to the invention. If the TK-inhibting herbicides A and/or the herbicidal compounds B as described herein have ionizable functional groups, they can also be employed in the form of their agriculturally acceptable salts. Suitable are, in general, the salts of those cations and the acid addition salts of those acids whose cations and anions, respectively, have no adverse effect on the activity of the active compounds. Preferred cations are the ions of the alkali metals, preferably of lithium, sodium and potassium, of the alkaline earth metals, preferably of calcium and magnesium, and of the transition metals, preferably of manganese, copper, zinc and iron, further ammonium and substituted ammonium in which one to four hydrogen atoms are replaced by C_1-C_4-alkyl, hydroxy-C_1-C_4-alkyl, C_1-C_4-alkoxy-C_1-C_4-alkyl, hydroxy-C_1-C_4-alkoxy-C_1-C_4-alkyl, phenyl or benzyl, preferably ammonium, methylammonium, isopropylammonium, dimethylammonium, diisopropylammonium, trimethylammonium, heptylammonium, dodecylammonium, tetradecylammonium, tetramethylammonium, tetraethylammonium, tetrabutylammonium, 2-hydroxyethyl-ammonium (olamine salt), 2-(2-hydroxyeth-1-oxy)eth-1-ylammonium (diglycolamine salt), di(2-hydroxyeth-1-yl)ammonium (diolamine salt), tris(2-hydroxyethyl)ammonium (trolamine salt), tris(2-hydroxypropyl)ammonium, benzyltrimethylammonium, benzyltriethylammonium, N,N,N-trimethylethanolammonium (choline salt), furthermore phosphonium ions, sulfonium ions, preferably tri(C_1-C_4-alkyl)sulfonium, such as trimethylsulfonium, and sulfoxonium ions, preferably tri(C_1-C_4-alkyl)sulfoxonium, and finally the salts of polybasic amines such as N,N-bis-(3-aminopropyl)methylamine and diethylenetriamine. Anions of useful acid addition salts are primarily chloride, bromide, fluoride, iodide, hydrogensulfate, methylsulfate, sulfate, dihydrogenphosphate, hydrogenphosphate, nitrate, bicarbonate, carbonate, hexafluorosilicate, hexafluorophosphate, benzoate and also the anions of C_1-C_4-alkanoic acids, preferably formate, acetate, propionate and butyrate.

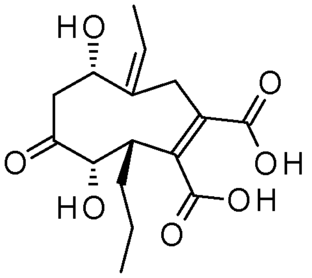
The TK-inhibting herbicides A and/or the herbicidal compounds B as described herein having a carboxyl group can be employed in the form of the acid, in the form of an agriculturally suitable salt as mentioned above or else in the form of an agriculturally acceptable derivative, for example as amides, such as mono- and di-C_1-C_6-alkylamides or arylamides, as esters, for example as allyl esters, propargyl esters, C_1-C_{10}-alkyl esters, alkoxyalkyl esters, tefuryl ((tetrahydrofuran-2-yl)methyl) esters and also as thioesters, for example as C_1-C_{10}-alkylthio esters. Preferred mono- and di-C_1-C_6-alkylamides are the methyl and the dimethylamides. Preferred arylamides are, for example, the anilides and the 2-chloroanilides. Preferred alkyl esters are, for example, the methyl, ethyl, propyl, isopropyl, butyl, isobutyl, pentyl, mexyl (1-methylhexyl), meptyl (1-methylheptyl), heptyl, octyl or isooctyl (2-ethylhexyl) esters. Preferred C_1-C_4-alkoxy-C_1-C_4-alkyl esters are the straight-chain or branched C_1-C_4-alkoxy ethyl esters, for example the 2-methoxyethyl, 2-ethoxyethyl, 2-butoxyethyl (butotyl), 2-butoxypropyl or 3-butoxypropyl ester. An example of a straight-chain or branched C_1-C_{10}-alkylthio ester is the ethylthio ester.

Examples of TK inhibiting herbicides which can be used according to the present invention are compounds having the Formula I.a1, I.a2, I.a3, or I.a4, known to the skilled artisan as (Hydroxy-) Cornexisthin.

Cornexistin (I.a1) and its corresponding dibasic acid (I.a2) are known from EP-A 0 290 193:

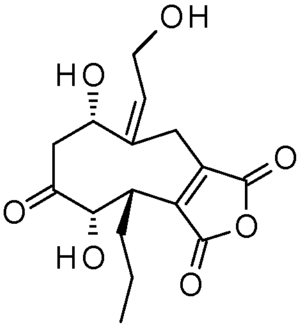


I.a1

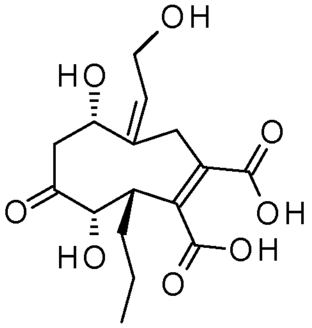


I.a2

Hydroxycornexistin (I.a3) and its corresponding dibasic acid (I.a4) are known from US 5,424,278:



I.a3



I.a4

The herbicidal activity of these compounds is also known from the cited references. JP-A 1990-256 602 discloses mixtures of cornexistin (I.a1) and its dibasic acid (I.a2) with certain herbicides.

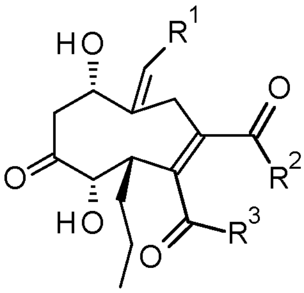
Nevertheless, there is still room for improvement, e.g. regarding activity, scope of activity and compatibility with useful plants of the known herbicidal compositions.

It is an object of the present invention to provide methods of using herbicidal compositions which are highly active against unwanted harmful plants. At the same time, the compositions should have good compatibility with the plants of the present invention. In addition, the compositions useful for the invention should have a broad spectrum of activity. A further object of the present invention is reducing the application rates of active ingredients.

This and further objects are achieved by the herbicidal compositions below.

Accordingly, in one aspect of the invention the TK-inhibiting herbicide usefuld for the present invention refers to a herbicidal composition comprising:

A) at least one herbicidally active compound of formula I (herbicide A)



I

wherein

R^1 is CH_3 or CH_2OH and

R^2 and R^3 together with the neighbouring carbon atoms form a dihydro-2,5-dioxofuran ring or

R^2 and R^3 are OH;

including agriculturally acceptable salts and derivatives thereof;

The TK-inhibiting herbicides described above that are useful to carry out the present invention are often best applied in conjunction with one or more other herbicides to obtain control of a wider variety of undesirable vegetation. For example, TK-inhibiting herbicides may further be used in conjunction with additional herbicides to which the crop plant is naturally tolerant, or to which it is resistant via expression of one or more additional transgenes as mentioned supra. When used in conjunction with other targeting herbicides, the TK-inhibiting herbicides, to which the plant of the present invention had been made resistant or tolerant, can be formulated with the other herbicide or herbicides, tank mixed with the other herbicide or herbicides, or applied sequentially with the other herbicide or herbicides.

Suitable components for mixtures are, for example, selected from the herbicides of class b1) to b15)

B) herbicides of class b1) to b15):
- b1) lipid biosynthesis inhibitors;
- b2) acetolactate synthase inhibitors (ALS inhibitors);
- b3) photosynthesis inhibitors;
- b4) protoporphyrinogen-IX oxidase inhibitors,
- b5) bleacher herbicides;
- b6) enolpyruvyl shikimate 3-phosphate synthase inhibitors (EPSP inhibitors);
- b7) glutamine synthetase inhibitors;
- b8) 7,8-dihydropteroate synthase inhibitors (DHP inhibitors);
- b9) mitosis inhibitors;
- b10) inhibitors of the synthesis of very long chain fatty acids (VLCFA inhibitors);

- b11) cellulose biosynthesis inhibitors;
- b12) decoupler herbicides;
- b13) auxinic herbicides;
- b14) auxin transport inhibitors; and
- b15) other herbicides selected from the group consisting of bromobutide, chlorflurenol, chlorflurenol-methyl, cinmethylin, cumyluron, dalapon, dazomet, difenzoquat, difenzoquat-metilsulfate, dimethipin, DSMA, dymron, endothal and its salts, etobenzanid, flamprop, flamprop-isopropyl, flamprop-methyl, flamprop-M-isopropyl, flamprop-M-methyl, flurenol, flurenol-butyl, flurprimidol, fosamine, fosamine-ammonium, indanofan, indaziflam, maleic hydrazide, mefluidide, metam, methiozolin (CAS 403640-27-7), methyl azide, methyl bromide, methyl-dymron, methyl iodide, MSMA, oleic acid, oxaziclomefone, pelargonic acid, pyributicarb, quinoclamine, triaziflam, tridiphane and 6-chloro-3-(2-cyclopropyl-6-methylphenoxy)-4-pyridazinol (CAS 499223-49-3) and its salts and esters;

including their agriculturally acceptable salts or derivatives.

The compositions according to the invention are suitable as herbicides as such or as appropriately formulated compositions (agrochemical compositions). As used herein, the term “agrochemical composition” refers to a composition comprising a pesticidally effective amount of at least one active ingredient and at least one auxiliary customary for agrochemical compositions.

The invention relates in particular to the use of compositions in the form of herbicidally active agrochemical compositions comprising a herbicidally effective amount of A) at least one compound of formula I (herbicide A) and B) at least one further compound selected from the herbicides of groups b1) to b15) (herbicide B), as defined above, and also at least one liquid and/or solid carrier and/or one or more surfactants and, if desired, one or more further auxiliaries customary for agrochemical compositions.

The invention also relates to compositions in the form of an agrochemical composition, which is a 1-component composition comprising at least one herbicide A and at least one further active compound selected from the herbicides B, and at least one solid or liquid carrier and/or one or more surfactants and, if desired, one or more further auxiliaries customary for agrochemical compositions.

The invention also relates to compositions in the form of an agrochemical composition, which is a 2-component composition comprising a first component comprising at least one herbicide A, a solid or liquid carrier and/or one or more surfactants, and a second component comprising at least one herbicide B, a solid or liquid carrier and/or one or more surfactants, where additionally both components may also comprise further auxiliaries customary for agrochemical compositions.

Moreover, the time frame, within which the desired herbicidal action can be achieved, may be expanded by the compositions according to the invention comprising at least one herbicide A and at least one herbicide B and optionally a safener C as defined below. This allows a more flexibly timed application of the compositions according to the invention in comparison with the single compounds.

Safeners are chemical compounds which prevent or reduce damage on useful plants without having a major impact on the herbicidal action of the herbicidal active components towards unwanted plants. Safeners can be applied before sowings (e.g. seed treatments), on shoots or seedlings as well as in the pre-emergence or post-emergence treatment of useful plants and their habitat.

The compositions according to the invention comprising both at least one herbicide A and at least one safener C as defined below also have good herbicidal activity against harmful plants and better compatibility with useful plants.

Surprisingly, compositions according to the invention comprising at least one herbicide A at least one herbicide B and at least one safener C have better herbicidal activity, i.e. better activity against harmful plants, than would have been expected based on the herbicidal activity observed for the individual compounds, or a broader activity spectrum, and show better compatibility with useful plants than compositions comprising only one herbicide A and one herbicide B.

Therefore, in one embodiment of the present invention the compositions comprise at least one herbicide A, at least one herbicide B and at least one safener C.

Examples of suitable safeners C are benoxacor, cloquintocet, cyometrinil, cyprosulfamide, dichlormid, dicyclonon, dietholate, fenchlorazole, fenclorim, flurazole, fluxofenim, furilazole, isoxadifen, mefenpyr, mephenate, naphthalic anhydride, oxabetrinil, 4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3), 2,2,5-trimethyl-3-(dichloroacetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4) and N-(2-methoxybenzoyl)-4-[(methylamino-carbonyl)amino]benzenesulfonamide (CAS 129531-12-0).

The safeners C, the herbicides A and the herbicides B can be applied simultaneously or in succession.

The term “cornexistin” means the compound of formula (I.a1) as well as agriculturally acceptable salts thereof.

The term “dibasic acid of cornexistin” means the compound of formula (I.a2) as well as agriculturally acceptable salts thereof.

The term “hydroxycornexistin” means the compound of formula (I.a3) as well as agriculturally acceptable salts thereof.

The term “dibasic acid of hydroxycornexistin” means the compound of formula (I.a4) as well as agriculturally acceptable salts thereof.

The compounds of formulae (I.a1) to (I.a4) as described herein are capable of forming geometrical isomers, for example E/Z isomers. Accordingly, the terms “cornexistin”, “dibasic acid of cornexistin”, “hydroxycornexistin” and “dibasic acid of hydroxycornexistin” also encompass the pure E or Z isomers and mixtures thereof.

The term “agriculturally acceptable salts” is used herein to mean in general, the salts of those cations and the acid addition salts of those acids whose cations and anions, respectively, have no adverse effect on the herbicidal activity of the dibasic acid of cornexistin and the dibasic acid of hydroxycornexistin.

Preferred cations are the ions of the alkali metals, preferably of lithium, sodium and potassium, of the alkaline earth metals, preferably of calcium and magnesium, and of the transition metals, preferably of manganese, copper, zinc and iron, further ammonium and substituted ammonium in which one to four hydrogen atoms are replaced by C_1-C_4-alkyl, hydroxy-C_1-C_4-alkyl, C_1-C_4-alkoxy-C_1-C_4-alkyl, hydroxy-C_1-C_4-alkoxy-C_1-C_4-alkyl, phenyl or benzyl, preferably ammonium, methylammonium, isopropylammonium, dimethylammonium, diisopropylammonium, trimethylammonium, heptylammonium, dodecylammonium, tetradecylammonium, tetramethylammonium, tetraethylammonium, tetrabutylammonium, 2 hydroxyethyl-ammonium (olamine salt), 2-(2-hydroxyeth-1-oxy)eth-1-ylammonium (diglycolamine salt), di(2-hydroxyeth-1-yl)-ammonium (diolamine salt), tris(2-hydroxyethyl)ammonium (trolamine salt), tris(2-hydroxy-propyl)ammonium, benzyltrimethylammonium, benzyltriethylammonium, N,N,N-trimethylethanol-ammonium (choline salt), furthermore phosphonium ions, sulfonium ions, preferably tri(C_1-C_4-alkyl)sulfonium, such as trimethylsulfonium, and sulfoxonium ions, preferably tri(C_1-C_4-alkyl)sulfoxonium, and finally the salts of polybasic amines such as N,N-bis-(3-amino-propyl)methylamine and diethylenetriamine.

Anions of useful acid addition salts are primarily chloride, bromide, fluoride, iodide, hydrogensulfate, methylsulfate, sulfate, dihydrogenphosphate, hydrogen¬phosphate, nitrate, bicarbonate, carbonate, hexafluorosilicate, hexafluorophosphate, benzoate and also the anions of C_1-C_4-alkanoic acids, preferably formate, acetate, propionate and butyrate.

If the herbicidal compounds B and/or D and/or safeners C as described herein are capable of forming geometrical isomers, for example E/Z isomers, it is possible to use both, the pure isomers and mixtures thereof, in the compositions according to the invention. If the herbicidal compounds B and/or D and/or safeners C as described herein have one or more centers of chirality and, as a consequence, are present as enantiomers or diastereomers, it is possible to use both, the pure enantiomers and diastereomers and their mixtures, in the compositions according to the invention. If the herbicidal compounds B and/or D and/or safeners C as described herein have ionizable functional groups, they can also be employed in the form of their agriculturally acceptable salts. Suitable are, in general, the salts of those cations and the acid addition salts of those acids whose cations and anions, respectively, have no adverse effect on the activity of the active compounds.

Preferred cations and anions are the ones listed for the compounds of formula I.

Herbicidal compounds A, B and/or D and/or safeners C as described herein having a carboxyl, hydroxy and/or an amino group can be employed as such or in form of an agriculturally suitable salt as mentioned above or else in the form of an agriculturally acceptable derivative, for example as amides, such as mono- and di-C_1-C_6-alkylamides or arylamides, as esters, for example as allyl esters, propargyl esters, C_1-C_{10}-alkyl esters, alkoxyalkyl esters, tefuryl ((tetrahydrofuran-2-yl)methyl) esters and also as thioesters, for example as C_1-C_{10}-alkylthio esters. Preferred mono- and di-C_1-C_6-alkylamides are the methyl and the dimethylamides. Preferred arylamides are, for example, the anilides and the 2-chloroanilides. Preferred alkyl esters are, for example, the methyl, ethyl, propyl, isopropyl, butyl, isobutyl, pentyl, mexyl (1-methylhexyl), meptyl (1-methylheptyl), heptyl, octyl or isooctyl (2-ethylhexyl) esters. Preferred C_1-C_4-alkoxy-C_1-C_4-alkyl esters are the straight-chain or branched C_1-C_4-alkoxy ethyl esters, for example the 2-methoxyethyl, 2-ethoxyethyl, 2-butoxyethyl (butotyl), 2-butoxypropyl or 3-butoxypropyl ester. An example of a straight-chain or branched C_1-C_{10}-alkylthio ester is the ethylthio ester.

Further embodiments of the invention are evident from the claims, the description and the examples. It is to be understood that the features mentioned above and still to be illustrated below of the subject matter of the invention can be applied not only in the combination given in each particular case but also in other combinations, without leaving the scope of the invention.

The preferred embodiments of the invention mentioned herein below have to be understood as being preferred either independently from each other or in combination with one another.

Preferably, the active ingredient A is selected from the following herbicides A:

a1) cornexistin (I.a1),
a2) the dibasic acid of cornexistin (I.a2),
a3) hydroxycornexistin (I.a3),
a4) the dibasic acid of hydroxycornexistin (I.a4),
a5) mixtures of cornexistin (I.a1) and the dibasic acid of cornexistin (I.a2),
a6) mixtures of hydroxycornexistin (I.a3) and the dibasic acid of hydroxycornexistin (I.a4)
a7) mixtures of cornexistin (I.a1) and hydroxycornexistin (I.a3),
a8) mixtures of cornexistin (I.a1) and the dibasic acid of hydroxycornexistin (I.a4),
a9) mixtures of the dibasic acid of cornexistin (I.a2) and hydroxycornexistin (I.a3),
a10) mixtures of the dibasic acid of cornexistin (I.a2) and the dibasic acid of hydroxycornexistin (I.a4) and

a11) mixtures cornexistin (I.a1), the dibasic acid of cornexistin (I.a2), hydroxycornexistin (I.a3) and the dibasic acid of hydroxycornexistin (I.a4)

including agriculturally acceptable salts and derivatives thereof.

Prefered are the compositions according to the present invention comprising at least one, preferably exactly one herbicide A and at least one herbicide B.

Further preferred are compositions according to the present invention comprising at least two, preferably exactly two herbicides A and at least one herbicide B.

In one preferred embodiment of the invention the compositions contain at least one inhibitor of the lipid biosynthesis (herbicide b1) through inhibition of acetylCoA carboxylase (hereinafter termed ACC herbicides). The ACC herbicides belong to the group A of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one inhibitor of photosynthesis (herbicide b3) on diverting the electron transfer in photosystem I in plants (so-called PSI inhibitors, group D of HRAC classification) and thus on an inhibition of photosynthesis.

According to a further embodiment of the invention the compositions contain at least one inhibitor of protoporphyrinogen-IX-oxidase (herbicide b4). The herbicidal activity of these compounds is based on the inhibition of the protoporphyrinogen-IX-oxidase. These inhibitors belong to the group E of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one bleacher-herbicide (herbicide b5). The herbicidal activity of these compounds is based on the inhibition of the carotenoid biosynthesis. These include compounds which inhibit carotenoid biosynthesis by inhibition of phytoene desaturase (so-called PDS inhibitors, group F1 of HRAC classification), compounds that inhibit the 4-hydroxyphenylpyruvate-dioxygenase (HPPD inhibitors, group F2 of HRAC classification), compounds that inhibit DOXsynthase (group F4 of HRAC class) and compounds which inhibit carotenoid biosynthesis by an unknown mode of action (bleacher - unknown target, group F3 of HRAC classification).

According to a further embodiment of the invention the compositions contain at least one EPSP synthase inhibitor (herbicide b6). The herbicidal activity of these compounds is based on the inhibition of enolpyruvyl shikimate 3-phosphate synthase, and thus on the inhibition of the amino acid biosynthesis in plants. These inhibitors belong to the group G of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one glutamine synthetase inhibitor (herbicide b7). The herbicidal activity of these compounds is based on the inhibition of glutamine synthetase, and thus on the inhibition of the aminoacid biosynthesis in plants. These inhibitors belong to the group H of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one DHP synthase inhibitor (herbicide b8). The herbicidal activity of these compounds is based on the inhibition of 7,8-dihydropteroate synthase. These inhibitors belong to the group I of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one cellulose biosynthesis inhibitor (herbicide b11). The herbicidal activity of these compounds is based on the inhibition of the biosynthesis of cellulose and thus on the inhibition of the synthesis of cell walls in plants. These inhibitors belong to the group L of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one decoupler herbicide (herbicide b12). The herbicidal activity of these compounds is based on the disruption of the cell membrane. These inhibitors belong to the group M of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one auxinic herbicide (herbicide b13). These include compounds that mimic auxins, i.e. plant hormones, and affect the growth of the plants. These compounds belong to the group O of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one auxin transport inhibitor (herbicide b14). The herbicidal activity of these compounds is based on the inhibition of the auxin transport in plants. These compounds belong to the group P of the HRAC classification system.

According to a further embodiment of the invention the compositions contain at least one other herbicide (herbicide b15) selected from the group consisting of bromobutide, chlorflurenol, chlorflurenol-methyl, cinmethylin, cumyluron, dalapon, dazomet, difenzoquat, difenzoquat-metilsulfate, dimethipin, DSMA, dymron, endothal and its salts, etobenzanid, flamprop, flamprop-isopropyl, flamprop-methyl, flamprop-M-isopropyl, flamprop-M-methyl, flurenol, flurenol-butyl, flurprimidol, fosamine, fosamine-ammonium, indanofan, maleic hydrazide, mefluidide, metam, methiozolin (CAS 403640-27-7), methyl azide, methyl bromide, methyl-daimuron, methyl iodide, MSMA, oleic acid, oxaziclomefone, pelargonic acid, pyributicarb, quinoclamine, and 6-chloro-3-(2-cyclopropyl-6-methylphenoxy)-4-pyridazinol (CAS 499223-49-3). These compounds have an unknown mode of action and belong to the group Z of the HRAC classification system.

As to the given mechanisms of action and classification of the active compounds, see e.g. "HRAC, Classification of Herbicides According to Mode of Action", (http://www.plantprotection.org/hrac/MOA.html).

Examples of herbicides B which can be used in combination with the TK-inhibiting herbicides according to the present invention are:

b1) from the group of the lipid biosynthesis inhibitors:
ACC-herbicides such as alloxydim, alloxydim-sodium, butroxydim, clethodim, clodinafop, clodinafop-propargyl, cycloxydim, cyhalofop, cyhalofop-butyl, diclofop, diclofop-methyl, fenoxaprop, fenoxaprop-ethyl, fenoxaprop-P, fenoxaprop-P-ethyl, fluazifop, fluazifop-butyl, fluazifop-P, fluazifop-P-butyl, haloxyfop, haloxyfop-methyl, haloxyfop-P, haloxyfop-P-methyl, metamifop, pinoxaden, profoxydim, propaquizafop, quizalofop, quizalofop-ethyl, quizalofop-tefuryl, quizalofop-P, quizalofop-P-ethyl, quizalofop-P-tefuryl, sethoxydim, tepraloxydim, tralkoxydim,
4-(4'-Chloro-4-cyclopropyl-2'-fluoro[1,1'-biphenyl]-3-yl)-5-hydroxy-2,2,6,6-tetramethyl-2H-pyran-3(6H)-one (CAS 1312337-72-6); 4-(2',4'-Dichloro-4-cyclopropyl[1,1'-biphenyl]-3-yl)-5-hydroxy-2,2,6,6-tetramethyl-2H-pyran-3(6H)-one (CAS 1312337-45-3); 4-(4'-Chloro-4-ethyl-2'-fluoro[1,1'-biphenyl]-3-yl)-5-hydroxy-2,2,6,6-tetramethyl-2H-pyran-3(6H)-one (CAS 1033757-93-5); 4-(2',4'-Dichloro-4-ethyl[1,1'-biphenyl]-3-yl)-2,2,6,6-tetramethyl-2H-pyran-3,5(4H,6H)-dione (CAS 1312340-84-3); 5-(Acetyloxy)-4-(4'-chloro-4-cyclopropyl-2'-fluoro[1,1'-biphenyl]-3-yl)-3,6-dihydro-2,2,6,6-tetramethyl-2H-pyran-3-one (CAS 1312337-48-6); 5-(Acetyloxy)-4-(2′,4'-dichloro-4-cyclopropyl- [1,1'-biphenyl]-3-yl)-3,6-dihydro-2,2,6,6-tetramethyl-2H-pyran-3-one; 5-(Acetyloxy)-4-(4'-chloro-4-ethyl-2'-fluoro[1,1'-biphenyl]-3-yl)-3,6-dihydro-2,2,6,6-tetramethyl-2H-pyran-3-one (CAS 1312340-82-1); 5-(Acetyloxy)-4-(2',4'-dichloro-4-ethyl[1,1'-biphenyl]-3-yl)-3,6-dihydro-2,2,6,6-tetramethyl-2H-pyran-3-one (CAS 1033760-55-2); 4-(4'-Chloro-4-cyclopropyl-2'-fluoro[1,1'-biphenyl]-3-yl)-5,6-dihydro-2,2,6,6-tetramethyl-5-oxo-2H-pyran-3-yl carbonic acid methyl ester (CAS 1312337-51-1); 4-(2′,4'-Dichloro -4-cyclopropyl- [1,1'-biphenyl]-3-yl)-5,6-dihydro-2,2,6,6-tetramethyl-5-oxo-2H-pyran-3-yl carbonic acid methyl ester; 4-(4'-Chloro-4-ethyl-2'-fluoro[1,1'-biphenyl]-3-yl)-5,6-dihydro-2,2,6,6-tetramethyl-5-oxo-2H-pyran-3-yl carbonic acid methyl ester (CAS 1312340-83-2); 4-(2',4'-Dichloro-4-ethyl[1,1'-biphenyl]-3-yl)-5,6-dihydro-2,2,6,6-tetramethyl-5-oxo-2H-pyran-3-yl carbonic acid methyl ester (CAS 1033760-58-5); and non ACC herbicides such as benfuresate, butylate, cycloate, dalapon, dimepiperate, EPTC, esprocarb, ethofumesate, flupropanate, molinate, orbencarb, pebulate, prosulfocarb, TCA, thiobencarb, tiocarbazil, triallate and vernolate;

b2) from the group of the ALS inhibitors:
sulfonylureas such as amidosulfuron, azimsulfuron, bensulfuron, bensulfuron-methyl, chlorimuron, chlorimuron-ethyl, chlorsulfuron, cinosulfuron, cyclosulfamuron, ethametsulfuron, ethametsulfuron-methyl, ethoxysulfuron, flazasulfuron, flucetosulfuron, flupyrsulfuron, flupyrsulfuron-methyl-sodium, foramsulfuron, halosulfuron, halosulfuron-methyl, imazosulfuron, iodosulfuron, iodosulfuron-methyl-sodium, iofensulfuron, iofensulfuron-sodium, mesosulfuron, metazosulfuron, metsulfuron, metsulfuron-methyl,

nicosulfuron, orthosulfamuron, oxasulfuron, primisulfuron, primisulfuron-methyl, propyrisulfuron, prosulfuron, pyrazosulfuron, pyrazosulfuron-ethyl, rimsulfuron, sulfometuron, sulfometuron-methyl, sulfosulfuron, thifensulfuron, thifensulfuron-methyl, triasulfuron, tribenuron, tribenuron-methyl, trifloxysulfuron, triflusulfuron, triflusulfuron-methyl and tritosulfuron,
imidazolinones such as imazamethabenz, imazamethabenz-methyl, imazamox, imazapic, imazapyr, imazaquin and imazethapyr, triazolopyrimidine herbicides and sulfonanilides such as cloransulam, cloransulam-methyl, diclosulam, flumetsulam, florasulam, metosulam, penoxsulam, pyrimisulfan and pyroxsulam,
pyrimidinylbenzoates such as bispyribac, bispyribac-sodium, pyribenzoxim, pyriftalid, pyriminobac, pyriminobac-methyl, pyrithiobac, pyrithiobac-sodium, 4-[[[2-[(4,6-dimethoxy-2-pyrimidinyl)oxy]phenyl]methyl]amino]-benzoic acid-1-methylethyl ester (CAS 420138-41-6), 4-[[[2-[(4,6-dimethoxy-2-pyrimidinyl)oxy]phenyl]methyl]amino]-benzoic acid propyl ester (CAS 420138-40-5), N-(4-bromophenyl)-2-[(4,6-dimethoxy-2-pyrimidinyl)oxy]benzenemethanamine (CAS 420138-01-8),
sulfonylaminocarbonyl-triazolinone herbicides such as flucarbazone, flucarbazone-sodium, propoxycarbazone, propoxycarbazone-sodium, thiencarbazone and thiencarbazone-methyl; and triafamone;
among these, a preferred embodiment of the invention relates to those compositions comprising at least one imidazolinone herbicide;

b3) from the group of the photosynthesis inhibitors:
amicarbazone, inhibitors of the photosystem II, e.g. triazine herbicides, including of chlorotriazine, triazinones, triazindiones, methylthiotriazines and pyridazinones such as ametryn, atrazine, chloridazone, cyanazine, desmetryn, dimethametryn,hexazinone, metribuzin, prometon, prometryn, propazine, simazine, simetryn, terbumeton, terbuthylazin, terbutryn and trietazin, aryl urea such as chlorobromuron, chlorotoluron, chloroxuron, dimefuron, diuron, fluometuron, isoproturon, isouron, linuron, metamitron, methabenzthiazuron, metobenzuron, metoxuron, monolinuron, neburon, siduron, tebuthiuron and thiadiazuron, phenyl carbamates such as desmedipham, karbutilat, phenmedipham, phenmedipham-ethyl, nitrile herbicides such as bromofenoxim, bromoxynil and its salts and esters, ioxynil and its salts and esters, uraciles such as bromacil, lenacil and terbacil, and bentazon and bentazon-sodium, pyridate, pyridafol, pentanochlor and propanil and inhibitors of the photosystem I such as diquat, diquat-dibromide, paraquat, paraquat-dichloride and paraquat-dimetilsulfate. Among these, a preferred embodiment of the invention relates to those compositions comprising at least one aryl urea herbicide. Among these, likewise a preferred embodiment of the invention relates to those compositions comprising at least one triazine herbicide. Among these, likewise a preferred embodiment of the invention relates to those compositions comprising at least one nitrile herbicide;

b4) from the group of the protoporphyrinogen-IX oxidase inhibitors:

acifluorfen, acifluorfen-sodium, azafenidin, bencarbazone, benzfendizone, bifenox, butafenacil, carfentrazone, carfentrazone-ethyl, chlomethoxyfen, cinidon-ethyl, fluazolate, flufenpyr, flufenpyr-ethyl, flumiclorac, flumiclorac-pentyl, flumioxazin, fluoroglycofen, fluoroglycofen-ethyl, fluthiacet, fluthiacet-methyl, fomesafen, halosafen, lactofen, oxadiargyl, oxadiazon, oxyfluorfen, pentoxazone, profluazol, pyraclonil, pyraflufen, pyraflufen-ethyl, saflufenacil, sulfentrazone, thidiazimin, tiafenacil, ethyl [3-[2-chloro-4-fluoro-5-(1-methyl-6-trifluoromethyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-3-yl)phenoxy]-2-pyridyloxy]acetate (CAS 353292-31-6; S-3100, N-ethyl-3-(2,6-dichloro-4-trifluoromethylphenoxy)-5-methyl-1*H*-pyrazole-1-carboxamide (CAS 452098-92-9), N-tetrahydrofurfuryl-3-(2,6-dichloro-4-trifluoromethylphenoxy)-5-methyl-1*H*-pyrazole-1-carboxamide (CAS 915396-43-9), N-ethyl-3-(2-chloro-6-fluoro-4-trifluoromethylphenoxy)-5-methyl-1*H*-pyrazole-1-carboxamide (CAS 452099-05-7), N-tetrahydrofurfuryl-3-(2-chloro-6-fluoro-4-trifluoromethylphenoxy)-5-methyl-1*H*-pyrazole-1-carboxamide (CAS 452100-03-7), 3-[7-fluoro-3-oxo-4-(prop-2-ynyl)-3,4-dihydro-2H-benzo[1,4]oxazin-6-yl]-1,5-dimethyl-6-thioxo-[1,3,5]triazinan-2,4-dione, 1,5-dimethyl-6-thioxo-3-(2,2,7-trifluoro-3-oxo-4-(prop-2-ynyl)-3,4-dihydro-2H-benzo[b][1,4]oxazin-6-yl)-1,3,5-triazinane-2,4-dione (CAS 1258836-72-4), 2-(2,2,7-Trifluoro-3-oxo-4-prop-2-ynyl-3,4-dihydro-2H-benzo[1,4]oxazin-6-yl)-4,5,6,7-tetrahydro-isoindole-1,3-dione, 1-Methyl-6-trifluoromethyl-3-(2,2,7-trifluoro-3-oxo-4-prop-2-ynyl-3,4-dihydro-2H-benzo[1,4]oxazin-6-yl)-1H-pyrimidine-2,4-dione (CAS 1304113-05-0), methyl (*E*)-4-[2-chloro-5-[4-chloro-5-(difluoromethoxy)-1*H*-methyl-pyrazol-3-yl]-4-fluoro-phenoxy]-3-methoxy-but-2-enoate [CAS 948893-00-3*]*, and 3-[7-Chloro-5-fluoro-2-(trifluoromethyl)-1H-benzimidazol-4-yl]-1-methyl-6-(trifluoromethyl)-1H-pyrimidine-2,4-dione (CAS 212754-02-4);

b5) from the group of the bleacher herbicides:
PDS inhibitors: beflubutamid, diflufenican, fluridone, flurochloridone, flurtamone, norflurazon, picolinafen, and 4-(3-trifluoromethylphenoxy)-2-(4-trifluoromethylphenyl)-pyrimidine (CAS 180608-33-7), HPPD inhibitors: benzobicyclon, benzofenap, clomazone, isoxaflutole, mesotrione, pyrasulfotole, pyrazolynate, pyrazoxyfen, sulcotrione, tefuryltrione, tembotrione, topramezone and bicyclopyrone, bleacher, unknown target: aclonifen, amitrole and flumeturon;

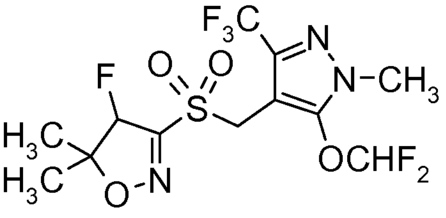
b6) from the group of the EPSP synthase inhibitors:
glyphosate, glyphosate-isopropylammonium, glyposate-potassium and glyphosate-trimesium (sulfosate);

b7) from the group of the glutamine synthase inhibitors:
bilanaphos (bialaphos), bilanaphos-sodium, glufosinate, glufosinate-P and glufosinate-ammonium;

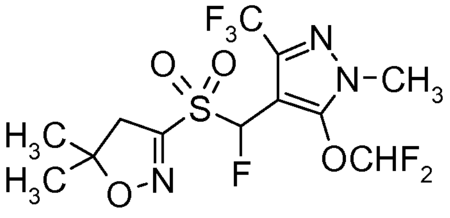
b8) from the group of the DHP synthase inhibitors:
asulam;

b9) from the group of the mitosis inhibitors:
compounds of group K1: dinitroanilines such as benfluralin, butralin, dinitramine, ethalfluralin, fluchloralin, oryzalin, pendimethalin, prodiamine and trifluralin, phosphoramidates such as amiprophos, amiprophos-methyl, and butamiphos, benzoic acid herbicides such as chlorthal, chlorthal-dimethyl, pyridines such as dithiopyr and thiazopyr, benzamides such as propyzamide and tebutam; compounds of group K2: chlorpropham, propham and carbetamide, among these, compounds of group K1, in particular dinitroanilines are preferred;

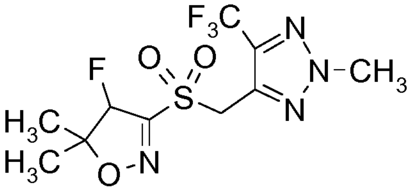
b10) from the group of the VLCFA inhibitors:
chloroacetamides such as acetochlor, alachlor, butachlor, dimethachlor, dimethenamid, dimethenamid-P, metazachlor, metolachlor, metolachlor-S, pethoxamid, pretilachlor, propachlor, propisochlor and thenylchlor, oxyacetanilides such as flufenacet and mefenacet, acetanilides such as diphenamid, naproanilide and napropamide, tetrazolinones such fentrazamide, and other herbicides such as anilofos, cafenstrole, fenoxasulfone, ipfencarbazone, piperophos, pyroxasulfone and isoxazoline compounds of the formulae II.1, II.2, II.3, II.4, II.5, II.6, II.7, II.8 and II.9



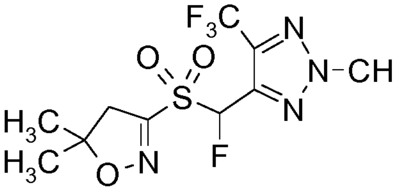
II.1



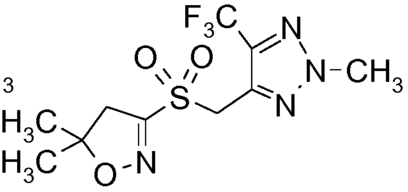
II.2



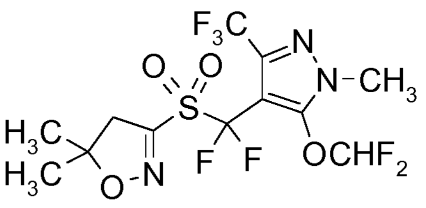
II.3



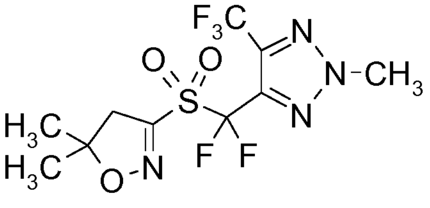
II.4



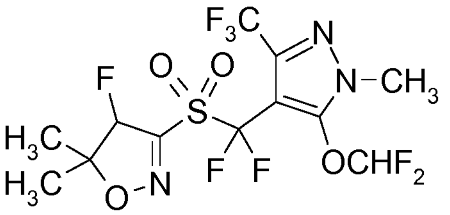
II.5



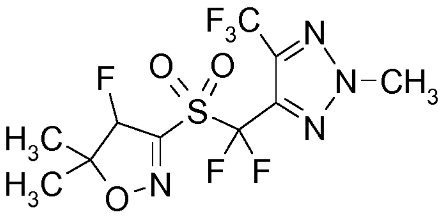
II.6



II.7



II.8



II.9

the isoxazoline compounds of the formula (I)I are known in the art, e.g. from WO 2006/024820, WO 2006/037945, WO 2007/071900 and WO 2007/096576;

among the VLCFA inhibitors, preference is given to chloroacetamides and oxyacetamides;

b11) from the group of the cellulose biosynthesis inhibitors:
chlorthiamid, dichlobenil, flupoxam, indaziflam, triaziflam, isoxaben and 1-Cyclohexyl-5-pentafluorphenyloxy-1^4-[1,2,4,6]thiatriazin-3-ylamine;

b12) from the group of the decoupler herbicides:
dinoseb, dinoterb and DNOC and its salts;

b13) from the group of the auxinic herbicides:
2,4-D and its salts and esters such as clacyfos, 2,4-DB and its salts and esters, aminocyclopyrachlor and its salts and esters, aminopyralid and its salts such as aminopyralid-tris(2-hydroxypropyl)ammonium and its esters, benazolin, benazolin-ethyl, chloramben and its salts and esters, clomeprop, clopyralid and its salts and esters, dicamba and its salts and esters, dichlorprop and its salts and esters, dichlorprop-P and its salts and esters, fluroxypyr, fluroxypyr-butometyl, fluroxypyr-meptyl, halauxifen and its salts and esters (CAS 943832-60-8); MCPA and its salts and esters, MCPA-thioethyl, MCPB and its salts and esters, MCPP and its salts and esters, mecoprop and its salts and esters, mecoprop-P and its salts and esters, picloram and its salts and esters, quinclorac, quinmerac, TBA (2,3,6) and its salts and esters and triclopyr and its salts and esters;

b14) from the group of the auxin transport inhibitors: diflufenzopyr, diflufenzopyr-sodium, naptalam and naptalam-sodium;

b15) from the group of the other herbicides: bromobutide, chlorflurenol, chlorflurenol-methyl, cinmethylin, cumyluron, cyclopyrimorate (CAS 499223-49-3) and its salts and esters, dalapon, dazomet, difenzoquat, difenzoquat-metilsulfate, dimethipin, DSMA, dymron, endothal and its salts, etobenzanid, flamprop, flamprop-isopropyl, flamprop-methyl, flamprop-M-isopropyl, flamprop-M-methyl, flurenol, flurenol-butyl, flurprimidol, fosamine, fosamine-ammonium, indanofan, indaziflam, maleic hydrazide, mefluidide, metam, methiozolin (CAS 403640-27-7), methyl azide, methyl bromide, methyl-dymron, methyl iodide, MSMA, oleic acid, oxaziclomefone, pelargonic acid, pyributicarb, quinoclamine, triaziflam and tridiphane..

Most preferred herbicides B that can be used in combination with herbicides A are given in table 3 below:

Table 3

	Herbicide B
B.1	clethodim
B.2	clodinafop-propargyl
B.3	cycloxydim
B.4	pinoxaden
B.5	sethodydim
B.6	paraquat
B.7	carfentrazone-ethyl
B.8	flumioxazin
B.9	lactofen
B.10	oxadiargyl
B.11	oxyflurofen
B.12	saflufenacil
B.13	sulfentrazone
B.14	1,5-dimethyl-6-thioxo-3-(2,2,7-trifluoro-3-oxo-4-(prop-2-ynyl)-3,4-dihydro-2H-benzo[b][1,4]oxazin-6-yl)-1,3,5-triazinane-2,4-dione
B.15	benzobicylone
B.16	bicyclopyrone
B.17	clomazone
B.18	difufenican
B.19	isoxaflutole
B.20	mesotrione
B.21	norflurazon
B.22	picolinafen
B.23	sulcotrione
B.24	tembotrione
B.25	topramezone
B.26	glyphosate
B.27	glufosinate
B.28	indaziflam
B.29	isoxaben
B.30	dinoseb
B.31	2,4D
B.32	dicamba
B.33	fluroxypyr
B.34	MCPA
B.35	quinmerac
B.36	quinclorac
B.37	diflufenzopyr
B.38	daimuron
B.39	pelargonic acid

Any of the herbicides B.1 to B.39 having a carboxylic, amino and/or hydroxyl group are to be understood as to be employed as such and/or in form of agriculturally acceptable salts thereof and/or in form of agriculturally acceptable esters or amides thereof.

According to a preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide B.

According to another preferred embodiment of the invention, the composition comprises at least two, preferably exactly two herbicides B different from each other.

According to another preferred embodiment of the invention, the composition comprises at least three, preferably exactly three herbicides B different from each other.

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A and at least one, preferably exactly one, herbicide B.

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A and at least two, preferably exactly two, herbicides B different from each other.

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A and at least three, preferably exactly three, herbicides B different from each other.

According to one preferred embodiment of the invention, the composition comprises at least one herbicide B selected from compounds of groups b1), b3), b4), b5), b6), b7), b8), b11), b12), b13), b14) and b15), preferably selected from compounds of groups b1), b3), b4), b5), b6), b7), b8), b11), b12), b13), b14) or b15).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A, preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b1), in particular selected from clethodim (B.1), clodinafop-propargyl (B.2), cycloxydim (B.3), pinoxaden (B.4) and sethodydim (B.5).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b3), in particular paraquat and agriculturally acceptable salts thereof (B.6).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b4), in particular selected from carfentrazone-ethyl (B.7), flumioxazin (B.8), lactofen (B.9), oxadiargyl (B.10), oxyflurofen (B.11), saflufenacil (B.12), sulfentrazone (B.13) and 1,5-dimethyl-6-thioxo-3-(2,2,7-trifluoro-3-oxo-4-(prop-2-ynyl)-3,4-dihydro-2H-benzo[b][1,4]oxazin-6-yl)-1,3,5-triazinane-2,4-dione (CAS 1258836-72-4) (B.14).

According to another preferred embodiment of the invention, the composition comprises at leats one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b5), preferably selected from benzobicyclone (B.15), bicyclopyrone (B.16), clomazone (B.17), diflufenican (B.18), isoxaflutole (B.19), mesotrione (B.20), norflurazon (B.21), picolinafen (B.22), sulcotrione (B.23), tembotrione (B.24), and topramezone (B.25), in particular selected from benzobicyclone (B.15), isoxaflutole (B.19), mesotrione (B.20) and topramezone (B.25).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b6), in particular glyphosate and agriculturally acceptable salts thereof (B.26).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b7), in particular glufosinate and agriculturally acceptable salts thereof (B.27).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), at least one, preferably exactly one herbicide B selected from the compounds of group b11), in particular selected from indaziflam (B.28) and isoxaben (B.29).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), at least one, preferably exactly one herbicide B selected from the compounds of group b12), in particular dinoseb (B.30).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b13), in particular from 2,4D and agriculturally acceptable salts thereof (B.31), dicamba and agriculturally acceptable salts thereof (B.32), fluroxypyr (B.33), MCPA and agriculturally acceptable salts thereof (B.34), quinmerac and agriculturally acceptable salts thereof (B.35) and quinclorac and agriculturally acceptable salts thereof (B.36).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b14), in particular diflufenzopyr and agriculturally acceptable salts thereof (B.37).

According to another preferred embodiment of the invention, the composition comprises at least one, preferably exactly one herbicide A selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) and a11), most preferably selected from a1), a2), a3), a4), a5), a6), a7), a8), a9), a10) or a11), and at least one, preferably exactly one herbicide B selected from the compounds of group b15), in particular selected from dymron (= daimuron) (B.38) and pelargonic acid and agriculturally acceptable salts thereof (B.39).

In another embodiment, compositions according to the invention comprise at least two herbicides B, whereby a first herbicide B is selected from the compounds of groups b1), b3), b4), b5), b8), b11), b12, b13), b14 and b15), preferably selected from the compounds of groups b1), b3), b4), b5), b8), b11), b12, b13), b14 or b15) and a second herbicide B is selected from the compounds of groups b6) and b7), preferably selected from the compounds of groups b6) or b7).

Moreover, it may be useful to apply the TK-inhibiting herbicides, when used in combination with a compound B described SUPRA, in combination with safeners. Safeners are chemical compounds which prevent or reduce damage on useful plants without having a major impact on the herbicidal action of herbicides towards unwanted plants. They can be applied either before sowings (e.g. on seed treatments, shoots or seedlings) or in the pre-emergence application or post-emergence application of the useful plant.

Furthermore, the safeners C, the TK-inhibiting herbicides and/or the herbicides B can be applied simultaneously or in succession.

Suitable safeners are e.g. (quinolin-8-oxy)acetic acids, 1-phenyl-5-haloalkyl-1H-1,2,4-triazol-3-carboxylic acids, 1-phenyl-4,5-dihydro-5-alkyl-1H-pyrazol-3,5-dicarboxylic acids, 4,5-dihydro-5,5-diaryl-3-isoxazol carboxylic acids, dichloroacetamides, alpha-oximinophenylacetonitriles, acetophenonoximes, 4,6-dihalo-2-phenylpyrimidines, N-[[4-(aminocarbonyl)phenyl]sulfonyl]-2-benzoic amides, 1,8-naphthalic anhydride, 2-halo-4-(haloalkyl)-5-thiazol carboxylic acids, phosphorthiolates and N-alkyl-O-phenylcarbamates and their agriculturally acceptable salts and their agriculturally acceptable derivatives such amides, esters, and thioesters, provided they have an acid group.

Examples of preferred safeners C are benoxacor, cloquintocet, cyometrinil, cyprosulfamide, dichlormid, dicyclonon, dietholate, fenchlorazole, fenclorim, flurazole, fluxofenim, furilazole, isoxadifen, mefenpyr, mephenate, naphthalic anhydride, oxabetrinil, 4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3) and 2,2,5-trimethyl-3-(dichloroacetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4).

Especially preferred safeners C are benoxacor, cloquintocet, cyprosulfamide, dichlormid, fenchlorazole, fenclorim, flurazole, fluxofenim, furilazole, isoxadifen, mefenpyr, naphthalic anhydride, oxabetrinil, 4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3) and 2,2,5-trimethyl-3-(dichloroacetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4).

Particularly preferred safeners C are benoxacor, cloquintocet, cyprosulfamide, dichlormid, fenchlorazole, fenclorim, furilazole, isoxadifen, mefenpyr, naphtalic anhydride, 4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3), and 2,2,5-trimethyl-3-(dichloroacetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4).

Also preferred safeners C are benoxacor, cloquintocet, cyprosulfamide, dichlormid, fenchlorazole, fenclorim, furilazole, isoxadifen, mefenpyr, 4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3) and 2,2,5-trimethyl-3-(dichloroacetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4)..

Particularly preferred safeners C, which, as component C, are constituent of the composition according to the invention are the safeners C as defined above; in particular the safeners C.1 - C.12 listed below in table C:

Table C

	Safener C
C.1	benoxacor
C.2	cloquintocet
C.3	cyprosulfamide
C.4	dichlormid
C.5	fenchlorazole
C.6	fenclorim
C.7	furilazole
C.8	isoxadifen
C.9	mefenpyr
C.10	naphtalic acid anhydride
C.11	4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3)

C.12	2,2,5-trimethyl-3-(dichloro-acetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4)

The TK-inhibiting herbicides (compounds A) and the active compounds B of groups b1) to b15) and the active compounds C are known herbicides and safeners, see, for example, The Compendium of Pesticide Common Names (http://www.alanwood.net/pesticides/); Farm Chemicals Handbook 2000 volume 86, Meister Publishing Company, 2000; B. Hock, C. Fedtke, R. R. Schmidt, Herbizide [Herbicides], Georg Thieme Verlag, Stuttgart 1995; W. H. Ahrens, Herbicide Handbook, 7th edition, Weed Science Society of America, 1994; and K. K. Hatzios, Herbicide Handbook, Supplement for the 7th edition, Weed Science Society of America, 1998. 2,2,5-Trimethyl-3-(dichloroacetyl)-1,3-oxazolidine [CAS No. 52836-31-4] is also referred to as R-29148. 4-(Dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane [CAS No. 71526-07-3] is also referred to as AD-67 and MON 4660.

The assignment of the active compounds to the respective mechanisms of action is based on current knowledge. If several mechanisms of action apply to one active compound, this substance was only assigned to one mechanism of action.

Active compounds B and C having a carboxyl group can be employed in the form of the acid, in the form of an agriculturally suitable salt as mentioned above or else in the form of an agriculturally acceptable derivative in the compositions according to the invention.

In the case of dicamba, suitable salts include those, where the counterion is an agriculturally acceptable cation. For example, suitable salts of dicamba are dicamba-sodium, dicamba-potassium, dicamba-methylammonium, dicamba-dimethylammonium, dicamba-isopropylammonium, dicamba-diglycolamine, dicamba-olamine, dicamba-diolamine, dicamba-trolamine, dicamba-N,N-bis-(3-aminopropyl)methylamine and dicamba-diethylenetriamine. Examples of a suitable ester are dicamba-methyl and dicamba-butotyl.
Suitable salts of 2,4-D are 2,4-D-ammonium, 2,4-D-dimethylammonium, 2,4-D-diethylammonium, 2,4-D-diethanolammonium (2,4-D-diolamine), 2,4-D-triethanolammonium, 2,4-D-isopropylammonium, 2,4-D-triisopropanolammonium, 2,4-D-heptylammonium, 2,4-D-dodecylammonium, 2,4-D-tetradecylammonium, 2,4-D-triethylammonium, 2,4-D-tris(2-hydroxypropyl)ammonium, 2,4-D-tris(isopropyl)ammonium, 2,4-D-trolamine, 2,4-D-lithium, 2,4-D-sodium. Examples of suitable esters of 2,4-D are 2,4-D-butotyl, 2,4-D-2-butoxypropyl, 2,4-D-3-butoxypropyl, 2,4-D-butyl, 2,4-D-ethyl, 2,4-D-ethylhexyl, 2,4-D-isobutyl, 2,4-D-isooctyl, 2,4-D-isopropyl, 2,4-D-meptyl, 2,4-D-methyl, 2,4-D-octyl, 2,4-D-pentyl, 2,4-D-propyl, 2,4-D-tefuryl and clacyfos.
Suitable salts of 2,4-DB are for example 2,4-DB-sodium, 2,4-DB-potassium and 2,4-DB-dimethylammonium. Suitable esters of 2,4-DB are for example 2,4-DB-butyl and 2,4-DB-isoctyl.
Suitable salts of dichlorprop are for example dichlorprop-sodium, dichlorprop-potassium and dichlorprop-dimethylammonium. Examples of suitable esters of dichlorprop are dichlorprop-butotyl and dichlorprop-isoctyl.

Suitable salts and esters of MCPA include MCPA-butotyl, MCPA-butyl, MCPA-dimethylammonium, MCPA-diolamine, MCPA-ethyl, MCPA-thioethyl, MCPA-2-ethylhexyl, MCPA-isobutyl, MCPA-isoctyl, MCPA-isopropyl, MCPA-isopropylammonium, MCPA-methyl, MCPA-olamine, MCPA-potassium, MCPA-sodium and MCPA-trolamine.

A suitable salt of MCPB is MCPB sodium. A suitable ester of MCPB is MCPB-ethyl.

Suitable salts of clopyralid are clopyralid-potassium, clopyralid-olamine and clopyralid-tris-(2-hydroxypropyl)ammonium. Example of suitable esters of clopyralid is clopyralid-methyl.

Examples of a suitable ester of fluroxypyr are fluroxypyr-meptyl and fluroxypyr-2-butoxy-1-methylethyl, wherein fluroxypyr-meptyl is preferred.

Suitable salts of picloram are picloram-dimethylammonium, picloram-potassium, picloram-triisopropanolammonium, picloram-triisopropylammonium and picloram-trolamine. A suitable ester of picloram is picloram-isoctyl.

A suitable salt of triclopyr is triclopyr-triethylammonium. Suitable esters of triclopyr are for example triclopyr-ethyl and triclopyr-butotyl.

Suitable salts and esters of chloramben include chloramben-ammonium, chloramben-diolamine, chloramben-methyl, chloramben-methylammonium and chloramben-sodium.

Suitable salts and esters of 2,3,6-TBA include 2,3,6-TBA-dimethylammonium, 2,3,6-TBA-lithium, 2,3,6-TBA-potassium and 2,3,6-TBA-sodium.

Suitable salts and esters of aminopyralid include aminopyralid-potassium and aminopyralid-tris(2-hydroxypropyl)ammonium.

Suitable salts of glyphosate are for example glyphosate-ammonium, glyphosate-diammonium, glyphoste-dimethylammonium, glyphosate-isopropylammonium, glyphosate-potassium, glyphosate-sodium, glyphosate-trimesium as well as the ethanolamine and diethanolamine salts, preferably glyphosate-diammonium, glyphosate-isopropylammonium and glyphosate-trimesium (sulfosate).

A suitable salt of glufosinate is for example glufosinate-ammonium.

A suitable salt of glufosinate-P is for example glufosinate-P-ammonium.

Suitable salts and esters of bromoxynil are for example bromoxynil-butyrate, bromoxynil-heptanoate, bromoxynil-octanoate, bromoxynil-potassium and bromoxynil-sodium.

Suitable salts and esters of ioxonil are for example ioxonil-octanoate, ioxonil-potassium and ioxonil-sodium.

Suitable salts and esters of mecoprop include mecoprop-butotyl, mecoprop-dimethylammonium, mecoprop-diolamine, mecoprop-ethadyl, mecoprop-2-ethylhexyl, mecoprop-isoctyl, mecoprop-methyl, mecoprop-potassium, mecoprop-sodium and mecoprop-trolamine.

Suitable salts of mecoprop-P are for example mecoprop-P-butotyl, mecoprop-P-dimethylammonium, mecoprop-P-2-ethylhexyl, mecoprop-P-isobutyl, mecoprop-P-potassium and mecoprop-P-sodium.

A suitable salt of diflufenzopyr is for example diflufenzopyr-sodium.

A suitable salt of naptalam is for example naptalam-sodium.

Suitable salts and esters of aminocyclopyrachlor are for example aminocyclopyrachlor-dimethylammonium, aminocyclopyrachlor-methyl, aminocyclopyrachlor-

triisopropanolammonium, aminocyclopyrachlor-sodium and aminocyclopyrachlor-potassium.
A suitable salt of quinclorac is for example quinclorac-dimethylammonium.
A suitable salt of quinmerac is for example quinclorac-dimethylammonium.
A suitable salt of imazamox is for example imazamox-ammonium.
Suitable salts of imazapic are for example imazapic-ammonium and imazapic-isopropylammonium.
Suitable salts of imazapyr are for example imazapyr-ammonium and imazapyr-isopropylammonium.
A suitable salt of imazaquin is for example imazaquin-ammonium.
Suitable salts of imazethapyr are for example imazethapyr-ammonium and imazethapyr-isopropylammonium.
A suitable salt of topramezone is for example topramezone-sodium.

The preferred embodiments of the invention mentioned herein below have to be understood as being preferred either independently from each other or in combination with one another.

According to a preferred embodiment of the invention, the composition comprises as component B at least one, preferably exactly one herbicide B.

According to another preferred embodiment of the invention, the composition comprises at least two, preferably exactly two, herbicides B different from each other.

According to another preferred embodiment of the invention, the composition comprises at least three, preferably exactly three, herbicides B different from each other.

Here and below, the term “binary compositions” includes compositions comprising one or more, for example 1, 2 or 3, herbicides A and either one or more, for example 1, 2 or 3, herbicides B. Correspondingly, the term “ternary compositions” includes compositions comprising one or more, for example 1, 2 or 3, herbicides A, for example 1, 2 or 3, herbicides B and at least one herbicide D.

Further preferred embodiments relate to compositions which correspond to the binary and ternary compositions mentioned above and additionally comprise a safener C, in particular selected from the group consisting of benoxacor (C.1), cloquintocet (C.2), cyprosulfamide (C.3), dichlormid (C.4), fenchlorazole (C.5), fenclorim (C.6), furilazole (C.7), isoxadifen (C.8), mefenpyr (C.9), 4-(dichloroacetyl)-1-oxa-4-azaspiro[4.5]decane (MON4660, CAS 71526-07-3) (C.10) and 2,2,5-trimethyl-3-(dichloroacetyl)-1,3-oxazolidine (R-29148, CAS 52836-31-4) (C.11).

In binary compositions comprising at least one herbicide A and at least one herbicide B, the weight ratio of A:B is generally in the range of from 1:1000 to 1000:1, preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1.

In compositions comprising both at least one herbicide A, at least one herbicide B and at least one safener C, the weight ratio A:B is generally in the range of from 1:1000 to 1000:1, preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1; the weight ratio A:C is generally in the range of from 1:1000 to 1000:1, preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1; and the weight ratio B:C is generally in the range of from 1:1000 to 1000:1, preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1. The weight ratio (A + B):C is preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1.

In ternary compositions comprising at least one herbicide A, at least one herbicide B, and a herbicide D, the weight ratios A:B, A:D and B:D in each case are generally in the range of from 1:1000 to 1000:1, preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1. The weight ratio (A + B):D is preferably in the range of from 1:500 to 500:1, in particular in the range of from 1:250 to 250:1 and particularly preferably in the range of from 1:75 to 75:1.

It is generally preferred to use the compounds of the invention in combination with herbicides that are selective for the crop being treated and which complement the spectrum of weeds controlled by these compounds at the application rate employed. It is further generally preferred to apply the compounds of the invention and other complementary herbicides at the same time, either as a combination formulation or as a tank mix.

In another embodiment, the present invention refers to a method for identifying a TK-inhibiting herbicide by using a mutated TK encoded by a nucleic acid which comprises the nucleotide sequence of SEQ ID NO: 182 or 183, or a variant or derivative thereof.

Said method comprises the steps of:
a) generating a transgenic cell or plant comprising a nucleic acid encoding a mutated TK, wherein the mutated TK is expressed;
b) applying a TK-inhibiting herbicide to the transgenic cell or plant of a) and to a control cell or plant of the same variety;
c) determining the growth or the viability of the transgenic cell or plant and the control cell or plant after application of said TK-inhibiting herbicide, and
d) selecting “TK-inhibiting herbicides” which confer reduced growth to the control cell or plant as compared to the growth of the transgenic cell or plant.

As described above, the present invention teaches compositions and methods for increasing the TK-inhibiting tolerance of a crop plant or seed as compared to a wild-type variety of the plant or seed. In a preferred embodiment, the TK-inhibiting tolerance of a crop plant or seed is increased such that the plant or seed can withstand a TK-inhibiting herbicide application of preferably approximately 1-1000 g ai ha^{-1}, more preferably 1-200 g ai ha^{-1}, even more preferably 5-150 g ai ha^{-1}, and most preferably 10-100 g ai ha^{-1}. As used herein, to “withstand” a TK-inhibiting herbicide application means that the plant is either not killed or only moderately injured by such application. It will be understood by the person skilled in the art that the application rates may vary, depending on the environmental conditions such as temperature or humidity, and depending on the chosen kind of herbicide (active ingredient ai).

Post-emergent weed control methods useful in various embodiments hereof utilize about >0.3x application rates of TK-inhibiting herbicides; in some embodiments, this can be about, for example, >0.3x, >0.4x, >0.5x, >0.6x, >0.7x, >0.8x, >0.9x, or >lx of TK-inhibiting herbicides. In one embodiment, TK-inhibiting herbicides-tolerant plants of the present invention have tolerance to a post-emergant application of a TK-inhibiting herbicides at an amount of about 25 to about 200 g ai/ha. In some embodiments, wherein the TK-inhibiting herbicides-tolerant plant is a dicot (e.g. , soy, cotton), the post-emergant application of the TK-inhibiting herbicides is at an amount of about 50 g ai/ha. In another embodiment, wherein the TK-inhibiting herbicides-tolerant plant is a monocot (e.g., maize, rice, sorghum), the post-emergant application of the TK-inhibiting herbicides is at an amount of about 200 g ai/ha. In other embodiments, wherein the TK-inhibiting herbicides-tolerant plant is a Brassica (e.g., canola), the post-emergant application of the TK-inhibiting herbicides is at an amount of about 25 g ai/ha. In post-emergent weed control methods hereof, in some embodiments, the method can utilize TK-inhibiting herbicides application rates at about 7 to 10 days post-emergent. In another embodiment, the application rate can exceed lx TK-inhibiting herbicides; in some embodiments, the rate can be up to 4x TK-inhibiting herbicides, though more typically it will be about 2.5x or less, or about 2x or less, or about 1x or less.

Furthermore, the present invention provides methods that involve the use of at least one TK-inhibiting herbicide, optionally in combination with one or more herbicidal compounds B, and, optionally, a safener C, as described in detail supra.

In these methods, the TK-inhibiting herbicide can be applied by any method known in the art including, but not limited to, seed treatment, soil treatment, and foliar treatment. Prior to application, the TK-inhibiting herbicide can be converted into the customary formulations, for example solutions, emulsions, suspensions, dusts, powders, pastes and granules. The use form depends on the particular intended purpose; in each case, it should ensure a fine and even distribution of the compound according to the invention.

By providing plants having increased tolerance to TK-inhibiting herbicide, a wide variety of formulations can be employed for protecting plants from weeds, so as to enhance plant growth and reduce competition for nutrients. A TK-inhibiting herbicide can be used by itself for pre-emergence, post-emergence, pre-planting, and at-planting control of weeds in areas surrounding the crop plants described herein, or a TK-inhibiting herbicide formulation can be used that contains other additives. The TK-inhibiting herbicide can also be used as a seed treatment. Additives found in a TK-inhibiting herbicide formulation include other herbicides, detergents, adjuvants, spreading agents, sticking agents, stabilizing agents, or the like. The TK-inhibiting herbicide formulation can be a wet or dry preparation and can include, but is not limited to, flowable powders, emulsifiable concentrates, and liquid concentrates. The TK-inhibiting herbicide and herbicide formulations can be applied in accordance with conventional methods, for example, by spraying, irrigation, dusting, or the like.

Suitable formulations are described in detail in PCT/EP2009/063387 and PCT/EP2009/063386, which are incorporated herein by reference.

As disclosed herein, the TK nucleic acids of the invention find use in enhancing the herbicide tolerance of plants that comprise in their genomes a gene encoding a herbicide-tolerant wild-type or mutated TK protein. Such a gene may be an endogenous gene or a transgene, as described above. Additionally, in certain embodiments, the nucleic acids of the present invention can be stacked with any combination of polynucleotide sequences of interest in order to create plants with a desired phenotype. For example, the nucleic acids of the present invention may be stacked with any other polynucleotides encoding polypeptides having pesticidal and/or insecticidal activity, such as, for example, the Bacillus thuringiensis toxin proteins (described in U.S. Patent Nos. 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; and Geiser et al (1986) Gene 48: 109), 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), Glyphosate acetyl transferase (GAT), cytochrome P450 monooxygenase, phosphinothricin acetyltransferase (PAT), Acetohydroxyacid synthase (AHAS; EC 4.1.3.18, also known as acetolactate synthase or ALS), hydroxyphenyl pyruvate dioxygenase (HPPD), Phytoene desaturase (PD), Protoporphyrinogen oxidase (PPO) and dicamba degrading enzymes as disclosed in WO 02/068607, or phenoxyaceticacid- and phenoxypropionicacid-derivative degrading enzymes as disclosed in WO 2008141154 or WO 2005107437. The combinations generated can also include multiple copies of any one of the polynucleotides of interest.

Consequently, Herbicide-tolerant plants of the invention can be used in conjunction with an herbicide to which they are tolerant. Herbicides can be applied to the plants of the invention using any techniques known to those skilled in the art. Herbicides can be applied at any point in the plant cultivation process. For example, herbicides can be applied pre-planting, at planting, pre-emergence, post-emergence or combinations thereof. Herbicides may be applied to seeds and dried to form a layer on the seeds.

In some embodiments, seeds are treated with a safener, followed by a post- emergent application of a TK-inhibiting herbicides. In one embodiment, the post-emergent application of the TK-inhibiting herbicides is about 7 to 10 days following planting of safener-treated seeds. In some embodiments, the safener is cloquintocet, dichlormid, fluxofenim, or combinations thereof.

Methods of controlling weeds or undesired vegetation

In other aspects, the present invention provides a method for controlling weeds at a locus for growth of a plant or plant part thereof, the method comprising: applying a composition comprising a TK-inhibiting herbicides to the locus.

In some aspects, the present invention provides a method for controlling weeds at a locus for growth of a plant, the method comprising: applying an herbicide composition comprising TK-inhibiting herbicides to the locus; wherein said locus is: (a) a locus that contains: a plant or a seed capable of producing said plant; or (b) a locus that is to be after said applying is made to contain the plant or the seed; wherein the plant or the seed comprises in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

Herbicide compositions hereof can be applied, e.g., as foliar treatments, soil treatments, seed treatments, or soil drenches. Application can be made, e.g., by spraying, dusting, broadcasting, or any other mode known useful in the art.

In one embodiment, herbicides can be used to control the growth of weeds that may be found growing in the vicinity of the herbicide-tolerant plants invention. In embodiments of this type, an herbicide can be applied to a plot in which herbicide-tolerant plants of the invention are growing in vicinity to weeds. An herbicide to which the herbicide-tolerant plant of the invention is tolerant can then be applied to the plot at a concentration sufficient to kill or inhibit the growth of the weed. Concentrations of herbicide sufficient to kill or inhibit the growth of weeds are known in the art and are disclosed above.

In other embodiments, the present invention provides a method for controlling weeds in the vicinity of a TK-inhibiting herbicides-tolerant plant of the invention. The method comprises applying an effective amount of a TK-inhibiting herbicides to the weeds and to the auxinic herbicide- tolerant plant, wherein the plant has increased tolerance to auxinic herbicide when compared to a wild-type plant. In some embodiments, the TK-inhibiting herbicides-tolerant plants of the invention are preferably crop plants, including, but not limited to, sunflower, alfalfa, Brassica sp., soybean, cotton, safflower, peanut, tobacco, tomato, potato, wheat, rice, maize, sorghum, barley, rye, millet, and sorghum.

In other aspects, herbicide(s) (e.g., TK-inhibiting herbicides) can also be used as a seed treatment. In some embodiments, an effective concentration or an effective amount of herbicide(s), or a composition comprising an effective concentration or an effective amount of herbicide(s) can be applied directly to the seeds prior to or during the sowing of the seeds. Seed Treatment formulations may additionally comprise binders and optionally colorants.

Binders can be added to improve the adhesion of the active materials on the seeds after treatment. In one embodiments, suitable binders are block copolymers EO/PO surfactants but also polyvinylalcoholsl, polyvinylpyrrolidones, polyacrylates, polymethacrylates, polybutenes, polyisobutylenes, polystyrene, polyethyleneamines, polyethyleneamides, polyethyleneimines (Lupasol(R), Polymin(R)), polyethers, polyurethans, polyvinylacetate, tylose and copolymers derived from these polymers. Optionally, also colorants can be included in the formulation. Suitable colorants or dyes for seed treatment formulations are Rhodamin B, C.I. Pigment Red 112, C.I. Solvent Red 1 , pigment blue 15:4, pigment blue 15:3, pigment blue 15:2, pigment blue 15: 1, pigment blue 80, pigment yellow 1 , pigment yellow 13, pigment red 1 12, pigment red 48:2, pigment red 48: 1, pigment red 57: 1 , pigment red 53:1 , pigment orange 43, pigment orange 34, pigment orange 5, pigment green 36, pigment green 7, pigment white 6, pigment brown 25, basic violet 10, basic violet 49, acid red 51 , acid red 52, acid red 14, acid blue 9, acid yellow 23, basic red 10, basic red 108.

The term seed treatment comprises all suitable seed treatment techniques known in the art, such as seed dressing, seed coating, seed dusting, seed soaking, and seed pelleting. In one embodiment, the present invention provides a method of treating soil by the application, in particular into the seed drill: either of a granular formulation containing the TK-inhibiting herbicides as a composition/formulation (e.g., a granular formulation), with optionally one or more solid or liquid, agriculturally acceptable carriers and/or optionally with one or more agriculturally acceptable surfactants. This method is advantageously employed, for example, in seedbeds of cereals, maize, cotton, and sunflower.

The present invention also comprises seeds coated with or containing with a seed treatment formulation comprising TK-inhibiting herbicides and at least one other herbicide such as, e.g. , an AHAS-inhibitor selected from the group consisting of amidosulfuron, azimsulfuron, bensulfuron, chlorimuron, chlorsulfuron, cinosulfuron, cyclosulfamuron, ethametsulfuron, ethoxysulfuron, flazasulfuron, flupyrsulfuron, foramsulfuron, halosulfuron, imazosulfuron, iodosulfuron, mesosulfuron, metsulfuron, nicosulfuron, oxasulfuron, primisulfuron, prosulfuron, pyrazosulfuron, rimsulfuron, sulfometuron, sulfosulfuron, thifensulfuron, triasulfuron, tribenuron, trifloxysulfuron, triflusulfuron, tritosulfuron, imazamethabenz, imazamox, imazapic, imazapyr, imazaquin, imazethapyr, cloransulam, diclosulam, florasulam, flumetsulam, metosulam, penoxsulam, bispyribac, pyriminobac, propoxycarbazone, flucarbazone, pyribenzoxim, pyriftalid and pyrithiobac.

The term "coated with and/or containing" generally signifies that the active ingredient is for the most part on the surface of the propagation product at the time of application, although a greater or lesser part of the ingredient may penetrate into the propagation product, depending on the method of application. When the said propagation product is (re)planted, it may absorb the active ingredient.

In some embodiments, the seed treatment application with TK-inhibiting herbicides or with a formulation comprising the TK-inhibiting herbicides is carried out by spraying or dusting the seeds before sowing of the plants and before emergence of the plants.

In other embodiments, in the treatment of seeds, the corresponding formulations are applied by treating the seeds with an effective amount of TK-inhibiting herbicides or a formulation comprising the TK-inhibiting herbicides.

In other aspects, the present invention provides a method for combating undesired vegetation or controlling weeds comprising contacting the seeds of the TK-inhibiting herbicides-tolerant plants of the present invention before sowing and/or after pregermination with TK-inhibiting herbicides. The method can further comprise sowing the seeds, for example, in soil in a field or in a potting medium in greenhouse. The method finds particular use in combating undesired vegetation or controlling weeds in the immediate vicinity of the seed. The control of undesired vegetation is understood as the killing of weeds and/or otherwise retarding or inhibiting the normal growth of the weeds. Weeds, in the broadest sense, are understood as meaning all those plants which grow in locations where they are undesired.

The weeds of the present invention include, for example, dicotyledonous and monocotyledonous weeds. Dicotyledonous weeds include, but are not limited to, weeds of the genera: Sinapis, Lepiclium, Galium, Stellaria, Matricaria, Anthemis, Galinsoga, Chenopodium, Urtica, Senecio, Amaranthus, Portulaca, Xanthium, Convolvulus, Ipomoea, Polygonum, Sesbania, Ambrosia, Cirsium, Carduus, Sonchus, Solarium, Rorippa, Rotala, Lindernia, Lamium, Veronica, Abutilon, Emex, Datura, Viola, Galeopsis, Papaver, Centaurea, Trifolium, Ranunculus, and Taraxacum. Monocotyledonous weeds include, but are not limited to, weeds of the genera: Echinochloa, Setaria, Panicum, Digitaria, Phleum, Poa, Festuca, Eleusine, Brachiaria, Lolium, Bromus, Avena, Cyperus, Sorghum, Agropyron, Cynodon, Monochoria, Fimbristyslis, Sagittaria, Eleocharis, Scirpus, Paspalum, Ischaemum, Sphenoclea, Dactyloctenium, Agrostis, Alopecurus, and Apera.

In addition, the weeds of the present invention can include, for example, crop plants that are growing in an undesired location. For example, a volunteer maize plant that is in a field that predominantly comprises soybean plants can be considered a weed, if the maize plant is undesired in the field of soybean plants.

In other embodiments, in the treatment of seeds, the corresponding formulations are

applied by treating the seeds with an effective amount of TK-inhibiting herbicides or a formulation comprising the TK-inhibiting herbicides.
In still further aspects, treatment of loci, plants, plant parts, or seeds of the present invention comprises application of an agronomically acceptable composition that does not contain an A.I. In one embodiment, the treatment comprises application of an agronomically acceptable composition that does not contain a TK-inhibiting herbicides A.I. In some embodiments, the treatment comprises application of an agronomically acceptable composition that does not contain a TK-inhibiting herbicides A.L, wherein the composition comprises one or more of agronomically-acceptable carriers, diluents, excipients, plant growth regulators, and the like. In other embodiments, the treatment comprises application of an agronomically acceptable composition that does not contain a TK-inhibiting herbicides A.I., wherein the composition comprises an adjuvant. In one embodiment, the adjuvant is a surfactant, a spreader, a sticker, a penetrant, a drift-control agent, a crop oil, an emulsifier, a compatibility agent, or combinations thereof.

It should also be understood that the foregoing relates to preferred embodiments of the present invention and that numerous changes may be made therein without departing from the scope of the invention. The invention is further illustrated by the following examples, which are not to be construed in any way as imposing limitations upon the scope thereof. On the contrary, it is to be clearly understood that resort may be had to various other embodiments, modifications, and equivalents thereof, which, after reading the description herein, may suggest themselves to those skilled in the art without departing from the spirit of the present invention and/or the scope of the appended claims.

EXAMPLES

EXAMPLE 1: Sequencing and full length assembly of transketolase genes

Isolation of RNA and cDNA synthesis

Plant leaf tissue is harvested, frozen and grounded in liquid nitrogen and total RNA is extracted using an Ambion RNAqueous-Midi kit (AM1911, Ambion) with the Plant RNA Isolation Aid (AM9690, Ambion) as per manufacturer’s recommendation. The last elution is done with 10 ul of elution solution. To validate the quality of the extracted RNA 1 uL of the final product is run on a Bioanalyzer 2100 using the RNA 6000 Nano kit with the Plant RNA Nano method. The final solution, containing purified RNA, is stored at -80° C until library preparation.
For assembly of plant genes, an RNA sequencing experiment is performed. RNA sequencing libraries are produced using TruSeq RNA Sample preparation kits V2 (RS-122-2001) from Illumina according to the instructions of the manufacturer. Briefly, 1 µg of total RNA is first purified twice on a poly-dT column. During the second elution step, RNA is fragmented and primed for cDNA synthesis. The material is reverse transcribed, RNA is removed and the second strand is produced. After rendering the ends of the fragment blunt, 3’ ends are adenylated and Illumina sequencing-specific bar-coded adaptors are ligated at both ends of the fragments. The DNA fragments bearing adaptors at both ends are enriched by a 15 cycle PCR

amplification. Libraries are pooled prior to sequencing. The pooled libraries are first put on a flowcell using a TruSeq PE Cluster kit V3 (PE-401-3001) on the cBot and clusters are amplified on the device. Afterwards, the flowcell is transferred onto the Illumina Hiseq machine and the material on the flowcell is then sequenced using Illumina TruSeq SBS Kit V3 (FC-401-3001) as per manufacturer's recommendation.

The data produced by the Illumina Hiseq sequencer is first trimmed at both ends using a quality threshold of 15 using the FASTQC Quality Trimmer (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). These sequences are further analyzed to remove any Illumina adaptor sequences using CutAdapt (http://code.google.com/p/cutadapt/). Sequence reads are assembled using CLC bio algorithm (version 4.01). The transketolase full length gene sequences (SEQID NO 182 and 183) are identified by performing a BLAST search with *Arabidopsis thaliana* transketolase sequences as query and the CLC bio assembly as database.

EXAMPLE 2: Mutagenesis of the Gene Encoding for Transketolase

All nucleic acid coding sequence and all single and double mutants based on SEQ ID NO 183 were synthesized and cloned by Geneart (Geneart AG, Regensburg, Germany). Rational design mutants were synthesized by Geneart. Random transketolase gene libraries were synthesized by Geneart. Plasmids were isolated from *E. coli* TOP10 by performing a plasmid minpreparation and confirmed by DNA sequencing.

EXAMPLE 3: Expression and purification of recombinant wild-type and mutant Transketolase

Clones in pET27b vector were transformed into BL21(DE3)-pLysS strain of *E. coli*. Cells were grown in 250 - 1000 mL of LB with 100 μgmL−1 of ampicillin or kanamycin, shaking overnight at 37 °C. The cells were harvested by centrifugation at 1600×g, washed with 0.09% NaCl, and stored at −80 °C. Cells were lysed using a French press at 140 MPa in 50 mM sodium phosphate pH 7.5, 1 M NaCl, 5 mM imidazole, 5% glycerol, and 1 μg mL−1 leupeptin. Following lysis, 0.5 U of Benzonase (Novagen, EMD Chemicals, Inc., Gibbstown, NJ) and PMSF (final concentration of 1 mM) were added. Cell debris was removed by centrifugation at 3000×g. His-tagged transketolase proteins were purified on a nickel activated Hitrap Chelating HP column (GE Healthcare Bio-Sciences Corp., Piscataway, NJ) equilibrated with 20 mM sodium phosphate pH 8.0, 50 mM NaCl, 5 mM imidazole, 5 mM MgCl2, 0.1mM EDTA, and 17% glycerol. Target ptotein is eluted with 250 mM imidazole. The active protein was desalted on a PD-10 column (GE Healthcare Bio-Sciences Corp., Piscataway, NJ) equilibrated with a 20 mM sodium phosphate buffer, pH 7.5, 5 mM MgCl2, 1 mM EDTA and 17% glycerol.

EXAMPLE 4: Wild-type and mutant Transketolase Enzyme and Inhibitor Assay

Assays were measured for 1 hour at 340 nm using a standard plate reader. Before starting the reactions with Xylulose-5-Phosphat, the reaction mixture was incubated for 15 minutes at room temperature. Both Cornexistin and Hydroxycornexistin were used as inhibitors of transketolase activity. Dose response curves were measured by titrating either of the inhibitors. On the basis of dose response curves, IC50 values were calculated and use to evaluate in vitro tolerance to either inhibitor. (IC50 is defined as the concentration of inhibitor needed to induce 50% loss in enzyme activity as compared to a similar reaction without inhibitor.)

The results are shown in Table 4

Table 4

Substitution Position 1	Substitution Position 2	Substitution Position 3	Relative Activity	IC50 (M) Cornexistin	Tolerance Factor (wild type = 1,0)
-	-	-	4,00	4,46E-06	1,0
M463A	L544C	-	0,41	6,58E-04	147,5
V514N	L544V	-	4,41	>1,00E-4	>22,4
T511S	V514N	L544V	5,10	>1,00E-4	>22,4
M463C	L544V	-	0,71	>1,00E-4	>22,4
M463G	L544S	-	0,47	>1,00E-4	>22,4
M463G	L544T	-	0,74	>1,00E-4	>22,4
M463G	L544V	-	0,49	>1,00E-4	>22,4
M463G	L544C	-	0,50	>1,00E-4	>22,4
M463G	V514N	L544T	0,49	>1,00E-4	>22,4
M463G	V514N	L544C	0,38	>1,00E-4	>22,4
M463G	V514N	L544S	0,32	>1,00E-4	>22,4
M463G	V514N	L544V	0,44	>1,00E-4	>22,4
M463G	T511S	L544T	1,18	>1,00E-4	>22,4
M463G	T511S	L544C	0,36	>1,00E-4	>22,4
M463G	T511S	L544S	0,49	>1,00E-4	>22,4
M463C	T511S	L544T	1,47	>1,00E-4	>22,4
M463C	T511S	L544C	0,52	>1,00E-4	>22,4
M463C	T511S	L544S	0,30	>1,00E-4	>22,4
M463C	V514N	L544T	2,37	>1,00E-4	>22,4
M463C	V514N	L544C	0,20	>1,00E-4	>22,4
M463C	V514N	L544S	0,36	>1,00E-4	>22,4
M463C	V514N	L544V	0,33	>1,00E-4	>22,4
Y337D	-	-	4,99	1,61E-04	36,2
M463C	L544T	-	2,89	1,57E-04	35,3
M463C	L544A	-	0,49	1,00E-04	22,4
M463A	L544A	-	0,29	1,00E-04	22,4
M463A	L544C	-	0,18	1,00E-04	22,4
M463A	L544S	-	0,10	1,00E-04	22,4
M463C	L544C	-	2,01	9,07E-05	20,4
Y337I	-	-	5,33	8,32E-05	18,7
M463L	-	-	4,54	8,03E-05	18,0
S460A	-	-	5,73	7,94E-05	17,8
L544C	-	-	2,94	7,78E-05	17,5
V514N	-	-	9,83	7,68E-05	17,2
M463K	-	-	6,01	7,29E-05	16,4

S461M	-	-	0,04	6,16E-05	13,8
Y337N	-	-	5,58	6,01E-05	13,5
Y337C	-	-	7,78	5,99E-05	13,4
M463G	-	-	9,75	5,94E-05	13,3
S460E	-	-	5,10	5,82E-05	13,1
F512C	-	-	0,60	5,76E-05	12,9
Y337L	-	-	5,11	5,60E-05	12,6
L544V	-	-	8,18	5,30E-05	11,9
M463N	-	-	5,09	5,18E-05	11,6
L458I	-	-	4,98	4,98E-05	11,2
L544V	-	-	5,84	4,98E-05	11,2
L544I	-	-	5,31	4,68E-05	10,5
S342D	A343P	-	0,79	4,61E-05	10,3
L544G	-	-	0,55	4,03E-05	9,0
S460L	-	-	5,58	3,90E-05	8,8
S342D	-	-	0,84	3,87E-05	8,7
L458V	-	-	6,41	3,65E-05	8,2
L544H	-	-	0,09	3,51E-05	7,9
M463A	L544S	-	0,11	3,46E-05	7,8
V514H	-	-	0,07	3,44E-05	7,7
M463A	L544T	-	0,39	3,42E-05	7,7
M463Y	-	-	0,91	3,18E-05	7,1
F513L	-	-	5,39	3,03E-05	6,8
S460G	-	-	7,30	2,98E-05	6,7
M463C	L544S	-	1,55	2,83E-05	6,3
S461A	-	-	0,47	2,82E-05	6,3
L544Q		-	0,57	2,78E-05	6,2
I267C	-	-	0,36	2,66E-05	6,0
S460Q	-	-	7,25	2,64E-05	5,9
A343V	-	-	1,28	2,59E-05	5,8
M463F	-	-	1,22	2,59E-05	5,8
Y337F	-	-	5,52	2,51E-05	5,6
S460T	-	-	5,40	2,48E-05	5,6
A343S	-	-	2,27	2,47E-05	5,5
L458M	-	-	5,25	2,45E-05	5,5
Y337E	-	-	7,53	2,40E-05	5,4
L458C	-	-	1,90	2,38E-05	5,3
V514C	-	-	10,36	2,28E-05	5,1
Y337T	-	-	5,36	2,26E-05	5,1
M463I	-	-	0,18	2,26E-05	5,1
T511S	-	-	6,42	2,24E-05	5,0
Y337A	-	-	5,51	2,09E-05	4,7

Y337H	-	-	8,34	1,98E-05	4,4
S460N	-	-	5,44	1,95E-05	4,4
V514A	-	-	8,70	1,86E-05	4,2
S342T	-	-	5,25	1,86E-05	4,2
S460F	-	-	5,54	1,80E-05	4,0
I267L	-	-	1,95	1,79E-05	4,0
A459S	-	-	7,00	1,78E-05	4,0
F513M	-	-	3,48	1,77E-05	4,0
L458S	-	-	0,04	1,74E-05	3,9
V514M	-	-	3,85	1,71E-05	3,8
A459C	-	-	1,80	1,64E-05	3,7
V514T	-	-	1,54	1,61E-05	3,6
A343T	-	-	0,81	1,58E-05	3,5
F513Y	-	-	4,64	1,55E-05	3,5
Y337V	-	-	0,87	1,55E-05	3,5
L458A	-	-	0,96	1,55E-05	3,5
F513I	-	-	1,22	1,53E-05	3,4
V514D	-	-	1,49	1,51E-05	3,4
F513V	-	-	2,43	1,47E-05	3,3
Y337Q	-	-	10,62	1,42E-05	3,2
S460M	-	-	5,14	1,39E-05	3,1
A459T	-	-	5,73	1,39E-05	3,1
S342A	-	-	5,90	1,38E-05	3,1
V514L	-	-	0,71	1,38E-05	3,1
Y337M	-	-	1,17	1,33E-05	3,0
S342N	A343P	-	3,85	1,32E-05	3,0
Y337S	-	-	1,09	1,15E-05	2,6
F515A	-	-	0,11	1,15E-05	2,6
Y337W	-	-	4,09	1,12E-05	2,5
F512A	-	-	0,07	1,11E-05	2,5
F515M	-	-	0,14	1,11E-05	2,5
I267T	-	-	0,93	1,07E-05	2,4
I267D	-	-	0,05	1,06E-05	2,4
M463H	-	-	0,26	1,02E-05	2,3
A459I	-	-	0,04	1,01E-05	2,3
S342T	A343P	-	4,97	1,00E-05	2,3
A343C	-	-	3,70	9,75E-06	2,2
V514Q	-	-	1,22	9,71E-06	2,2
S342Q	-	-	5,13	9,68E-06	2,2
V514I	-	-	2,02	9,20E-06	2,1
F513T	-	-	0,10	9,08E-06	2,0
V514F	-	-	0,13	8,65E-06	1,9

A343I	-	-	0,66	8,59E-06	1,9
F512M	-	-	5,40	8,48E-06	1,9
S460C	-	-	5,67	8,40E-06	1,9
A459G	-	-	0,42	7,63E-06	1,7
S342V	-	-	2,55	7,52E-06	1,7
A343M	-	-	0,86	7,14E-06	1,6
V514E	-	-	0,10	7,12E-06	1,6
F512H	-	-	0,58	6,86E-06	1,5
V514S	-	-	1,40	6,68E-06	1,5
A343P	-	-	5,80	6,67E-06	1,5
S342N	-	-	4,96	6,58E-06	1,5
S460V	-	-	4,80	6,51E-06	1,5
S342I	-	-	1,34	6,23E-06	1,4
I267A	-	-	0,29	6,10E-06	1,4
S342C	-	-	5,07	5,57E-06	1,2
F513C	-	-	0,96	5,13E-06	1,2
F513A	-	-	0,54	5,02E-06	1,1
S342F	-	-	0,78	4,96E-06	1,1
S342C	A343P	-	3,03	4,64E-06	1,0
S342H	A343P	-	2,27	4,46E-06	1,0
S342M	-	-	4,57	4,36E-06	1,0
S342K	A343P	-	2,14	4,29E-06	1,0
I267V	-	-	0,81	4,23E-06	0,9
F512Q	-	-	0,68	4,05E-06	0,9
L458T	-	-	0,98	3,88E-06	0,9
S460I	-	-	6,53	3,69E-06	0,8
F515Y	-	-	0,74	1,67E-06	0,4
S461N	-	-	0,00	-	-
T511V	-	-	0,00	-	-
M463S	-	-	0,00	-	-
T511A	-	-	0,00	-	-
T511F	-	-	0,00	-	-
F513Q	-	-	0,00	-	-
F513S	-	-	0,00	-	-
S461Q	-	-	0,00	-	-
F515Q	-	-	0,00	-	-
F515C	-	-	0,00	-	-
T511C	-	-	0,00	-	-
F515I	-	-	0,00	-	-
M463V	L544A	-	0,00	-	-
F512E	-	-	0,00	-	-
T511N	-	-	0,00	-	-

F515H	-	-	0,00	-	-
M463V	L544C	-	0,00	-	-
L544E	-	-	0,00	-	-
T511E	-	-	0,00	-	-
F513E	-	-	0,00	-	-
M463V	L544T	-	0,00	-	-
F515N	-	-	0,00	-	-
F515D	-	-	0,00	-	-
T511F	-	-	0,00	-	-
T511L	-	-	0,00	-	-
F515T	-	-	0,00	-	-
F515E	-	-	0,00	-	-
V514Y	-	-	0,00	-	-
F513D	-	-	0,00	-	-
F515S	-	-	0,00	-	-
F515L	-	-	0,00	-	-
T511H	-	-	0,00	-	-
F513H	-	-	0,00	-	-
T511D	-	-	0,00	-	-
T511I	-	-	0,00	-	-
T511Q	-	-	0,00	-	-
F515V	-	-	0,00	-	-
T511M	-	-	0,00	-	-
F513N	-	-	0,00	-	-
M463V	L544S	-	0,00	-	-
T511K	-	-	0,00	-	-
I267F	-	-	0,06	-	-
I267K	-	-	0,02	-	-
I267M	-	-	0,37	-	-
I267N	-	-	0,20	-	-
I267R	-	-	0,02	-	-
L458D	-	-	0,07	-	-
L458F	-	-	0,04	-	-
L458N	-	-	0,24	-	-
A343L	-	-	0,02	-	-
M463A	-	-	5,03	-	-
M463C	-	-	6,05	-	-
M463D	-	-	0,28	-	-
A459V	-	-	0,01	-	-
S461C	-	-	0,14	-	-
S461D	-	-	0,00	-	-
S461I	-	-	0,01	-	-

S461T	-	-	0,77	-	-
M463V	-	-	0,80	-	-
F512I	-	-	1,13	-	-
F512L	-	-	2,85	-	-
F512N	-	-	0,10	-	-
F512T	-	-	0,31	-	-
F512V	-	-	0,63	-	-
F512Y	-	-	0,67	-	-
L544A	-	-	0,88	-	-
L544D	-	-	0,16	-	-
L544F	-	-	0,15	-	-
L544M	-	-	0,80	-	-
L544N	-	-	0,81	-	-
L544S	-	-	0,85	-	-
L544T	-	-	2,07	-	-
I267W	-	-	0,08	-	-
I267S	-	-	0,37	-	-
S460D	-	-	2,30	-	-
S461V	-	-	0,11	-	-
F512S	-	-	0,03	-	-
M463S	-	-	1,95	-	-
M463T	-	-	0,00	-	-
F512D	-	-	0,05	-	-

Table 5

Assay components:		200µl Assay volumes	final concentrations
2 mM	Inhibitor in 100%DMSO	10	1,0E-4 M
Mix A			
50 mM	Tris/HCl pH 7,7	84,7	50 mM
0,10%	Cocarboxylase (TPP)	3,4	0,002%
10 mM	ß-NADH	10	0,5 mM
300 mM	$MgCl_2$	10	15 mM
10 units/ml (1:100)	a-GDH/TPI	3,4	0,17 units/ml
10 mM	D-Ribose 5-Phosphat	15	0,75 mM
	Transketolase	enzyme concentration dependent	
		140µl	
Reaction Start:			
	Xylulose-5-Phosphat	depends on substrate	
50 mM	Tris/HCl pH 7,7	ad. 60 µL	50 mM
		60 µl	
Indicator Reaction			
Assay Components:		200µl Volumes	Final Concentrations
Mix A			
2mM	Inhibitor	10	1,0E-4 M
50mM	Tris/HCl pH 7,7	Ad.140	50mM
0,10%	Cocarboxylase (TPP)	3,4	0,002%
10 mM	ß-NADH	10	0,5 mM
300mM	$MgCl_2$	10	15mM
1units/ml (1:1000)	a-GDH/TPI	10	0,05 units/mL
		140µl	
Reaction Start:			
2,94 mM	Glycerinaldehyd 3-Phosphat	51	0,75mM
50mM	Tris/HCl pH 7,7	9	
		60µl	

EXAMPLE 5: Demonstration of herbicide tolerance in a transient tobacco expression system

Transient expression of transketolase genes (e.g. SEQ ID NO: 1, 10, 104, 130 and 156) are done as described previously (Voinnet O., *et al.*, 2003, The Plant Journal 33, 949-956). In brief, for transient transformation of tobacco leaves or stable *Arabidopsis thaliana* transformation, Wildtype or mutated TKL sequences encoding TKL polypeptides comprising SEQ ID NO: 1, 10, 104, 130 and 156, are synthesized (Life Technologies, USA) and cloned with standard cloning techniques as described in Sambrook *et al.* (Molecular cloning (2001) Cold Spring Harbor Laboratory Press) in a binary vector containing resistance marker gene cassette (AHAS) and TKL sequence in between ubiquitin promoter (PcUbi) and nopaline synthase terminator (NOS) sequence. In addition, a plastid targeting peptide (amino acid 1- 59 from swiss prot gene id FENR_SPIOL) with a linker sequence CSSAAA was fused to SEQ ID NO: 104, 130 and 156. Young leaves of *Nicotiana benthamiana* are infiltrated with transgenic *Agrobacterium* suspension (strain: pGV 2260, OD^{600} of 1.0) harbouring binary vector constructs with a transketolase gene. Two to three days after infiltration punches of leave discs (0.75 cm in diameter) are transferred to 6-well plates with medium containing herbicide of interest in different concentrations. Multi well plates are incubated in a growth chamber at 22°C, 75% relative humidity and 110 µmol Phot * m^{-2} * s^{-1} with 14 : 10 h light : dark photoperiod. Herbicide effect are monitored by visual inspection and analysis of photosynthetic yield by imaging PAM (Walz, Effeltrich, Germany) 24h, 48h and 96h after treatment. Tolerance factors are calculated based on IC_{50} values of PSII yield inhibition of transformed *versus* empty vector-transformed leave discs. IC_{50} of PSII yield inhibition in empty vector-transformed leaf discs treated with Cornexistin for 24h or 48 h was measured with $4.7*10^{-6}$ M or $1.0*10^{-6}$ M, respectively.

Table 6: Tolerance values as observed in transient expression system, 24h after treatment.

Gene	***Seq ID***	*Cornexistin*
LEMPA_TKL wt	10	2.9
SPIOL_TKL wt	1	3.6
Pd_TKL1	104	5.9
Pd_TKL2	130	2.8
Pd_TKL3	156	4.8
SPIOL TKL S461T	1*	4.5
SPIOL TKL L544T	1*	2.1

* with respective mutation as stated in column 'gene'

EXAMPLE 6: Maize whole plant transformation and herbicide tolerance testing: Immature embryos were transformed according to the procedure outlined in Peng et al. (WO2006/136596). Plants were tested for the presence of the T-DNA by Taqman analysis with the target being the *nos* terminator which is present in all constructs. Healthy looking plants were sent to the greenhouse for hardening out and subsequent spray testing.

The plants were individually transplanted into MetroMix 360 soil in 4" pots. Once they had been in the greenhouse (day/night cycle of 27°C /21°C with 14 hour day length supported by 600W high pressure sodium lights), they were allowed to grow for 14 days. They were then sprayed with a herbicide treatment. Herbicide injury evaluations were taken 7 and 14 days after treatment to look for injury to new growth points and overall plant health. The top survivors were transplanted into gallon pots filled with MetroMix 360 for seed production.

EXAMPLE 7: Soybean transformation and herbicide tolerance testing:

Soybean cv Jake was transformed as previously described[1]. After regeneration, transformants were transplanted to soil in small pots, placed in growth chambers (16 hr day/ 8 hr night; 25°C day/ 23°C night; 65% relative humidity; 130-150 µE m^{-2} s^{-1}) and subsequently tested for the presence of the T-DNA via Taqman analysis. After a few weeks, healthy, transgenic positive, single copy events were transplanted to larger pots and allowed to grow in the growth chamber. An optimal shoot for cutting was about 3-4” tall, with at least two nodes present. Each cutting was taken from the original transformant (mother plant) and dipped into rooting hormone powder (indole-3-butyric acid, IBA). The cutting was then placed in oasis wedges inside a bio-dome. The mother plant was taken to maturity in the greenhouse and harvested for seed. Wild type cuttings were also taken simultaneously to serve as negative controls. The cuttings were kept in the bio-dome for 5-7 days and then transplanted to 3” pots and then acclimated in the growth chamber for two more days. Subsequently, the cuttings were transferred to the greenhouse, acclimated for approximately 4 days, and then subjected to herbicide spray tests. Herbicide injury evaluations were taken at 7 and 14 days after treatment.

References

1. Hong HP, Zhang H, Olhoft P, Hill S, Wiley H, Toren E, Hillebrand H, Jones T, Cheng M. Organogenic callus as the target for plant regeneration and transformation via *Agrobacterium* in soybean (*Glycine max* (L.) Merr.). In Vitro Cell. Dev. Biol.-Plant 2007: 43: 558-568

EXAMPLE 8. Engineering herbicide tolerant plants

Herbicide tolerant soybean (Glycine max) or corn (Zea mays) plants are generated as described by Olhoft et al. (US patent 2009/0049567). For transformation of soybean or Arabidopsis thaliana, transketolase genes are cloned with standard cloning techniques as described in Sambrook et al. (Molecular cloning (2001) Cold Spring Harbor Laboratory Press) in a binary vector containing resistance marker gene cassette (AHAS) and transketolase sequence (marked as GOI) in between ubiquitin promoter (PcUbi) and nopaline synthase terminator (NOS) sequence. For corn transformation, transketolase sequences are cloned with standard cloning techniques as described in Sambrook et al. (Molecular cloning (2001) Cold Spring Harbor Laboratory Press) in a binary vector containing resistance marker gene cassette (AHAS) and transketolase sequence (marked as GOI) in between corn ubiquitin promoter (ZmUbi) and nopaline synthase terminator (NOS) sequence. Binary plasmids are introduced to Agrobacterium tumefaciens for plant transformation. Plasmid constructs are introduced into soybean’s axillary meristem cells at the primary node of seedling explants via Agrobacterium-mediated transformation. After inoculation and co-cultivation with Agrobacteria, the explants are transferred to shoot introduction media without selection for one week. The explants were subsequently transferred to a shoot induction medium with 1-3 μM imazapyr (Arsenal) for 3 weeks to select for transformed cells. Explants with healthy callus/shoot pads at the primary node are then transferred to shoot elongation medium containing 1-3 μM imazapyr until a shoot elongated or the explant died. Transgenic plantlets are rooted, subjected to TaqMan analysis for the presence of the transgene, transferred to soil and grown to maturity in greenhouse. Transformation of corn plants are done by a method described by McElver and Singh (WO 2008/124495). Plant transformation vector constructs containing transketolase sequences are introduced into maize immature embryos via Agrobacterium-mediated transformation.

Transformed cells are selected in selection media supplemented with 0.5-1.5 μM imazethapyr for 3-4 weeks. Transgenic plantlets are regenerated on plant regeneration media and rooted afterwards. Transgenic plantlets are subjected to TaqMan analysis for the presence of the transgene before being transplanted to potting mixture and grown to maturity in greenhouse. Arabidopsis thaliana are transformed with transketolase sequences by floral dip method as described by McElver and Singh (WO 2008/124495). Transgenic Arabidopsis plants are subjected to TaqMan analysis for analysis of the number of integration loci. Transformation of Oryza sativa (rice) are done by protoplast transformation as described by Peng et al. (US 6653529)

EXAMPLE 9: Demonstration of herbicide tolerance

T0 or T1 transgenic plant of soybean, corn, and rice containing transketolase sequences are tested for improved tolerance to TK-inhibiting herbicides in greenhouse studies and mini-plot studies.For the pre-emergence treatment, the herbicides are applied directly after sowing by means of finely distributing nozzles. The containers are irrigated gently to promote germination and growth and subsequently covered with transparent plastic hoods until the plants have rooted. This cover causes uniform germination of the test plants, unless this has been impaired by the herbicides. For post emergence treatment, the test plants are first grown to a height of 3 to 15 cm, depending on the plant habit, and only then treated with the herbicides. For this purpose, the test plants are either sown directly and grown in the same containers, or they are first grown separately and transplanted into the test containers a few days prior to treatment.

For testing of T0 plants, cuttings can be used. In the case of soybean plants, an optimal shoot for cutting is about 7.5 to 10 cm tall, with at least two nodes present. Each cutting is taken from the original transformant (mother plant) and dipped into rooting hormone powder (indole-3-butyric acid, IBA). The cutting is then placed in oasis wedges inside a bio-dome. Wild type cuttings are also taken simultaneously to serve as controls. The cuttings are kept in the bio-dome for 5-7 days and then transplanted to pots and then acclimated in the growth chamber for two more days. Subsequently, the cuttings are transferred to the greenhouse, acclimated for approximately 4 days, and then subjected to spray tests as indicated. Depending on the species, the plants are kept at 10-25°C or 20-35°C. The test period extends over 3 weeks. During this time, the plants are tended and their response to the individual treatments is evaluated. Herbicide injury evaluations are taken at 2 and 3 weeks after treatment. Plant injury is rated on a scale of 0% to 100%, 0% being no injury and 100% being complete death.

Transgenic Arabidopsis thaliana plants are assayed for improved tolerance to TK-inhibiting herbicides in 48-well plates. Therefore, T2 seeds are surface sterilized by stirring for 5 min in ethanol + water (70+30 by volume), rinsing one time with ethanol + water (70+30 by volume) and two times with sterile, deionized water. The seeds are resuspended in 0.1% agar dissolved in water (w/v) Four to five seeds per well are plated on solid nutrient medium consisting of half-strength murashige skoog nutrient solution, pH 5.8 (Murashige and Skoog (1962) Physiologia Plantarum 15: 473-497). Compounds are dissolved in dimethylsulfoxid

(DMSO) and added to the medium prior solidification (final DMSO concentration 0.1%). Multi well plates are incubated in a growth chamber at 22°C, 75% relative humidity and 110 µmol Phot * m^{-2} * s^{-1} with 14 : 10 h light : dark photoperiod. Growth inhibition is evaluated seven to ten days after seeding in comparison to wild type plants (Table 6). Additionally, transgenic T1 Arabidopsis plants can be tested for improved tolerance to herbicides in greenhouse studies with TK-inhibiting herbicides such as cornexisthin.

Table 7: Relative tolerance rates of transgenic *Arabidopsis* plants as compared to a non-transgenic *Arabidopsis* plants (non-transgenic = 1.0), treated with Cornexistin. Growth inhibition is evaluated seven to ten days after seeding in comparison to wild type plants.

Gene	***Seq ID***	*Cornexistin*
LEMPA_TKL wt	10	9
SPIOL_TKL wt	1	4.8
Pd_TKL1	104	28
Pd_TKL2	130	8.6
Pd_TKL3	156	10

EXAMPLE 10: Sequence Analysis.

Leaf tissue was collected from clonal plants separated for transplanting and analyzed as individuals. Genomic DNA was extracted using a Wizard® 96 Magnetic DNA Plant System kit (Promega, US Patent Nos. 6,027,945 & 6,368,800) as directed by the manufacturer. Isolated DNA was PCR amplified using the appropriate forward and reverse primer. PCR amplification was performed using Hotstar Taq DNA Polymerase (Qiagen) using touchdown thermocycling program as follows: 96ºC for 15 min, followed by 35 cycles (96ºC, 30 sec; 58ºC – 0.2 ºC per cycle, 30 sec; 72ºC, 3 min and 30 sec), 10 min at 72ºC. PCR products were verified for concentration and fragment size via agarose gel electrophoresis. Dephosphorylated PCR products were analyzed by direct sequence using the PCR primers (DNA Landmarks, or Entelechon). Chromatogram trace files (.scf) were analyzed for mutation relative to the wild-type gene using Vector NTI Advance 10™ (Invitrogen). Based on sequence information, mutations were identified in several individuals. Sequence analysis was performed on the representative chromatograms and corresponding AlignX alignment with default settings and edited to call secondary peaks.

Claims:

1. A plant or plant part comprising a polynucleotide encoding a wildtype or mutated transketolase (TK) polypeptide, the expression of said polynucleotide confers to the plant or plant part tolerance to TK-inhibiting herbicides.

2. The plant or plant part of claim 1, wherein the polynucleotide encoding the wildtype or mutated TK polypeptide comprises the nucleic acid sequence set forth in SEQ ID NO: 182 or 183, or a homologue, variant or derivative thereof.

3. The plant or plant part of any of claims 1 to 2, wherein the wildtype or mutated TK polypeptide is a functional variant having, over the full-length of the variant, at least about 60%, illustratively, at least about 80%, 90%, 95%, 98%, 99% or more amino acid sequence identity to SEQ ID NO: 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, 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, or 181.

4. The plant or plant part of any of claims 1 to 3, wherein the mutated TK refers to a TK polypeptide comprising the sequence of SEQ ID NO: 1, an orthologue, paralogue, or homologue thereof, wherein the amino acid sequence differs from the wildtype amino acid sequence at one or more positions corresponding to positions 265, 267, 337, 342, 343, 458, 459, 460, 461, 463, 511, 512, 513, 514, 515, 544 of SEQ ID NO: 1.

5. A seed capable of germination into a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

6. A plant cell of or capable of regenerating a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides, wherein the plant cell comprises the

polynucleotide operably linked to a promoter.

7. A plant cell comprising a polynucleotide operably linked to a promoter operable in a cell, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

8. A plant product prepared from a plant or plant part comprising in at least some of its cells cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

9. A progeny or descendant plant derived from a plant comprising in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, wherein the progeny or descendant plant comprises in at least some of its cells the polynucleotide operably linked to the promoter, the expression of the wildtype or mutated TK polypeptide conferring to the progeny or descendant plant tolerance to the TK-inhibiting herbicides.

10. A method for controlling weeds at a locus for growth of a plant, the method comprising: (a) applying an herbicide composition comprising TK-inhibiting herbicides to the locus; and (b) planting a seed at the locus, wherein the seed is capable of producing a plant that comprises in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

11. The method of claim 10, wherein herbicide composition is applied to the weeds and to the plant produced by the seed.

12. A method of producing a plant having tolerance to TK-inhibiting herbicide, the method comprising regenerating a plant from a plant cell transformed with transformed with a recombinant polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

13. A method of producing a progeny plant having tolerance to TK-inhibiting herbicide, the method comprising: crossing a first TK-inhibiting herbicide-tolerant plant with a second plant to produce a TK-inhibiting herbicide- tolerant progeny plant, wherein the first plant and the progeny plant comprise in at least some of their cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides.

14. A plant or plant part comprising in at least some of its cells a recombinant polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides, wherein the plant or plant part further exhibits a second or third herbicide-tolerant trait.

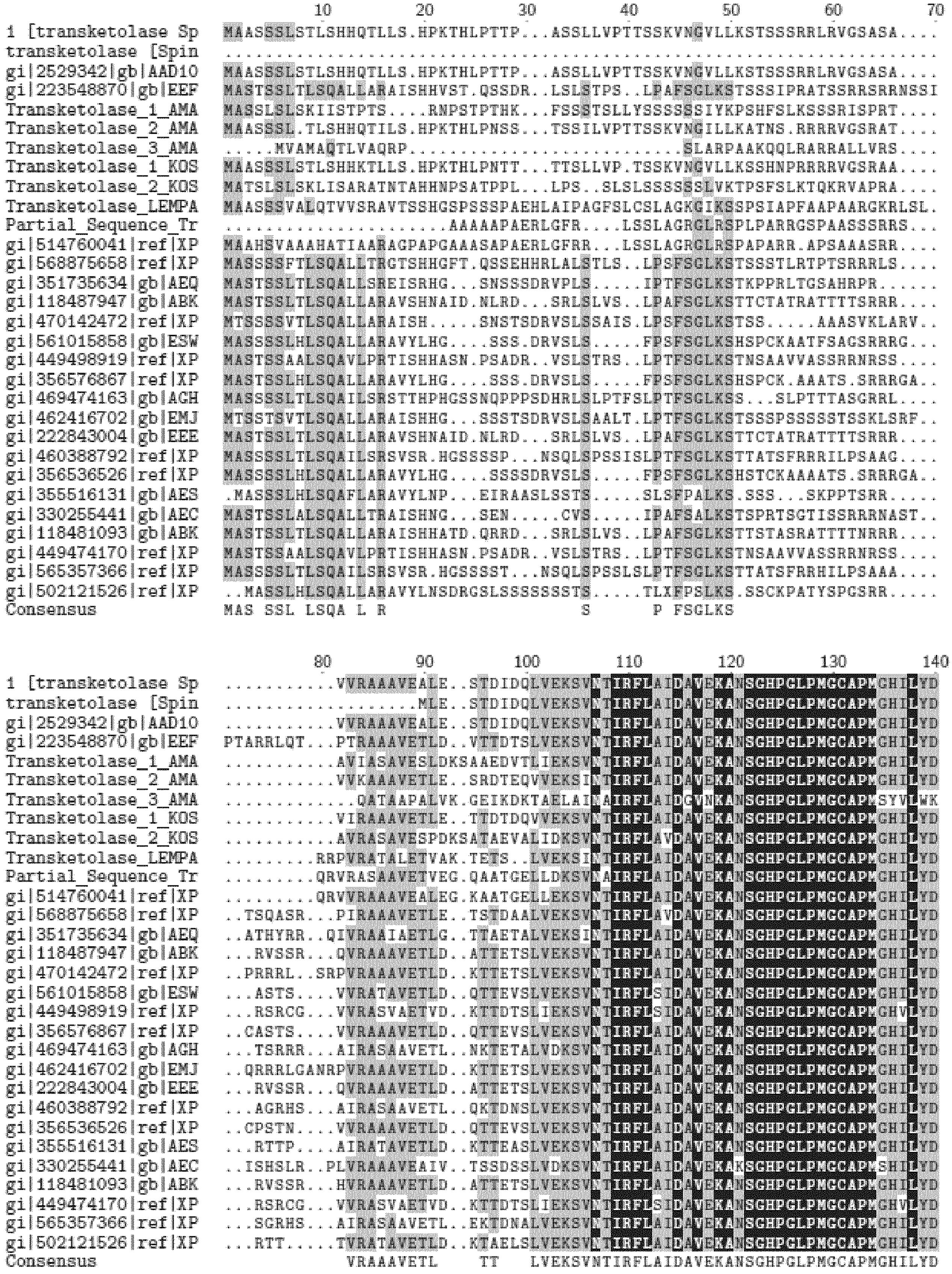
15. A method for producing a plant product from the plant of claim 1, the method comprising processing the plant or a plant part thereof to obtain the plant product.

16. The method of claim 15, wherein the plant product is fodder, seed meal, oil, or seed-treatment-coated seeds.

17. A plant product obtained from a plant or plant part thereof, wherein the plant or plant part comprises in at least some of its cells a polynucleotide operably linked to a promoter operable in plant cells, the promoter capable of expressing a wildtype or mutated TK polypeptide encoded by the polynucleotide, the expression of the wildtype or mutated TK polypeptide conferring to the plant tolerance to TK-inhibiting herbicides, wherein the plant or plant part further exhibits a second or third herbicide-tolerant trait.

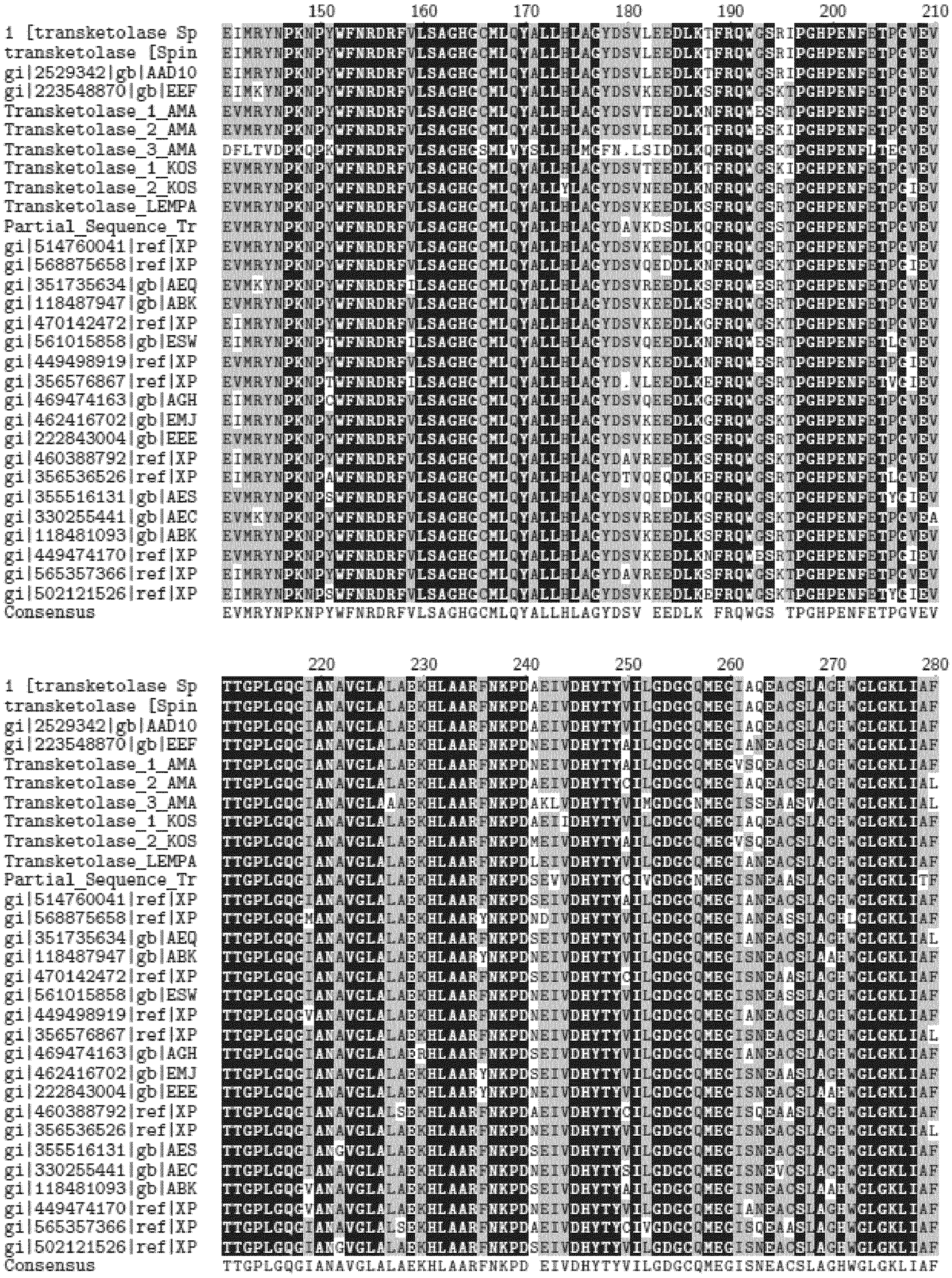
18. The plant product of claim 17, wherein the product is fodder, seed meal, oil, or seed-treatment-coated seed.

Figure 1A



```
                             10        20        30        40        50        60        70
1 [transketolase Sp  MAASSSLSTLSHHQTLLS.HPKTHLPTTP...ASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASA....
transketolase [Spin  ......................................................................
gi|2529342|gb|AAD10  MAASSSLSTLSHHQTLLS.HPKTHLPTTP...ASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASA....
gi|223548870|gb|EEF  MASTSSLTLSQALLARAISHHVST.QSSDR..LSLSTPS..LPAFSGLKSTSSSIPRATSSRSRRNSSI
Transketolase_1_AMA  MASSLSLSKIISTPTS....RNPSTPTHK...FSSSTSLLYSSSSSSIYKPSHFSLKSSSRISPRT....
Transketolase_2_AMA  MAASSSL.TLSHHQTILS.HPKTHLPNSS...TSSILVPTTSSKVNGILLKATNS.RRRRVGSRAT....
Transketolase_3_AMA  .....MVAMAQTLVAQRP..........................SLARPAAKQQLRARRALLVRS....
Transketolase_1_KOS  MAASSSLSTLSHHKTLLS.HPKTHLPNTT...TTSLLVP.TSSKVNGVLLKSSHNPRRRRVGSRAA....
Transketolase_2_KOS  MATSLSLSKLISARATNTAHHNPSATPPL...LPS..SLSLSSSSSSLVKTPSFSLKTQKRVAPRA....
Transketolase_LEMPA  MAASSSVALQTVVSRAVTSSHGSPSSSPAEHLAIPAGFSLCSLAGKGIKSSPSIAPFAAPAARGKRLSL.
Partial_Sequence_Tr  ......................AAAAAPAERLGFR....LSSLAGRGLRSPLPARRGSPAASSSRRS...
gi|514760041|ref|XP  MAAHSVAAAHATIAARAGPAPGAAASAPAERLGFRR...LSSLAGRGLRSPAPARR.APSAAASRR....
gi|568875658|ref|XP  MASSSSFTLSQALLTRGTSHHGFT.QSSEHHRLALSTLS..LPSFSGLKSTSSSTLRTPTSRRRLS....
gi|351735634|gb|AEQ  MASTSSLTLSQALLSREISRHG...SNSSSDRVPLS.....IPTFSGLKSTKPPRLTGSAHRPR......
gi|118487947|gb|ABK  MASTSSLTLSQALLARAVSHNAID.NLRD...SRLSLVS..LPAFSGLKSTTCTATRATTTTSRRR....
gi|470142472|ref|XP  MTSSSSVTLSQALLARAISH.....SNSTSDRVSLSSAIS.LPSFSGLKSTSS.....AAASVKLARV..
gi|561015858|gb|ESW  MASSSSLHLSQALLARAVYLHG....SSS.DRVSLS.....FPSFSGLKSHSPCKAATFSAGSRRRG...
gi|449498919|ref|XP  MASTSSAALSQAVLPRTISHHASN.PSADR..VSLSTRS..LPTFSGLKSTNSAAVVASSRRNRSS....
gi|356576867|ref|XP  MASTSSLHLSQALLARAVYLHG....SSS.DRVSLS.....FPSFSGLKSHSPCK.AAATS.SRRRGA..
gi|469474163|gb|AGH  MASSSSLTLSQAILSRSTTHPHGSSNQPPPSDHRLSLPTFSLPTFSGLKSS...SLPTTTASGRRL....
gi|462416702|gb|EMJ  MTSSTSVTLSQALLARAISHHG...SSSTSDRVSLSAALT.LPTFSGLKSTSSSPSSSSSTSSKLSRF..
gi|222843004|gb|EEE  MASTSSLTLSQALLARAVSHNAID.NLRD...SRLSLVS..LPAFSGLKSTTCTATRATTTTSRRR....
gi|460388792|ref|XP  MASSSSLTLSQAILSRSVSR.HGSSSSP...NSQLSPSSISLPTFSGLKSTTATSFRRRILPSAAG....
gi|356536526|ref|XP  MASSSSLHLSQALLARAVYLHG....SSSSDRVSLS.....FPSFSGLKSHSTCKAAAATS.SRRRGA..
gi|355516131|gb|AES  .MASSSLHLSQAFLARAVYLNP...EIRAASLSSTS.....SLSFPALKS.SSS...SKPPTSRR.....
gi|330255441|gb|AEC  MASTSSLALSQALLTRAISHNG...SEN.....CVS.....IPAFSALKSTSPRTSGTISSRRRNAST..
gi|118481093|gb|ABK  MASTSSLTLSQALLARAISHHATD.QRRD...SRLSLVS..LPAFSGLKSTTSTASRATTTTNRRR....
gi|449474170|ref|XP  MASTSSAALSQAVLPRTISHHASN.PSADR..VSLSTRS..LPTFSGLKSTNSAAVVASSRRNRSS....
gi|565357366|ref|XP  MASSSSLTLSQAILSRSVSR.HGSSSST...NSQLSPSSLSLPTFSGLKSTTATSFRRHILPSAAA....
gi|502121526|ref|XP  ..MASSLHLSQALLARAVYLNSDRGSLSSSSSSSTS.....TLXFPSLKS.SSCKPATYSPGSRR.....
Consensus            MAS SSL LSQA L R                   S      P FSGLKS

                             80        90       100       110       120       130       140
1 [transketolase Sp  ...........VVRAAAVEALE..STDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
transketolase [Spin  ...................MLE..STDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|2529342|gb|AAD10  ...........VVRAAAVEALE..STDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|223548870|gb|EEF  PTARRLQT...PTRAAAVETLD..VTTDTSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
Transketolase_1_AMA  ...........AVIASAVESLDKSAAEDVTLIEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
Transketolase_2_AMA  ...........VVKAAAVETLE..SRDTEQVVEKSINTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
Transketolase_3_AMA  .............QATAAPALVK.GEIKDKTAELAINAIRFLAIDGVNKANSGHPGLPMGCAPMSYVLWK
Transketolase_1_KOS  ...........VIRAAAVETLE..TTDTDQVVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
Transketolase_2_KOS  ...........AVRASAVESPDKSATAEVALIDKSVNTIRFLAVDAVEKANSGHPGLPMGCAPMGHILYD
Transketolase_LEMPA  .........RRPVRATALETVAK.TETS..LVEKSINTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
Partial_Sequence_Tr  .........QRVRASAAVETVEG.QAATGELLDKSVNAIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|514760041|ref|XP  .........QRVVRAAAVEALEG.KAATGELLEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|568875658|ref|XP  ..TSQASR...PIRAAAVETLE..TSTDAALVEKSVNTIRFLAVDAVEKANSGHPGLPMGCAPMGHILYD
gi|351735634|gb|AEQ  ..ATHYRR..QIVRAAIAETLG..TTAETALVEKSINTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|118487947|gb|ABK  ...RVSSR...QVRAAAVETLD..ATTETSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|470142472|ref|XP  ..PRRRL..SRPVRAAAVETLD..KTTETSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|561015858|gb|ESW  ...ASTS....VVRATAVETLD..QTTEVSLVEKSVNTIRFLSIDAVEKANSGHPGLPMGCAPMGHILYD
gi|449498919|ref|XP  ...RSRCG...VVRASVAETVD..KTTDTSLIEKSVNTIRFLSIDAVEKANSGHPGLPMGCAPMGHVLYD
gi|356576867|ref|XP  ..CASTS....VVRAAAVETLD..QTTEVSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|469474163|gb|AGH  ...TSRRR...AIRASAAVETL..NKTETALVDKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|462416702|gb|EMJ  ..QRRRLGANRPVRAAAVETLD..KTTETSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|222843004|gb|EEE  ...RVSSR...QVRAAAVETLD..ATTETSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|460388792|ref|XP  ...AGRHS...AIRASAAVETL..QKTDNSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|356536526|ref|XP  ..CPSTN....VVRAAAVETLD..QTTEVSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|355516131|gb|AES  ...RTTP....AIRATAVETLD..KTTEASLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|330255441|gb|AEC  ..ISHSLR..PLVRAAAVEAIV..TSSDSSLVDKSVNTIRFLAIDAVEKAKSGHPGLPMGCAPMSHILYD
gi|118481093|gb|ABK  ...RVSSR...HVRAAAVETLD..ATTETSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|449474170|ref|XP  ...RSRCG...VVRASVAETVD..KTTDTSLIEKSVNTIRFLSIDAVEKANSGHPGLPMGCAPMGHVLYD
gi|565357366|ref|XP  ...SGRHS...AIRASAAVETL..EKTDNALVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
gi|502121526|ref|XP  ...RTT.....TVRATAVETLD..KTAELSLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
Consensus                        VRAAAVETL    TT   LVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYD
```

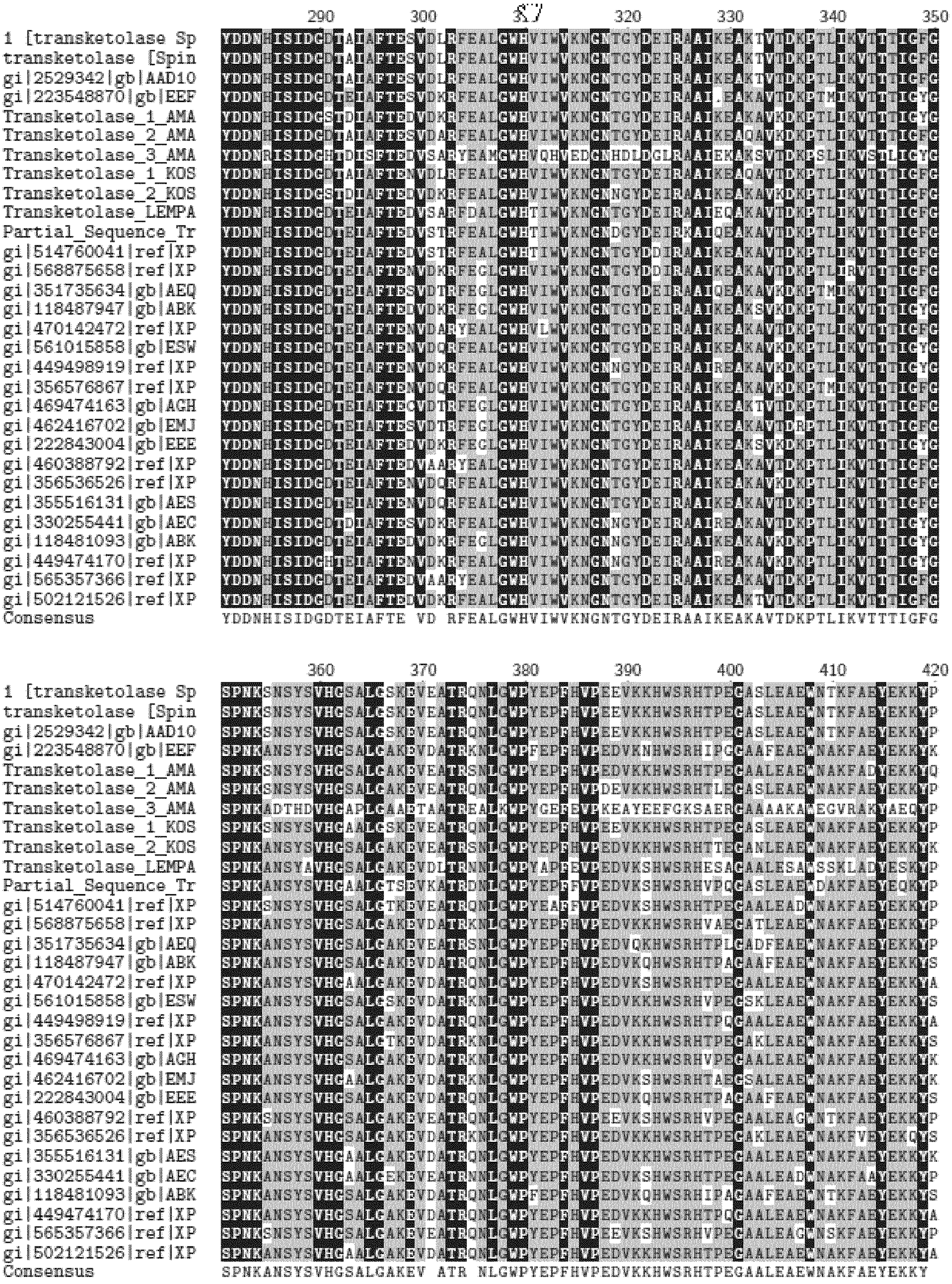
Figure 1A continued



```
                               150       160       170       180       190       200       210
1 [transketolase Sp    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEV
transketolase [Spin    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEV
gi|2529342|gb|AAD10    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEV
gi|223548870|gb|EEF    EIMKYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSKTPGHPENFETPGVEV
Transketolase_1_AMA    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVTEEDLKNFRQWESRTPGHPENFETPGVEV
Transketolase_2_AMA    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWESKIPGHPENFETPGVEV
Transketolase_3_AMA    DFLTVDPKQPKWFNRDRFVLSAGHGSMLVYSLLHLMGFN.LSIDDLKQFRQWGSKTPGHPENFLTEGVEV
Transketolase_1_KOS    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVTEEDLKTFRQWGSKIPGHPENFETPGVEV
Transketolase_2_KOS    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLYLAGYDSVNEEDLKNFRQWGSRTPGHPENFETPGIEV
Transketolase_LEMPA    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSRTPGHPENFETPGVEV
Partial_Sequence_Tr    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVKDSDLKQFRQWGSSTPGHPENFETPGVEV
gi|514760041|ref|XP    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKQFRQWGSRTPGHPENFETPGVEV
gi|568875658|ref|XP    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVQEDDLKNFRQWGSKTPGHPENFETPGIEV
gi|351735634|gb|AEQ    EVMKYNPKNPYWFNRDRFILSAGHGCMLQYALLHLAGYDSVREEDLKSFRQWESRTPGHPENFETPGVEV
gi|118487947|gb|ABK    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSRTPGHPENFETPGVEV
gi|470142472|ref|XP    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKGFRQWGSKTPGHPENFETPGVEV
gi|561015858|gb|ESW    EIMRYNPKNPTWFNRDRFILSAGHGCMLQYALLHLAGYDSVQEEDLKNFRQWESRTPGHPENFETLGVEV
gi|449498919|ref|XP    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKNFRQWESRTPGHPENFETPGIEV
gi|356576867|ref|XP    EVMRYNPKNPTWFNRDRFILSAGHGCMLQYALLHLAGYD.VLEEDLKEFRQWGSRTPGHPENFETVGIEV
gi|469474163|gb|AGH    EIMRYNPKNPCWFNRDRFVLSAGHGCMLQYALLHLAGYDSVQEEDLKGFRQWGSKTPGHPENFETPGVEV
gi|462416702|gb|EMJ    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKGFRQWGSKTPGHPENFETPGVEV
gi|222843004|gb|EEE    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSRTPGHPENFETPGVEV
gi|460388792|ref|XP    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVREEDLKSFRQWGSKTPGHPENFETPGVEV
gi|356536526|ref|XP    EIMRYNPKNPAWFNRDRFVLSAGHGCMLQYALLHLAGYDTVQEQDLKEFRQWGSRTPGHPENFETLGVEV
gi|355516131|gb|AES    EVMRYNPKNPSWFNRDRFVLSAGHGCMLQYALLHLAGYDSVQEDDLKQFRQWGSKTPGHPENFETYGIEV
gi|330255441|gb|AEC    EVMKYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVREEDLKSFRQWGSKTPGHPENFETPGVEA
gi|118481093|gb|ABK    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSRTPGHPENFETPGVEV
gi|449474170|ref|XP    EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKNFRQWESRTPGHPENFETPGIEV
gi|565357366|ref|XP    EIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVREEDLKSFRQWGSKTPGHPENFETPGVEV
gi|502121526|ref|XP    EIMRYNPKNPSWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKEFRQWGSKTPGHPENFETYGIEV
Consensus              EVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSV EEDLK FRQWGS TPGHPENFETPGVEV
```

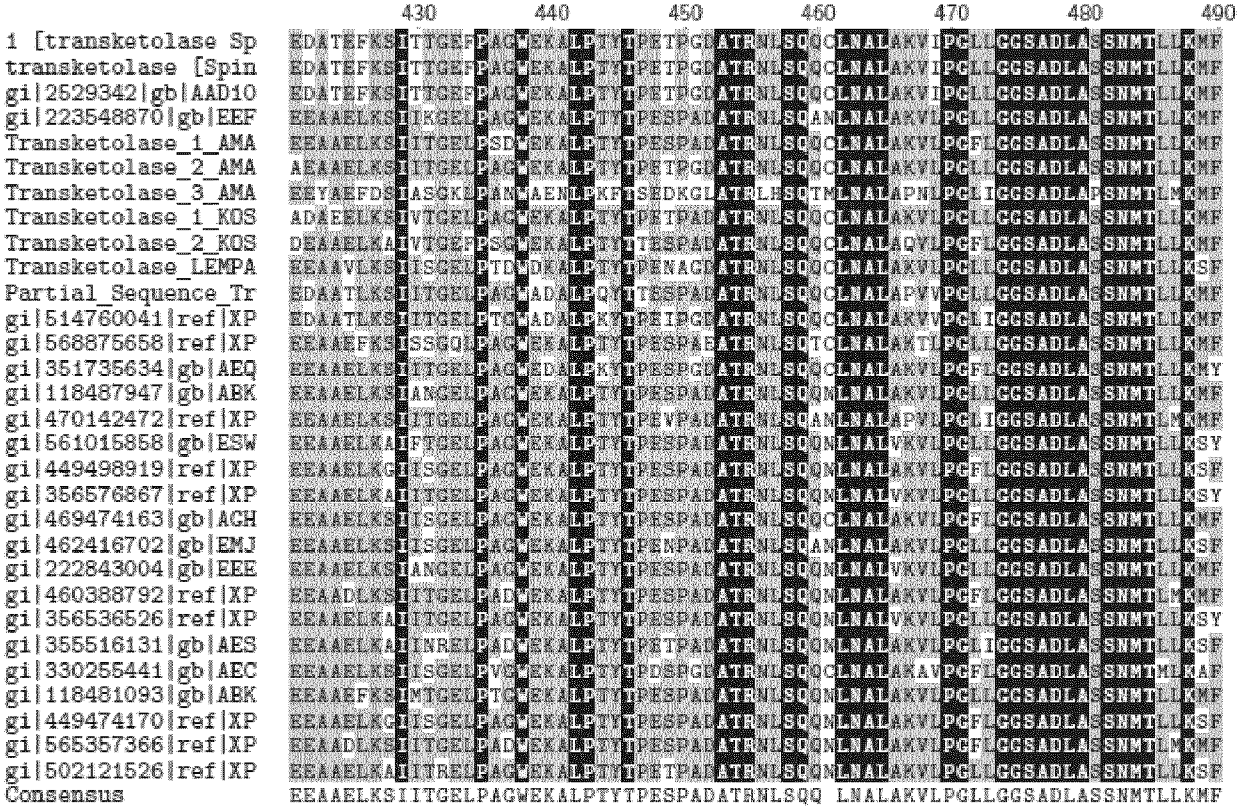
```
                               220       230       240       250       260       270       280
1 [transketolase Sp    TTGPLGQGIANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAF
transketolase [Spin    TTGPLGQGIANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAF
gi|2529342|gb|AAD10    TTGPLGQGIANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAF
gi|223548870|gb|EEF    TTGPLGQGIANAVGLALAEKHLAARFNKPDNEIVDHYTYAILGDGCQMEGIANEACSLAGHWGLGKLIAF
Transketolase_1_AMA    TTGPLGQGIANAVGLALAEKHLAARFNKPDNEIVDHYTYAILGDGCQMEGVSQEACSLAGHWGLGKLIAF
Transketolase_2_AMA    TTGPLGQGIANAVGLALAEKHLAARFNKPDAEIVDHYTYCILGDGCQMEGIAQEACSLAGHWGLGKLIAL
Transketolase_3_AMA    TTGPLGQGIANAVGLAAAEKHLAARFNKPDAKLVDHYTYVIMGDGCNMEGISSEAASVAGHWGLGKLIAL
Transketolase_1_KOS    TTGPLGQGIANAVGLALAEKHLAARFNKPDAEIIDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAF
Transketolase_2_KOS    TTGPLGQGIANAVGLALAEKHLAARFNKPDMEIVDHYTYAILGDGCQMEGVSQEACSLAGHWGLGKLIAF
Transketolase_LEMPA    TTGPLGQGIANAVGLALAEKHLAARFNKPDLEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAF
Partial_Sequence_Tr    TTGPLGQGIANAVGLALAEKHLAARFNKPDSEVVDHYTYCIVGDGCNMEGISNEAASLAGHWGLGKLITF
gi|514760041|ref|XP    TTGPLGQGIANAVGLALAEKHLAARFNKPDSEIVDHYTYAILGDGCQMEGIANEACSLAGHWGLGKLIAF
gi|568875658|ref|XP    TTGPLGQGMANAVGLALAEKHLAARYNKPDNDIVDHYTYVILGDGCQMEGIANEASSLAGHLGLGKLIAF
gi|351735634|gb|AEQ    TTGPLGQGIANAVGLALAEKHLAARFNKPDSEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAL
gi|118487947|gb|ABK    TTGPLGQGIANAVGLALAEKHLAARYNKPDNEIVDHYTYVILGDGCQMEGISNEACSLAAHWGLGKLIAF
gi|470142472|ref|XP    TTGPLGQGIANAVGLALAEKHLAARFNKPDSEIVDHYTYCILGDGCQMEGISNEAASLAGHWGLGKLIAF
gi|561015858|gb|ESW    TTGPLGQGIANAVGLALAEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGISNEASSLAGHWGLGKLIAF
gi|449498919|ref|XP    TTGPLGQGVANAVGLALAEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAF
gi|356576867|ref|XP    TTGPLGQGIANAVGLALAEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAL
gi|469474163|gb|AGH    TTGPLGQGIANAVGLALAERHLAARFNKPDSEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAF
gi|462416702|gb|EMJ    TTGPLGQGIANAVGLALAEKHLAARYNKPDSEIVDHYTYVILGDGCQMEGISNEAASLAGHWGLGKLIAF
gi|222843004|gb|EEE    TTGPLGQGIANAVGLALAEKHLAARYNKPDNEIVDHYTYVILGDGCQMEGISNEACSLAAHWGLGKLIAF
gi|460388792|ref|XP    TTGPLGQGIANAVGLALSEKHLAARFNKPDAEIVDHYTYCILGDGCQMEGISQEAASLAGHWGLGKLIAF
gi|356536526|ref|XP    TTGPLGQGIANAVGLALAEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAL
gi|355516131|gb|AES    TTGPLGQGIANGVGLALAEKHLAARFNKPDSEIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAF
gi|330255441|gb|AEC    TTGPLGQGIANAVGLALAEKHLAARFNKPDNEIVDHYTYSILGDGCQMEGISNEVCSLAGHWGLGKLIAF
gi|118481093|gb|ABK    TTGPLGQGVANAVGLALAEKHLAARFNKPDSEIVDHYTYAILGDGCQMEGISNEACSLAAHWGLGKLIAF
gi|449474170|ref|XP    TTGPLGQGVANAVGLALAEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAF
gi|565357366|ref|XP    TTGPLGQGIANAVGLALSEKHLAARFNKPDAEIVDHYTYCIVGDGCQMEGISQEAASLAGHWGLGKLIAF
gi|502121526|ref|XP    TTGPLGQGIANGVGLALAEKHLAARFNKPDSEIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAF
Consensus              TTGPLGQGIANAVGLALAEKHLAARFNKPD EIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAF
```

Figure 1A continued

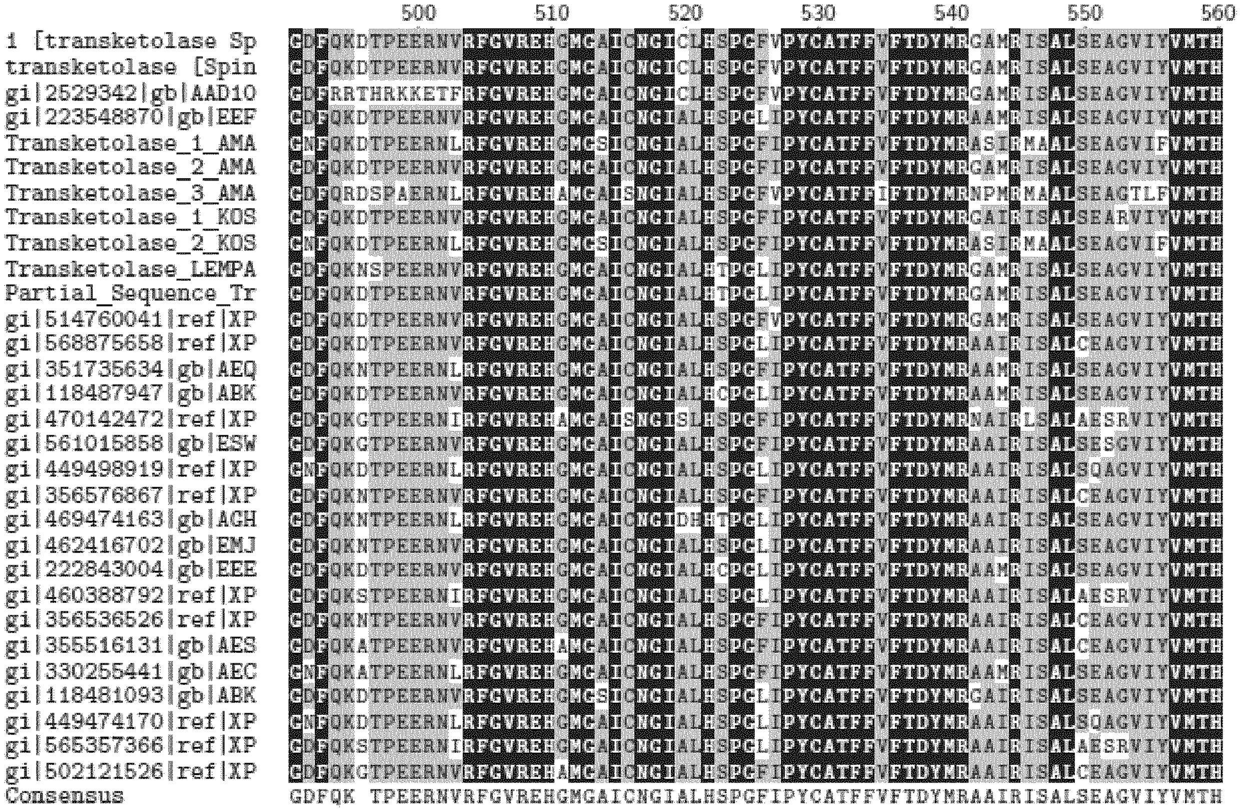


```
                     290       300       310       320       330       340       350
1 [transketolase Sp  YDDNHISIDGDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFG
transketolase [Spin  YDDNHISIDGDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFG
gi|2529342|gb|AAD10  YDDNHISIDGDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFG
gi|223548870|gb|EEF  YDDNHISIDGDTEIAFTESVDKRFEALGWHVIWVKNGNTGYDEIRAAI.EAKAVTDKPTMIKVTTTIGYG
Transketolase_1_AMA  YDDNHISIDGSTDIAFTEDVDKRFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVKDKPTLIKVTTTIGYG
Transketolase_2_AMA  YDDNHISIDGDTAIAFTESVDARFEALGWHVIWVKNGNTGYDEIRAAIKEAQAVKDKPTLIKVTTTIGFG
Transketolase_3_AMA  YDDNRISIDGHTDISFTEDVSARYEAMGWHVQHVEDGNHDLDGLRAAIEKAKSVTDKPSLIKVSTLIGYG
Transketolase_1_KOS  YDDNHISIDGDTAIAFTENVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAQAVTDKPTLIKVTTTIGFG
Transketolase_2_KOS  YDDNHISIDGSTDIAFTEDVDKRFEALGWHVIWVKNGNNGYDEIRAAIKEAKAVKDKPTLIKVTTTIGFG
Transketolase_LEMPA  YDDNHISIDGDTEIAFTEDVSARFDALGWHTIWVKNGNTGYDEIRAAIEQAKAVTDKPTLIKVTTTIGFG
Partial_Sequence_Tr  YDDNHISIDGDTEIAFTEDVSTRFEALGWHTIWVKNGNDGYDEIRKAIQEAKAVTDKPTLIKVTTTIGFG
gi|514760041|ref|XP  YDDNHISIDGDTEIAFTEDVSTRFEALGWHTIWVKNGNTGYDDIRAAIKEAKAVTDKPTLIKVTTTIGFG
gi|568875658|ref|XP  YDDNHISIDGDTEIAFTENVDKRFEGLGWHVIWVKNGNTGYDDIRAAIKEAKAVTDKPTLIRVTTTIGFG
gi|351735634|gb|AEQ  YDDNHISIDGDTEIAFTESVDTRFEGLGWHVIWVKNGNTGYDEIRAAIQEAKAVKDKPTMIKVTTTIGFG
gi|118487947|gb|ABK  YDDNHISIDGDTEIAFTEDVDKRFEGLGWHVIWVKNGNTGYDEIRAAIKEAKSVKDKPTLIKVTTTIGYG
gi|470142472|ref|XP  YDDNHISIDGDTEIAFTENVDARYEALGWHVLWVKNGNTGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFG
gi|561015858|gb|ESW  YDDNHISIDGDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVKDKPTLIKVTTTIGYG
gi|449498919|ref|XP  YDDNHISIDGDTEIAFTENVDKRFEALGWHVIWVKNGNNGYDEIRAAIREAKAVKDKPTLIKVTTTIGYG
gi|356576867|ref|XP  YDDNHISIDGDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVKDKPTMIKVTTTIGFG
gi|469474163|gb|AGH  YDDNHISIDGDTEIAFTECVDTRFEGLGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFG
gi|462416702|gb|EMJ  YDDNHISIDGDTEIAFTESVDTRFEGLGWHVIWVKNGNTGYDEIRAAIKEAKAVTDRPTLIKVTTTIGFG
gi|222843004|gb|EEE  YDDNHISIDGDTEIAFTEDVDKRFEGLGWHVIWVKNGNTGYDEIRAAIKEAKSVKDKPTLIKVTTTIGYG
gi|460388792|ref|XP  YDDNHISIDGDTEIAFTEDVAARYEALGWHVIWVKNGNTGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFG
gi|356536526|ref|XP  YDDNHISIDGDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVKDKPTLIKVTTTIGFG
gi|355516131|gb|AES  YDDNHISIDGDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFG
gi|330255441|gb|AEC  YDDNHISIDGDTDIAFTESVDKRFEALGWHVIWVKNGNNGYDEIRAAIREAKAVTDKPTLIKVTTTIGYG
gi|118481093|gb|ABK  YDDNHISIDGDTEIAFTEDVDKRFEGLGWHVIWVKNGNNGYDEIRAAIKEAKAVKDKPTLIKVTTTIGYG
gi|449474170|ref|XP  YDDNHISIDGHTEIAFTENVDKRFEALGWHVIWVKNGNNGYDEIRAAIREAKAVKDKPTLIKVTTTIGYG
gi|565357366|ref|XP  YDDNHISIDGDTEIAFTEDVAARYEALGWHVIWVKNGNTGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFG
gi|502121526|ref|XP  YDDNHISIDGDTEIAFTEDVDKRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFG
Consensus            YDDNHISIDGDTEIAFTE VD RFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFG

                     360       370       380       390       400       410       420
1 [transketolase Sp  SPNKSNSYSVHGSALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYP
transketolase [Spin  SPNKSNSYSVHGSALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYP
gi|2529342|gb|AAD10  SPNKSNSYSVHGSALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYP
gi|223548870|gb|EEF  SPNKANSYSVHGSALGAKEVDATRKNLGWPFEPFHVPEDVKNHWSRHIPGGAAFEAEWNAKFAEYEKKYK
Transketolase_1_AMA  SPNKSNSYSVHGSALGAKEVEATRSNLGWPYEPFHVPEDVKKHWSRHTPEGAALEAEWNAKFADYEKKYQ
Transketolase_2_AMA  SPNKSNSYSVHGSALGAKEVEATRQNLGWPYEPFHVPDEVKKHWSRHTLEGASLEAEWNAKFAEYEKKYP
Transketolase_3_AMA  SPNKADTHDVHGAPLGAAETAATREALKWPYGEFEVPKEAYEEFGKSAERGAAAAKAWEGVRAKYAEQYP
Transketolase_1_KOS  SPNKSNSYSVHGAALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNAKFAEYEKKYP
Transketolase_2_KOS  SPNKANSYSVHGSALGAKEVEATRSNLGWPYEPFHVPEDVKKHWSRHTTEGANLEAEWNAKFAEYEKKYK
Transketolase_LEMPA  SPNKANSYAVHGSALGAKEVDLTRNNLGWPYAPFEVPEDVKSHWSRHESAGAALESAWSSKLADYESKYP
Partial_Sequence_Tr  SPNKANSYSVHGAALGTSEVKATRDNLGWPYEPFFVPEDVKSHWSRHVPQGASLEAEWDAKFAEYEQKYP
gi|514760041|ref|XP  SPNKSNSYSVHGSALGTKEVEATRQNLGWPYEAFFVPEDVKSHWSRHTPEGAALEADWNAKFAEYEKKYP
gi|568875658|ref|XP  SPNKANSYSVHGSALGAKEVDATRKNLGWPYEPFHVPEDVKKHWSRHVAEGATLEAEWNAKFAEYEKKYP
gi|351735634|gb|AEQ  SPNKANSYSVHGSALGAKEVEATRSNLGWPYEPFHVPEDVQKHWSRHTPLGADFEAEWNAKFAEYEKKYP
gi|118487947|gb|ABK  SPNKANSYSVHGSALGAKEVDATRQNLGWPYEPFHVPEDVKQHWSRHTPAGAAFEAEWNAKFAEYEKKYS
gi|470142472|ref|XP  SPNKANSYSVHGAALGAKEVDATRQNLGWPYEPFHVPEDVKSHWSRHTPEGAALEAEWNAKFAEYEKKYA
gi|561015858|gb|ESW  SPNKANSYSVHGSALGSKEVDATRKNLGWPYEPFHVPEDVKKHWSRHVPEGSKLEAEWNAKFAEYEKKYS
gi|449498919|ref|XP  SPNKANSYSVHGSALGAKEVDATRQNLGWPYEPFHVPEDVKKHWSRHTPQGAALEAEWNAKFAEYEKKYA
gi|356576867|ref|XP  SPNKANSYSVHGSALGTKEVDATRKNLGWPYEPFHVPEDVKKHWSRHTPEGAKLEAEWNAKFAEYEKKYS
gi|469474163|gb|AGH  SPNKANSYSVHGSALGAKEVDATRKNLGWPYEPFHVPEDVKKHWSRHVPEGAALEAEWNAKFAEYEKKYK
gi|462416702|gb|EMJ  SPNKANSYSVHGAALGAKEVDATRKNLGWPYEPFHVPEDVKSHWSRHTAEGSALEAEWNAKFAEYEKKYK
gi|222843004|gb|EEE  SPNKANSYSVHGSALGAKEVDATRQNLGWPYEPFHVPEDVKQHWSRHTPAGAAFEAEWNAKFAEYEKKYS
gi|460388792|ref|XP  SPNKSNSYSVHGSALGAKEVEATRQNLGWPYEPFHVPEEVKSHWSRHVPEGAALEAGWNTKFAEYEKKYP
gi|356536526|ref|XP  SPNKANSYSVHGSALGAKEVDATRKNLGWPYEPFHVPEDVKKHWSRHTPEGAKLEAEWNAKFVEYEKQYS
gi|355516131|gb|AES  SPNKANSYSVHGAALGAKEVDATRQNLGWPYEPFHVPEDVKKHWSRHTPEGAALEAEWNAKFAEYEKKYK
gi|330255441|gb|AEC  SPNKANSYSVHGAALGEKEVEATRNNLGWPYEPFHVPEDVKSHWSRHTPEGAALEADWNAKFAAYEKKYP
gi|118481093|gb|ABK  SPNKANSYSVHGSALGAKEVDATRQNLGWPFEPFHVPEDVKQHWSRHIPAGAAFEAEWNTKFAEYEKKYS
gi|449474170|ref|XP  SPNKANSYSVHGSALGAKEVDATRQNLGWPYEPFHVPEDVKKHWSRHTPQGAALEAEWNAKFAEYEKKYA
gi|565357366|ref|XP  SPNKSNSYSVHGSALGAKEVEATRQNLGWPYEPFHVPEEVKSHWSRHVPEGAALEAGWNSKFAEYEKKYP
gi|502121526|ref|XP  SPNKANSYSVHGAALGAKEVDATRQNLGWPYEPFHVPEDVKKHWSRHTPEGAALEAEWNAKFAEYEKKYA
Consensus            SPNKANSYSVHGSALGAKEV ATR NLGWPYEPFHVPEDVKKHWSRHTPEGAALEAEWNAKFAEYEKKY
```

Figure 1A continued

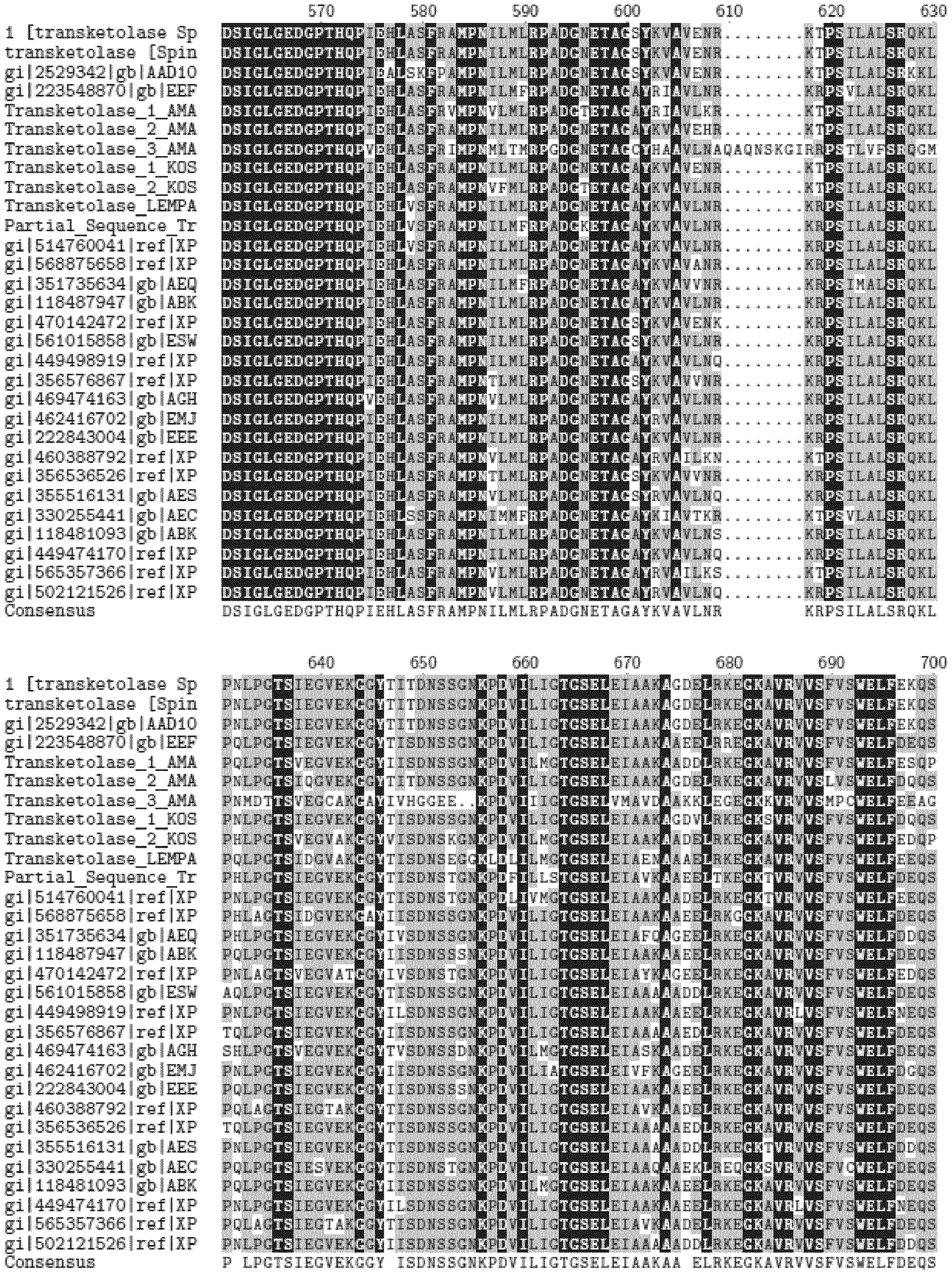


```
                                430          440          450          460          470          480          490
1 [transketolase Sp     EDATEFKSITTGEFPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMF
transketolase [Spin     EDATEFKSITTGEFPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMF
gi|2529342|gb|AAD10     EDATEFKSITTGEFPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMF
gi|223548870|gb|EEF     EEAAELKSIIKGELPAGWEKALPTYTPESPADATRNLSQANLNALAKVLPGLLGGSADLASSNMTLLKMF
Transketolase_1_AMA     EEAAELKSIITGELPSDWEKALPTYTPESPADATRNLSQQCLNALAKVLPGFLGGSADLASSNMTLLKMF
Transketolase_2_AMA     AEAAELKSIITGELPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVLPGLLGGSADLASSNMTLLKMF
Transketolase_3_AMA     EEYAEFDSIASGKLPANWAENLPKFTSEDKGLATRLHSQTMLNALAPNLPGLIGGSADLAPSNMTLMKMF
Transketolase_1_KOS     ADAEELKSIVTGELPAGWEKALPTYTPETPADATRNLSQQCLNALAKVLPGLLGGSADLASSNMTLLKMF
Transketolase_2_KOS     DEAAELKAIVTGEFPSGWEKALPTYTTESPADATRNLSQQCLNALAQVLPGFLGGSADLASSNMTLLKMF
Transketolase_LEMPA     EEAAVLKSIISGELPTDWDKALPTYTPENAGDATRNLSQQCLNALAKVLPGLLGGSADLASSNMTLLKSF
Partial_Sequence_Tr     EDAATLKSIITGELPAGWADALPQYTTESPADATRNLSQQCLNALAPVVPGLLGGSADLASSNMTLLKMF
gi|514760041|ref|XP     EDAATLKSIITGELPTGWADALPKYTPEIPGDATRNLSQQCLNALAKVVPGLIGGSADLASSNMTLLKMF
gi|568875658|ref|XP     EEAAEFKSISSGQLPAGWEKALPTYTPESPAEATRNLSQTCLNALAKTLPGLLGGSADLASSNMTLLKMF
gi|351735634|gb|AEQ     EEAAELKSIITGELPAGWEDALPKYTPESPGDATRNLSQQCLNALAKVLPGFLGGSADLASSNMTLLKMY
gi|118487947|gb|ABK     EEAAELKSIANGELPAGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGLLGGSADLASSNMTLLKMF
gi|470142472|ref|XP     EEAAELKSIITGELPAGWEKALPTYTPEVPADATRNLSQANLNALAPVLPGLIGGSADLASSNMTLMKMF
gi|561015858|gb|ESW     EEAAELKAIFTGELPAGWEKALPTYTPESPADATRNLSQQNLNALVKVLPGLLGGSADLASSNMTLLKSY
gi|449498919|ref|XP     EEAAELKGIISGELPAGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTLLKSF
gi|356576867|ref|XP     EEAAELKAIITGELPAGWEKALPTYTPESPADATRNLSQQNLNALVKVLPGLLGGSADLASSNMTLLKSY
gi|469474163|gb|ACH     EEAAELKSIISGELPAGWEKALPTYTPESPADATRNLSQQCLNALAKVLPGFLGGSADLASSNMTLLKMF
gi|462416702|gb|EMJ     EEAAELKSIISGELPAGWEKALPTYTPENPADATRNLSQANLNALAKVLPGLLGGSADLASSNMTLLKSF
gi|222843004|gb|EEE     EEAAELKSIANGELPAGWEKALPTYTPESPADATRNLSQQNLNALVKVLPGLLGGSADLASSNMTLLKMF
gi|460388792|ref|XP     EEAADLKSIITGELPADWEKALPTYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTLMKMF
gi|356536526|ref|XP     EEAAELKAIITGELPAGWEKALPTYTPESPADATRNLSQQNLNALVKVLPGLLGGSADLASSNMTLLKSY
gi|355516131|gb|AES     EEAAELKAIINRELPADWEKALPTYTPETPADATRNLSQQNLNALAKVLPGLIGGSADLASSNMTLLKSF
gi|330255441|gb|AEC     EEAAELKSIISGELPVGWEKALPTYTPDSPGDATRNLSQQCLNALAKAVPGFLGGSADLASSNMTMLKAF
gi|118481093|gb|ABK     EEAAEFKSIMTGELPTGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGLLGGSADLASSNMTLLKMF
gi|449474170|ref|XP     EEAAELKGIISGELPAGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTLLKSF
gi|565357366|ref|XP     EEAADLKSIITGELPADWEKALPTYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTLMKMF
gi|502121526|ref|XP     EEAAELKAIITRELPAGWEKALPTYTPETPADATRNLSQQNLNALAKVLPGLLGGSADLASSNMTLLKSF
Consensus               EEAAELKSIITGELPAGWEKALPTYTPESPADATRNLSQQ LNALAKVLPGLLGGSADLASSNMTLLKMF
```



```
                                500          510          520          530          540          550          560
1 [transketolase Sp     GDFQKDTPEERNVRFGVREHGMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
transketolase [Spin     GDFQKDTPEERNVRFGVREHGMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
gi|2529342|gb|AAD10     GDFRRTHRKKETFRFGVREHGMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
gi|223548870|gb|EEF     GDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTH
Transketolase_1_AMA     GNFQKDTPEERNLRFGVREHGMGSICNGIALHSPGFIPYCATFFVFTDYMRASIRMAALSEAGVIFVMTH
Transketolase_2_AMA     GDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
Transketolase_3_AMA     GDFQRDSPAERNLRFGVREHAMGAISNGIALHSPGFVPYCATFFIFTDYMRNPMRMAALSEAGTLFVMTH
Transketolase_1_KOS     GDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRGAIRISALSEARVIYVMTH
Transketolase_2_KOS     GNFQKDTPEERNLRFGVREHGMGSICNGIALHSPGFIPYCATFFVFTDYMRASIRMAALSEAGVIFVMTH
Transketolase_LEMPA     GDFQKNSPEERNVRFGVREHGMGAICNGIALHTPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
Partial_Sequence_Tr     GDFQKDTPEERNVRFGVREHGMGAICNGIALHTPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
gi|514760041|ref|XP     GDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTH
gi|568875658|ref|XP     GDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALCEAGVIYVMTH
gi|351735634|gb|AEQ     GDFQKNTPEERNLRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTH
gi|118487947|gb|ABK     GDFQKDTPEERNVRFGVREHGMGAICNGIALHCPGLIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTH
gi|470142472|ref|XP     GDFQKGTPEERNIRFGVREHAMGAISNGISLHSPGFIPYCATFFVFTDYMRNAIRLSALAESRVIYVMTH
gi|561015858|gb|ESW     GDFQKGTPEERNVRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAIRISALSESGVIYVMTH
gi|449498919|ref|XP     GNFQKDTPEERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALSQAGVIYVMTH
gi|356576867|ref|XP     GDFQKNTPEERNVRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAIRISALCEAGVIYVMTH
gi|469474163|gb|ACH     GDFQKNTPEERNLRFGVREHGMGAICNGIDHHTPGLIPYCATFFVFTDYMRAAIRISALSEAGVIYVMTH
gi|462416702|gb|EMJ     GDFQKNTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALSEAGVIYVMTH
gi|222843004|gb|EEE     GDFQKDTPEERNVRFGVREHGMGAICNGIALHCPGLIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTH
gi|460388792|ref|XP     GDFQKSTPEERNIRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALAESRVIYVMTH
gi|356536526|ref|XP     GDFQKNTPEERNVRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAIRISALCEAGVIYVMTH
gi|355516131|gb|AES     GDFQKATPEERNVRFGVREHAMGAICNGIALHSPGFIPYCATFFVFTDYMRAAIRISALCEAGVIYVMTH
gi|330255441|gb|AEC     GNFQKATPEERNLRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTH
gi|118481093|gb|ABK     GDFQKDTPEERNVRFGVREHGMGSICNGIALHSPGLIPYCATFFVFTDYMRGAIRISALSEAGVIYVMTH
gi|449474170|ref|XP     GNFQKDTPEERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALSQAGVIYVMTH
gi|565357366|ref|XP     GDFQKSTPEERNIRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALAESRVIYVMTH
gi|502121526|ref|XP     GDFQKGTPEERNVRFGVREHAMGAICNGIALHSPGFIPYCATFFVFTDYMRAAIRISALCEAGVIYVMTH
Consensus               GDFQK TPEERNVRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAIRISALSEAGVIYVMTH
```

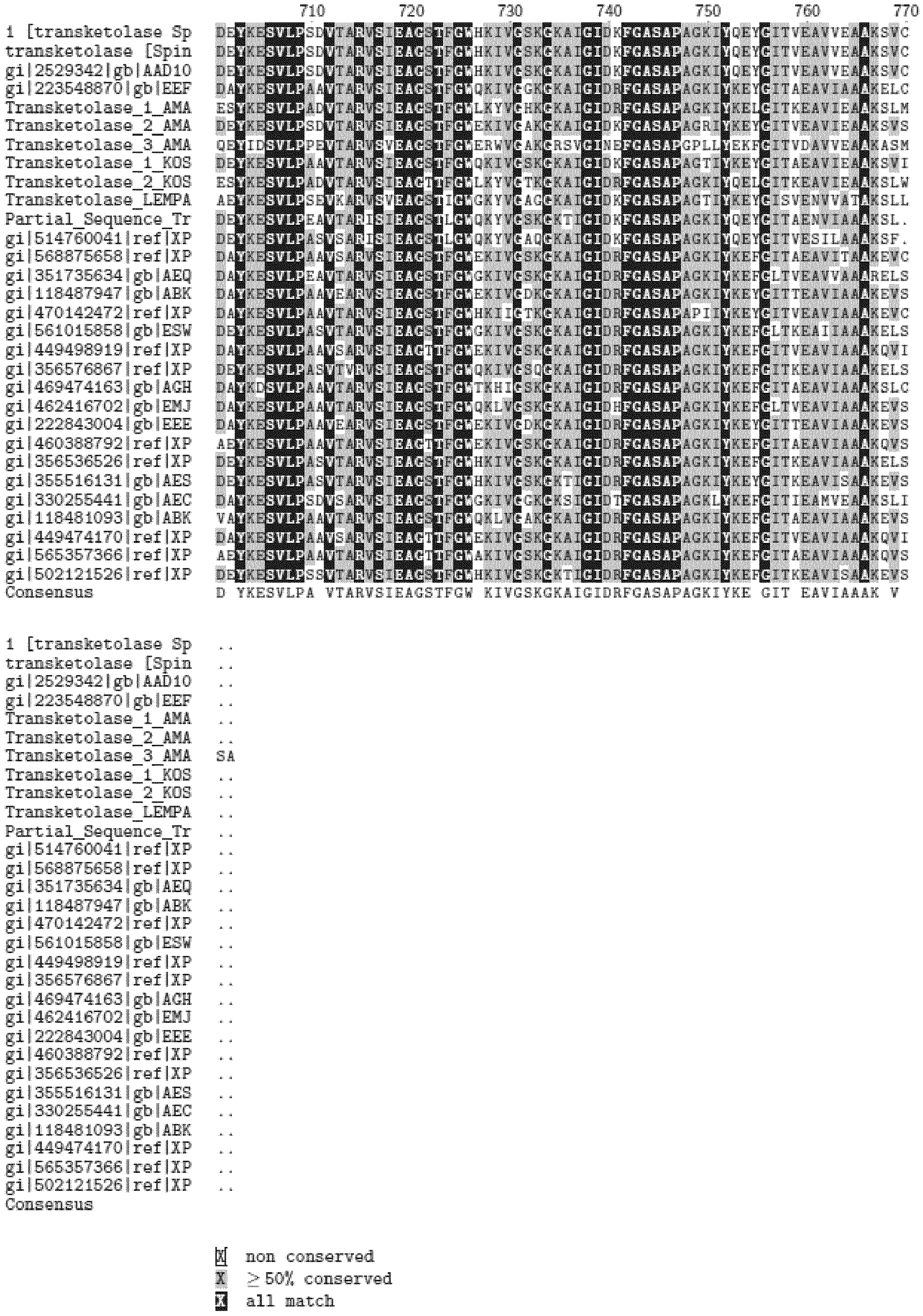
Figure 1A continued



```
                                 570       580       590       600       610       620       630
1 [transketolase Sp     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVENR........KTPSILALSRQKL
transketolase [Spin     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVENR........KTPSILALSRQKL
gi|2529342|gb|AAD10     DSIGLGEDGPTHQPIEALSKFPAMPNILMLRPADGNETAGSYKVAVENR........KTPSILALSRKKL
gi|223548870|gb|EEF     DSIGLGEDGPTHQPIEHLASFRAMPNILMFRPADGNETAGAYRIAVLNR........KRPSVLALSRQKL
Transketolase_1_AMA     DSIGLGEDGPTHQPIEHLASFRVMPNVLMLRPADGTETAGAYRIAVLKR........KTPSILALSRQKL
Transketolase_2_AMA     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVEHR........KTPSILALSRQKL
Transketolase_3_AMA     DSIGLGEDGPTHQPVEHLASFRIMPNMLTMRPGDGNETAGCYHAAVLNAQAQNSKGIRRPSTLVFSRQGM
Transketolase_1_KOS     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVENR........KTPSILALSRQKL
Transketolase_2_KOS     DSIGLGEDGPTHQPIEHLASFRAMPNVFMLRPADGTETAGAYKVAVLNR........KTPSILALSRQKL
Transketolase_LEMPA     DSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNR........KRPSILALSRQKL
Partial_Sequence_Tr     DSIGLGEDGPTHQPIEHLVSFRAMPNILMFRPADGKETAGAYKVAVLNR........KRPSILALSRQKL
gi|514760041|ref|XP     DSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNR........KRPSILALSRQKL
gi|568875658|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVANR........KRPSILALSRQKL
gi|351735634|gb|AEQ     DSIGLGEDGPTHQPIEHLASFRAMPNILMFRPADGNETAGAYKVAVVNR........KRPSIMALSRQKL
gi|118487947|gb|ABK     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNR........KRPSILALSRQKL
gi|470142472|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVENK........KRPSILALSRQKL
gi|561015858|gb|ESW     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVLNR........KRPSILALSRQKL
gi|449498919|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNQ........KRPSILALSRQKL
gi|356576867|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNTLMLRPADGNETAGSYKVAVVNR........KRPSILALSRQKL
gi|469474163|gb|ACH     DSIGLGEDGPTHQPVEHLASFRAMPNVLMLRPADGNETAGAYKVAVLNR........KRPSILALSRQKL
gi|462416702|gb|EMJ     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYRVAVLNR........KRPSILALSRQKL
gi|222843004|gb|EEE     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNR........KRPSILALSRQKL
gi|460388792|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNVLMLRPADGNETAGAYRVAILKN........KTPSILALSRQKL
gi|356536526|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNTLMLRPADGNETAGSYKVAVVNR........KRPSILALSRQKL
gi|355516131|gb|AES     DSIGLGEDGPTHQPIEHLASFRAMPNVLMLRPADGNETAGSYRVAVLNQ........KRPSILALSRQKL
gi|330255441|gb|AEC     DSIGLGEDGPTHQPIEHLSSFRAMPNIMMFRPADGNETAGAYKIAVTKR........KTPSVLALSRQKL
gi|118481093|gb|ABK     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNS........KRPSILALSRQKL
gi|449474170|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNQ........KRPSILALSRQKL
gi|565357366|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNVLMLRPADGNETAGAYRVAILKS........KTPSILALSRQKL
gi|502121526|ref|XP     DSIGLGEDGPTHQPIEHLASFRAMPNVLMLRPADGNETAGAYRVAVLNQ........KRPSILALSRQKL
Consensus               DSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNR        KRPSILALSRQKL
```

```
                                 640       650       660       670       680       690       700
1 [transketolase Sp     PNLPGTSIEGVEKGGYTITDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQS
transketolase [Spin     PNLPGTSIEGVEKGGYTITDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQS
gi|2529342|gb|AAD10     PNLPGTSIEGVEKGGYTITDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQS
gi|223548870|gb|EEF     PQLPGTSIEGVEKGGYTISDNSSGNKPDVILIGTGSELEIAAKAAEELRREGKAVRVVSFVSWELFDEQS
Transketolase_1_AMA     PQLPGTSVEGVEKGGYIISDNSSGNKPDVILMGTGSELEIAAKAADDLRKEGKAVRVVSFVSWELFESQP
Transketolase_2_AMA     PNLPGTSIQGVEKGGYTITDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSLVSWELFDQQS
Transketolase_3_AMA     PNMDTTSVEGCAKGAYIVHGGEE..KPDVIIIGTGSELVMAVDAAKKLEGEGKKVRVVSMPCWELFEEAG
Transketolase_1_KOS     PNLPGTSIEGVEKGGYTISDNSSGNKPDVILIGTGSELEIAAKAGDVLRKEGKSVRVVSFVSWELFDQQS
Transketolase_2_KOS     PHLPGTSVEGVAKGGYVISDNSKGNKPDVILMGTGSELEIAAKAADELRKEGKAVRVVSFVSWELFEDQP
Transketolase_LEMPA     PQLPGTSIDGVAKGGYTISDNSEGGKLDLILMGTGSELEIAENAAAELRKEGKAVRVVSFVSWELFEEQS
Partial_Sequence_Tr     PHLPGTSIEGVEKGGYTISDNSTGNKPDFILLSTGSELEIAVKAAEELTKEGKTVRVVSFVSWELFDEQS
gi|514760041|ref|XP     PNLPGTSIEGVEKGGYTISDNSTGNKPDLIVMGTGSELEIAAKAADELRKEGKTVRVVSFVSWELFEEQS
gi|568875658|ref|XP     PHLAGTSIDGVEKGAYIISDNSSGNKPDVILIGTGSELEIAAKAAEELRKGGKAVRVVSFVSWELFDEQS
gi|351735634|gb|AEQ     PHLPGTSIEGVEKGGYIVSDNSSGNKPDVILIGTGSELEIAFQAGEELRKEGKAVRVVSFVSWELFDDQS
gi|118487947|gb|ABK     PQLPGTSIEGVEKGGYIISDNSSSNKPDVILIGTGSELEIAAKAAEELRKEGKAVRVVSFVSWELFDEQS
gi|470142472|ref|XP     PNLAGTSVEGVATGGYIVSDNSTGNKPDVILIGTGSELEIAYKAGEELRKEGKAVRVVSFVSWELFEDQS
gi|561015858|gb|ESW     AQLPGTSIEGVEKGGYTISDNSSGNKPDVILIGTGSELEIAAAAADDLRKEGKAVRVVSFVSWELFDEQS
gi|449498919|ref|XP     PNLPGTSIEGVEKGGYILSDNSSGNKPDVILIGTGSELEIAAKAAEELRKEGKAVRLVSFVSWELFNEQS
gi|356576867|ref|XP     TQLPGTSIEGVEKGGYIISDNSSGNKPDVILIGTGSELEIAAAAAEDLRKEGKAVRVVSFVSWELFDEQS
gi|469474163|gb|ACH     SHLPGTSVEGVEKGGYTVSDNSSDNKPDVILMGTGSELEIASKAADELRKEGKAVRVVSFVSWELFDEQS
gi|462416702|gb|EMJ     PNLPGTSIEGVEKGGYIISDNSSGNKPDVILIATGSELEIVFKAGEELRKEGKAVRVVSFVSWELFDGQS
gi|222843004|gb|EEE     PQLPGTSIEGVEKGGYIISDNSSSNKPDVILIGTGSELEIAAKAAEELRKEGKAVRVVSFVSWELFDEQS
gi|460388792|ref|XP     PQLAGTSIEGTAKGGYTISDNSSGNKPDVILIGTGSELEIAVKAADELRKEGKAVRVVSFVSWELFDEQS
gi|356536526|ref|XP     TQLPGTSIEGVEKGGYTISDNSSGNKPDVILIGTGSELEIAAAAAEDLRKEGKAVRVVSFVSWELFDEQS
gi|355516131|gb|AES     PNLPGTSIEGVEKGGYTISDNSSGNKPDVILIGTGSELEIAAAAADDLRKEGKTVRVVSFVSWELFDDQS
gi|330255441|gb|AEC     PQLPGTSIESVEKGGYTISDNSTGNKPDVILIGTGSELEIAAQAAEKLREQGKSVRVVSFVCWELFDEQS
gi|118481093|gb|ABK     PQLPGTSIEGVEKGGYIISDNSSGNKPDVILMGTGSELEIAAKAAEELRKEGKAVRVVSFVSWELFDEQS
gi|449474170|ref|XP     PNLPGTSIEGVEKGGYILSDNSSGNKPDVILIGTGSELEIAAKAAEELRKEGKAVRLVSFVSWELFNEQS
gi|565357366|ref|XP     PQLAGTSIEGTAKGGYTISDNSSGNKPDVILIGTGSELEIAVKAADELRKEGKAVRVVSFVSWELFDEQS
gi|502121526|ref|XP     PNLPGTSIEGVEKGGYIISDNSSGNKPDVILIGTGSELEIAAAAADDLRKEGKAVRVVSFVSWELFDDQS
Consensus               P LPGTSIEGVEKGGY ISDNSSGNKPDVILIGTGSELEIAAKAA ELRKEGKAVRVVSFVSWELFDEQS
```

Figure 1A continued



```
                              710       720       730       740       750       760       770
1 [transketolase Sp   DEYKESVLPSDVTARVSIEAGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC
transketolase [Spin   DEYKESVLPSDVTARVSIEAGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC
gi|2529342|gb|AAD10   DEYKESVLPSDVTARVSIEAGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC
gi|223548870|gb|EEF   DAYKESVLPAAVTARVSIEAGSTFGWQKIVGGKGKAIGIDRFGASAPAGKIYKEYGITAEAVIAAAKELC
Transketolase_1_AMA   ESYKESVLPADVTARVSIEAGSTFGWLKYVGHKGKAIGIDRFGASAPAGKIYKELGITKEAVIEAAKSLM
Transketolase_2_AMA   DEYKESVLPSDVTARVSIEAGSTFGWEKIVGAKGKAIGIDKFGASAPAGRIYKEYGITVEAVIEAAKSVS
Transketolase_3_AMA   QEYIDSVLPPEVTARVSVEAGSTFGWERWVGAKGRSVGINEFGASAPGPLLYEKFGITVDAVVEAAKASM
Transketolase_1_KOS   DEYKESVLPAAVTARVSIEAGSTFGWQKIVGSKGKAIGIDKFGASAPAGTIYKEYGITAEAVIEAAKSVI
Transketolase_2_KOS   ESYKESVLPADVTARVSIEAGTTFGWLKYVGTKGKAIGIDRFGASAPAGKIYQELGITKEAVIEAAKSLW
Transketolase_LEMPA   AEYKESVLPSEVKARVSVEAGSTIGWGKYVGAGGKAIGIDKFGASAPAGTIYKEYGISVENVVATAKSLL
Partial_Sequence_Tr   DEYKESVLPEAVTARISIEAGSTLGWQKYVGSKGKTIGIDKFGASAPAGKIYQEYGITAENVIAAAKSL.
gi|514760041|ref|XP   DEYKESVLPASVSARISIEAGSTLGWQKYVGAQGKAIGIDKFGASAPAGKIYQEYGITVESILAAAKSF.
gi|568875658|ref|XP   DAYKESVLPAAVSARVSIEAGSTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITAEAVITAAKEVC
gi|351735634|gb|AEQ   DAYKESVLPEAVTARVSIEAGSTFGWGKIVGSKGKAIGIDRFGASAPAGKIYKEFGLTVEAVVAAARELS
gi|118487947|gb|ABK   DAYKESVLPAAVEARVSIEAGSTFGWEKIVGDKGKAIGIDRFGASAPAGKIYKEYGITTEAVIAAAKEVS
gi|470142472|ref|XP   DAYKESVLPAAVTARVSIEAGSTFGWHKIIGTKGKAIGIDRFGASAPAPIIYKEYGITVEAVIAAAKEVC
gi|561015858|gb|ESW   DEYKESVLPASVTARVSIEAGSTFGWGKIVGSKGKAIGIDRFGASAPAGKIYKEFGLTKEAIIAAAKELS
gi|449498919|ref|XP   DAYKESVLPAAVSARVSIEAGTTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITVEAVIAAAKQVI
gi|356576867|ref|XP   DEYKESVLPASVTVRVSIEAGSTFGWQKIVGSQGKAIGIDRFGASAPAGKIYKEFGITKEAVIAAAKELS
gi|469474163|gb|AGH   DAYKDSVLPAAVTARVSIEAGSTFGWTKHIGSKGKAIGIDRFGASAPAGKIYKEFGITAEAVIAAAKSLC
gi|462416702|gb|EMJ   DAYKESVLPAAVTARVSIEAGSTFGWQKLVGSKGKAIGIDHFGASAPAGKIYKEFGLTVEAVIAAAKEVS
gi|222843004|gb|EEE   DAYKESVLPAAVEARVSIEAGSTFGWEKIVGDKGKAIGIDRFGASAPAGKIYKEYGITTEAVIAAAKEVS
gi|460388792|ref|XP   AEYKESVLPAAVTARVSIEAGTTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITAEAVIAAAKQVS
gi|356536526|ref|XP   DEYKESVLPASVTARVSIEAGSTFGWHKIVGSKGKAIGIDRFGASAPAGKIYKEFGITKEAVIAAAKELS
gi|355516131|gb|AES   DEYKESVLPASVTARVSIEAGSTFGWHKIVGSKGKTIGIDRFGASAPAGKIYKEYGITKEAVISAAKEVS
gi|330255441|gb|AEC   DAYKESVLPSDVSARVSIEAGSTFGWGKIVGGKGKSIGIDTFGASAPAGKLYKEFGITIEAMVEAAKSLI
gi|118481093|gb|ABK   VAYKESVLPAAVTARVSIEAGSTFGWQKLVGAKGKAIGIDRFGASAPAGKIYKEFGITAEAVIAAAKEVS
gi|449474170|ref|XP   DAYKESVLPAAVSARVSIEAGTTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITVEAVIAAAKQVI
gi|565357366|ref|XP   AEYKESVLPAAVTARVSIEAGTTFGWAKIVGSKGKAIGIDRFGASAPAGKIYKEFGITAEAVIAAAKQVS
gi|502121526|ref|XP   DEYKESVLPSSVTARVSIEAGSTFGWHKIVGSKGKTIGIDRFGASAPAGKIYKEFGITKEAVISAAKEVS
Consensus             D YKESVLPA VTARVSIEAGSTFGW KIVGSKGKAIGIDRFGASAPAGKIYKE GIT EAVIAAAK V

1 [transketolase Sp   ..
transketolase [Spin   ..
gi|2529342|gb|AAD10   ..
gi|223548870|gb|EEF   ..
Transketolase_1_AMA   ..
Transketolase_2_AMA   ..
Transketolase_3_AMA   SA
Transketolase_1_KOS   ..
Transketolase_2_KOS   ..
Transketolase_LEMPA   ..
Partial_Sequence_Tr   ..
gi|514760041|ref|XP   ..
gi|568875658|ref|XP   ..
gi|351735634|gb|AEQ   ..
gi|118487947|gb|ABK   ..
gi|470142472|ref|XP   ..
gi|561015858|gb|ESW   ..
gi|449498919|ref|XP   ..
gi|356576867|ref|XP   ..
gi|469474163|gb|AGH   ..
gi|462416702|gb|EMJ   ..
gi|222843004|gb|EEE   ..
gi|460388792|ref|XP   ..
gi|356536526|ref|XP   ..
gi|355516131|gb|AES   ..
gi|330255441|gb|AEC   ..
gi|118481093|gb|ABK   ..
gi|449474170|ref|XP   ..
gi|565357366|ref|XP   ..
gi|502121526|ref|XP   ..
Consensus
```

X non conserved

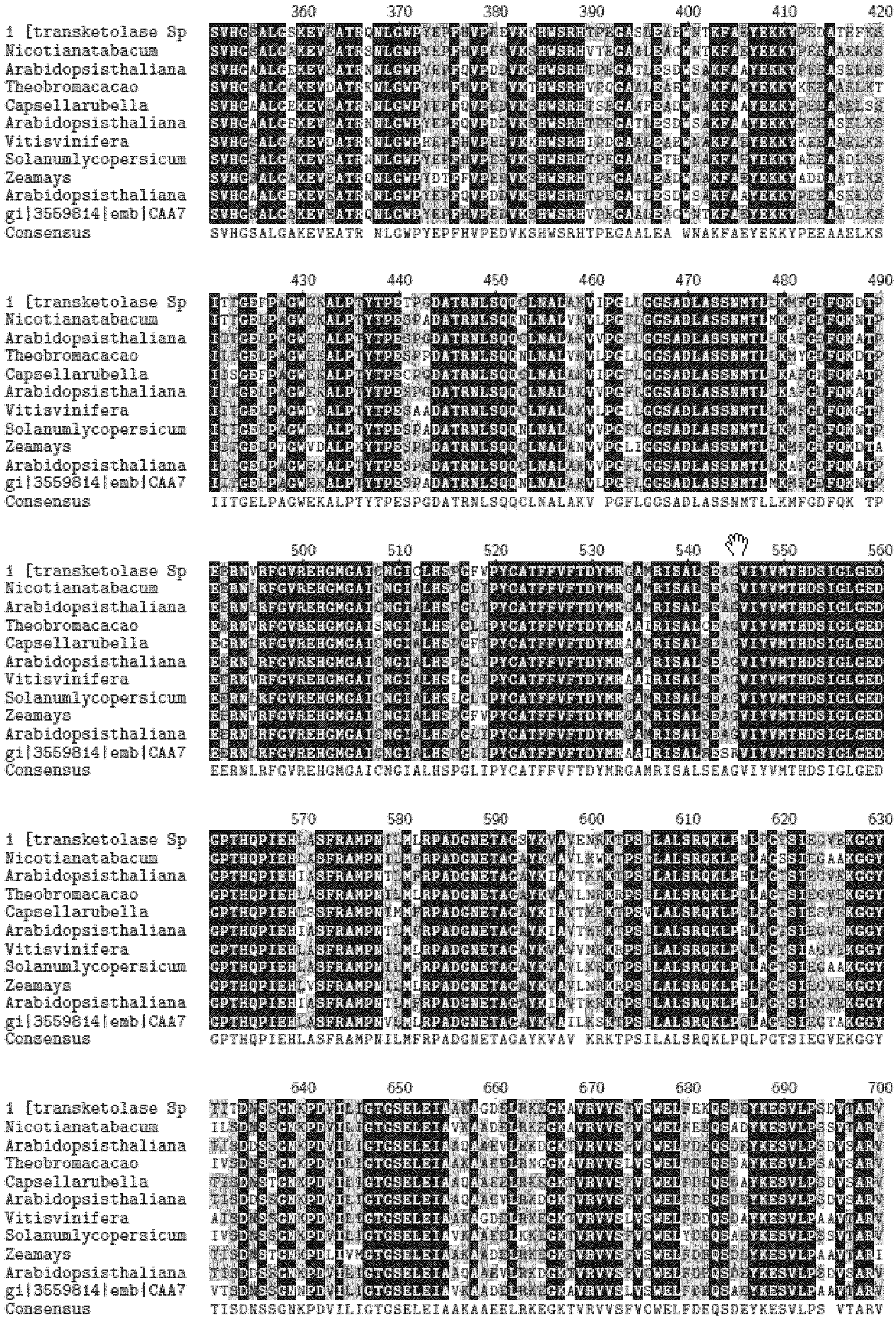
X ≥ 50% conserved

X all match

Figure 1B

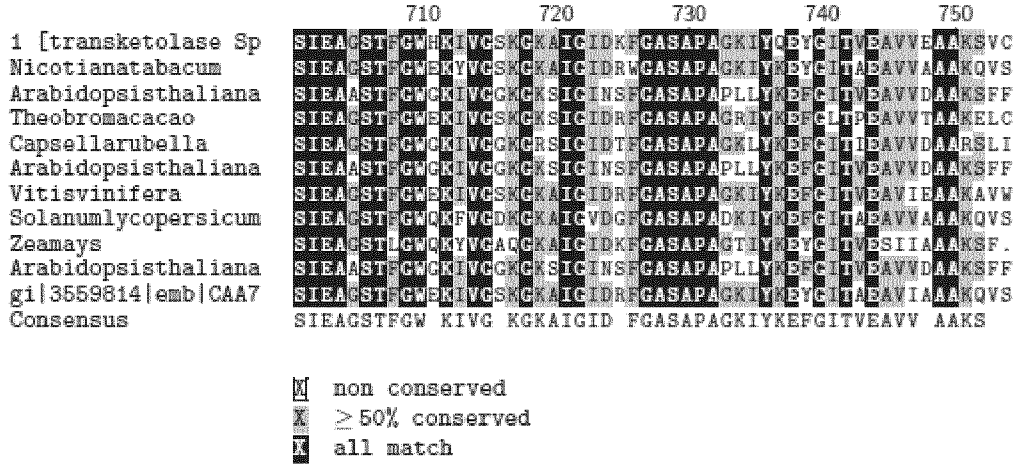
```
                               10        20        30        40        50        60        70
i [transketolase Sp     MAASSSLSTLSHHQTLLSHPKTHLPTTPAS......SLLVPTTSSKVNGVLLKSTSSSRRLR.....VGS
Nicotianatabacum        .MASSSSLTLS..QAILSRSVPRHGSASS..SQLSPSSL...TFSGLKSNPNITTSRRRTPSSAAAAAVV
Arabidopsisthaliana     .MASTSSLALS..QALLARAISHHGSDQRG.........SLPAFSGLKSTGSRASASSRRRIAQSMTKNR
Theobromacacao          .MASTSSTTLS..QALLARAISYHGSTQSSDHRVSLSTLSLPTFSGLKSTTSRASAFRRRPP.....VRS
Capsellarubella         .MASTSSLALS..QALLTRTISRDGSGNRF.........SVPAFSGLKSTFPRTISTRRGCVSTSSHSFR
Arabidopsisthaliana     .MASTSSLALS..QALLARAISHHGSDQRG.........SLPAFSGLKSTGSRASASSRRRIAQSMTKNR
Vitisvinifera           MAAASASSSLY....LLPRTLNHHSSTPHP......NRLSLSG.......LTPKTQALRTRH.....GRR
Solanumlycopersicum     .MASSSSLTLS..QAIFSPSLPRHGSSSS.....SSPSISFSTFSGLKSTPFTSSHRRILPS...TTVTK
Zeamays                 ......................................................................
Arabidopsisthaliana     .MASTSSLALS..QALLARAISHHGSDQRG.........SLPAFSGLKSTGSRASASSRRRIAQSMTKNR
gi|3559814|emb|CAA7     .MASSS.LPLS..QAILSRSIPRHGSSSSTNSQFSPSSLSLPTFSGLKSTTSTTFRRRTLPS....PVAV
Consensus                MAS SSL LS  QALL R I  HGS             SLP FSGLKST        R

                               80        90       100       110       120       130       140
i [transketolase Sp     AS.AVVRAAAVEALESTDID.QLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEIMRYNPK
Nicotianatabacum        RSPAIRASAATETIEKTETA..LVDKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPK
Arabidopsisthaliana     SLRPLVRAAAVETVEPTTDS.SIVDKSVNSIRFLAIDAVEKAKSGHPGLPMGCAPMAHILYDEVMRYNPK
Theobromacacao          YQNRQVRAAAVETIGTAAET.SLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPK
Capsellarubella         ..PDLVRAAAVEAIETTTDS.SLVDKSVNTIRFLAIDAVEKAKSGHPGLPMGCAPMSHILYDEFMRYNPK
Arabidopsisthaliana     SLRPLVRAAAVETVEPTTDS.SIVDKSVNSIRFLAIDAVEKAKSGHPGLPMGCAPMAHILYDEVMRYNPK
Vitisvinifera           AR.ATVRAAAVETLQKADTS..LVEKSINTIRFLSIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPS
Solanumlycopersicum     QHFSVRASSAVETLEKTDAA..IVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMKYNPK
Zeamays                 ........GAVETLQGKAATGELLEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHVLYDEVMRYNPK
Arabidopsisthaliana     SLRPLVRAAAVETVEPTTDS.SIVDKSVNSIRFLAIDAVEKAKSGHPGLPMGCAPMAHILYDEVMRYNPK
gi|3559814|emb|CAA7     RSPEIRASAAVETLEKTDNA..LVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPIGHILYDEIMRYNPK
Consensus                    VRAAAVET E T     LVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPK

                              150       160       170       180       190       200       210
i [transketolase Sp     NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEVTTGPLGQG
Nicotianatabacum        NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVREEDLKSFRQWGSKTPGHPENFETPGVEVTTGPLGQG
Arabidopsisthaliana     NPYWFNRDRFVLSAGHGCMLLYALLHLAGYDSVQEEDLKQFRQWGSKTPGHPENFETPGIEVTTGPLGQG
Theobromacacao          NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKNFRQWGSKTPGHPENFETLGVEVTTGPLGQG
Capsellarubella         NPYWFNRDRFVLSAGHGCMLHYALLHLAGYDSVREEDLKSFRQWGSKTPGHPENFETPGVEVTTGPLGQG
Arabidopsisthaliana     NPYWFNRDRFVLSAGHGCMLLYALLHLAGYDSVQEEDLKQFRQWGSKTPGHPENFETPGIEVTTGPLGQG
Vitisvinifera           NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSRTPGHPENFETPGVEVTTGPLGQG
Solanumlycopersicum     NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVQEDDLKSFRQWGSKIPGHPENFETPGVEVTTGPLGQG
Zeamays                 NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKQFRQWGSRTPGHPENFETPGVEVTTGPLGQG
Arabidopsisthaliana     NPYWFNRDRFVLSAGHGCMLLYALLHLAGYDSVQ.EDLKQFRQWGSKTPGHPENFETPGIEVTTGPLGQG
gi|3559814|emb|CAA7     NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVREEDLKSFRQWGSKTPGHPENFETPGVEVTTGPLGQG
Consensus               NPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSV EEDLK FRQWGSKTPGHPENFETPGVEVTTGPLGQG

                              220       230       240       250       260       270       280
i [transketolase Sp     IANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAFYDDNHISI
Nicotianatabacum        IANAVGLALVEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGISQEACSLAGHWGLGKLIAFYDDNHISI
Arabidopsisthaliana     IANAVGLALAEKHLAARFNKPDAEVVDHYTYAILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISI
Theobromacacao          VANAVGLALAEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISI
Capsellarubella         IANAVGLALAEKHLAARFNKPDNEIVDHYTYSILGDGCQMEGISNEVCSLAAHWGLGKLIAFYDDNHISI
Arabidopsisthaliana     IANAVGLALAEKHLAARFNKPDAEVVDHYTYAILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISI
Vitisvinifera           IANAVGLAAVEKHLAARFNKPDNEIVDHYTYVILGDGCQMEGISNEVCSLAGHWGLGKLIAFYDDNHISI
Solanumlycopersicum     IANAVGLAVAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISI
Zeamays                 IANAVGLALAEKHLAARFNKPDSEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISI
Arabidopsisthaliana     IANAVGLALAEKHLAARFNKPDAEVVDHYTYAILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISI
gi|3559814|emb|CAA7     IANAVGLALSEKHLAARFNKPGAEIVDHYTYCILGDGCQMEGISQEAASLAGHWGLGKLIAFYDDNHISI
Consensus               IANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISI

                              290       300       310       320       330       340       350
i [transketolase Sp     DGDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFGSPNKSNSY
Nicotianatabacum        DGDTEIAFTEDVGARFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTMIKVTTTIGFGSPNKANSY
Arabidopsisthaliana     DGDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGYGSPNKANSY
Theobromacacao          DGDTEIAFTESVDKRFEGLGWHVIWVKNGNTGYDDIRAAIKEAKAVKDKPTLIKLTTTIGYGSPNKANSY
Capsellarubella         DGDTDIAFTESVDKRFEALGWHVIWVKNGNNGYDEIRAAIREAKAVKDKPTLIKVTTTIGYGSPNKANSY
Arabidopsisthaliana     DGDTEIAFTENVDQLFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGYGSPNKANSY
Vitisvinifera           DGDTEIAFTESVDTRFEGLGWHVIWVKNGNTGYDEIRAAIKEAKAVKDKPTLIKVTTTIGYGSPNKANSY
Solanumlycopersicum     DGDTEIAFTEDVSARFEALGWHVIWVKNGNTGYDEIRAAIKEAKSVKDKPTMIKVTTTIGFGSPNKANSY
Zeamays                 DGDTEIAFTEDVSTRFEALGWHTIWVKNGNTGYDDIRAAIKEAKAVTDKPTLIKVTTTIGFGSPNKANSY
Arabidopsisthaliana     DGDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGYGSPNKANSY
gi|3559814|emb|CAA7     DGDTEIAFTEDVGARFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVTDKPTMIKVTTTIGFGSPNKANSY
Consensus               DGDTEIAFTE VD RFEALGWHVIWVKNGNTGYDEIRAAIKEAK VTDKPTLIKVTTTIGYGSPNKANSY
```

Figure 1B continued



```
                                   360       370       380       390       400       410       420
1 [transketolase Sp        SVHGSALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYPEDATEFKS
Nicotianatabacum           SVHGSALGAKEVEATRSNLGWPYEPFHVPEDVKSHWSRHVTEGAALEAGWNTKFAEYEKKYPEEAAELKS
Arabidopsisthaliana        SVHGAALGEKEVEATRNNLGWPYEPFQVPDDVKSHWSRHTPEGATLESDWSAKFAAYEKKYPEEASELKS
Theobromacacao             SVHGSALGAKEVDATRKNLGWPYEPFHVPEDVKTHWSRHVPQGAALEAEWNAKFAEYEKKYKEEAAELKT
Capsellarubella            SVHGAALGEKEVEATRNNLGWPYEPFQVPEDVKSHWSRHTSEGAAFEADWNAKFAAYEKKYPEEAAELSS
Arabidopsisthaliana        SVHGAALGEKEVEATRNNLGWPYEPFQVPDDVKSHWSRHTPEGATLESDWSAKFAAYEKKYPEEASELKS
Vitisvinifera              SVHGSALGAKEVDATRKNLGWPHEPFHVPEDVKKHWSRHIPDGAALEAEWNAKFAEYEKKYKEEAAELKS
Solanumlycopersicum        SVHGSALGAKEVEATRNNLGWPYEPFHVPEDVKSHWSRHTPEGAALETEWNAKFAEYEKKYAEEAADLKS
Zeamays                    SVHGSALGAKEVEATRQNLGWPYDTFFVPEDVKSHWSRHTPEGAALEADWNAKFAEYEKKYADDAATLKS
Arabidopsisthaliana        SVHGAALGEKEVEATRNNLGWPYEPFQVPDDVKSHWSRHTPEGATLESDWSAKFAAYEKKYPEEASELKS
gi|3559814|emb|CAA7        SVHGSALGAKEVEATRQNLGWPYEPFHVPEDVKSHWSRHVPEGAALEAGWNTKFAEYEKKYPEEAADLKS
Consensus                  SVHGSALGAKEVEATR NLGWPYEPFHVPEDVKSHWSRHTPEGAALEA WNAKFAEYEKKYPEEAAELKS

                                   430       440       450       460       470       480       490
1 [transketolase Sp        ITTGEFPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMFGDFQKDTP
Nicotianatabacum           ITTGELPAGWEKALPTYTPESPADATRNLSQQNLNALVKVLPGFLGGSADLASSNMTLMKMFGDFQKNTP
Arabidopsisthaliana        IITGELPAGWEKALPTYTPESPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTLLKAFGDFQKATP
Theobromacacao             IITGELPAGWEKALPTYTPESPPDATRNLSQQNLNALVKVLPGLLGGSADLASSNMTLLKMYGDFQKDTP
Capsellarubella            IISGEFPAGWEKALPTYTPECPGDATRNLSQQCLNALAKVIPGFLGGSADLASSNMTLLKAFGNFQKATP
Arabidopsisthaliana        IITGELPAGWEKALPTYTPESPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTLLKAFGDFQKATP
Vitisvinifera              IITGELPAGWDKALPTYTPESAADATRNLSQQCLNALAKVLPGLLGGSADLASSNMTLLKMFGDFQKGTP
Solanumlycopersicum        IITGELPAGWEKALPTYTPESPADATRNLSQQNLNALAKVVPGFLGGSADLASSNMTLLKMFGDFQKNTP
Zeamays                    IITGELPTGWVDALPKYTPESPGDATRNLSQQCLNALANVVPGLIGGSADLASSNMTLLKMFGDFQKDTA
Arabidopsisthaliana        IITGELPAGWEKALPTYTPESPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTLLKAFGDFQKATP
gi|3559814|emb|CAA7        IITGELPAGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTLMKMFGDFQKNTP
Consensus                  IITGELPAGWEKALPTYTPESPGDATRNLSQQCLNALAKV PGFLGGSADLASSNMTLLKMFGDFQK TP

                                   500       510       520       530       540       550       560
1 [transketolase Sp        EERNVRFGVREHGMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
Nicotianatabacum           EERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
Arabidopsisthaliana        EERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
Theobromacacao             EERNVRFGVREHGMGAISNGIALHSPGLIPYCATFFVFTDYMRAAIRISALCEAGVIYVMTHDSIGLGED
Capsellarubella            EGRNLRFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTHDSIGLGED
Arabidopsisthaliana        EERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
Vitisvinifera              EERNVRFGVREHGMGAICNGIALHSLGLIPYCATFFVFTDYMRAAIRISALSEAGVIYVMTHDSIGLGED
Solanumlycopersicum        EERNLRFGVREHGMGAICNGIALHSLGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
Zeamays                    EERNVRFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
Arabidopsisthaliana        EERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED
gi|3559814|emb|CAA7        EERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAIRISALSESRVIYVMTHDSIGLGED
Consensus                  EERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGED

                                   570       580       590       600       610       620       630
1 [transketolase Sp        GPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVENRKTPSILALSRQKLPNLPGTSIEGVEKGGY
Nicotianatabacum           GPTHQPIEHLASFRAMPNILMFRPADGNETAGAYKVAVLKWKTPSILALSRQKLPQLAGSSIEGAAKGGY
Arabidopsisthaliana        GPTHQPIEHIASFRAMPNTLMFRPADGNETAGAYKIAVTKRKTPSILALSRQKLPHLPGTSIEGVEKGGY
Theobromacacao             GPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPQLAGTSIEGVEKGGY
Capsellarubella            GPTHQPIEHLSSFRAMPNIMMFRPADGNETAGAYKIAVTKRKTPSVLALSRQKLPQLPGTSIESVEKGGY
Arabidopsisthaliana        GPTHQPIEHIASFRAMPNTLMFRPADGNETAGAYKIAVTKRKTPSILALSRQKLPHLPGTSIEGVEKGGY
Vitisvinifera              GPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVVNRKRPSILALSRQKLPQLPGTSIAGVEKGGY
Solanumlycopersicum        GPTHQPIEHLASFRAMPNILMFRPADGNETAGAYKVAVLKRKTPSILALSRQKLPQLAGTSIEGAAKGGY
Zeamays                    GPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLPGTSIEGVEKGGY
Arabidopsisthaliana        GPTHQPIEHIASFRAMPNTLMFRPADGNETAGAYKIAVTKRKTPSILALSRQKLPHLPGTSIEGVEKGGY
gi|3559814|emb|CAA7        GPTHQPIEHLASFRAMPNVLMLRPADGNETAGAYKVAILKSKTPSILALSRQKLPQLAGTSIEGTAKGGY
Consensus                  GPTHQPIEHLASFRAMPNILMFRPADGNETAGAYKVAV KRKTPSILALSRQKLPQLPGTSIEGVEKGGY

                                   640       650       660       670       680       690       700
1 [transketolase Sp        TITDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQSDEYKESVLPSDVTARV
Nicotianatabacum           ILSDNSSGNKPDVILIGTGSELEIAVKAADELRKEGKAVRVVSFVCWELFEEQSADYKESVLPSSVTARV
Arabidopsisthaliana        TISDDSSGNKPDVILIGTGSELEIAAQAAEVLRKDGKTVRVVSFVCWELFDEQSDEYKESVLPSDVSARV
Theobromacacao             IVSDNSSGNKPDVILIGTGSELEIAAKAAEELRNGGKAVRVVSLVSWELFDEQSDAYKESVLPSAVSARV
Capsellarubella            TISDNSTGNKPDVILIGTGSELEIAAQAAEELRKEGKTVRVVSFVCWELFDEQSDAYKESVLPSDVSARV
Arabidopsisthaliana        TISDDSSGNKPDVILIGTGSELEIAAQAAEVLRKDGKTVRVVSFVCWELFDEQSDEYKESVLPSDVSARV
Vitisvinifera              AISDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKTVRVVSLVSWELFDDQSDAYKESVLPAAVTARV
Solanumlycopersicum        IVSDNSSGNKPDVILIGTGSELEIAVKAAEELKKEGKTVRVVSFVCWELYDEQSAEYKESVLPSSVTARV
Zeamays                    TISDNSTGNKPDLIVMGTGSELEIAAKAADELRKEGKTVRVVSFVSWELFDEQSDEYKESVLPAAVTARI
Arabidopsisthaliana        TISDDSSGNKPDVILIGTGSELEIAAQAAEVLRKDGKTVRVVSFVCWELFDEQSDEYKESVLPSDVSARV
gi|3559814|emb|CAA7        VTSDNSSGNNPDVILIGTGSELEIAVKAADELRKEGKAVRVVSLVSWELFDEQSAEYKESVLPAAVTARV
Consensus                  TISDNSSGNKPDVILIGTGSELEIAAKAAEELRKEGKTVRVVSFVCWELFDEQSDEYKESVLPS VTARV
```

Figure 1B continued

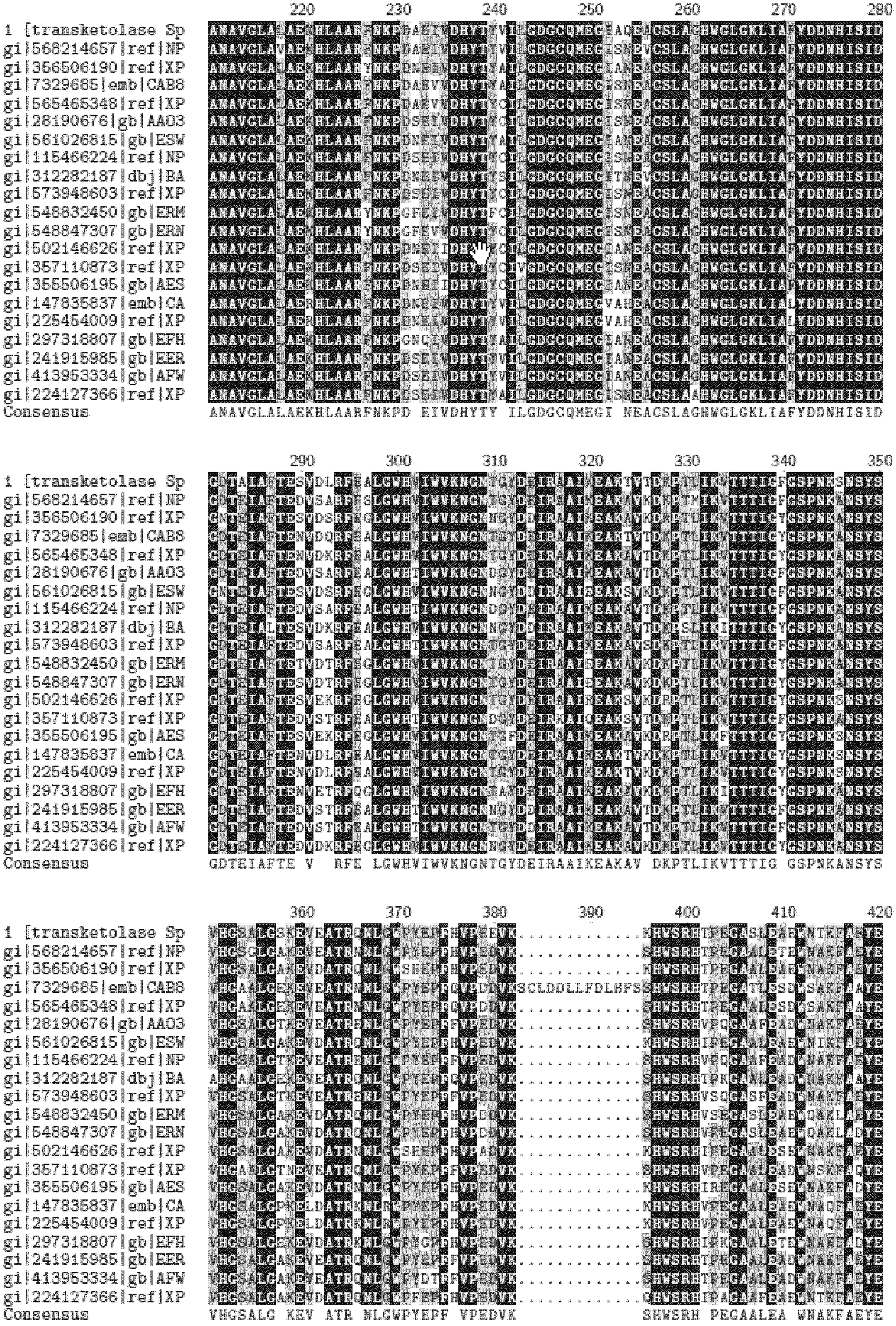


```
                                          710       720       730       740       750
1 [transketolase Sp    SIEAGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC
Nicotianatabacum       SIEAGSTFGWEKYVGSKGKAIGIDRWGASAPAGKIYKEYGITAEAVVAAAKQVS
Arabidopsisthaliana    SIEAASTFGWGKIVGGKGKSIGINSFGASAPAPLLYKEFGITVEAVVDAAKSFF
Theobromacacao         SIEAGSTFGWEKIVGSKGKSIGIDRFGASAPAGRIYKEFGLTPEAVVTAAKELC
Capsellarubella        SIEAGSTFGWGKIVGGKGRSIGIDTFGASAPAGKLYKEFGITIEAVVDAARSLI
Arabidopsisthaliana    SIEAASTFGWGKIVGGKGKSIGINSFGASAPAPLLYKEFGITVEAVVDAAKSFF
Vitisvinifera          SIEAGSTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITVEAVIEAAKAVW
Solanumlycopersicum    SIEAGSTFGWQKFVGDKGKAIGVDGFGASAPADKIYKEFGITAEAVVAAAKQVS
Zeamays                SIEAGSTLGWQKYVGAQGKAIGIDKFGASAPAGTIYKEYGITVESIIAAAKSF.
Arabidopsisthaliana    SIEAASTFGWGKIVGGKGKSIGINSFGASAPAPLLYKEFGITVEAVVDAAKSFF
gi|3559814|emb|CAA7    SIEAGSTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEYGITAEAVIAAAKQVS
Consensus              SIEAGSTFGW KIVG KGKAIGID FGASAPAGKIYKEFGITVEAVV AAKS

                       X  non conserved
                       X  ≥50% conserved
                       X  all match
```

Figure 1C

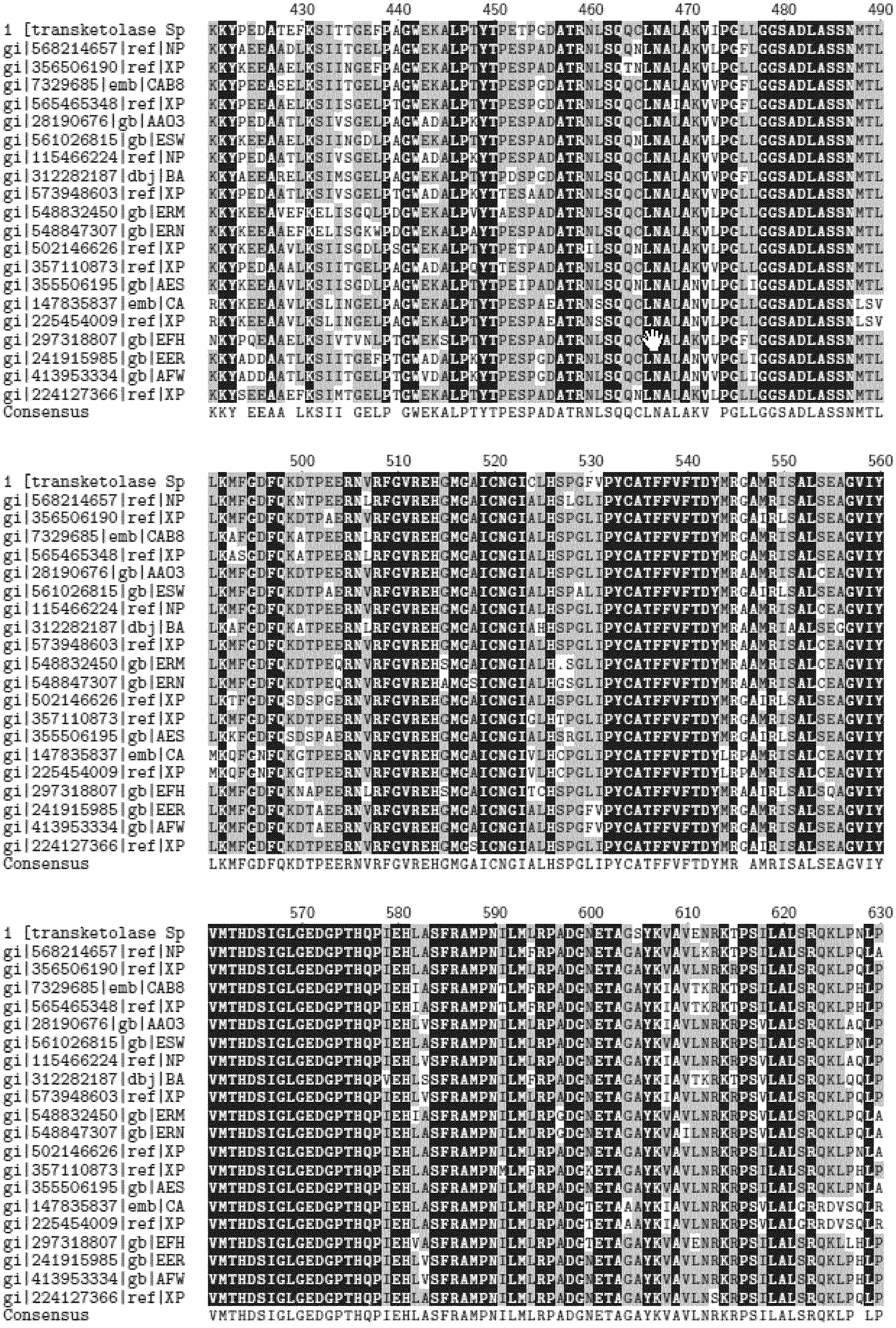
```
                              10        20        30        40        50        60        70
1 [transketolase Sp   ....MAASSSLSTLSHHQTLLSHPKTHLPTTPASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSAS.....
gi|568214657|ref|NP   ....MASSSS..LTLSQVIFSPSLPRHGSSSSSSPSLSFSTFSGLKST...PFTSSHRRILPSTTVTKQQ
gi|356506190|ref|XP   .....MASSTSSISLS.QALLRPTTTSLPSSSSS..........LRLTPTLR.PTPP..ALSLRR...RL
gi|7329685|emb|CAB8   ....MASTSS..LALSQALLARAISHHGSDQRGS....LPAFSGLKSTGSRASASSRRRIAQSMTKNRSL
gi|565465348|ref|XP   ....MASTSS..LALSQALLARAISHHGSDQRGS....LPAFTGLKSTGSRASASSRRRIAQSMNKNRSL
gi|28190676|gb|AA03   ....MAAHS...VAAAHATIAARAGAAAPAP.APPERLGFRLSALAGRGLRSPLPPRRGAPSASASRRRH
gi|561026815|gb|ESW   .....MASS.SSLSLS.QALLRPTPLSS.SHCPS..........LRVSPSLR.PAPPTAPLSLRR...RA
gi|115466224|ref|NP   ......................................................................
gi|312282187|dbj|BA   ....MASTSSS.LAISQALLGRAISQNGSDKCVS....IPAFSGLKSTSPRTTFSSRR.IAS..TNSHSL
gi|573948603|ref|XP   ......................................................................
gi|548832450|gb|ERM   MASSSLGMYQAPLAGAPAAAERRSSSSATLPLPLPSKPNTPAILLNNNFLRGTKVVLSRRSRARKPSTSM
gi|548847307|gb|ERN   .........................MKAVLQTTISN........................SRGRRP.WSS
gi|502146626|ref|XP   .....MSSSTSSLSLTNPHLTRLTS...TTRLSS..........LPSTISLRSPTTPSTHLSLSHTHLRR
gi|357110873|ref|XP   ....MAAHS...VAAAHATMAAPAGAASSACSAPAERLGFRLSSLAGRGLR..LPSRPSAASSSSSRRTN
gi|355506195|gb|AES   ....MSSSSSTSLYLTNPHLTRHTSSPSTTRLSS..........LPSTITLNRTPTPTTHLSQSH.HLRH
gi|147835837|emb|CA   ....MASSSS..LSISQAIVAGALNHRGSTNSSD..RLPLRSFTIPTFSGLKSTTSAASCPSRLGSARRR
gi|225454009|ref|XP   ....MASSSS..LSISQAIVAGALNHRGSTNSSD..RLPLRSFTIPTFSGLKSTTSAASCPSRLGSARRR
gi|297318807|gb|EFH   .....MASLTSSSQAS..ILARPSSNRSSPPLPP..........TDLHVSLSLPSFSALRASFTR..SSS
gi|241915985|gb|EER   ......................................................................
gi|413953334|gb|AFW   ......................................................................
gi|224127366|ref|XP   ......................................................................
Consensus                      S

                              80        90       100       110       120       130       140
1 [transketolase Sp   ..AVVRAAAVEALES.TDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEIMRYNPKN
gi|568214657|ref|NP   F.SVRASAAVET.LEKTDAA.IVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMKYNPKN
gi|356506190|ref|XP   TTSIRAAASVKTVEKATADAALVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHVLYDETMKYNPKN
gi|7329685|emb|CAB8   R.PLVRAAAVETVEPTTDSS.IVDKSVNSIRFLAIDAVEKAKSGHPGLPMGCAPMAHILYDEVMRYNPKN
gi|565465348|ref|XP   R.PLVRAAAVETVEPTTDSS.IVDKSVNSIRFLAIDAVEKAKSGHPGLPMGCAPMAHILYDEVMRYNPKN
gi|28190676|gb|AA03   NNRVRAAA.VETLEGQAATGALLEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN
gi|561026815|gb|ESW   AP.VRASVPLKTVEKATSDIALVEKSVNTIRFLAIDSVEKANSGHPGLPMGCAPMGHVLYDEAMKYNPKN
gi|115466224|ref|NP   ...VRAAA.VETLEGQAATGALLEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN
gi|312282187|dbj|BA   R.PLVRAAAVET...KTESS.LVEKSVNTIRFLAIDAVEKAKSGHPGLPMGCAPMSHILFDEVMRYNPKN
gi|573948603|ref|XP   .......A.VETLEGQAATGALLEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN
gi|548832450|gb|ERM   QAVVRASAAVETLD.QTSTTELIEKSINTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN
gi|548847307|gb|ERN   RVVVRA.AAVETLD.VTSTIALIEKSINTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN
gi|502146626|ref|XP   TSVKSSAVVTQTSENSTADLSLVEKSVNTIRFLAIDSVEKANSGHPGLPMGCAPMGHVIYDEVLRYNPKN
gi|357110873|ref|XP   ..RVRAAASVETVQGQAATGALLDKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN
gi|355506195|gb|AES   TAIHASVSAPPS...TTTDSSLVEKSVNTIRFLAVDSVEKANSGHPGLPMGCAPMGHVLYDEVMRYNPKN
gi|147835837|emb|CA   LCRRLAVQAASVETLEKTETTLIEKSVNTIRFLAVDSVEKANSGHPGLPMGCAPMGHVLYDEFMKFNPKN
gi|225454009|ref|XP   LCRRLAVQAASVETLEKTETTLIEKSVNTIRFLAVDSVEKANSGHPGLPMGCAPMGHVLYDEFMKFNPKN
gi|297318807|gb|EFH   PTLRRRRRLLLPISAAATDVSLVDKSVNTIRFLAIDAVEKAKSGHPGLPMGCAPMGHILFDEVMQYNPKN
gi|241915985|gb|EER   .................................................MGCAPMGHILYDEVMRYNPKN
gi|413953334|gb|AFW   .................................................MGCAPMGHVLYDEVMRYNPKN
gi|224127366|ref|XP   .................................................MGCAPMGHILYDEVMRYNPKN
Consensus                        T         L EKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEVMRYNPKN

                             150       160       170       180       190       200       210
1 [transketolase Sp   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEVTTGPLGQGI
gi|568214657|ref|NP   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVQEDDLKSFRQWGSRIPGHPENFETPGVEVTTGPLGQGI
gi|356506190|ref|XP   PFWFNRDRFVLSAGHGCMLQYALLHLAGFDSVKEEDLREFRQWGSRTPGHPENFETPGIEVTTGPLGQGI
gi|7329685|emb|CAB8   PYWFNRDRFVLSAGHGCMLLYALLHLAGYDSVQEEDLKQFRQWGSKTPGHPENFETPGIEVTTGPLGQGI
gi|565465348|ref|XP   PYWFNRDRFVLSAGHGCMLLYSLLHLAGYDSVQEEDLKNFRQWGSKTPGHPENFETPGIEVTTGPLGQGI
gi|28190676|gb|AA03   PYWFNRDRFILSAGHGCMLQYALLHLAGYDAVLEEDLKQFRQWGSKTPGHPENFETPGVEVTTGPLGQGI
gi|561026815|gb|ESW   PKWFNRDRFVLSAGHGCMLQYALLHLAGFDSVKEEDLRQFRQWGSRTPGHPENFETPGVEVTTGPLGQGI
gi|115466224|ref|NP   PYWFNRDRFILSAGHGCMLQYALLHLAGYDAVLEEDLKQFRQWGSKTPGHPENFETPGVEVTTGPLGQGI
gi|312282187|dbj|BA   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVREEDLKSFRQWRSKTPGHPENFETPGVEVTTGPLGQGI
gi|573948603|ref|XP   PYWFNRDRFILSAGHGCMLQYALLHLAGYDAVMEDDLKQFRQWGSRTPGHPENFETPGVEVTTGPLGQGI
gi|548832450|gb|ERM   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSKTPGHPENFETPGIEVTTGPLGQGV
gi|548847307|gb|ERN   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSKTPGHPENFETRGIEVTTGPLGQGI
gi|502146626|ref|XP   PKWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKAFRQWESKTPGHPENFETPGVEVTTGPLGQGI
gi|357110873|ref|XP   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVKEADLKQFRQWGSSTPGHPENFETPGVEVTTGPLGQGI
gi|355506195|gb|AES   PFWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKVEDLKQFRQWESRTPGHPENFETPGIEVTTGPLGQGI
gi|147835837|emb|CA   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVREEDLKSFRQWGSTTPGHPENFETPGVEVTTGPLGQGI
gi|225454009|ref|XP   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVREEDLKSFRQWGSTTPGHPENFETPGVEVTTGPLGQGI
gi|297318807|gb|EFH   PYWFNRDRFILSAGHGCMLQYALLHLAGYDSVREDDLKNFRQWGSRTPGHPENFETPGVEVTTGPLGQGI
gi|241915985|gb|EER   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKQFRQWGSKTPGHPENFETPGVEVTTGPLGQGI
gi|413953334|gb|AFW   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKQFRQWGSRTPGHPENFETPGVEVTTGPLGQGI
gi|224127366|ref|XP   PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLKSFRQWGSRTPGHPENFETPGVEVTTGPLGQGV
Consensus             PYWFNRDRFVLSAGHGCMLQYALLHLAGYDSV EEDLK FRQWGS TPGHPENFETPGVEVTTGPLGQGI
```

Figure 1C continued



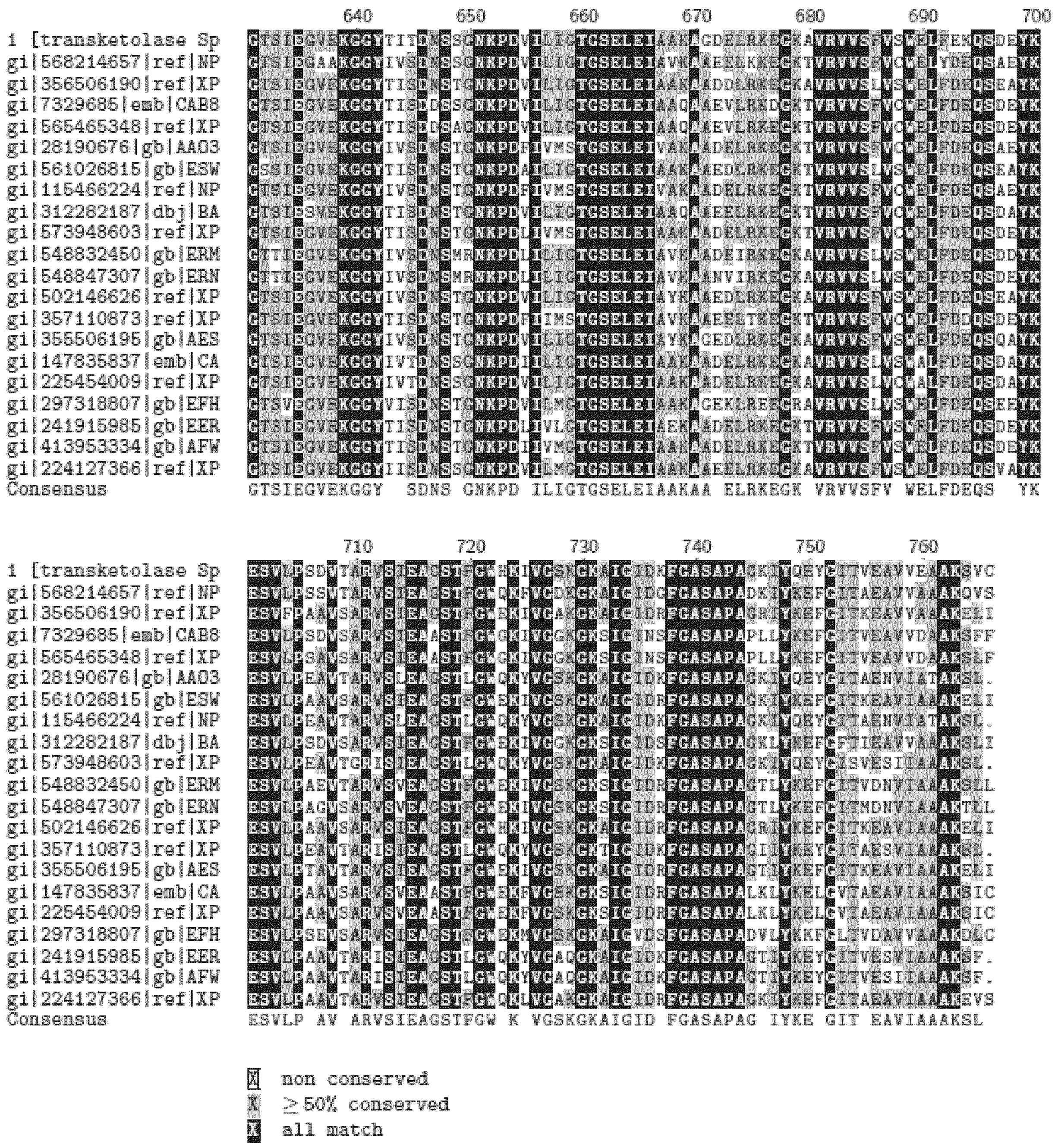
```
                               220       230       240       250       260       270       280
1 [transketolase Sp     ANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAFYDDNHISID
gi|568214657|ref|NP     ANAVGLAVAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGISNEVCSLAGHWGLGKLIAFYDDNHISID
gi|356506190|ref|XP     ANAVGLALAEKHLAARYNKPDNEIVDHYTYAILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|7329685|emb|CAB8     ANAVGLALAEKHLAARFNKPDAEVVDHYTYAILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|565465348|ref|XP     ANAVGLALAEKHLAARFNKPDAEVVDHYTYCILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|28190676|gb|AA03     ANAVGLALAEKHLAARFNKPDSEIVDHYTYCILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|561026815|gb|ESW     ANAVGLALAEKHLAARFNKPDNEIVDHYTYAILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|115466224|ref|NP     ANAVGLALAEKHLAARFNKPDSEIVDHYTYCILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|312282187|dbj|BA     ANAVGLALAEKHLAARFNKPDNEIVDHYTYSILGDGCQMEGITNEVCSLAGHWGLGKLIAFYDDNHISID
gi|573948603|ref|XP     ANAVGLALAEKHLAARFNKPDSEIVDHYTYCILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|548832450|gb|ERM     ANAVGLALAEKHLAARYNKPGFEIVDHYTFCILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|548847307|gb|ERN     ANAVGLALAEKHLAARYNKPGFEVVDHYTYCILGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|502146626|ref|XP     ANAVGLALAEKHLAARFNKPDNEIIDH[illegible]YCILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|357110873|ref|XP     ANAVGLALAEKHLAARFNKPDSEIVDHYTYCIVGDGCQMEGISNEACSLAGHWGLGKLIAFYDDNHISID
gi|355506195|gb|AES     ANAVGLALAEKHLAARFNKPDNEIIDHYTYCILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|147835837|emb|CA     ANAVGLALAERHLAARFNKPDNEIVDHYTYVILGDGCQMEGVAHEACSLAGHWGLGKLIALYDDNHISID
gi|225454009|ref|XP     ANAVGLALAERHLAARFNKPDNEIVDHYTYVILGDGCQMEGVAHEACSLAGHWGLGKLIALYDDNHISID
gi|297318807|gb|EFH     ANAVGLALAEKHLAARFNKPGNQIVDHYTYAILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|241915985|gb|EER     ANAVGLALAEKHLAARFNKPDSEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|413953334|gb|AFW     ANAVGLALAEKHLAARFNKPDSEIVDHYTYVILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISID
gi|224127366|ref|XP     ANAVGLALAEKHLAARFNKPDSEIVDHYTYAILGDGCQMEGISNEACSLAAHWGLGKLIAFYDDNHISID
Consensus               ANAVGLALAEKHLAARFNKPD EIVDHYTY ILGDGCQMEGI NEACSLAGHWGLGKLIAFYDDNHISID

                               290       300       310       320       330       340       350
1 [transketolase Sp     GDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFGSPNKSNSYS
gi|568214657|ref|NP     GDTEIAFTEDVSARFESLGWHVIWVKNGNTGYDEIRAAIKEAKAVKDKPTMIKVTTTIGFGSPNKANSYS
gi|356506190|ref|XP     GNTEIAFTESVDSRFEGLGWHVIWVKNGNNGYDDIRAAIKEAKAVKDKPTLIKVTTTIGYGSPNKANSYS
gi|7329685|emb|CAB8     GDTEIAFTENVDQRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGYGSPNKANSYS
gi|565465348|ref|XP     GDTEIAFTENVDKRFEALGWHVIWVKNGNTGYDEIRAAIKEAKAVTDKPTLIKVTTTIGYGSPNKANSYS
gi|28190676|gb|AA03     GDTEIAFTEDVSARFEALGWHTIWVKNGNDGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFGSPNKANSYS
gi|561026815|gb|ESW     GNTEIAFTESVDSRFEGLGWHVIWVKNGNNGYDDIRAAIEEAKSVKDKPTLIKVTTTIGFGSPNKANSYS
gi|115466224|ref|NP     GDTEIAFTEDVSARFEALGWHTIWVKNGNDGYDEIRAAIKEAKAVTDKPTLIKVTTTIGFGSPNKANSYS
gi|312282187|dbj|BA     GDTEIALTESVDKRFEALGWHVIWVKNGNNGYDDIRAAIKEAKAVTDKPSLIKITTTIGYGSPNKANSYS
gi|573948603|ref|XP     GDTEIAFTEDVSARFEALGWHTIWVKNGNTGYDEIRAAIKEAKAVSDKPTLIKVTTTIGFGSPNKANSYS
gi|548832450|gb|ERM     GDTEIAFTETVDTRFEGLGWHVIWVKNGNTGYDEIRAAIEEAKAVKDKPTLIKVTTTIGYGSPNKANSYS
gi|548847307|gb|ERN     GDTEIAFTESVDTRFEGLGWHVIWVKNGNTGYDEIRAAIEEAKAVKDKPTLIKVTTTIGYGSPNKANSYS
gi|502146626|ref|XP     GDTEIAFTESVEKRFEGLGWHVIWVKNGNTGYDEIRAAIREAKSVKDRPTLIKVTTTIGYGSPNKSNSYS
gi|357110873|ref|XP     GDTEIAFTEDVSTRFEALGWHTIWVKNGNDGYDEIRKAIQEAKSVTDKPTLIKVTTTIGFGSPNKANSYS
gi|355506195|gb|AES     GDTEIAFTESVEKRFEGLGWHVIWVKNGNTGFDEIRAAIKEAKAVKDRPTLIKFTTTIGYGSPNKSNSYS
gi|147835837|emb|CA     GDTEIAFTENVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVKDKPTLIKVTTTIGYGSPNKSNSYS
gi|225454009|ref|XP     GDTEIAFTENVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVKDKPTLIKVTTTIGYGSPNKSNSYS
gi|297318807|gb|EFH     GDTEIAFTENVETRFQGLGWHVIWVKNGNTAYDEIRAAIKEAKAVKDKPTLIKITTTIGYGSPNKANSYS
gi|241915985|gb|EER     GDTEIAFTEDVSTRFEALGWHTIWVKNGNNGYDDIRAAIKEAKAVTDKPTLIKVTTTIGFGSPNKANSYS
gi|413953334|gb|AFW     GDTEIAFTEDVSTRFEALGWHTIWVKNGNTGYDDIRAAIKEAKAVTDKPTLIKVTTTIGFGSPNKANSYS
gi|224127366|ref|XP     GDTEIAFTEDVDKRFEGLGWHVIWVKNGNNGYDEIRAAIKEAKAVKDKPTLIKVTTTIGYGSPNKANSYS
Consensus               GDTEIAFTE V  RFE LGWHVIWVKNGNTGYDEIRAAIKEAKAV DKPTLIKVTTTIG GSPNKANSYS

                               360       370       380       390       400       410       420
1 [transketolase Sp     VHGSALGSKEVEATRQNLGWPYEPFHVPEEVK.............KHWSRHTPEGASLEAEWNTKFAEYE
gi|568214657|ref|NP     VHGSGLGAKEVEATRNNLGWPYEPFHVPEDVK.............SHWSRHTPEGAALETEWNAKFAEYE
gi|356506190|ref|XP     VHGSALGAKEVDATRQNLGWSHEPFHVPEDVK.............KHWSRHTPEGAALEAEWNAKFAEYE
gi|7329685|emb|CAB8     VHGAALGEKEVEATRNNLGWPYEPFQVPDDVKSCLDDLLFDLHFSSHWSRHTPEGATLESDWSAKFAAYE
gi|565465348|ref|XP     VHGAALGEKEVEATRNNLGWPYEPFQVPDDVK.............SHWSRHTPEGAALESDWSAKFAAYE
gi|28190676|gb|AA03     VHGSALGTKEVEATRENLGWPYEPFFVPEDVK.............SHWSRHVPQGAAFEADWNAKFAEYE
gi|561026815|gb|ESW     VHGSALGAKEVDATRQNLGWPYEPFHVPEDVK.............KHWSRHIPEGAALEAEWNIKFAEYE
gi|115466224|ref|NP     VHGSALGTKEVEATRENLGWPYEPFFVPEDVK.............SHWSRHVPQGAAFEADWNAKFAEYE
gi|312282187|dbj|BA     AHGAALGEKEVEATRQNLGWPYEPFQVPEDVK.............SHWSRHTPKGAALEADWNAKFAAYE
gi|573948603|ref|XP     VHGSALGTKEVEATRENLGWPYEPFFVPEDVK.............SHWSRHVSQGASFEADWNAKFAEYE
gi|548832450|gb|ERM     VHGSALGAKEVDATRQNLGWPYEPFHVPDDVK.............SHWSRHVSEGASLEAEWQAKLAEYE
gi|548847307|gb|ERN     VHGSALGAKEVDATRQNLGWPYEPFHVPDDVK.............SHWSRHVPEGASLEAEWQAKLADYE
gi|502146626|ref|XP     VHGSALGAKEVDATRNNLGWSHEPFHVPADVK.............KHWSRHIPEGAALESEWNAKFAEYE
gi|357110873|ref|XP     VHGAALGTNEVEATRQNLGWPYEPFFVPEDVK.............SHWSRHVPEGAALEADWNSKFAQYE
gi|355506195|gb|AES     VHGSALGAKEVDATRNNLGWPYEPFHVPEDVK.............KHWSRHIREGAALESEWNAKFADYE
gi|147835837|emb|CA     VHGSALGPKELDATRKNLRWPYEPFHVPEDVK.............KHWSRHVPEGAALEAEWNAQFAEYE
gi|225454009|ref|XP     VHGSALGPKELDATRKNLRWPYEPFHVPEDVK.............KHWSRHVPEGAALEAEWNAQFAEYE
gi|297318807|gb|EFH     VHGSALGEKEVDATRKNLGWPYGPFHVPEDVK.............SHWSRHIPKGAALETEWNAKFADYE
gi|241915985|gb|EER     VHGSALGAKEVEATRQNLGWPYEPFFVPEDVK.............SHWSRHTPEGAALEADWNAKFAEYE
gi|413953334|gb|AFW     VHGSALGAKEVEATRQNLGWPYDTFFVPEDVK.............SHWSRHTPEGAALEADWNAKFAEYE
gi|224127366|ref|XP     VHGSALGAKEVDATRQNLGWPFEPFHVPEDVK.............QHWSRHIPAGAAFEAEWNTKFAEYE
Consensus               VHGSALG KEV ATR NLGWPYEPF VPEDVK             SHWSRH PEGAALEA WNAKFAEYE
```

Figure 1C continued



```
                           430       440       450       460       470       480       490
1 [transketolase Sp    KKYPEDATEFKSITTGEFPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTL
gi|568214657|ref|NP    KKYAEEAADLKSIITGELPAGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGFLGGSADLASSNMTL
gi|356506190|ref|XP    KKYKEEAAELKSIINGEFPAGWEKALPTYTPESPADATRNLSQTNLNALAKVLPGLLGGSADLASSNMTL
gi|7329685|emb|CAB8    KKYPEEASELKSIITGELPAGWEKALPTYTPESPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTL
gi|565465348|ref|XP    KKYPEEAAELKSIISGELPTGWEKALPTYTPESPGDATRNLSQQCLNAIAKVVPGFLGGSADLASSNMTL
gi|28190676|gb|AAO3    KKYPEDAATLKSIVSGELPAGWADALPKYTPESPADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTL
gi|561026815|gb|ESW    KKYKEEAAELKSIINGDLPAGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGLLGGSADLASSNMTL
gi|115466224|ref|NP    KKYPEDAATLKSIVSGELPAGWADALPKYTPESPADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTL
gi|312282187|dbj|BA    KKYAEEARELKSIMSGELPAGWEKALPTYTPDSPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTL
gi|573948603|ref|XP    KKYPEDAATLKSIVSGELPTGWADALPKYTTESAADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTL
gi|548832450|gb|ERM    KKYKEEAVEFKELISGQLPDGWEKALPVYTAESPADATRNLSQQCLNALAKVLPGLLGGSADLASSNMTL
gi|548847307|gb|ERN    KKYKEEAAEFKELISGKWPDGWEKALPAYTPESPADATRNLSQQCLNALAKVLPGLLGGSADLASSNMTL
gi|502146626|ref|XP    KKYKEEAAVLKSIISGDLPSGWEKALPTYTPETPADATRILSQQNLNALAKVLPGLLGGSADLASSNMTL
gi|357110873|ref|XP    KKYPEDAAALKSIITGELPAGWADALPQYTTESPADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTL
gi|355506195|gb|AES    KKYKEEAAVLKSIISGDLPAGWEKALPTYTPEIPADATRNLSQQNLNALANVLPGLIGGSADLASSNMTL
gi|147835837|emb|CA    RKYKEEAAVLKSLINGELPAGWEKALPTYTPESPAEATRNSSQQCLNALANVLPGLLGGSADLASSNLSV
gi|225454009|ref|XP    RKYKEEAAVLKSLINGELPAGWEKALPTYTPESPAEATRNSSQQCLNALANVLPGLLGGSADLASSNLSV
gi|297318807|gb|EFH    NKYPQEAAELKSIVTVNLPTGWEKSLPTYTPESPADATRNLSQQC[illegible]ALAKVLPGFLGGSADLASSNMTL
gi|241915985|gb|EER    KKYADDAATLKSIITGEFPTGWADALPKYTPESPGDATRNLSQQCLNALANVVPGLIGGSADLASSNMTL
gi|413953334|gb|AFW    KKYADDAATLKSIITGELPTGWVDALPKYTPESPGDATRNLSQQCLNALANVVPGLIGGSADLASSNMTL
gi|224127366|ref|XP    KKYSEEAAEFKSIMTGELPTGWEKALPTYTPESPADATRNLSQQNLNALAKVLPGLLGGSADLASSNMTL
Consensus              KKY EEAA LKSII GELP GWEKALPTYTPESPADATRNLSQQCLNALAKV PGLLGGSADLASSNMTL

                           500       510       520       530       540       550       560
1 [transketolase Sp    LKMFGDFQKDTPEERNVRFGVREHGMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|568214657|ref|NP    LKMFGDFQKNTPEERNLRFGVREHGMGAICNGIALHSLGLIPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|356506190|ref|XP    LKMFGDFQKDTPAERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAIRLSALSEAGVIY
gi|7329685|emb|CAB8    LKAFGDFQKATPEERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|565465348|ref|XP    LKASGDFQKATPEERNLRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|28190676|gb|AAO3    LKMFGDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEAGVIY
gi|561026815|gb|ESW    LKMFGDFQKDTPAERNVRFGVREHGMGAICNGIALHSPALIPYCATFFVFTDYMRGAIRLSALSEAGVIY
gi|115466224|ref|NP    LKMFGDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEAGVIY
gi|312282187|dbj|BA    LKAFGDFQKATPEERNLRFGVREHGMGAICNGIAHHSPGLIPYCATFFVFTDYMRAAMRIAALSEGGVIY
gi|573948603|ref|XP    LKMFGDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEAGVIY
gi|548832450|gb|ERM    LKMFGDFQKDTPEQRNVRFGVREHSMGAICNGIALH.SGLIPYCATFFVFTDYMRAAMRISALCEAGVIY
gi|548847307|gb|ERN    LKMFGDFQKDTPEQRNVRFGVREHAMGSICNGIALHGSGLIPYCATFFVFTDYMRAAMRISALCEAGVIY
gi|502146626|ref|XP    LKTFGDFQSDSPGERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAIRLSALSEAGVIY
gi|357110873|ref|XP    LKMFGDFQKDTPEERNVRFGVREHGMGAICNGIGLHTPGLIPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|355506195|gb|AES    LKKFGDFQSDSPAERNVRFGVREHGMGAICNGIALHSRGLIPYCATFFVFTDYMRGAIRLSALSEAGVIY
gi|147835837|emb|CA    MKQFGNFQKGTPEERNVRFGVREHGMGAICNGIVLHCPGLIPYCATFFVFTDYLRPAMRISALCEAGVIY
gi|225454009|ref|XP    MKQFGNFQKGTPEERNVRFGVREHGMGAICNGIVLHCPGLIPYCATFFVFTDYLRPAMRISALCEAGVIY
gi|297318807|gb|EFH    LKMFGDFQKNAPEERNLRFGVREHSMGAICNGITCHSPGLIPYCATFFVFTDYMRAAIRLSALSQAGVIY
gi|241915985|gb|EER    LKMFGDFQKDTAEERNVRFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|413953334|gb|AFW    LKMFGDFQKDTAEERNVRFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIY
gi|224127366|ref|XP    LKMFGDFQKDTPEERNVRFGVREHGMGSICNGIALHSPGLIPYCATFFVFTDYMRGAIRISALSEAGVIY
Consensus              LKMFGDFQKDTPEERNVRFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMR AMRISALSEAGVIY

                           570       580       590       600       610       620       630
1 [transketolase Sp    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVENRKTPSILALSRQKLPNLP
gi|568214657|ref|NP    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMFRPADGNETAGAYKVAVLKRKTPSILALSRQKLPQLA
gi|356506190|ref|XP    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPQLP
gi|7329685|emb|CAB8    VMTHDSIGLGEDGPTHQPIEHIASFRAMPNTLMFRPADGNETAGAYKIAVTKRKTPSILALSRQKLPHLP
gi|565465348|ref|XP    VMTHDSIGLGEDGPTHQPIEHIASFRAMPNTLMFRPADGNETAGAYKIAVTKRKTPSILALSRQKLPHLP
gi|28190676|gb|AAO3    VMTHDSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKIAVLNRKRPSVLALSRQKLAQLP
gi|561026815|gb|ESW    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPNLP
gi|115466224|ref|NP    VMTHDSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKIAVLNRKRPSVLALSRQKLAQLP
gi|312282187|dbj|BA    VMTHDSIGLGEDGPTHQPVEHLSSFRAMPNILMFRPADGNETAGAYKIAVTKRKTPSVLALSRQKLQQLP
gi|573948603|ref|XP    VMTHDSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKIAVLNRKRPSVLALSRQKLPQLP
gi|548832450|gb|ERM    VMTHDSIGLGEDGPTHQPIEHIASFRAMPNILMLRPCDGNETAGAYKVAVLNRKRPSVLALSRQKLPQLA
gi|548847307|gb|ERN    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPCDGNETAGAYKVAILNRKRPSVLALSRQKLPQLA
gi|502146626|ref|XP    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPNLA
gi|357110873|ref|XP    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNMLMFRPADGKETAGAYKVAVLNRKRPSILALSRQKLPHLP
gi|355506195|gb|AES    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPNLA
gi|147835837|emb|CA    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGTETAAAYKIAVLNRKRPSVLALGRRDVSQLR
gi|225454009|ref|XP    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGTETAAAYKIAVLNRKRPSVLALGRRDVSQLR
gi|297318807|gb|EFH    VMTHDSIGLGEDGPTHQPIEHVASFRAMPNILMLRPADGTETAGAYKVAVENRKRPSILALSRQKLLHLP
gi|241915985|gb|EER    VMTHDSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLP
gi|413953334|gb|AFW    VMTHDSIGLGEDGPTHQPIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLP
gi|224127366|ref|XP    VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNSKRPSILALSRQKLPQLP
Consensus              VMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLP LP
```

Figure 1C continued



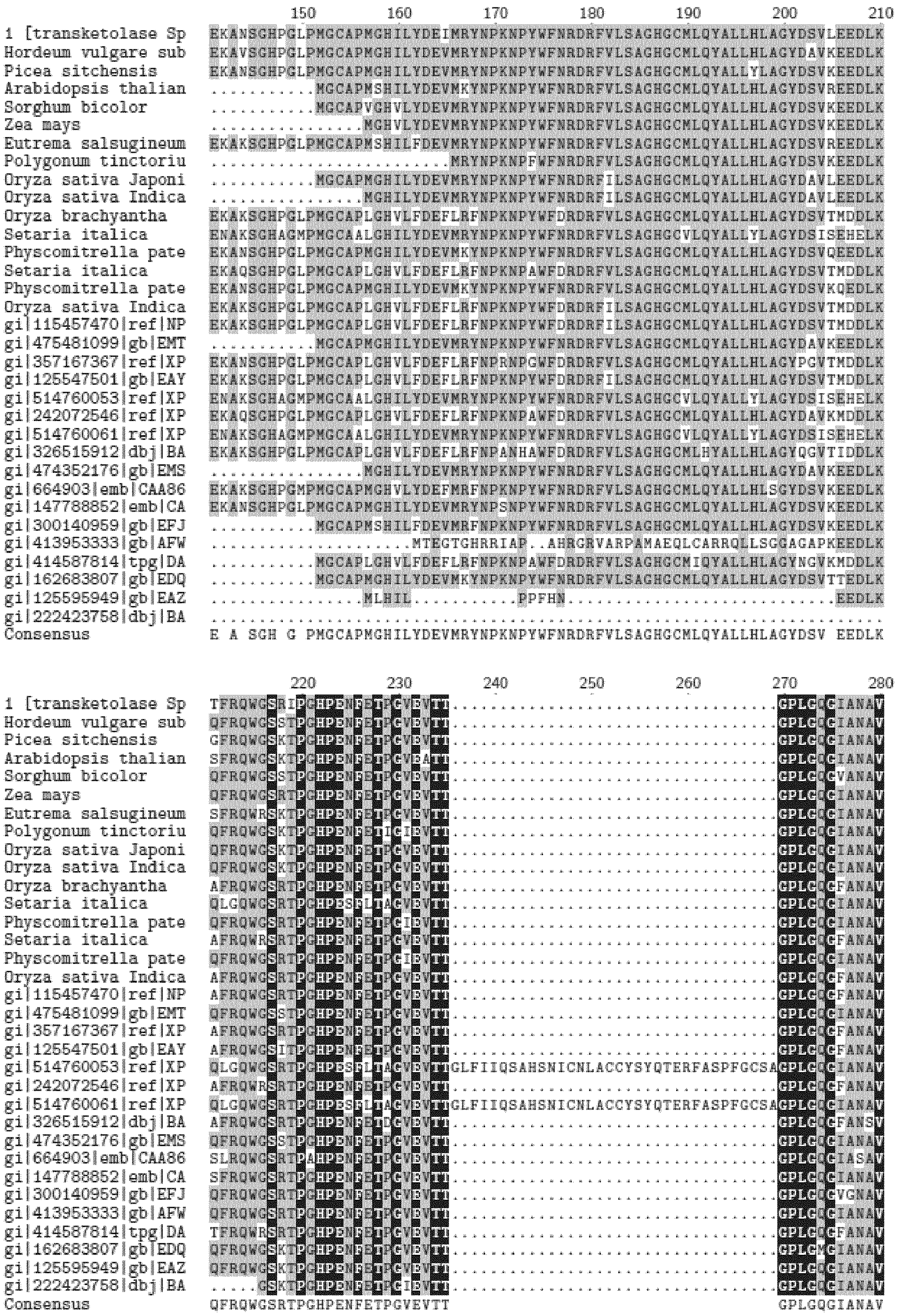
```
                              640       650       660       670       680       690       700
i [transketolase Sp  GTSIEGVEKGGYTITDNSSGNKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQSDEYK
gi|568214657|ref|NP  GTSIEGAAKGGYIVSDNSSGNKPDVILIGTGSELEIAVKAAEELKKEGKTVRVVSFVCWELYDEQSAEYK
gi|356506190|ref|XP  GTSIEGVEKGGYTISDNSTGNKPDVILIGTGSELEIAAKAADDLRKEGKAVRVVSLVSWELFDEQSEAYK
gi|7329685|emb|CAB8  GTSIEGVEKGGYTISDDSSGNKPDVILIGTGSELEIAAQAAEVLRKDGKTVRVVSFVCWELFDEQSDEYK
gi|565465348|ref|XP  GTSIEGVEKGGYTISDDSAGNKPDVILIGTGSELEIAAQAAEVLRKEGKTVRVVSFVCWELFDEQSDEYK
gi|28190676|gb|AAO3  GTSIEGVEKGGYIVSDNSTGNKPDFIVMSTGSELEIVAKAADELRKEGKTVRVVSFVCWELFDEQSAEYK
gi|561026815|gb|ESW  GSSIEGVEKGGYTISDNSTGNKPDAILIGTGSELEIAAKAAEDLRKEGKTVRVVSLVSWELFDEQSEAYK
gi|115466224|ref|NP  GTSIEGVEKGGYIVSDNSTGNKPDFIVMSTGSELEIVAKAADELRKEGKTVRVVSFVCWELFDEQSAEYK
gi|312282187|dbj|BA  GTSIESVEKGGYTISDNSTGNKPDVILIGTGSELEIAAQAAEELRKEGKTVRVVSFVCWELFDEQSDAYK
gi|573948603|ref|XP  GTSIEGVEKGGYTISDNSTGNKPDLIVMSTGSELEIAAKAADELRKEGKTVRVVSFVCWELFDEQSDEYK
gi|548832450|gb|ERM  GTTIEGVEKGGYIVSDNSMRNKPDLILIGTGSELEIAVKAADEIRKEGKAVRVVSLVSWELFDEQSDDYK
gi|548847307|gb|ERN  GTTIEGVEKGGYIVSDNSMRNKPDLILIGTGSELEIAVKAANVIRKEGKAVRVVSLVSWELFDEQSDEYK
gi|502146626|ref|XP  GTSIEGVEKGGYIVSDNSTGNKPDVILIGTGSELEIAYKAAEDLRKEGKAVRVVSFVSWELFDEQSEAYK
gi|357110873|ref|XP  GTSIEGVEKGGYTISDNSTGNKPDFIIMSTGSELEIAVKAAEELTKEGKTVRVVSFVCWELFDDQSDEYK
gi|355506195|gb|AES  GTSIEGVEKGGYIVSDNSTGNKPDVILIGTGSELEIAYKAGEDLRKEGKAVRVVSFVSWELFDEQSQAYK
gi|147835837|emb|CA  GTSIEGVEKGGYIVTDNSSGNKPDIILIGTGSELEIAAKAADELRKEGKAVRVVSLVSWALFDEQSDAYK
gi|225454009|ref|XP  GTSIEGVEKGGYIVTDNSSGNKPDVILIGTGSELEIAAKAADELRKEGKAVRVVSLVCWALFDEQSDAYK
gi|297318807|gb|EFH  GTSVEGVEKGGYVISDNSTGNKPDVILMGTGSELEIAAKAGEKLREEGRAVRVVSLVSWELFDEQSEEYK
gi|241915985|gb|EER  GTSIEGVEKGGYTISDNSTGNKPDLIVLGTGSELEIAEKAADELRKEGKTVRVVSFVSWELFDEQSDEYK
gi|413953334|gb|AFW  GTSIEGVEKGGYTISDNSTGNKPDIIVMGTGSELEIAAKAADELRKEGKTVRVVSFVSWELFDEQSDEYK
gi|224127366|ref|XP  GTSIEGVEKGGYIISDNSSGNKPDVILMGTGSELEIAAKAAEELRKEGKAVRVVSFVSWELFDEQSVAYK
Consensus            GTSIEGVEKGGY  SDNS GNKPD ILIGTGSELEIAAKAA ELRKEGK VRVVSFV WELFDEQS  YK

                              710       720       730       740       750       760
i [transketolase Sp  ESVLPSDVTARVSIEAGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC
gi|568214657|ref|NP  ESVLPSSVTARVSIEAGSTFGWQKFVGDKGKAIGIDGFGASAPADKIYKEFGITAEAVVAAAKQVS
gi|356506190|ref|XP  ESVFPAAVSARVSIEAGSTFGWEKIVGAKGKAIGIDRFGASAPAGRIYKEFGITKEAVVAAAKELI
gi|7329685|emb|CAB8  ESVLPSDVSARVSIEAASTFGWGKIVGGKGKSIGINSFGASAPAPLLYKEFGITVEAVVDAAKSFF
gi|565465348|ref|XP  ESVLPSAVSARVSIEAASTFGWGKIVGGKGKSIGINSFGASAPAPLLYKEFGITVEAVVDAAKSLF
gi|28190676|gb|AAO3  ESVLPEAVTARVSLEAGSTLGWQKYVGSKGKAIGIDKFGASAPAGKIYQEYGITAENVIATAKSL.
gi|561026815|gb|ESW  ESVLPAAVSARVSIEAGSTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITKEAVIAAAKELI
gi|115466224|ref|NP  ESVLPEAVTARVSLEAGSTLGWQKYVGSKGKAIGIDKFGASAPAGKIYQEYGITAENVIATAKSL.
gi|312282187|dbj|BA  ESVLPSDVSARVSIEAGSTFGWEKIVGGKGKSIGIDSFGASAPAGKLYKEFGFTIEAVVAAAKSLI
gi|573948603|ref|XP  ESVLPEAVTGRISIEAGSTLGWQKYVGSKGKAIGIDKFGASAPAGKIYQEYGISVESIIAAAKSL.
gi|548832450|gb|ERM  ESVLPAEVTARVSVEAGSTFGWEKIVGSKGKSIGIDRFGASAPAGTLYKEFGITVDNVIAAAKSLL
gi|548847307|gb|ERN  ESVLPAGVSARVSVEAGSTFGWEKIVGSKGKSIGIDRFGASAPAGTLYKEFGITMDNVIAAAKTLL
gi|502146626|ref|XP  ESVLPAAVSARVSIEAGSTFGWHKIVGSKGKAIGIDRFGASAPAGRIYKEFGITKEAVIAAAKELI
gi|357110873|ref|XP  ESVLPEAVTARISIEAGSTLGWQKYVGSKGKTIGIDKFGASAPAGIIYKEYGITAESVIAAAKSL.
gi|355506195|gb|AES  ESVLPTAVTARVSIEAGSTFGWEKIVGSKGKAIGIDRFGASAPAGTIYKEFGITKEAVIAAAKELI
gi|147835837|emb|CA  ESVLPAAVSARVSVEAASTFGWEKFVGSKGKSIGIDRFGASAPALKLYKELGVTAEAVIAAAKSIC
gi|225454009|ref|XP  ESVLPAAVSARVSVEAASTFGWEKFVGSKGKSIGIDRFGASAPALKLYKELGVTAEAVIAAAKSIC
gi|297318807|gb|EFH  ESVLPSEVSARVSIEAGSTFGWEKMVGSKGKAIGVDSFGASAPADVLYKKFGLTVDAVVAAAKDLC
gi|241915985|gb|EER  ESVLPAAVTARISIEAGSTLGWQKYVGAQGKAIGIDKFGASAPAGTIYKEYGITVESVIAAAKSF.
gi|413953334|gb|AFW  ESVLPAAVTARISIEAGSTLGWQKYVGAQGKAIGIDKFGASAPAGTIYKEYGITVESIIAAAKSF.
gi|224127366|ref|XP  ESVLPAAVTARVSIEAGSTFGWQKLVGAKGKAIGIDRFGASAPAGKIYKEFGITAEAVIAAAKEVS
Consensus            ESVLP AV ARVSIEAGSTFGW K VGSKGKAIGID FGASAPAG IYKE GIT EAVIAAAKSL
```

X non conserved
X ≥50% conserved
X all match

Figure 1D

```
                             10        20        30        40        50        60        70
1 [transketolase Sp  ............................................MAASSSLSTLSHHQTLLSHPKTHLPT
Hordeum vulgare sub  ......................................................................
Picea sitchensis     ...................................MASSFTVSRASAGSVTATALNNAIQSERLSSSASP
Arabidopsis thalian  ......................................................................
Sorghum bicolor      ......................................................................
Zea mays             ......................................................................
Eutrema salsugineum  ..............................................MASTSSSLAISQALLGRAISQNGS
Polygonum tinctoriu  ......................................................................
Oryza sativa Japoni  ......................................................................
Oryza sativa Indica  ......................................................................
Oryza brachyantha    ............................................................MATRAP....
Setaria italica      MILVYFRKMHLQIYTFPLTSCSFQVAISHLCARLLTHTNHLITNMTVHSIAGAYTTIGSHFSSVPCGRAP
Physcomitrella pate  ...................................MAAARCNVAAVAGAVAPAAVR..............
Setaria italica      ..........................................................MAAPTPTT....
Physcomitrella pate  ......................................................................
Oryza sativa Indica  ..........................................................MATATQAP....
gi|115457470|ref|NP  ..........................................................MATATQAP....
gi|475481099|gb|EMT  ......................................................................
gi|357167367|ref|XP  .........................................................MARMPTPIP....
gi|125547501|gb|EAY  ..........................................................MATATQAP....
gi|514760053|ref|XP  MILVYFRKMHLQIYTFPLTSCSFQVAISHLCARLLTHTNHLITNMTVHSIAGAYTTIGSHFSSVPCGRAP
gi|242072546|ref|XP  ..........................................................MVTPTSTTPVST
gi|514760061|ref|XP  ............................................MTVHSIAGAYTTIGSHFSSVPCGRAP
gi|326515912|dbj|BA  ..........................................................MATPTPAP....
gi|474352176|gb|EMS  ......................................................................
gi|664903|emb|CAA86  ......................................................................
gi|147788852|emb|CA  ............................................MAAASASSSLX....LLPRTLNHHSS
gi|300140959|gb|EFJ  ......................................................................
gi|413953333|gb|AFW  ......................................................................
gi|414587814|tpg|DA  ......................................................................
gi|162683807|gb|EDQ  ......................................................................
gi|125595949|gb|EAZ  ......................................................................
gi|222423758|dbj|BA  ......................................................................
Consensus

                             80        90       100       110       120       130       140
1 [transketolase Sp  TPASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASAVVRAAAVEALESTDIDQLVEKSVNTIRFLAIDAV
Hordeum vulgare sub  ...................................NPHPTPPTSPSPPSATGELLDKSVNAIRFLAIDAV
Picea sitchensis     AGFGFSKFAGLKLRSNGKVACSSRSSSQRRLRSSAPKVVRAAAVETLEAISTGLIEKSINAIRFLAIDAV
Arabidopsis thalian  ......................................................................
Sorghum bicolor      ......................................................................
Zea mays             ......................................................................
Eutrema salsugineum  DKCVSIPAFSGLKSTSPRTTFSSRRIASTNSHSLRPLVRAAAVETKT...ESSLVEKSVNTIRFLAIDAV
Polygonum tinctoriu  ......................................................................
Oryza sativa Japoni  ......................................................................
Oryza sativa Indica  ......................................................................
Oryza brachyantha    .TSPAAAGASGRASRLLFLR.CPSTRFRLAAG.....RSATALRAQP..AAAEVVEQSVNTIRFLAVDAV
Setaria italica      AEHIGLRHLSALSSCSLHSPLFSHCSPVSSPSTKCAVHAIAIETLERNAASGQLLEKCVNTIRFLAIDAV
Physcomitrella pate  ......................GTAQARNARKGKKVSVRATASVETAQKTDNALVEKSVNTIRFLAIDAV
Setaria italica      ..SAAAAPGSGRG.VVRLLRSSAGSRLPFAAAVRPGGRPATALRARAPPAAPKLVEQSVNTIRFLAVDAV
Physcomitrella pate  ........................MPARNARRGKQVSVRATAAVEIAQKTDDALVEKSVNTIRFLAIDAV
Oryza sativa Indica  .TSPAAA..SGRGSVLPLLRGRPSTRVRLAVGR....RPQTALRAQPP.GAAEVVEQSVNTIRFLAVDSV
gi|115457470|ref|NP  .TSPAAA..SGRGSVLPLLRGRPSTRVRLAVGR....RPQTALRAQPP.GAAEVVEQSVNTIRFLAVDSV
gi|475481099|gb|EMT  ......................................................................
gi|357167367|ref|XP  .TTFASSVASGHGLLLVRGRRSTRAARALSLGT.PGGRSGTAIHSSRQPAAAELVEQSVNTIRFLAVDAV
gi|125547501|gb|EAY  .TSPAAA..SGRGSVLPLLRGRPSTRVRLAVGR....RPQTALRAQPP.GAAEVVEQSVNTIRFLAVDSV
gi|514760053|ref|XP  AEHIGLRHLSALSSCSLHSPLFSHCSPVSSPSTKCAVHAIAIETLERNAASGQLLEKCVNTIRFLAIDAV
gi|242072546|ref|XP  ASAAAPAPGSGRGGVLRLLRYSACSLLPSAAAARSGARLTTALRARTQPAEPELVEQSVNTIRFLAVDAV
gi|514760061|ref|XP  AEHIGLRHLSALSSCSLHSPLFSHCSPVSSPSTKCAVHAIAIETLERNAASGQLLEKCVNTIRFLAIDAV
gi|326515912|dbj|BA  .MSPASKAAALASGGFLVGRARVSLGRRPALR..............AQPAAAELVERSVNTIRFLAVDAV
gi|474352176|gb|EMS  ......................................................................
gi|664903|emb|CAA86  .........................................MAPKTTLIAEPELVSKSVNTIRFLAIDAV
gi|147788852|emb|CA  TPHPNRLSLSG.......LTPKTQALRTRHGRRARATVRAAAVETLQKADTS.LVEKSINTIRFLSIDAV
gi|300140959|gb|EFJ  ......................................................................
gi|413953333|gb|AFW  ......................................................................
gi|414587814|tpg|DA  ......................................................................
gi|162683807|gb|EDQ  ......................................................................
gi|125595949|gb|EAZ  ......................................................................
gi|222423758|dbj|BA  ......................................................................
Consensus                                                                   E  VNTIRFLA D V
```

Figure 1D continued



```
                               150       160       170       180       190       200       210
1 [transketolase Sp     EKANSGHPGLPMGCAPMGHILYDEIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVLEEDLK
Hordeum vulgare sub     EKAVSGHPGLPMGCAPMGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVKEEDLK
Picea sitchensis        EKANSGHPGLPMGCAPMGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLYLAGYDSVKEEDLK
Arabidopsis thalian     ...........MGCAPMSHILYDEVMKYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVREEDLK
Sorghum bicolor         ...........MGCAPVGHVLYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLK
Zea mays                ................MGHVLYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLK
Eutrema salsugineum     EKAKSGHPGLPMGCAPMSHILFDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVREEDLK
Polygonum tinctoriu     .........................MRYNPKNPFWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLK
Oryza sativa Japoni     ...........MGCAPMGHILYDEVMRYNPKNPYWFNRDRFILSAGHGCMLQYALLHLAGYDAVLEEDLK
Oryza sativa Indica     ................MGHILYDEVMRYNPKNPYWFNRDRFILSAGHGCMLQYALLHLAGYDAVLEEDLK
Oryza brachyantha       EKAKSGHPGLPMGCAPLGHVLFDEFLRFNPKNPYWFDRDRFVLSAGHGCMLQYALLHLAGYDSVTMDDLK
Setaria italica         ENAKSGHAGMPMGCAALGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCVLQYALLYLAGYDSISEHELK
Physcomitrella pate     EKANSGHPGLPMGCAPMGHILYDEVMKYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVQEEDLK
Setaria italica         EKAQSGHPGLPMGCAPLGHVLFDEFLRFNPKNPAWFDRDRFVLSAGHGCMLQYALLHLAGYDSVTMDDLK
Physcomitrella pate     EKANSGHPGLPMGCAPMGHILYDEVMKYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKQEDLK
Oryza sativa Indica     EKAKSGHPGLPMGCAPLGHVLFDEFLRFNPKNPYWFDRDRFILSAGHGCMLQYALLHLAGYDSVTMDDLK
gi|115457470|ref|NP     EKAKSGHPGLPMGCAPLGHVLFDEFLRFNPKNPYWFDRDRFILSAGHGCMLQYALLHLAGYDSVTMDDLK
gi|475481099|gb|EMT     ...........MGCAPMGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVKEEDLK
gi|357167367|ref|XP     EKANSGHPGLPMGCAPLGHVLFDEFLRFNPRNPGWFDRDRFVLSAGHGCMLQYALLHLAGYPGVTMDDLK
gi|125547501|gb|EAY     EKAKSGHPGLPMGCAPLGHVLFDEFLRFNPKNPYWFDRDRFILSAGHGCMLQYALLHLAGYDSVTMDDLK
gi|514760053|ref|XP     ENAKSGHAGMPMGCAALGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCVLQYALLYLAGYDSISEHELK
gi|242072546|ref|XP     EKAQSGHPGLPMGCAPLGHVLFDEFLRFNPKNPAWFDRDRFVLSAGHGCMLQYALLHLAGYDAVKMDDLK
gi|514760061|ref|XP     ENAKSGHAGMPMGCAALGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCVLQYALLYLAGYDSISEHELK
gi|326515912|dbj|BA     EKAKSGHPGLPMGCAPLGHVLFDEFLRFNPANHAWFDRDRFVLSAGHGCMLHYALLHLAGYQGVTIDDLK
gi|474352176|gb|EMS     ................MGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDAVKEEDLK
gi|664903|emb|CAA86     EKAKSGHPGMPMGCAPMGHVLYDEFMRFNPKNPYWFNRDRFVLSAGHGCMLQYALLHLSGYDSVKEEDLK
gi|147788852|emb|CA     EKANSGHPGLPMGCAPMGHILYDEVMRYNPSNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLK
gi|300140959|gb|EFJ     ...........MGCAPMSHILFDEVMRFNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVKEEDLK
gi|413953333|gb|AFW     .....................MTEGTGHRRIAP..AHRGRVARPAMAEQLCARRQLLSGGAGAPKEEDLK
gi|414587814|tpg|DA     ...........MGCAPLGHVLFDEFLRFNPKNPAWFDRDRFVLSAGHGCMIQYALLHLAGYNGVKMDDLK
gi|162683807|gb|EDQ     ...........MGCAPMGHILYDEVMKYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSVTTEDLK
gi|125595949|gb|EAZ     ................MLHIL..........PPFHN.............................EEDLK
gi|222423758|dbj|BA     ......................................................................
Consensus               E A SGH G PMGCAPMGHILYDEVMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLAGYDSV EEDLK

                               220       230       240       250       260       270       280
1 [transketolase Sp     TFRQWGSRIPGHPENFETPGVEVTT..................................GPLGQGIANAV
Hordeum vulgare sub     QFRQWGSSTPGHPENFETPGVEVTT..................................GPLGQGIANAV
Picea sitchensis        GFRQWGSKTPGHPENFETPGVEVTT..................................GPLGQGIANAV
Arabidopsis thalian     SFRQWGSKTPGHPENFETPGVEATT..................................GPLGQGIANAV
Sorghum bicolor         QFRQWGSSTPGHPENFETPGVEVTT..................................GPLGQGVANAV
Zea mays                QFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGIANAV
Eutrema salsugineum     SFRQWRSKTPGHPENFETPGVEVTT..................................GPLGQGIANAV
Polygonum tinctoriu     QFRQWGSKTPGHPENFETIGIEVTT..................................GPLGQGIANAV
Oryza sativa Japoni     QFRQWGSKTPGHPENFETPGVEVTT..................................GPLGQGIANAV
Oryza sativa Indica     QFRQWGSKTPGHPENFETPGVEVTT..................................GPLGQGIANAV
Oryza brachyantha       AFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
Setaria italica         QLGQWGSRTPGHPESFLTAGVEVTT..................................GPLGQGIANAV
Physcomitrella pate     QFRQWGSRTPGHPENFETPGIEVTT..................................GPLGQGIANAV
Setaria italica         AFRQWRSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
Physcomitrella pate     QFRQWGSRTPGHPENFETPGIEVTT..................................GPLGQGIANAV
Oryza sativa Indica     AFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
gi|115457470|ref|NP     AFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
gi|475481099|gb|EMT     QFRQWGSSTPGHPENFETPGVEVTT..................................GPLGQGIANAV
gi|357167367|ref|XP     AFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
gi|125547501|gb|EAY     AFRQWGSITPGHPENFETPGVEVTT..................................GPLGQGFANAV
gi|514760053|ref|XP     QLGQWGSRTPGHPESFLTAGVEVTTGLFIIQSAHSNICNLACCYSYQTERFASPFGCSAGPLGQGIANAV
gi|242072546|ref|XP     AFRQWRSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
gi|514760061|ref|XP     QLGQWGSRTPGHPESFLTAGVEVTTGLFIIQSAHSNICNLACCYSYQTERFASPFGCSAGPLGQGIANAV
gi|326515912|dbj|BA     AFRQWGSRTPGHPENFETDGVEVTT..................................GPLGQGFANSV
gi|474352176|gb|EMS     QFRQWGSSTPGHPENFETPGVEVTT..................................GPLGQGIANAV
gi|664903|emb|CAA86     SLRQWGSRTPAHPENFETPGVEVTT..................................GPLGQGIASAV
gi|147788852|emb|CA     SFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGIANAV
gi|300140959|gb|EFJ     QFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGVGNAV
gi|413953333|gb|AFW     QFRQWGSRTPGHPENFETPGVEVTT..................................GPLGQGIANAV
gi|414587814|tpg|DA     TFRQWRSRTPGHPENFETPGVEVTT..................................GPLGQGFANAV
gi|162683807|gb|EDQ     QFRQWGSKTPGHPENFETPGVEVTT..................................GPLGMGIANAV
gi|125595949|gb|EAZ     QFRQWGSKTPGHPENFETPGVEVTT..................................GPLGQGIANAV
gi|222423758|dbj|BA     .....GSKTPGHPENFETPGIEVTT..................................GPLGQGIANAV
Consensus               QFRQWGSRTPGHPENFETPGVEVTT                                  GPLGQGIANAV
```

Figure 1D continued

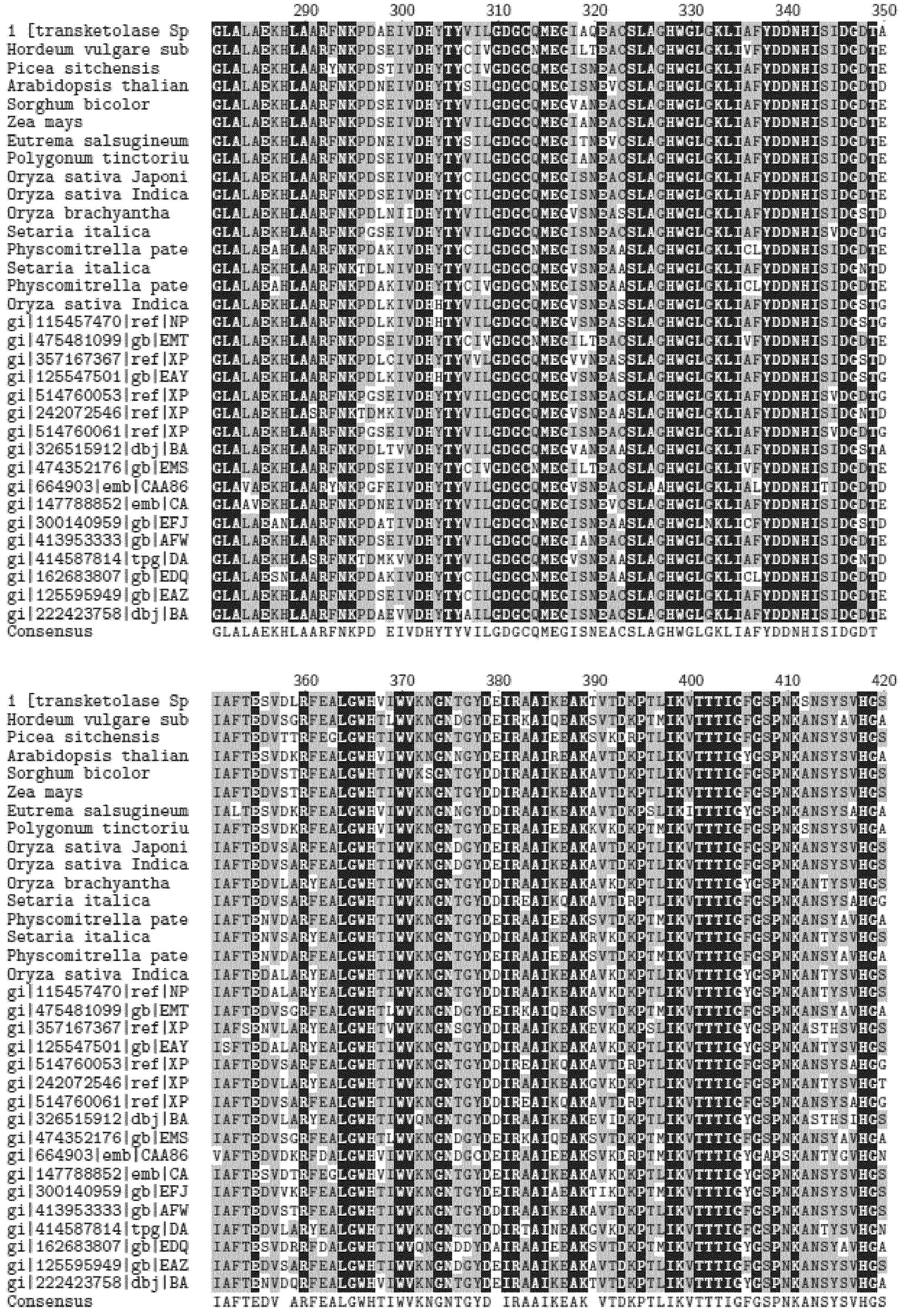
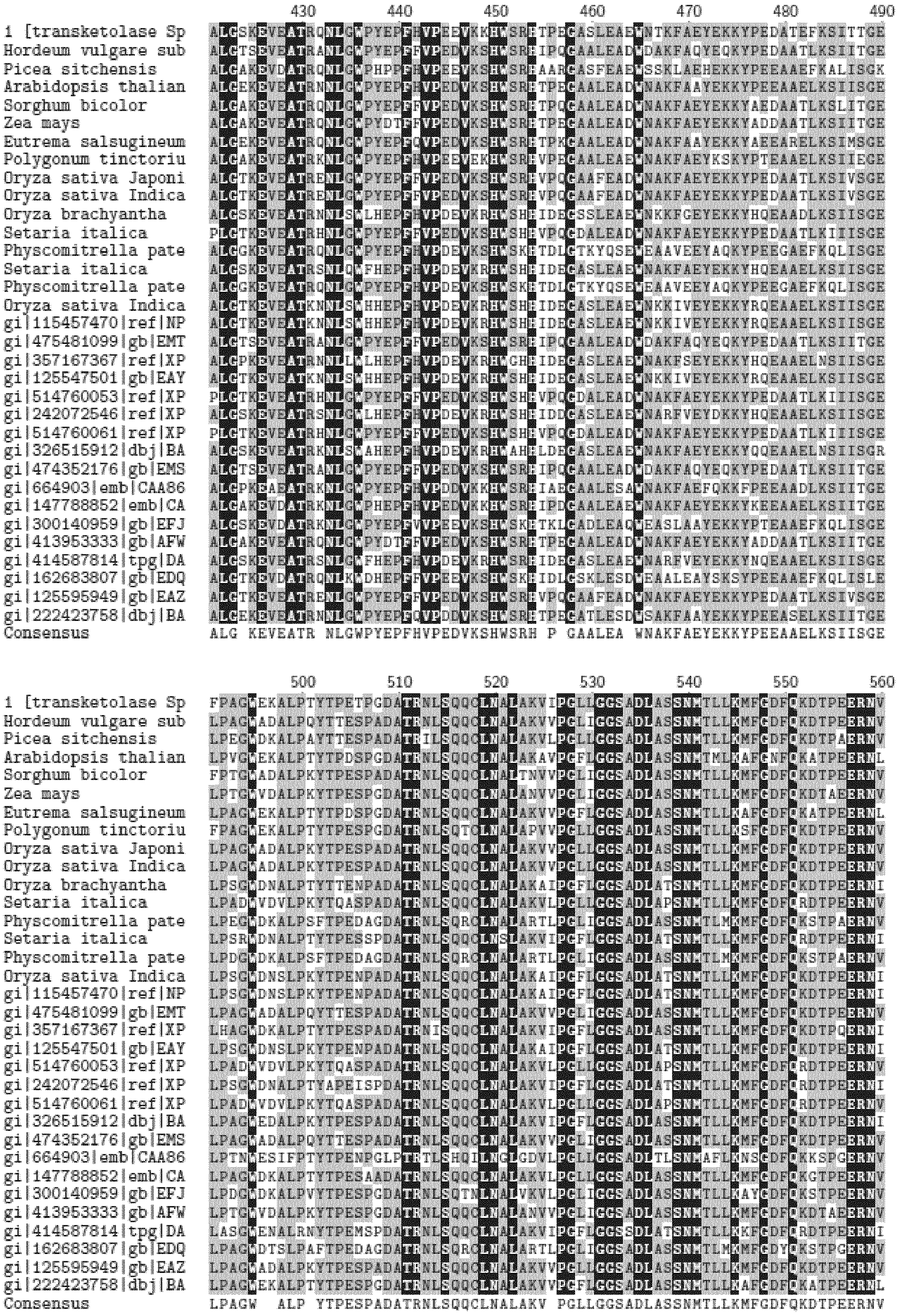
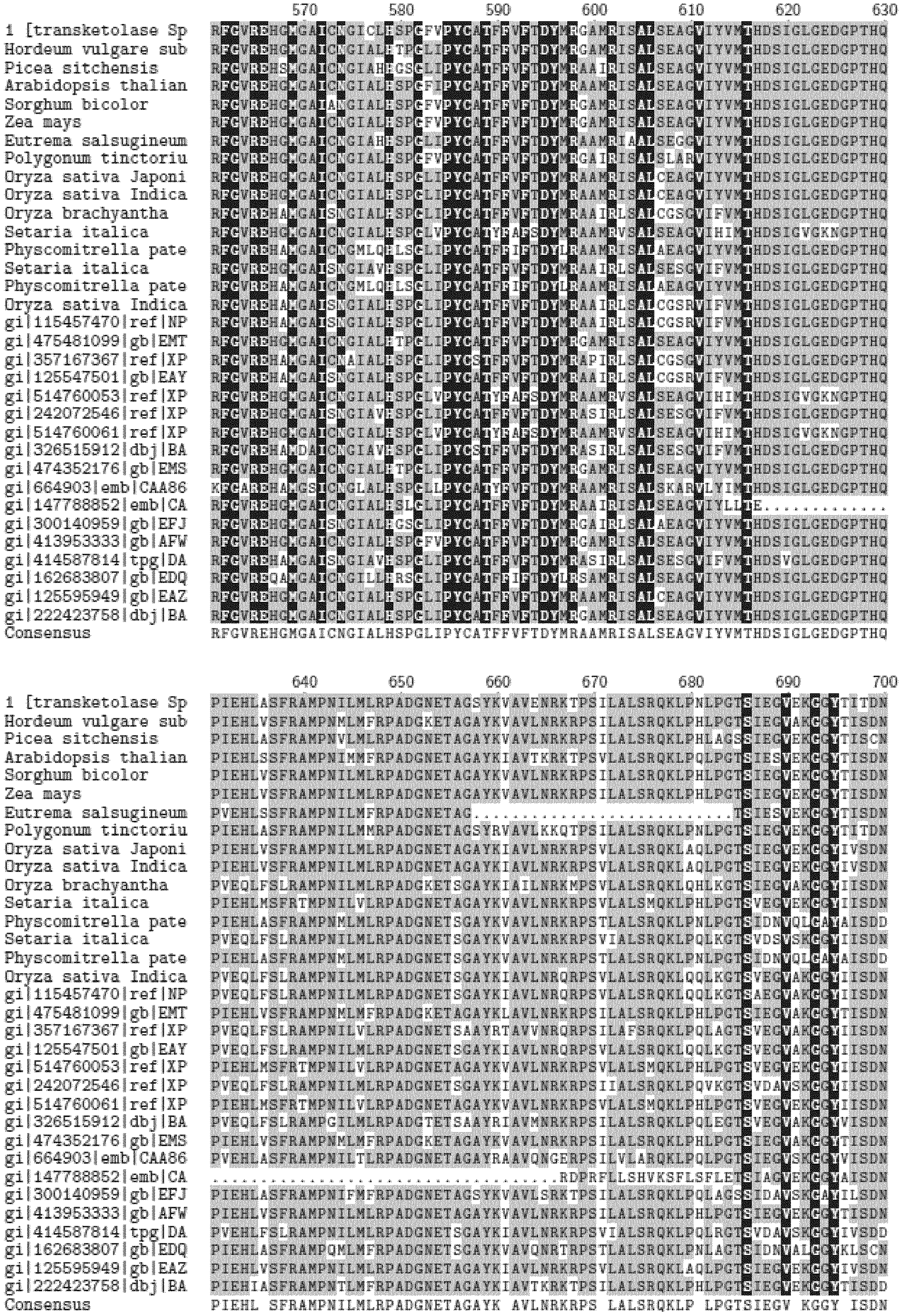


Figure 1D continued



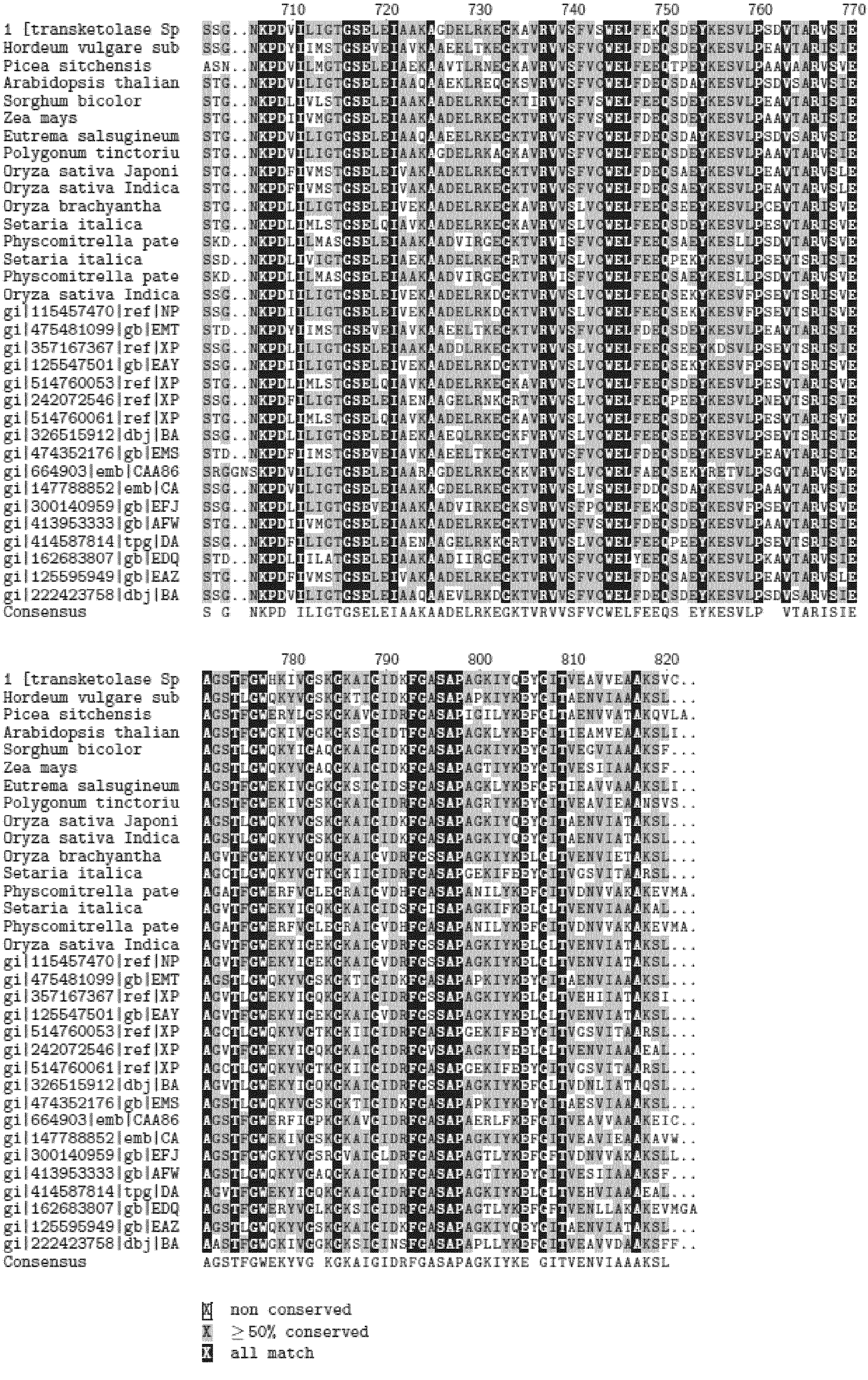
```
                                 430       440       450       460       470       480       490
1 [transketolase Sp    ALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYPEDATEFKSITTGE
Hordeum vulgare sub    ALGTSEVEATRANLGWPYEPFFVPEDVKSHWSRHIPQGAALEADWDAKFAQYEQKYPEDAATLKSIITGE
Picea sitchensis       ALGAKEVDATRQNLGWPHPPFHVPEEVKSHWSRHAARGASFEAEWSSKLAEHEKKYPEEAAEFKALISGK
Arabidopsis thalian    ALGEKEVEATRNNLGWPYEPFHVPEDVKSHWSRHTPEGAALEADWNAKFAAYEKKYPEEAAELKSIISGE
Sorghum bicolor        ALGAKEVEATRQNLGWPYEPFFVPEDVKSHWSRHTPQGAALEADWNAKFAEYEKKYAEDAATLKSLITGE
Zea mays               ALGAKEVEATRQNLGWPYDTFFVPEDVKSHWSRHTPEGAALEADWNAKFAEYEKKYADDAATLKSIITGE
Eutrema salsugineum    ALGEKEVEATRQNLGWPYEPFQVPEDVKSHWSRHTPKGAALEADWNAKFAAYEKKYAEEARELKSIMSGE
Polygonum tinctoriu    ALGAKEVEATRKNLGWPYEPFHVPEEVEKHWSRHVPEGAALEAEWNAKFAEYKSKYPTEAAELKSIIEGE
Oryza sativa Japoni    ALGTKEVEATRENLGWPYEPFFVPEDVKSHWSRHVPQGAAFEADWNAKFAEYEKKYPEDAATLKSIVSGE
Oryza sativa Indica    ALGTKEVEATRENLGWPYEPFFVPEDVKSHWSRHVPQGAAFEADWNAKFAEYEKKYPEDAATLKSIVSGE
Oryza brachyantha      ALGSKEVEATRNNLSWLHEPFHVPDEVKRHWSHHIDEGSSLEAEWNKKFGEYEKKYHQEAADLKSIISGE
Setaria italica        PLGTKEVEATRHNLGWPYEPFFVPEDVKSHWSHHVPQGDALEADWNAKFAEYEKKYPEDAATLKIIISGE
Physcomitrella pate    ALGGKEVEATRQNLGWPYEPFHVPDEVKSHWSKHTDLGTKYQSEWEAAVEEYAQKYPEEGAEFKQLISGE
Setaria italica        ALGSKEVEATRSNLQWFHEPFHVPDEVKRHWSHHIDEGASLEAEWNAKFAEYEKKYHQEAAELKSIISGE
Physcomitrella pate    ALGGKEVEATRQNLGWPYEPFHVPDEVKSHWSKHTDLGTKYQSEWEAAVEEYAQKYPEEGAEFKQLISGE
Oryza sativa Indica    ALGTKEVEATKNNLSWHHEPFHVPDEVKRHWSHHIDEGASLEAEWNKKIVEYEKKYRQEAAELKSIISGE
gi|115457470|ref|NP    ALGTKEVEATKNNLSWHHEPFHVPDEVKRHWSHHIDEGASLEAEWNKKIVEYEKKYRQEAAELKSIISGE
gi|475481099|gb|EMT    ALGTSEVEATRANLGWPYEPFFVPEDVKSHWSRHIPQGAALEADWDAKFAQYEQKYPEDAATLKSIITGE
gi|357167367|ref|XP    ALGPKEVEATRNNLLWLHEPFHVPDEVKRHWGHHIDEGASLEAEWNAKFSEYEKKYHQEAAELNSIISGE
gi|125547501|gb|EAY    ALGTKEVEATKNNLSWHHEPFHVPDEVKRHWSHHIDEGASLEAEWNKKIVEYEKKYRQEAAELKSIISGE
gi|514760053|ref|XP    PLGTKEVEATRHNLGWPYEPFFVPEDVKSHWSHHVPQGDALEADWNAKFAEYEKKYPEDAATLKIIISGE
gi|242072546|ref|XP    ALGSKEVEATRSNLGWLHEPFHVPDEVKRHWSHHIDDGASLEAEWNARFVEYDKKYHQEAAELKSIISGE
gi|514760061|ref|XP    PLGTKEVEATRHNLGWPYEPFFVPEDVKSHWSHHVPQGDALEADWNAKFAEYEKKYPEDAATLKIIISGE
gi|326515912|dbj|BA    ALGSKEVEATRKNLSWAHEPFHVPDEVKRHWAHHLDEGASLEAEWNAKFAEYEKKYQQEAAELNSIISGR
gi|474352176|gb|EMS    ALGTSEVEATRANLGWPYEPFFVPEDVKSHWSRHIPQGAALEADWDAKFAQYEQKYPEDAATLKSIITGE
gi|664903|emb|CAA86    ALGPKEAEATRKNLGWPYEPFHVPDDVKKHWSRHIAEGAALESAWNAKFAEFQKKFPEEAADLKSIITGE
gi|147788852|emb|CA    ALGAKEVDATRKNLGWPHEPFHVPEDVKKHWSRHIPDGAALEAEWNAKFAEYEKKYKEEAAELKSIITGE
gi|300140959|gb|EFJ    ALGSKEVDATRKNLGWPYEPFVVPEEVKSHWSKHTKLGADLEAQWEASLAAYEKKYPTEAAEFKQLISGE
gi|413953333|gb|AFW    ALGAKEVEATRQNLGWPYDTFFVPEDVKSHWSRHTPEGAALEADWNAKFAEYEKKYADDAATLKSIITGE
gi|414587814|tpg|DA    ALGSKEVEATRSNLGWFHEPFHVPDEVKRHWSRHIDEGASLEAEWNARFVEYEKKYNQEAAELKSIISGE
gi|162683807|gb|EDQ    ALGTKEVDATRQNLKWDHEPFFVPEEVKSHWSKHIDLGSKLESDWEAALEAYSKSYPEEAAEFKQLISLE
gi|125595949|gb|EAZ    ALGTKEVEATRENLGWPYEPFFVPEDVKSHWSRHVPQGAAFEADWNAKFAEYEKKYPEDAATLKSIVSGE
gi|222423758|dbj|BA    ALGEKEVEATRNNLGWPYEPFQVPDDVKSHWSRHTPEGATLESDWSAKFAAYEKKYPEEASELKSIITGE
Consensus              ALG KEVEATR NLGWPYEPFHVPEDVKSHWSRH P GAALEA WNAKFAEYEKKYPEEAAELKSIISGE

                                 500       510       520       530       540       550       560
1 [transketolase Sp    FPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMFGDFQKDTPEERNV
Hordeum vulgare sub    LPAGWADALPQYTTESPADATRNLSQQCLNALAKVVPGLIGGSADLASSNMTLLKMFGDFQKDTPEERNV
Picea sitchensis       LPEGWDKALPAYTTESPADATRILSQQCLNALAKVLPGLLGGSADLASSNMTLLKMFGDFQKDTPAERNV
Arabidopsis thalian    LPVGWEKALPTYTPDSPGDATRNLSQQCLNALAKAVPGFLGGSADLASSNMTMLKAFGNFQKATPEERNL
Sorghum bicolor        FPTGWADALPKYTPESPADATRNLSQQCLNALTNVVPGLIGGSADLASSNMTLLKMFGDFQKDTPEERNV
Zea mays               LPTGWVDALPKYTPESPGDATRNLSQQCLNALANVVPGLIGGSADLASSNMTLLKMFGDFQKDTAEERNV
Eutrema salsugineum    LPAGWEKALPTYTPDSPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTLLKAFGDFQKATPEERNL
Polygonum tinctoriu    FPAGWEKALPTYTPESPGDATRNLSQTCLNALAPVVPGLLGGSADLASSNMTLLKSFGDFQKDTPEERNV
Oryza sativa Japoni    LPAGWADALPKYTPESPADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTLLKMFGDFQKDTPEERNV
Oryza sativa Indica    LPAGWADALPKYTPESPADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTLLKMFGDFQKDTPEERNV
Oryza brachyantha      LPSGWDNALPTYTTENPADATRNLSQQCLNALAKAIPGFLGGSADLATSNMTLLKMFGDFQKDTPEERNI
Setaria italica        LPADWVDVLPKYTQASPADATRNLSQQCLNALAKVLPGLLGGSADLAPSNMTLLKMFGDFQRDTPEERNV
Physcomitrella pate    LPEGWDKALPSFTPEDAGDATRNLSQRCLNALARTLPGLIGGSADLASSNMTLMKMFGDFQKSTPAERNV
Setaria italica        LPSRWDNALPTYTPESSPDATRNLSQQCLNSLAKVIPGFLGGSADLATSNMTLLKMFGDFQRDTPEERNI
Physcomitrella pate    LPDGWDKALPSFTPEDAGDATRNLSQRCLNALARTLPGLIGGSADLASSNMTLMKMFGDFQKSTPAERNV
Oryza sativa Indica    LPSGWDNSLPKYTPENPADATRNLSQQCLNALAKAIPGFLGGSADLATSNMTLLKMFGDFQKDTPEERNI
gi|115457470|ref|NP    LPSGWDNSLPKYTPENPADATRNLSQQCLNALAKAIPGFLGGSADLATSNMTLLKMFGDFQKDTPEERNI
gi|475481099|gb|EMT    LPAGWADALPQYTTESPADATRNLSQQCLNALAKVVPGLIGGSADLASSNMTLLKMFGDFQKDTPEERNV
gi|357167367|ref|XP    LHAGWDKALPTYTPESPADATRNISQQCLNALAKVIPGFLGGSADLASSNMTLLKMFGDFQKDTPQERNI
gi|125547501|gb|EAY    LPSGWDNSLPKYTPENPADATRNLSQQCLNALAKAIPGFLGGSADLATSNMTLLKMFGDFQKDTPEERNI
gi|514760053|ref|XP    LPADWVDVLPKYTQASPADATRNLSQQCLNALAKVLPGLLGGSADLAPSNMTLLKMFGDFQRDTPEERNV
gi|242072546|ref|XP    LPSGWDNALPTYAPEISPDATRNLSQQCLNALAKVIPGFLGGSADLATSNMTLLKMFGDFQRDTPEERNI
gi|514760061|ref|XP    LPADWVDVLPKYTQASPADATRNLSQQCLNALAKVLPGLLGGSADLAPSNMTLLKMFGDFQRDTPEERNV
gi|326515912|dbj|BA    LPAGWEDALPKYTPESPADATRNLSQQCLNALAKVIPGFLGGSADLASSNMTLLKMFGDFQKDTPEERNI
gi|474352176|gb|EMS    LPAGWADALPQYTTESPADATRNLSQQCLNALAKVVPGLIGGSADLASSNMTLLKMFGDFQKDTPEERNV
gi|664903|emb|CAA86    LPTNWESIFPTYTPENPGLPTRTLSHQILNGLGDVLPGLLGGSADLTLSNMAFLKNSGDFQKKSPGERNV
gi|147788852|emb|CA    LPAGWDKALPTYTPESAADATRNLSQQCLNALAKVLPGLLGGSADLASSNMTLLKMFGDFQKGTPEERNV
gi|300140959|gb|EFJ    LPDGWDKALPVYTPESPGDATRNLSQTNLNALVKVLPGLIGGSADLASSNMTLLKAYGDFQKSTPEERNV
gi|413953333|gb|AFW    LPTGWVDALPKYTPESPGDATRNLSQQCLNALANVVPGLIGGSADLASSNMTLLKMFGDFQKDTAEERNV
gi|414587814|tpg|DA    LASGWENALRNYTPEMSPDATRNLSQQCLNALAKVIPGFLGGSSDLATSNMTLLKKFGDFQRDTPEERNI
gi|162683807|gb|EDQ    LPAGWDTSLPAFTPEDAGDATRNLSQRCLNALARTLPGLIGGSADLASSNMTLMKMFGDYQKSTPGERNV
gi|125595949|gb|EAZ    LPAGWADALPKYTPESPADATRNLSQQCLNALAKVVPGLLGGSADLASSNMTLLKMFGDFQKDTPEERNV
gi|222423758|dbj|BA    LPAGWEKALPTYTPESPGDATRNLSQQCLNALAKVVPGFLGGSADLASSNMTLLKAFGDFQKATPEERNL
Consensus              LPAGW  ALP YTPESPADATRNLSQQCLNALAKV PGLLGGSADLASSNMTLLKMFGDFQKDTPEERNV
```

Figure 1D continued



```
                               570       580       590       600       610       620       630
1 [transketolase Sp    RFGVREHGMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Hordeum vulgare sub    RFGVREHGMGAICNGIALHTPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Picea sitchensis       RFGVREHSMGAICNGIAHHGSGLIPYCATFFVFTDYMRAAIRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Arabidopsis thalian    RFGVREHGMGAICNGIALHSPGFIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Sorghum bicolor        RFGVREHGMGAIANGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Zea mays               RFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Eutrema salsugineum    RFGVREHGMGAICNGIAHHSPGLIPYCATFFVFTDYMRAAMRIAALSEGGVIYVMTHDSIGLGEDGPTHQ
Polygonum tinctoriu    RFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAIRISALSLARVIYVMTHDSIGLGEDGPTHQ
Oryza sativa Japoni    RFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEAGVIYVMTHDSIGLGEDGPTHQ
Oryza sativa Indica    RFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEAGVIYVMTHDSIGLGEDGPTHQ
Oryza brachyantha      RFGVREHAMGAISNGIALHSPGLIPYCATFFVFTDYMRAAIRLSALCGSGVIFVMTHDSIGLGEDGPTHQ
Setaria italica        RFGVREHGMGAICNGIALHSPGLVPYCATYFAFSDYMRAAMRVSALSEAGVIHIMTHDSIGVGKNGPTHQ
Physcomitrella pate    RFGVREHAMGAICNGMLQHLSGLIPYCATFFIFTDYLRAAMRISALAEAGVIYVMTHDSIGLGEDGPTHQ
Setaria italica        RFGVREHGMGAISNGIAVHSPGLIPYCATFFVFTDYMRAAIRLSALSESGVIFVMTHDSIGLGEDGPTHQ
Physcomitrella pate    RFGVREHAMGAICNGMLQHLSGLIPYCATFFIFTDYLRAAMRISALAEAGVIYVMTHDSIGLGEDGPTHQ
Oryza sativa Indica    RFGVREHAMGAISNGIALHSPGLIPYCATFFVFTDYMRAAIRLSALCGSRVIFVMTHDSIGLGEDGPTHQ
gi|115457470|ref|NP    RFGVREHAMGAISNGIALHSPGLIPYCATFFVFTDYMRAAIRLSALCGSRVIFVMTHDSIGLGEDGPTHQ
gi|475481099|gb|EMT    RFGVREHGMGAICNGIALHTPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
gi|357167367|ref|XP    RFGVREHAMGAICNAIALHSPGLIPYCSTFFVFTDYMRAPIRLSALCGSGVIYVMTHDSIGLGEDGPTHQ
gi|125547501|gb|EAY    RFGVREHAMGAISNGIALHSPGLIPYCATFFVFTDYMRAAIRLSALCGSRVIFVMTHDSIGLGEDGPTHQ
gi|514760053|ref|XP    RFGVREHGMGAICNGIALHSPGLVPYCATYFAFSDYMRAAMRVSALSEAGVIHIMTHDSIGVGKNGPTHQ
gi|242072546|ref|XP    RFGVREHGMGAISNGIAVHSPGLIPYCATFFVFTDYMRASIRLSALSESGVIFVMTHDSIGLGEDGPTHQ
gi|514760061|ref|XP    RFGVREHGMGAICNGIALHSPGLVPYCATYFAFSDYMRAAMRVSALSEAGVIHIMTHDSIGVGKNGPTHQ
gi|326515912|dbj|BA    RFGVREHAMDAICNGIAVHSPGLIPYCSTFFVFTDYMRASIRLSALSESGVIFVMTHDSIGLGEDGPTHQ
gi|474352176|gb|EMS    RFGVREHGMGAICNGIALHTPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
gi|664903|emb|CAA86    KFGAREHAMGSICNGLALHSPGLLPYCATYFVFTDYMRAAMRISALSKARVLYIMTHDSIGLGEDGPTHQ
gi|147788852|emb|CA    RFGVREHGMGAICNGIALHSLGLIPYCATFFVFTDYMRAAIRISALSEAGVIYLLTE.............
gi|300140959|gb|EFJ    RFGVREHGMGAISNGIALHGSGLIPYCATFFVFTDYMRGAIRLSALAEAGVIYVMTHDSIGLGEDGPTHQ
gi|413953333|gb|AFW    RFGVREHGMGAICNGIALHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
gi|414587814|tpg|DA    RFGVREHAMGAISNGIAVHSPGLIPYCATFFVFTDYMRASIRLSALSESGVIFVMTHDSVGLGEDGPTHQ
gi|162683807|gb|EDQ    RFGVREQAMGAICNGILLHRSGLIPYCATFFIFTDYLRSAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
gi|125595949|gb|EAZ    RFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEAGVIYVMTHDSIGLGEDGPTHQ
gi|222423758|dbj|BA    RFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ
Consensus              RFGVREHGMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALSEAGVIYVMTHDSIGLGEDGPTHQ

                               640       650       660       670       680       690       700
1 [transketolase Sp    PIEHLASFRAMPNILMLRPADGNETAGSYKVAVENRKTPSILALSRQKLPNLPGTSIEGVEKGGYTITDN
Hordeum vulgare sub    PIEHLVSFRAMPNMLMFRPADGKETAGAYKVAVLNRKRPSILALSRQKLPHLPGTSIEGVAKGGYTISDN
Picea sitchensis       PIEHLASFRAMPNVLMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLAGSSIEGVEKGGYTISCN
Arabidopsis thalian    PIEHLSSFRAMPNIMMFRPADGNETAGAYKIAVTKRKTPSVLALSRQKLPQLPGTSIESVEKGGYTISDN
Sorghum bicolor        PIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLPGTSIEGVEKGGYTISDN
Zea mays               PIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLPGTSIEGVEKGGYTISDN
Eutrema salsugineum    PVEHLSSFRAMPNILMFRPADGNETAG...........................TSIESVEKGGYTISDN
Polygonum tinctoriu    PIEHLASFRAMPNILMMRPADGNETAGSYRVAVLKKQTPSILALSRQKLPNLPGTSIEGVEKGGYTITDN
Oryza sativa Japoni    PIEHLVSFRAMPNILMLRPADGNETAGAYKIAVLNRKRPSVLALSRQKLAQLPGTSIEGVEKGGYIVSDN
Oryza sativa Indica    PIEHLVSFRAMPNILMLRPADGNETAGAYKIAVLNRKRPSVLALSRQKLAQLPGTSIEGVEKGGYIVSDN
Oryza brachyantha      PVEQLFSLRAMPNILMLRPADGKETSGAYKIAILNRKMPSVLALSRQKLQHLKGTSIEGVAKGGYIISDN
Setaria italica        PIEHLMSFRTMPNILVLRPADGNETAGAYKVAVLNRKRPSVLALSMQKLPHLPGTSVEGVEKGGYIISDN
Physcomitrella pate    PIEHLASFRAMPNMLMLRPADGNETSGAYKVAVLNRKRPSTLALSRQKLPNLPGTSIDNVQLGAYAISDD
Setaria italica        PVEQLFSLRAMPNILMLRPADGNETSGAYKVAVLNRKRPSVIALSRQKLPQLKGTSVDSVSKGGYIISDN
Physcomitrella pate    PIEHLASFRAMPNMLMLRPADGNETSGAYKVAVLNRKRPSTLALSRQKLPNLPGTSIDNVQLGAYAISDD
Oryza sativa Indica    PVEQLFSLRAMPNILMLRPADGNETSGAYKIAVLNRQRPSVLALSRQKLQQLKGTSVEGVAKGGYIISDN
gi|115457470|ref|NP    PVEQLFSLRAMPNILMLRPADGNETSGAYKIAVLNRQRPSVLALSRQKLQQLKGTSAEGVAKGGYIISDN
gi|475481099|gb|EMT    PIEHLVSFRAMPNMLMFRPADGKETAGAYKLAVLNRKRPSILALSRQKLPHLPGTSIEGVAKGGYTISDN
gi|357167367|ref|XP    PVEQLFSLRAMPNILVLRPADGNETSAAYRTAVVNRQRPSILAFSRQKLPQLAGTSVEGVAKGGYIISDN
gi|125547501|gb|EAY    PVEQLFSLRAMPNILMLRPADGNETSGAYKIAVLNRQRPSVLALSRQKLQQLKGTSVEGVAKGGYIISDN
gi|514760053|ref|XP    PIEHLMSFRTMPNILVLRPADGNETAGAYKVAVLNRKRPSVLALSMQKLPHLPGTSVEGVEKGGYIISDN
gi|242072546|ref|XP    PVEQLFSLRAMPNILMLRPADGNETSGAYKIAVLNRKRPSIIALSRQKLPQVKGTSVDAVSKGGYIISDN
gi|514760061|ref|XP    PIEHLMSFRTMPNILVLRPADGNETAGAYKVAVLNRKRPSVLALSMQKLPHLPGTSVEGVEKGGYIISDN
gi|326515912|dbj|BA    PVEQLFSLRAMPGILMLRPADGTETSAAYRIAVMNRKRPSILALSRQKLPQLEGTSVEGVAKGGYVISDN
gi|474352176|gb|EMS    PIEHLVSFRAMPNMLMFRPADGKETAGAYKVAVLNRKRPSILALSRQKLPHLPGTSIEGVAKGGYTISDN
gi|664903|emb|CAA86    PVEHLASFRAMPNILTLRPADGNETAGAYRAAVQNGERPSILVLARQKLPQLPGTSIEGVSKGGYVISDN
gi|147788852|emb|CA    ....................................RDPRFLLSHVKSFLSFLETSIAGVEKGGYAISDN
gi|300140959|gb|EFJ    PIEHLASFRAMPNIFMFRPADGNETAGSYKVAVLSRKTPSILALSRQKLPQLAGSSIDAVSKGAYILSDN
gi|413953333|gb|AFW    PIEHLVSFRAMPNILMLRPADGNETAGAYKVAVLNRKRPSILALSRQKLPHLPGTSIEGVEKGGYTISDN
gi|414587814|tpg|DA    PVEHLFSLRAMPNILMLRPADGNETSGAYKIAVLNRKRPSVIALSRQKLPQLRGTSVDAVSKGGYIVSDD
gi|162683807|gb|EDQ    PIEHLASFRAMPQMLMFRPADGNETSGAYKVAVQNRTRPSTLALSRQKLPNLAGTSIDNVALGGYKLSCN
gi|125595949|gb|EAZ    PIEHLVSFRAMPNILMLRPADGNETAGAYKIAVLNRKRPSVLALSRQKLAQLPGTSIEGVEKGGYIVSDN
gi|222423758|dbj|BA    PIEHIASFRAMPNTLMFRPADGNETAGAYKIAVTKRKTPSILALSRQKLPHLPGTSIEGVEKGGYTISDD
Consensus              PIEHL SFRAMPNILMLRPADGNETAGAYK AVLNRKRPS LALSRQKLP LPGTSIEGV KGGY ISDN
```

Figure 1D continued



```
                                710       720       730       740       750       760       770
1 [transketolase Sp   SSG..NKPDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQSDEYKESVLPSDVTARVSIE
Hordeum vulgare sub   SSG..NKPDYIIMSTGSEVEIAVKAAEELTKEGKTVRVVSFVCWELFDEQSDEYKESVLPEAVTARISIE
Picea sitchensis      ASN..NKPDVILMGTGSELEIAEKAAVTLRNEGKAVRVVSFVCWELFEEQTPEYKESVLPAAVAARVSVE
Arabidopsis thalian   STG..NKPDVILIGTGSELEIAAQAAEKLREQGKSVRVVSFVCWELFDEQSDAYKESVLPSDVSARVSIE
Sorghum bicolor       STG..NKPDLIVLSTGSELEIAAKAADELRKEGKTIRVVSFVSWELFEEQSDEYKESVLPEAVTARISIE
Zea mays              STG..NKPDIIVMGTGSELEIAAKAADELRKEGKTVRVVSFVSWELFDEQSDEYKESVLPAAVTARISIE
Eutrema salsugineum   STG..NKPDVILIGTGSELEIAAQAAEELRKEGKTVRVVSFVCWELFDEQSDAYKESVLPSDVSARVSIE
Polygonum tinctoriu   STG..NKPDVILIGTGSELEIAAKAGDELRKAGKAVRVVSFVCWELFEEQSDEYKESVLPAAVTARVSIE
Oryza sativa Japoni   STG..NKPDFIVMSTGSELEIVAKAADELRKEGKTVRVVSFVCWELFDEQSAEYKESVLPEAVTARVSLE
Oryza sativa Indica   STG..NKPDFIVMSTGSELEIVAKAADELRKEGKTVRVVSFVCWELFDEQSAEYKESVLPEAVTARVSLE
Oryza brachyantha     STG..NKPDLILIGTGSELEIVEKAADELRKEGKAVRVVSLVCWELFEEQSEEYKESVLPCEVTARISVE
Setaria italica       STG..NKPDLIMLSTGSELQIAVKAADELRKEGKAVRVVSLVCWELFEEQSDEYKESVLPESVTARISVE
Physcomitrella pate   SKD..NKPDLILMASGSELEIAAKAADVIRGEGKTVRVISFVCWELFEEQSAEYKESLLPSDVTARVSVE
Setaria italica       SSD..NKPDLIVIGTGSELEIAEKAADELRKEGRTVRVVSLVCWELFEEQPEKYKESVLPSEVTSRISIE
Physcomitrella pate   SKD..NKPDLILMASGSELEIAAKAADVIRGEGKTVRVISFVCWELFEEQSAEYKESLLPSDVTARVSVE
Oryza sativa Indica   SSG..NKPDIILIGTGSELEIVEKAADELRKDGKTVRVVSLVCWELFEEQSEKYKESVFPSEVTSRISVE
gi|115457470|ref|NP   SSG..NKPDIILIGTGSELEIVEKAADELRKDGKTVRVVSLVCWELFEEQSEKYKESVFPSEVTSRISVE
gi|475481099|gb|EMT   STD..NKPDYIIMSTGSEVEIAVKAAEELTKEGKTVRVVSFVCWELFDEQSDEYKESVLPEAVTARISIE
gi|357167367|ref|XP   SSG..NKPDLILIGTGSELEIAAKAADDLRKEGKTVRVVSLVCWELFEEQSEEYKDSVLPSEVTSRISIE
gi|125547501|gb|EAY   SSG..NKPDIILIGTGSELEIVEKAADELRKDGKTVRVVSLVCWELFEEQSEKYKESVFPSEVTSRISVE
gi|514760053|ref|XP   STG..NKPDLIMLSTGSELQIAVKAADELRKEGKAVRVVSLVCWELFEEQSDEYKESVLPESVTARISVE
gi|242072546|ref|XP   SSG..NKPDFILIGTGSELEIAENAAGELRNKGRTVRVVSLVCWELFEEQPEEYKESVLPNEVTSRISIE
gi|514760061|ref|XP   STG..NKPDLIMLSTGSELQIAVKAADELRKEGKAVRVVSLVCWELFEEQSDEYKESVLPESVTARISVE
gi|326515912|dbj|BA   SSG..NKPDLILIGTGSELEIAEKAAEQLRKEGKFVRVVSLVCWELFEEQSEEYKESVLPSEVTSRISIE
gi|474352176|gb|EMS   STD..NKPDFIIMSTGSEVEIAVKAAEELTKEGKTVRVVSFVCWELFDEQSDEYKESVLPEAVTARISIE
gi|664903|emb|CAA86   SRGGNSKPDVILIGTGSELEIAARAGDELRKEGKKVRVVSLVCWELFAEQSEKYRETVLPSGVTARVSVE
gi|147788852|emb|CA   SSG..NKPDVILIGTGSELEIAAKAGDELRKEGKTVRVVSLVSWELFDDQSDAYKESVLPAAVTARVSIE
gi|300140959|gb|EFJ   SSG..NKPDLILIGTGSEVEIAAKAADVIRKEGKSVRVVSFPCWELFEKQSDEYKESVFPSEVTARVSVE
gi|413953333|gb|AFW   STG..NKPDIIVMGTGSELEIAAKAADELRKEGKTVRVVSFVSWELFDEQSDEYKESVLPAAVTARISIE
gi|414587814|tpg|DA   SSG..NKPDFILIGTGSELEIAENAAGELRKKGRTVRVVSLVCWELFEEQPEEYKESVLPSEVTSRISIE
gi|162683807|gb|EDQ   STD..NKPDLIILATGSELEIAAKAADIIRGEGKTVRVVSFVCWELYEEQSAEYKESVLPKAVTARVSIE
gi|125595949|gb|EAZ   STG..NKPDFIVMSTGSELEIVAKAADELRKEGKTVRVVSFVCWELFDEQSAEYKESVLPEAVTARVSLE
gi|222423758|dbj|BA   SSG..NKPDVILIGTGSELEIAAQAAEVLRKDGKTVRVVSFVCWELFDEQSDEYKESVLPSDVSARVSIE
Consensus             S G  NKPD ILIGTGSELEIAAKAADELRKEGKTVRVVSFVCWELFEEQS EYKESVLP  VTARISIE

                                780       790       800       810       820
1 [transketolase Sp   AGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC..
Hordeum vulgare sub   AGSTLGWQKYVGSKGKTIGIDKFGASAPAPKIYKEYGITAENVIAAAKSL...
Picea sitchensis      AGSTFGWERYLGSKGKAVGIDRFGASAPIGILYKEFGLTAENVVATAKQVLA.
Arabidopsis thalian   AGSTFGWGKIVGGKGKSIGIDTFGASAPAGKLYKEFGITIEAMVEAAKSLI..
Sorghum bicolor       AGSTLGWQKYIGAQGKAIGIDKFGASAPAGKIYKEYGITVEGVIAAAKSF...
Zea mays              AGSTLGWQKYVGAQGKAIGIDKFGASAPAGTIYKEYGITVESIIAAAKSF...
Eutrema salsugineum   AGSTFGWEKIVGGKGKSIGIDSFGASAPAGKLYKEFGFTIEAVVAAAKSLI..
Polygonum tinctoriu   AGSTFGWEKIVGSKGKAIGIDRFGASAPAGRIYKEYGITVEAVIEAANSVS..
Oryza sativa Japoni   AGSTLGWQKYVGSKGKAIGIDKFGASAPAGKIYQEYGITAENVIATAKSL...
Oryza sativa Indica   AGSTLGWQKYVGSKGKAIGIDKFGASAPAGKIYQEYGITAENVIATAKSL...
Oryza brachyantha     AGVTFGWEKYVGQKGKAIGVDRFGSSAPAGKIYKELGLTVENVIETAKSL...
Setaria italica       AGCTLGWQKYVGTKGKIIGIDRFGASAPGEKIFEEYGITVGSVITAARSL...
Physcomitrella pate   AGATFGWERFVGLEGRAIGVDHFGASAPANILYKEFGITVDNVVAKAKEVMA.
Setaria italica       AGVTFGWEKYIGQKGKAIGIDSFGISAPAGKIFKELGLTVENVIAAAKAL...
Physcomitrella pate   AGATFGWERFVGLEGRAIGVDHFGASAPANILYKEFGITVDNVVAKAKEVMA.
Oryza sativa Indica   AGVTFGWEKYIGEKGKAIGVDRFGSSAPAGKIYKELGLTVENVIATAKSL...
gi|115457470|ref|NP   AGVTFGWEKYIGEKGKAIGVDRFGSSAPAGKIYKELGLTVENVIATAKSL...
gi|475481099|gb|EMT   AGSTLGWQKYVGSKGKTIGIDKFGASAPAPKIYKEYGITAENVIAAAKSL...
gi|357167367|ref|XP   AGVTLGWEKYIGQKGKAIGIDRFGSSAPAGKIYKELGLTVEHIIATAKSI...
gi|125547501|gb|EAY   AGVTFGWEKYIGEKGKAIGVDRFGSSAPAGKIYKELGLTVENVIATAKSL...
gi|514760053|ref|XP   AGCTLGWQKYVGTKGKIIGIDRFGASAPGEKIFEEYGITVGSVITAARSL...
gi|242072546|ref|XP   AGVTFGWEKYIGQKGKAIGIDRFGVSAPAGKIYEELGLTVENVIAAAEAL...
gi|514760061|ref|XP   AGCTLGWQKYVGTKGKIIGIDRFGASAPGEKIFEEYGITVGSVITAARSL...
gi|326515912|dbj|BA   AGVTLGWEKYIGQKGKAIGIDRFGSSAPAGKIYKEFGLTVDNLIATAQSL...
gi|474352176|gb|EMS   AGSTLGWQKYVGSKGKTIGIDKFGASAPAPKIYKEYGITAESVIAAAKSL...
gi|664903|emb|CAA86   AGSTFGWERFIGPKGKAVGIDRFGASAPAERLFKEFGITVEAVVAAAKEIC..
gi|147788852|emb|CA   AGSTFGWEKIVGSKGKAIGIDRFGASAPAGKIYKEFGITVEAVIEAAKAVW..
gi|300140959|gb|EFJ   AGSTFGWGKYVGSRGVAIGLDRFGASAPAGTLYKEFGFTVDNVVAKAKSLL..
gi|413953333|gb|AFW   AGSTLGWQKYVGAQGKAIGIDKFGASAPAGTIYKEYGITVESIIAAAKSF...
gi|414587814|tpg|DA   AGVTFGWEKYIGQKGKAIGIDRFGASAPAGKIYKELGLTVEHVIAAAEAL...
gi|162683807|gb|EDQ   AGSTFGWERYVGLKGKSIGIDRFGASAPAGTLYKEFGFTVENLLAKAKEVMGA
gi|125595949|gb|EAZ   AGSTLGWQKYVGSKGKAIGIDKFGASAPAGKIYQEYGITAENVIATAKSL...
gi|222423758|dbj|BA   AASTFGWGKIVGGKGKSIGINSFGASAPAPLLYKEFGITVEAVVDAAKSFF..
Consensus             AGSTFGWEKYVG KGKAIGIDRFGASAPAGKIYKE GITVENVIAAAKSL
```

X non conserved

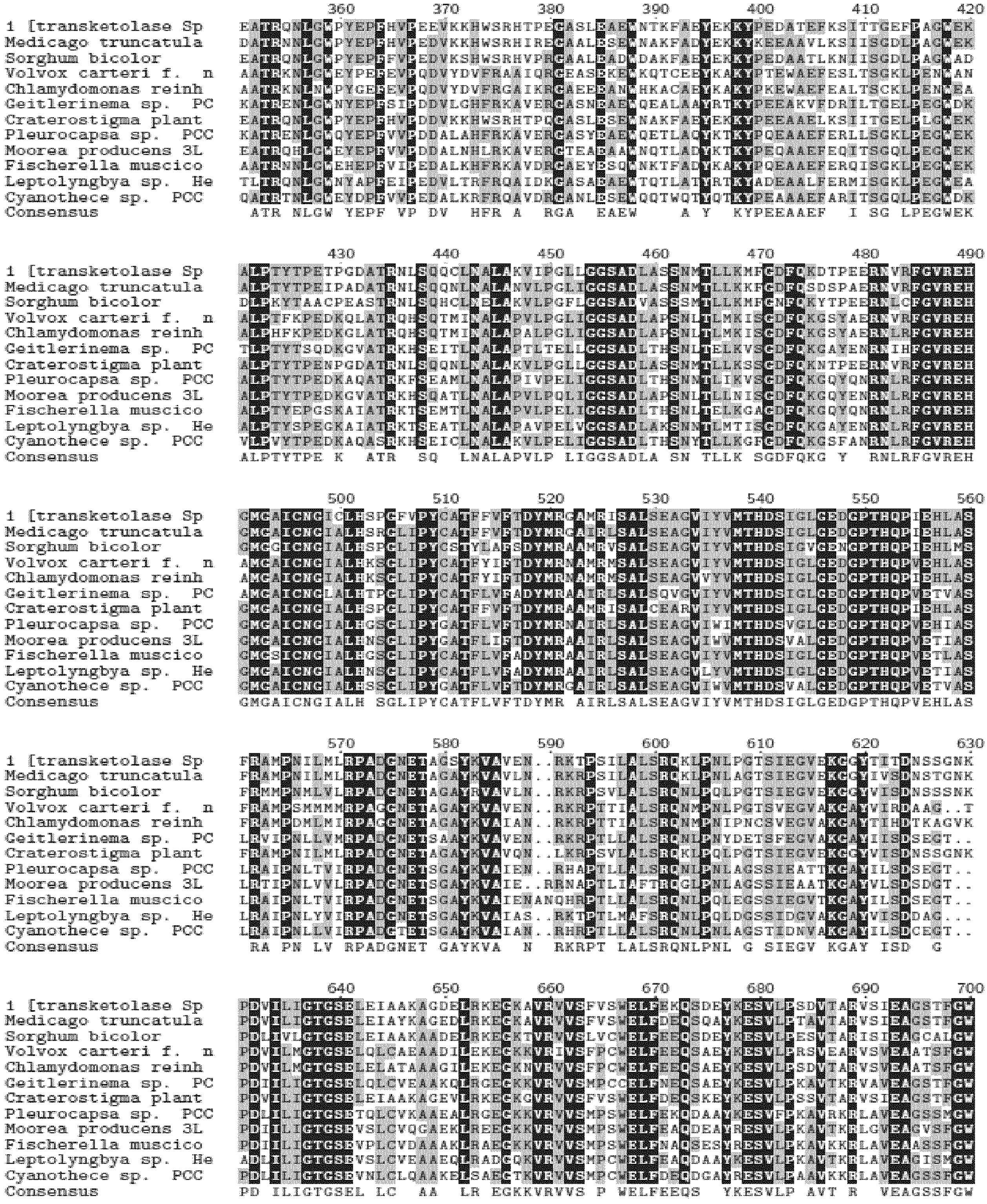
X ≥ 50% conserved

X all match

Figure 1E

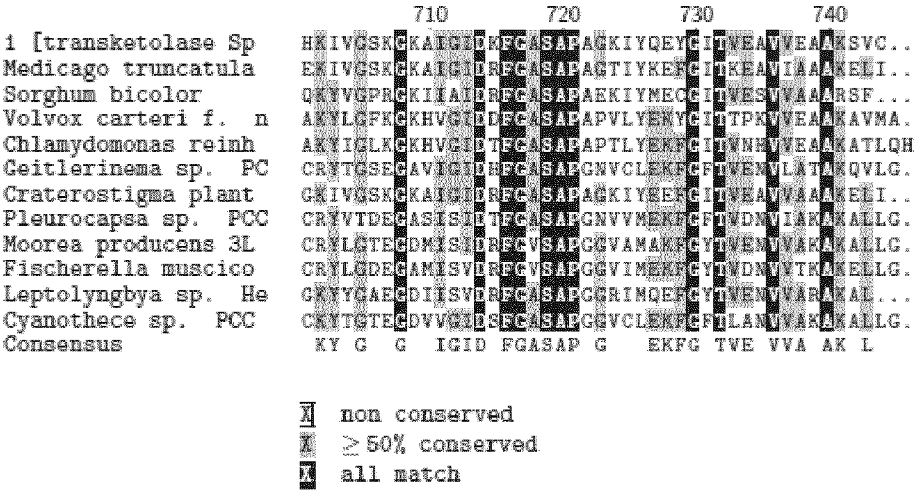
```
                             10        20        30        40        50        60        70
1 [transketolase Sp  MAASSSLSTLSHHQTLLSHPKTHLPTTPASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASAVVRAAAVE
Medicago truncatula  ......................................................................
Sorghum bicolor      ......................................................................
Volvox carteri f.  n ..........................MQTMMKQRAPAAAGKQAQALPRVAPKIGRARSIVLAQAAPATAK
Chlamydomonas reinh  ..........................MQTMLKQRCQPAVGKQAKAVPAVAPKVGRARNVVVAQAAPAAAK
Geitlerinema sp.  PC ...............................................................MAVAT..
Craterostigma plant  ......................................................................
Pleurocapsa sp.  PCC ...............................................................MVVAT..
Moorea producens 3L  ...............................................................MAVAT..
Fischerella muscico  ...............................................................MAVAT..
Leptolyngbya sp.  He ...............................................................MAVAT..
Cyanothece sp.  PCC  ...............................................................MAVATT.
Consensus                                                                             A AT

                             80        90       100       110       120       130       140
1 [transketolase Sp  ALESTDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEIMRYNPKNPYWFNRDRFVLS
Medicago truncatula  ......................................................................
Sorghum bicolor      ........................................MGIVNYLYDETN..................
Volvox carteri f.  n VDKPAISRDLVDKCINAIRFLAIDAVNKSKSGHPGMPMGCAPMGYLLWNEVMKYNPKNPDWFNRDRFVLS
Chlamydomonas reinh  AAAPSISRDEVEKCINAIRFLAIDAINKSKSGHPGMPMGCAPMGYVLWNEVMKYNPKNPDFFNRDRFVLS
Geitlerinema sp.  PC .......QSLDELCINAIRFLAIDAVEKAKSGHPGLPMGAAPMAFVLWDRFMRFNPKNPTWFNRDRFVLS
Craterostigma plant  ......................................................................
Pleurocapsa sp.  PCC .......QSLEERCINSIRFLAIDAVEKAKSGHPGLPMGAAPMAFVLWDRFMRYNPKNPKWFNRDRFVLS
Moorea producens 3L  .......QSLEELCINSIRFLAIDAVEKAKSGHPGLPMGAAPMAFVLWDRFMRYNPKNPNWFNRDRFVLS
Fischerella muscico  .......QTLEELCINSIRFLAIDAVEKAKSGHPGLPMGAAPMAFVLWDRFMRFNPKNPKWFNRDRFVLS
Leptolyngbya sp.  He .......QSLETQCINAIRFLAIDAVEKAKSGHPGLPMGAAPMAFVLWDQFMRFNPKNPQWFNRDRFVLS
Cyanothece sp.  PCC  .......QSLEELCINSIRFLAIDAVEKAKSGHPGLPMGAAPMAFVLWDQFMRFNPKNPKWFNRDRFVLS
Consensus                     L E CIN IRFLAIDAVEKAKSGHPGLPMG APM  VLWD  MR NPKNP WFNRDRFVLS

                            150       160       170       180       190       200       210
1 [transketolase Sp  AGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEVTTGPLGQGIANAVGLALAEKH
Medicago truncatula  ....................MVEDLKQFRQWESRTPGHPENFETPGIEVTTGPLGQGIANAVGLALAEKH
Sorghum bicolor      ..............SSFPKEKVQELKQFSQLGSKTPGHPENFVTPGVEVTTGPLGQGITNAVGLALAEKH
Volvox carteri f.  n AGHGSMFQYAMLHLTGYDSVPLNEVKQFRQWGSLTPGHPENFVTPGVEVTTGPLGQGICNAVGLAVAEAH
Chlamydomonas reinh  AGHGSMFQYSMMHLTGYDSVPLDQIKQFRQWNSLTPGHPENFVTPGVEVTTGPLGQGICNAVGLAVAEAH
Geitlerinema sp.  PC AGHGCMLQYALLHLTGYDSVTLDDIKSFRQWGSRTPGHPENFETLGVEVTTGPLGQGIANAVGLAMAEAH
Craterostigma plant  ......................................................................
Pleurocapsa sp.  PCC AGHGCMLQYALLYLTGYDSVTLDEIKNFRQWGSKTPGHPENHITPGVEVTTGPLGQGIANAVGLAIAEAH
Moorea producens 3L  AGHGCMLQYALLYLTGYDGVTIDDLQQFRQWESVTPGHPENFETDGVEVTTGPLGQGIANGVGLAMAEAH
Fischerella muscico  AGHGCMLQYALLYLTGYDSVTLDDIKQFRQWGSKTPGHPENFETPGVEVTTGPLGQGIANAVGLAMAEAH
Leptolyngbya sp.  He AGHGCMLQYALLHLTGYDSVKLEDIKNFRQWGATTPGHPENFETLGVEVTTGPLGQGICNGVGLAMAEAH
Cyanothece sp.  PCC  AGHGCMLQYALLYLTGYDSVGLEDIKNFRQWGSRTPGHPENFETPGVEVTTGPLGQGIANAVGLAIAEAH
Consensus            AGHGCMLQYALL LTGYDSV L D K FRQWGS TPGHPENFETPGVEVTTGPLGQGIANAVGLA AEAH

                            220       230       240       250       260       270       280
1 [transketolase Sp  LAARFNKPDAE.IVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAFYDDNHISIDGDTAIAFTESV
Medicago truncatula  LAARFNKPDNE.IIDHYTYCILGDGCQMEGIANEACSLAGHWGLGKLIAFYDDNHISIDGDTEIAFTESV
Sorghum bicolor      LAGRFNKPDSE.VVNHYTYVILGDGCQMEGISSEACSLAGHWGLGKLIAFYDDNHISVDGDTEIAFTEDV
Volvox carteri f.  n LAARFNKPDCKPIVDHYTYCILGDGCMMEGISNEACSLAGHWGLGKLIALYDDNKISIDGHTDISFTEDV
Chlamydomonas reinh  LAARFNKPDVKPIVDHYTYCILGDGCMMEGISNEACSLAGHWGLGKLIALYDDNKISIDGHTDISFTEDV
Geitlerinema sp.  PC LAARFNKPDST.IVDHYTYVILGDGCNMEGISGEACSLAGHLGLGKLIALYDDNHISIDGSTDVAFTEDV
Craterostigma plant  .............VDHYTYCILGDGCQMEGVSNEACSIAAHWGLGKLIALYDDNHISIDGDTDIAFTEDV
Pleurocapsa sp.  PCC LAAKFNKPDCK.LVDHYTYVILGDGCNMEGISGEACSLAGHLGLGKLIALYDDNHISIDGSTDIAFTEDV
Moorea producens 3L  LAARFNKPDAT.IVDHYTYVILGDGCNMEGVSGEACSLAGHLGLGKLIALYDDNHISIDGSTDLAFTEDV
Fischerella muscico  LAAKFNKPDAK.IVDHYTYVILGDGCNMEGVSGEACSLAGHLGLGKLIALYDDNHISIDGSTDIAFTEDV
Leptolyngbya sp.  He LAARFNKPDCT.LVDHYTYVILGDGCNMEGVSGEACSLAGHLGLGKLIALYDDNHISIDGSTDIAFTEDV
Cyanothece sp.  PCC  LAAKFNKPDST.LVDHYTYVILGDGCNMEGVSGEACSLAGHLGLGKLIALYDDNHISIDGSTDIAFTEDV
Consensus            LAARFNKPD   IVDHYTYVILGDGC MEGIS EACSLAGH GLGKLIALYDDNHISIDG TDIAFTEDV

                            290       300       310       320       330       340       350
1 [transketolase Sp  DLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFGSPNKSNSYSVHGSALGSKEV
Medicago truncatula  EKRFEGLGWHVIWVKNGNTGFDEIRAAIKEAKAVKDRPTLIKFTTTIGYGSPNKSNSYSVHGSALGAKEV
Sorghum bicolor      IARFEAYGWHTLWVKNGNTGYDDIREAIKEAKAVTDKPTLVKVTTTIGYGSPNKANSYSVHGCALGTKEV
Volvox carteri f.  n AKRYEALGWHVIHVINGNTDIDGLRAAITQAKAVKDKPTLIKVSTLIGYGSPNKADTHDVHGAPLGPEET
Chlamydomonas reinh  AKRYEALGWHVIHVINGNTDVDGLRAAIAQAKAVKDKPTLIKVSTLIGYGSPNKADSHDVHGAPLGPDET
Geitlerinema sp.  PC SKRFEAYGWHVLHVENGNTDLAAIQKAIEEAKAVTDKPSMIKVTTTIGYGSPNKADTAGVHGAALGGDEV
Craterostigma plant  DKRFEALGWHVIWVKNGNNGYDKIRAAIKEAQAVKDKPTMIKITTTIGFGSPNKSNSYSVHGSALGAKEV
Pleurocapsa sp.  PCC SKRFEAYGWHVLHVEDGNKDIESITKAIEAAKAVTDKPSMIKVTTTIGYGSPNKANTAGVHGAALGSDEV
Moorea producens 3L  GKRFEAYGWHVQHVESGNTDLEAINDAIAAAKAVTDKPSLIKVTTTIGYGSPNKANSHAVHGAALGGDEV
Fischerella muscico  GKRFEAYGWHVQAVSDGDTNLEAIHKAIEAAKAVTDKPSLIKVRTTIGYGSPNKANTAGVHGAALGGDEV
Leptolyngbya sp.  He SKRFEAYGWHVLHVEDGNTDLSAIAKAIEEAKQVTDKPTMIKVTTTIGYGSPNKADTAGVHGAALGDDEI
Cyanothece sp.  PCC  GKRFEAYGWHVQHVENGNTDLAGIAQAIAAAKAVTERPSLIKVTTTIGFGSPNKANTADVHGAALGADEV
Consensus             KRFEAYGWHV HV NGNTD   I  AI  AKAVTDKPTLIKVTTTIGYGSPNKAN   VHGAALG DEV
```

Figure 1E continued



```
                                   360          370          380          390          400          410          420
1 [transketolase Sp     EATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYPEDATEFKSITTGEFPAGWEK
Medicago truncatula     DATRNNLGWPYEPFHVPEDVKKHWSRHIREGAALESEWNAKFADYEKKYKEEAAVLKSIISGDLPAGWEK
Sorghum bicolor         EATRQNLGWPYEPFFVPEDVKSHWSRHVPRGAALEADWDAKFAEYEKKYPEDAATLKNIISGDLPAGWAD
Volvox carteri f.  n    AATRKNLGWEYPEFEVPQDVYDVFRAAIQRGEASEKEWKQTCEEYKAKYPTEWAEFESLTSGKLPENWAN
Chlamydomonas reinh     AATRKNLNWPYGEFEVPQDVYDVFRGAIKRGAEEEANWHKACAEYKAKYPKEWAEFEALTSCKLPENWEA
Geitlerinema sp.  PC    KATRENLGWNYEPFSIPDDVLGHFRKAVERGASNEAEWQEALAAYRTKYPEEAKVFDRILTGELPEGWDK
Craterostigma plant     EATRQNLGWPYEPFHVPDDVKKHWSRHTPQGASLESEWNAKFAEYEKKYPEEAAELKSIITGELPLGWEK
Pleurocapsa sp.  PCC    KATRENLGWQYEPFVVPDDALAHFRKAVERGASYEAEWQETLAQYKTKYPQEAAEFERLLSGKLPEGWEK
Moorea producens 3L     EATRQHLGWEYEPFVVPDDALNHLRKAVERGTEAEAAWNQTLADYKTKYPEQAAEFEQITSGQLPEGWEK
Fischerella muscico     AATRNNLGWEHEPFVIPEDALKHFRKAVDRGAEYESQWNKTFADYKAKYPQEAAEFERQISGKLPEGWEK
Leptolyngbya sp.  He    TLTRQNLGWNYAPFEIPEDVLTRFRQAIDKGASAEAEWTQTLATYRTKYADEAALFERMISGKLPEGWEA
Cyanothece sp.  PCC     QATRTMLGWEYDPFVVPEDALKRFRQAVDRGANLESEWQQTWQTYQTKYPEAAAEFARITSGQLPEGWDK
Consensus                ATR NLGW YEPF VP DV  HFR A  RGA  EAEW    A Y  KYPEEAAEF  I SG LPEGWEK

                                   430          440          450          460          470          480          490
1 [transketolase Sp     ALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMFGDFQKDTPEERNVRFGVREH
Medicago truncatula     ALPTYTPEIPADATRNLSQQNLNALANVLPGLIGGSADLASSNMTLLKKFGDFQSDSPAERNVRFGVREH
Sorghum bicolor         DLPKYTAACPEASTRNLSQHCLNELAKVLPGFLGGSADVASSSMTLLKMFGNFQKYTPEERNLCFGVREH
Volvox carteri f.  n    ALPTFKPEDKQLATRQHSQTMINALAPVLPGLIGGSADLAPSNLTLMKISGDFQKGSYAERNVRFGVREH
Chlamydomonas reinh     ALPHFKPEDKGLATRQHSQTMINALAPALPGLIGGSADLAPSNLTLMKISGDFQKGSYAERNLRFGVREH
Geitlerinema sp.  PC    TLPTYTSQDKGVATRKHSEITLNALAPTLTELLGGSADLTHSNLTELKVSGDFQKGAYENRNIHFGVREH
Craterostigma plant     ALPTYTPENPGDATRNLSQQNLNALAKVLPGLLGGSADLASSNMTLLKSSGDFQKNTPEERNVRFGVREH
Pleurocapsa sp.  PCC    ALPTYTPEDKAQATRKFSEAMLNALAPIVPELIGGSADLTHSNNTLIKVSGDFQKGQYQNRNLRFGVREH
Moorea producens 3L     ALPTYTPEDKGVATRKHSQATLNALAPVLPQLIGGSADLAPSNLTLLNISGDFQKGQYENRNLRFGVREH
Fischerella muscico     ALPTYEPGSKAIATRKTSEMTLNALAPVLPELIGGSADLTHSNLTELKGAGDFQKGQYQNRNLRFGVREH
Leptolyngbya sp.  He    ALPTYSPEGKAIATRKTSEATLNALAPAVPELVGGSADLAKSNNTLMTISGDFQKGAYENRNLRFGVREH
Cyanothece sp.  PCC     VLPVYTPEDKAQASRKHSEICLNALAKVLPELIGGSADLTHSNYTLLKGFGDFQKGSFAMRNLRFGVREH
Consensus               ALPTYTPE K  ATR  SQ  LNALAPVLP LIGGSADLA SN TLLK SGDFQKG Y  RNLRFGVREH

                                   500          510          520          530          540          550          560
1 [transketolase Sp     GMGAICNGICLHSPGFVPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQPIEHLAS
Medicago truncatula     GMGAICNGIALHSRGLIPYCATFFVFTDYMRGAIRLSALSEAGVIYVMTHDSIGLGEDGPTHQPIEHLAS
Sorghum bicolor         GMGGICNGIALHSPGLIPYCSTYLAFSDYMRAAMRVSALSEAGVIYVMTHDSIGVGENGPTHQPIEHLMS
Volvox carteri f.  n    AMGAICNGIALHKSGLIPYCATFYIFTDYMRNAMRMSALSEAGVIYVMTHDSIGLGEDGPTHQPVEHLAS
Chlamydomonas reinh     AMGAICNGIALHKSGLIPYCATFYIFTDYMRNAMRNSALSEAGVVYVMTHDSIGLGEDGPTHQPIEHLAS
Geitlerinema sp.  PC    AMGAICNGLALHTPGLIPYCATFLVFADYMRAAIRLSALSQVGVIYVMTHDSIGLGEDGPTHQPVETVAS
Craterostigma plant     GMGAICNGIALHSPGLIPYCATFFVFTDYMRAAMRISALCEARVIYVMTHDSIGLGEDGPTHQPIEHLAS
Pleurocapsa sp.  PCC    GMGAICNGIALHGSGLIPYGATFLVFTDYMRNAIRLSALSEAGVIWIMTHDSVGLGEDGPTHQPVEHIAS
Moorea producens 3L     GMGAICNGIALHNSGLIPYGATFLIFTDYMRAAIRLSALSEAGVIWVMTHDSVALGEDGPTHQPVETIAS
Fischerella muscico     GMGSICNGIALHGSGLIPYCATFLVFADYMRAAIRLSALSEAGVIYVMTHDSIGLGEDGPTHQPVETLAS
Leptolyngbya sp.  He    GMGAICNGIALHNSGLIPYCATFLVFADYMRAAIRLSALSEAGVLYVMTHDSIGLGEDGPTHQPVETIAS
Cyanothece sp.  PCC     GMGAICNGIALHSSGLIPYGATFLVFTDYMRGAIRLSALSEAGVIWVMTHDSVALGEDGPTHQPVETVAS
Consensus               GMGAICNGIALH SGLIPYCATFLVFTDYMR AIRLSALSEAGVIYVMTHDSIGLGEDGPTHQPVEHLAS

                                   570          580          590          600          610          620          630
1 [transketolase Sp     FRAMPNILMLRPADGNETAGSYKVAVEN..RKTPSILALSRQKLPNLPGTSIEGVEKGGYTITDNSSGNK
Medicago truncatula     FRAMPNILMLRPADGNETAGAYKVAVLN..RKRPSILALSRQKLPNLAGTSIEGVEKGGYIVSDNSTGNK
Sorghum bicolor         FRMMPNMLVLRPADGNETAGAYRVAVLN..RKRPSVLALSRQNLPQLPGTSIEGVEKGGYVISDNSSSNK
Volvox carteri f.  n    FRAMPSMMMMRPAGGNETAGAYKVAVEN..RKRPTTIALSRQNMPNLPGTSVEGVAKGAYVIRDAAG..T
Chlamydomonas reinh     FRAMPDMLMIRPAGGNETAGAYKVAIAN..RKRPTTIALSRQNMPNIPNCSVEGVAKGAYTIHDTKAGVK
Geitlerinema sp.  PC    LRVIPNLLVMRPADGNETSAAYKVAVEN..RKRPTLLALSRQNLPNYDETSFEGVAKGAYIISDSEGT..
Craterostigma plant     FRAMPNILMLRPADGNETAGAYKVAVQN..LKRPSVLALSRQKLPQLPGTSIEGVEKGGYVISDNSSGNK
Pleurocapsa sp.  PCC    LRAIPNLTVIRPADGNETSGAYKVAIEN..RHAPTLLALSRQNLPNLAGSSIEATTKGAYILSDSEGT..
Moorea producens 3L     LRTIPNLVVLRPADGNETSGAYKVAIE..RRNAPTLIAFTRQGLPNLAGSSIEAATKGAYVLSDSDGT..
Fischerella muscico     LRAIPNLTVIRPADGNETSGAYKVAIENANQHRPTLLALSRQNLPQLEGSSIEGVTKGAYILSDSEGT..
Leptolyngbya sp.  He    LRAIPNLYVIRPADGNETSGAYKVAIAS..RKTPTLMAFSRQNLPQLDGSSIDGVAKGAYVISDDAG...
Cyanothece sp.  PCC     LRAIPNLLVIRPADGTETSGAYKVAIAN..RHRPTLLALSRQNLPNLAGSTIDNVAKGAYILSDCEGT..
Consensus                RA PN LV RPADGNET GAYKVA  N  RKRPT LALSRQNLPNL G SIEGV KGAY ISD   G

                                   640          650          660          670          680          690          700
1 [transketolase Sp     PDVILIGTGSELEIAAKAGDELRKEGKAVRVVSFVSWELFEKQSDEYKESVLPSDVTARVSIEAGSTFGW
Medicago truncatula     PDVILIGTGSELEIAYKAGEDLRKEGKAVRVVSFVSWELFDEQSQAYKESVLPTAVTARVSIEAGSTFGW
Sorghum bicolor         PDLIVLGTGSELEIAAKAADELRKEGKTVRVVSLVCWELFEEQSDEYKESVLPESVTARISIEAGCALGW
Volvox carteri f.  n    PDVILMGTGSELQLCAEAADILEKEGKKVRIVSFPCWELFEEQSAEYKESVLPRSVEARVSVEAATSFGW
Chlamydomonas reinh     PDVILMGTGSELELATAAAGILEKEGKNVRVVSFPCWELFEEQSAEYKESVLPSDVTARVSVEAATSFGW
Geitlerinema sp.  PC    PDIILIGTGSELQLCVEAAKQLRGEGKKVRVVSMPCCELFNEQSAEYKESVLPKAVTKRVAVEAGSTFGW
Craterostigma plant     PDVILIGTGSELEIAAKAGEVLRKEGKGVRVVSFVSWELFDEQSKEYKESVLPSSVTARVSIEAGSTFGW
Pleurocapsa sp.  PCC    PDLILIGTGSETQLCVKAAEALRGEGKKVRVVSMPSWELFEKQDAAYKESVFPKAVRKRLAVEAGSSMGW
Moorea producens 3L     PDIILIGTGSEVSLCVQGAEKLREEGKKVRVVSMPSWELFEAQDEAYRESVLPKAVTKRLGVEAGVSFGW
Fischerella muscico     PDIILIGTGSEVPLCVDAAAKLRAEGKKVRVVSMPSWELFNAQSESYRESVLPKAVKKRLAVEAASSFGW
Leptolyngbya sp.  He    ADLILIGTGSEVSLCVEAAEQLRADGQKVRVVSMPCWELFEAQDAAYKESVLPKAVTKRLAVEAGISMGW
Cyanothece sp.  PCC     PDLILIGTGSEVNLCLQAAKELSAEGTKVRVVSMPCWELFDEQDGAYRESVLPAAVKKRLAVEAGSSFGW
Consensus               PD ILIGTGSEL LC  AA  LR EGKKVRVVS P WELFEEQS  YKESVLP AVT R  VEAGSSFGW
```

Figure 1E continued



```
                                 710       720       730       740
1 [transketolase Sp      HKIVGSKGKAIGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC..
Medicago truncatula      EKIVGSKGKAIGIDRFGASAPAGTIYKEFGITKEAVIAAAKELI..
Sorghum bicolor          QKYVGPRGKIIAIDRFGASAPAEKIYMECGITVESVVAAARSF...
Volvox carteri f.  n     AKYLGFKGKHVGIDDFGASAPAPVLYEKYGITTPKVVEAAKAVMA.
Chlamydomonas reinh      AKYIGLKGKHVGIDTFGASAPAPTLYEKFGITVNHVVEAAKATLQH
Geitlerinema sp.  PC     CRYTGSEGAVIGIDHFGASAPGNVCLEKFGFTVENVLATAKQVLG.
Craterostigma plant      GKIVGSKGKAIGIDRFGASAPAGKIYEEFGITVEAVVAAAKELI..
Pleurocapsa sp.  PCC     CRYVTDEGASISIDTFGASAPGNVVMEKFGFTVDNVIAKAKALLG.
Moorea producens 3L      CRYLGTEGDMISIDRFGVSAPGGVAMAKFGYTVENVVAKAKALLG.
Fischerella muscico      CRYLGDEGAMISVDRFGVSAPGGVIMEKFGYTVDNVVTKAKELLG.
Leptolyngbya sp.  He     GKYYGAEGDIISVDRFGASAPGGRIMQEFGYTVENVVARAKAL...
Cyanothece sp.  PCC      CKYTGTEGDVVGIDSFGASAPGGVCLEKFGFTLANVVAKAKALLG.
Consensus                 KY G  G  IGID FGASAP G   EKFG TVE VVA AK L
```

X non conserved

X ≥ 50% conserved

X all match

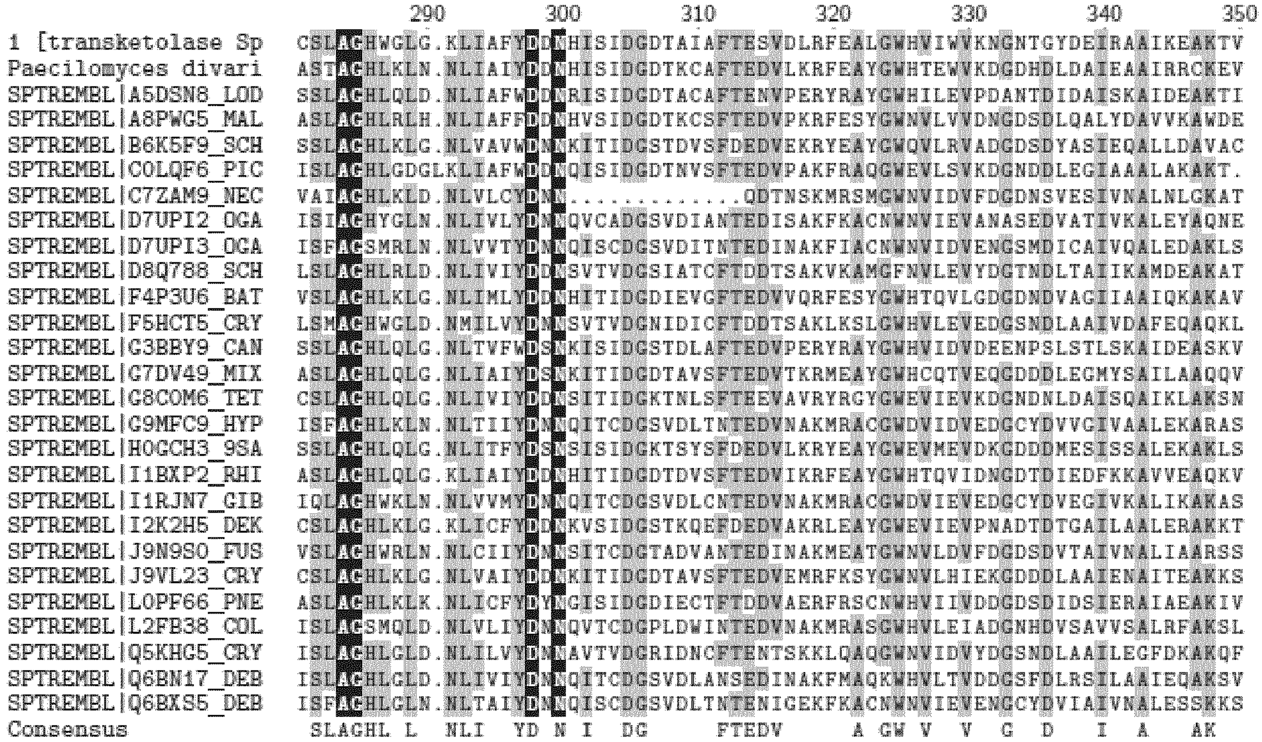
Figure 1F

```
                                10        20        30        40        50        60        70
1 [transketolase Sp    MAASSSLSTLSHHQTLLSHPKTHLPTTPASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASAVVRAAAV.
Paecilomyces divari    .................................................MGYTD................
SPTREMBL|A5DSN8_LOD    ..................................................MAPS................
SPTREMBL|A8PWG5_MAL    ...............................................MVFTPSE................
SPTREMBL|B6K5F9_SCH    ...............................................MVSTYTD................
SPTREMBL|C0LQF6_PIC    .....................................................M................
SPTREMBL|C7ZAM9_NEC    ..............................................MVPANS..................
SPTREMBL|D7UPI2_OGA    ..............................................MRIARAVSTSDY............
SPTREMBL|D7UPI3_OGA    ..............................................MTVSESSASLIQAKQD........
SPTREMBL|D8Q788_SCH    ..........................................MSSVVEAGALSGLSADPKVPAKG.....
SPTREMBL|F4P3U6_BAT    ................................................MPST..................
SPTREMBL|F5HCT5_CRY    ..........................................MSPVAVHNKHGDSASHGTALSKNPHVK.
SPTREMBL|G3BBY9_CAN    ..................................................MTKD................
SPTREMBL|G7DV49_MIX    .....................................MTPALAPLELLRNDSTEGDASIASAPYRKTRQ.
SPTREMBL|G8C0M6_TET    ................................................MSKFND................
SPTREMBL|G9MFC9_HYP    ..........................................MAPGILSTEKSDITKARNK.........
SPTREMBL|H0GCH3_9SA    ................................................MAQFSD................
SPTREMBL|I1BXP2_RHI    ................................................MTSA..................
SPTREMBL|I1RJN7_GIB    ..........................................MAPSAIAPETQYDHTALSN.........
SPTREMBL|I2K2H5_DEK    ..................................................MSQD................
SPTREMBL|J9N9S0_FUS    ..........................................MAPTLELNERAAVDKIVVKPLPSNGST.
SPTREMBL|J9VL23_CRY    ...............................................MANFSSN................
SPTREMBL|L0PF66_PNE    .................................................MFSRK................
SPTREMBL|L2FB38_COL    ..............................................MPPSENQCSFHIESVSSGSGQQPF
SPTREMBL|Q5KHG5_CRY    ..........................................MTATRNTNGHGELNREAHTTNK......
SPTREMBL|Q6BN17_DEB    ..........................................................MNQ.........
SPTREMBL|Q6BXS5_DEB    ..............................................MTVS.SKQHLADQGQ.........
Consensus

                                80        90       100       110       120       130       140
1 [transketolase Sp    .......................EALESTDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHI
Paecilomyces divari    .................................LDQLAVNTIRLLAVDATFKANAGHPGAPMGLAPAAHV
SPTREMBL|A5DSN8_LOD    .................................IDELSVSTIRGLSVDQVNAAQSGHPGAPLGLAPAAHA
SPTREMBL|A8PWG5_MAL    .................................IDLKSIATIRTLAADVVANAKSGHPGAPMGLAPLAHL
SPTREMBL|B6K5F9_SCH    .................................LDTKAVNTIRTLAVDTTCKSKSGHPGAPMGMAPAAHV
SPTREMBL|C0LQF6_PIC    .................................SDLLAINTIRLLAVDTVAKANSGHPGAPMGLAPAAHT
SPTREMBL|C7ZAM9_NEC    .............................RKLTADDIAVRDIRKLVIDCCRQNGGGHGGSAIGMAPLGVA
SPTREMBL|D7UPI2_OGA    .................................EHDQIIKYGRALVLDIVQQYDGGHPGSAMGGMALGIA
SPTREMBL|D7UPI3_OGA    ...............................QTKFDFMLKYYRNLIVDLVHNYKGGHGGGPNGMAAIGFA
SPTREMBL|D8Q788_SCH    ............................AIDPEIEKLAVNTVRVLGADLTQAFKGGHPGTVMGAAAIGVA
SPTREMBL|F4P3U6_BAT    .................................MDILAINTIRTLAADVVQKANSGHPGAPMGCAPMAHV
SPTREMBL|F5HCT5_CRY    ............................SSEVDTEKLVINAIRCLAADLCQQYKGGHPGTVMGAAAIATA
SPTREMBL|G3BBY9_CAN    .................................IDQLSINTIKTLAVDSVSAANSGHPGAPLALATLAHA
SPTREMBL|G7DV49_MIX    .......................PVKMTFEAHADDWTCINTIRTLSLDVVNKANSGHPGAPMGMAPVAHI
SPTREMBL|G8C0M6_TET    .................................IDRLCISTVRLLAVDAVQKANSGHPGAPLGLAPAMHV
SPTREMBL|G9MFC9_HYP    .................................EHELVLKTFRLLISDLCQQFNGGHPGGAIGMAAIGIA
SPTREMBL|H0GCH3_9SA    .................................IDKLAVSTLRLLSVDQVESAQSGHPGAPLGLAPVAHV
SPTREMBL|I1BXP2_RHI    .................................IDQKAINTIRVLAADVVRKADSGHPGAPMGLAPIAHT
SPTREMBL|I1RJN7_GIB    ......................VVKLIKNDDARTIDLALRQFRCLIADLCEQFKGGHPGGAMGMAAIGLA
SPTREMBL|I2K2H5_DEK    .................................LDTLAVNTIRVLSVDTVEKANSGHPGAPMGMAPAAHV
SPTREMBL|J9N9S0_FUS    ......................TVAGLNLKSPEKHDRVLKVFRAFIADLCQQFNGGHPGSAMGMAAIGIA
SPTREMBL|J9VL23_CRY    ..................................DITAVNTIRTLAADVVAKANSGHPGAPMGMAPAAHV
SPTREMBL|L0PF66_PNE    .................................ITEKCINTIRVLSVDITSKSNSGHPGAPMALAPCAHV
SPTREMBL|L2FB38_COL    LNGDKPSFVPPPPALGLNGHIDGQLKGWRETLNDDDFAVLALRNLIFDICNQNGGGHGGSAIGMAAIGVA
SPTREMBL|Q5KHG5_CRY    ............................VVSEKEEQLVLNTIRCLAADLCQQYKGGHPGTVMGASAIGIA
SPTREMBL|Q6BN17_DEB    .................................IEDLSLKTIRCLVSDLVQQYNGGHPGGAMGMAAIGIA
SPTREMBL|Q6BXS5_DEB    .................................IHDKVLKVFRLLIADLVQNFNGGHPGGAMDMAAIGIA
Consensus                                                      TIR L  D       GHPG  MG A
```

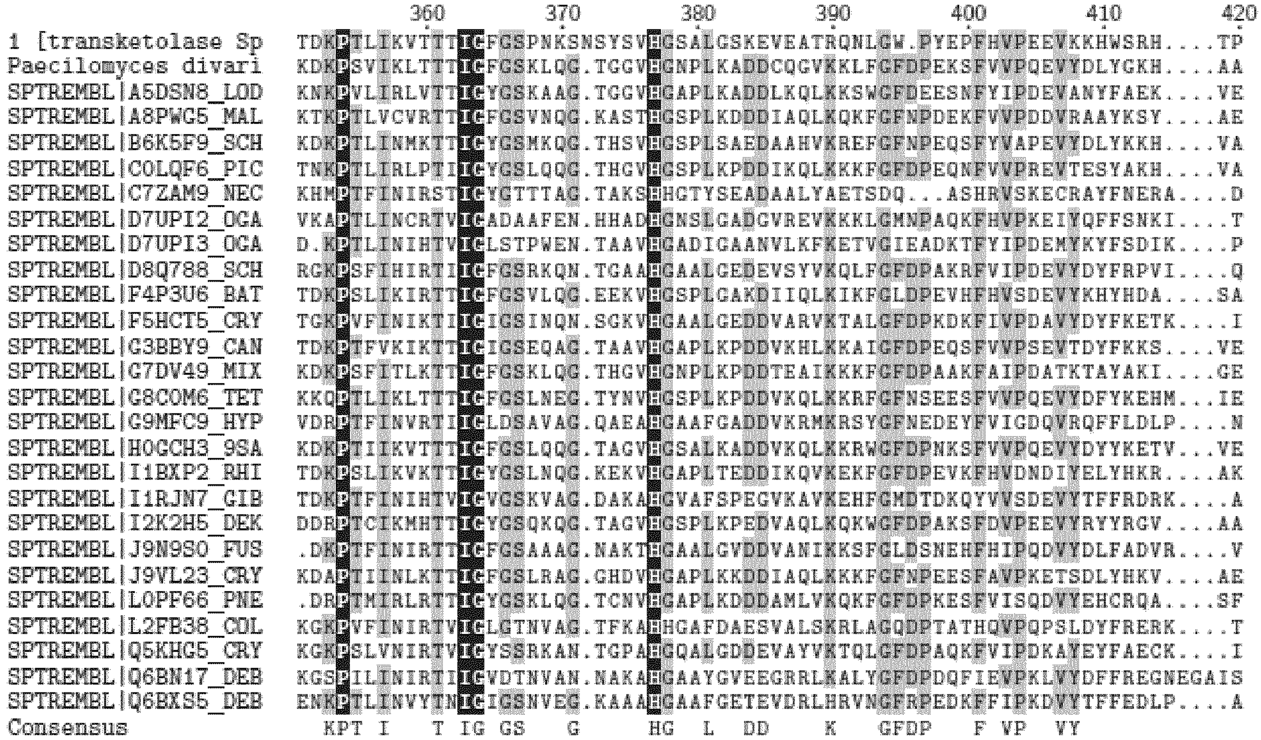
Figure 1F continued

```
                             150       160       170       180       190       200       210
1 [transketolase Sp   LYDEIMRYNPKNPYWFNRDRFVLSAGHGCMLQYALLHLACYDSVLEEDLKTFRQWGCSRI..........
Paecilomyces divari   LFNKFMTFNPKNPDWLNRDRFVLSNGHACMLQYALLHLFCYK.VSMDDLKSFRQLGSIT...........
SPTREMBL|A5DSN8_LOD   LWRKMR.FNPQDPKWINRDRFVLSNGHACALLYSLLVLYCYD.LTVDDLKQFRQLNSKT...........
SPTREMBL|A8PWG5_MAL   LWSRIMVADPKESHWINRDRFVLSNGHACALQYIMLHLMCYD.LTMDDLKAFRQLDSRT...........
SPTREMBL|B6K5F9_SCH   LFSRIMNFNPQKPRWINRDRFVLSNGHACVLQYALCHLLCYN.LTIDDLKNFRQLGSKT...........
SPTREMBL|C0LQF6_PIC   LFKQMR.FNPRNPAWINRDRFVLSNGHACALLYTMLFLYCYD.YTIDDLKSFRQLNSKT...........
SPTREMBL|C7ZAM9_NEC   LWRHTLRFSTKNPAWFDRDRFVLSNGHAAIFLYVMLHVACYPHMTLDELKMYADPKAFDHKTGTWKSTIC
SPTREMBL|D7UPI2_OGA   LWKYTMKYAPNDPTYFNRDRFVLSNGHVCLLQYVFQHFTCLKSMTMDQLKSYHSNDFHSHCP........
SPTREMBL|D7UPI3_OGA   LYKYVMKYNPENPSYFNRDRFILSNGHTCLFQYAFNHLVCYSHMTLEELKSYHSAEEESLCP........
SPTREMBL|D8Q788_SCH   LWRYQMRYNPKNPDWFNRDRFVLSAGHACLWQYIHLHLACYDAWTIDALKQYHNPDFG..IAAG......
SPTREMBL|F4P3U6_BAT   LFSNFICQDPQHPDWISRDRFVLSNGHGCVLQYIMLHMLCYN.LSMDDLKSFRQLNSRT...........
SPTREMBL|F5HCT5_CRY   LWKYSMRYNPANPDWINRDRFVLSAGHACLLQYIMLHLSCYSSWTLDQIKKYHAPTMDG.IAAG......
SPTREMBL|G3BBY9_CAN   IFQKTK.FTPNHPKWPNRDRFVLSNGHACALLYSILFLYCFD.YTLDDLKNFRQLNSKT...........
SPTREMBL|G7DV49_MIX   LWSRFMRCNPKSSHWPNRDRFVLSNGHACALQYVLLHLLCYK.LSMDDLKSFRQVDSTC...........
SPTREMBL|G8C0M6_TET   LCRDFMNVNHENVDWPNRDRFVLSNGHACALLYSTLHLLCYD.YTIDDLKRFRSVGSKT...........
SPTREMBL|G9MFC9_HYP   LWKYVMRYAPLQPSFFNRDRFVLSNGHTCLFQYTFLHLTCYRDMTMDQLKSYHSKRWDSICP........
SPTREMBL|H0GCH3_9SA   IFKQLR.CNPNNEHWINRDRFVLSNGHSCALLYSMLHLLCYD.YSIEDLRQFRQVNSRT...........
SPTREMBL|I1BXP2_RHI   LFTQFLHANPKNPYFINRDRFLLSNGHACVLQYILLHLLCYD.LSMDDLKQFRQLGSKT...........
SPTREMBL|I1RJN7_GIB   LWKYIMQYSPNDPDYFNRDRFVLSNGHTCLFQYTFLHLVCYKAMTFDHLKSYHSSNYDSLAP........
SPTREMBL|I2K2H5_DEK   LFSKMK.FNPKDPEWPNRDRFVLSNGHAVALLYTMLHLTCYD.LSMEDLKQFRQLGSKT...........
SPTREMBL|J9N9S0_FUS   LYKYVMRYSPNNCEYFNRDRFVLSNGHACLWQYLFMHLVCVKSMTLEQLKSYHSTKTDSLCP........
SPTREMBL|J9VL23_CRY   LFTRFMRFNSKNPKWINRDRFVLSNGHACALQYILLHLACYE.VSMEDLKQFRQIDSIT...........
SPTREMBL|L0PF66_PNE   LFNQHMIFDPNYPEWINRDRFVLSNGHACVLQYIILHLYCYK.VTMDDLKRFRQIDSIT...........
SPTREMBL|L2FB38_COL   LYKHVMRYNPNNPSWFDRDRFVLSNGHTAMFLYALNYLVCYDNFSMDEIRGYGSAKTNGYKT......IC
SPTREMBL|Q5KHG5_CRY   LWRYEMRYNPLNPDWFNRDRFVLSAGHACLFQYIFLHLSCYEAWTLDQIKMYHSPATSGSMAAG......
SPTREMBL|Q6BN17_DEB   LWKYILKYNPKNANWFNRDRFVLSNGHTCLFQYVFLHLVCYESFTMNQLKKYHAP.EVSQCA........
SPTREMBL|Q6BXS5_DEB   LWRYVLNYAPNNPTFFNRDRFVLSNGHTCLFQYAFHHLVCYQHMTMKQLKTYHSPNLESYCP........
Consensus             L    M  NP  P W NRDRFVLSNGHAC LQY  LHL GY   T D LK

                             220       230       240       250       260       270       280
1 [transketolase Sp   PGHPENFETPG.VEVTTGPLGQGIANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEA
Paecilomyces divari   PGHPEAH.DTPGVEVTTGPLGQGFANAVGLAIAQKHTAGVFNKPGYDLVNNYTYCFFGDGCAMEGIASEA
SPTREMBL|A5DSN8_LOD   PGHPERLDTPG.VEVTSGPLGQGISNAVGMAIAQKQLAATYNKPDFPISDSKVYAFMGDGCLMEGVSSEA
SPTREMBL|A8PWG5_MAL   PGHPEVHHTDG.VEVTTGPLGQGFANAVGLAIAQANAAGTYNKDNFDLFTNRTYVFLGDGCMQEGVASEA
SPTREMBL|B6K5F9_SCH   PGHPENHNPDLCIEVCTGPLGQGITNAVGLAIAQAHTAATYNKPGYDLINNRTFCFLGDGCLQEGVSSEA
SPTREMBL|C0LQF6_PIC   PGHPE.AELPG.VEVTTGPLGQGIANAVGLAIAQAQLAATYNKPNYELFSNYTYAFLGDGCLQEGVAQEA
SPTREMBL|C7ZAM9_NEC   HGHPEIE.VPG.VEVTTGPLGQGIANAVGLAIASKHLAATFNKPGYDMIQSRVYCSTGDGCLQEGVAMEA
SPTREMBL|D7UPI2_OGA   .GHPEIE.HDA.VEVTTGPLGQGIANSVGLAIATKNLAATYNKPGYNLVDNKTYCIVGDACLQEGPALEA
SPTREMBL|D7UPI3_OGA   .GHPEIE.HPA.IEVTTGPLGQGIANAVGMAVASKNLAATYNREGFPVVDNTIFCMVGDACLQEGPALEA
SPTREMBL|D8Q788_SCH   ..HPEIE.FPG.IELTTGLLGQGVANALGLAIASKNLAATYNKPSYPVVDNKIWCFVGDGCLQEGVGQEA
SPTREMBL|F4P3U6_BAT   PGHPEANHGVHGIEVSTGPLGQGISNAVGLAMAEAHMAATFNRPGFQVISNYTYGITGDGCLQEGVQAEA
SPTREMBL|F5HCT5_CRY   ..HPEIE.FPG.VELTTGLLGQGIANAVGLAVANKNMAATYNKDGFPIIQNKVWCFTGDGCLQEGVGQEA
SPTREMBL|G3BBY9_CAN   PGHPETPHCPG.VEVTTGPLGQGIANAVGIALAQKQFAATYNKPDYPIADARTFCIVGDGCLMEGISSEA
SPTREMBL|G7DV49_MIX   NGHPEVHHTDG.VEVTTGPLGQGFGNAVGLAMAQAQVCATFNKDGFSLFDNYTYMFTGDGCLQEGVASEA
SPTREMBL|G8C0M6_TET   PGHPE.YELPG.VELTTGPLGQGISNAVGFAIAQAHIAARYNKKDCEISTNYTYVFAGDGCMQEGVSSEA
SPTREMBL|G9MFC9_HYP   .GHPEIE.IEG.IEVTTGPLGQGVANAVGLAMATKNLAATYNRPGFDVVSNHTWCMIGDACLQEGVALEA
SPTREMBL|H0GCH3_9SA   PGHPE.FHSAG.VEITSGPLGQGISNAVGMAIAQANFAATYNEDGFPISDSYTFAIVGDGCLQEGVSSET
SPTREMBL|I1BXP2_RHI   PGHPER.IDTEGVEVTTGPLGQGISNAVGMAAAEAHLAATFNRPGYNIIDNFVYVFLGDGCLQEGVASEA
SPTREMBL|I1RJN7_GIB   .GHPEIE.HEG.IEVTTGPLGQGIANAVGLAMATKNLAATYNKPNFELVNNMTYCMIGDACLQEGVALEA
SPTREMBL|I2K2H5_DEK   PGHPE.AHTPG.IEVTTGPLGQGXANAVGLAIAQANVAATYNKPDFVVSDNYTYCFFGDGCMMEGISHEA
SPTREMBL|J9N9S0_FUS   .GHPEIE.NEG.VEVTTGPLGQGVANAVGLAMATKNLCATYNKPGYELVNNMTWCMIGDACLQEGVGLEA
SPTREMBL|J9VL23_CRY   PGHPEVGVTPG.IEVTTGPLGQGISNAVGLAIAQAHMCAVFNKENFSLIDNYTYCFLGDGCLQEGVASEA
SPTREMBL|L0PF66_PNE   PGHPEAS.VTPGIEVTTGPLGQGFANAVGLAIAQTHLAAVFNRPGFDIISNYTYCFFGDGCAMEGIFYEA
SPTREMBL|L2FB38_COL   HAHPEIE.VPG.VEVTTGPLGQGIANAVGLAIASKHLAARYNQPGLDVVQSRIYCMTGDGCLMEGVALEA
SPTREMBL|Q5KHG5_CRY   ..HPEIE.YPG.IEVTTGPLGQGISNAVGMAIASKQLAATYNREGLDIVDNKIWCFTGDGCLQEGVGQEA
SPTREMBL|Q6BN17_DEB   .GHPEIE.FEG.IEVTTGPLGQGIANAVGLAIASKNLAANYNKPDLDLVDNKIYCMVGDACIQEGVGLEA
SPTREMBL|Q6BXS5_DEB   .GHPEIE.HPG.IEVTTGPLGQGITNAVGLAIASKNLQATYNKPEFPVVSNHTFCIVGDACLQEGVALES
Consensus              GHPE     G  EVTTGPLGQGIANAVGLAIA    AATYNKP      N TYC  GDGCLQEGV  EA
```

Figure 1F continued

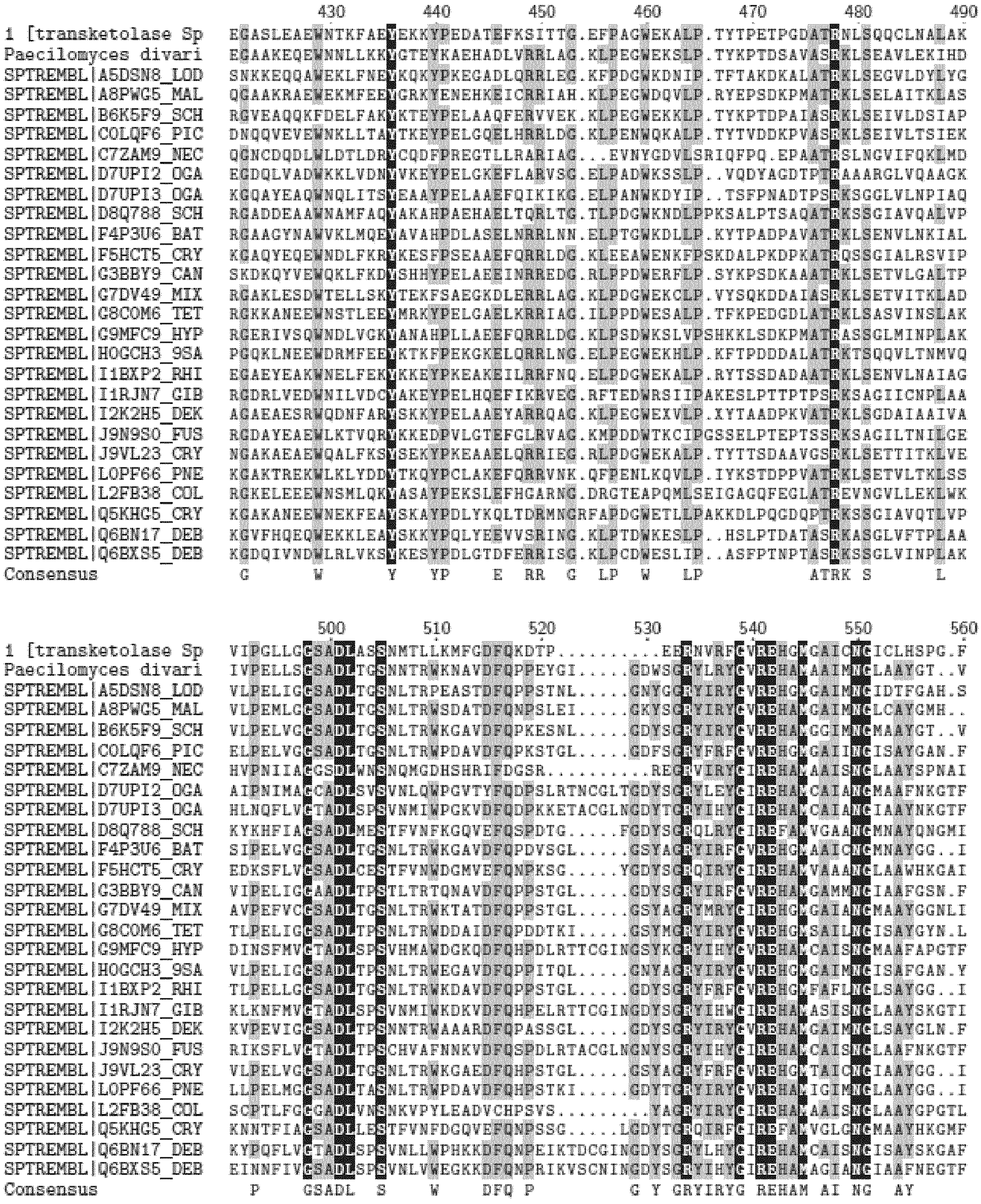


```
                               290       300       310       320       330       340       350
1 [transketolase Sp      CSLAGHWGLG.KLIAFYDDNHISIDGDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTV
Paecilomyces divari      ASTAGHLKLN.NLIAIYDDNHISIDGDTKCAFTEDVLKRFEAYGWHTEWVKDGDHDLDAIEAAIRRCKEV
SPTREMBL|A5DSN8_LOD      SSLAGHLQLD.NLIAFWDDNRISIDGDTACAFTENVPERYRAYGWHILEVPDANTDIDAISKAIDEAKTI
SPTREMBL|A8PWG5_MAL      ASLAGHLRLH.NLIAFFDDNHVSIDGDTKCSFTEDVPKRFESYGWNVLVVDNGDSDLQALYDAVVKAWDE
SPTREMBL|B6K5F9_SCH      SSLAGHLKLG.NLVAVWDNNKITIDGSTDVSFDEDVEKRYEAYGWQVLRVADGDSDYASIEQALLDAVAC
SPTREMBL|C0LQF6_PIC      ISLAGHLGDGLKLIAFWDDNQISIDGDTNVSFTEDVPAKFRAQGWEVLSVKDGNDDLEGIAAALAKAKT.
SPTREMBL|C7ZAM9_NEC      VAIAGHLKLD.NLVLCYDNN.............QDTNSKMRSMGWNVIDVFDGDNSVESIVNALNLGKAT
SPTREMBL|D7UPI2_OGA      ISIAGHYGLN.NLIVLYDNNQVCADGSVDIANTEDISAKFKACNWNVIEVANASEDVATIVKALEYAQNE
SPTREMBL|D7UPI3_OGA      ISFAGSMRLN.NLVVTYDNNQISCDGSVDITNTEDINAKFIACNWNVIDVENGSMDICAIVQALEDAKLS
SPTREMBL|D8Q788_SCH      LSLAGHLRLD.NLIVIYDDNSVTVDGSIATCFTDDTSAKVKAMGFNVLEVYDGTNDLTAIIKAMDEAKAT
SPTREMBL|F4P3U6_BAT      VSLAGHLKLG.NLIMLYDDNHITIDGDIEVGFTEDVVQRFESYGWHTQVLGDGDNDVAGIIAAIQKAKAV
SPTREMBL|F5HCT5_CRY      LSMAGHWGLD.NMILVYDNNSVTVDGNIDICFTDDTSAKLKSLGWHVLEVEDGSNDLAAIVDAFEQAQKL
SPTREMBL|G3BBY9_CAN      SSLAGHLQLG.NLTVFWDSNKISIDGSTDLAFTEDVPERYRAYGWHVIDVDEENPSLSTLSKAIDEASKV
SPTREMBL|G7DV49_MIX      ASLAGHLQLG.NLIAIYDSNKITIDGDTAVSFTEDVTKRMEAYGWHCQTVEQGDDDLEGMYSAILAAQQV
SPTREMBL|G8C0M6_TET      CSLAGHLQLG.NLIVIYDDNSITIDGKTNLSFTEEVAVRYRGYGWEVIEVKDGNDNLDAISQAIKLAKSN
SPTREMBL|G9MFC9_HYP      ISFAGHLKLN.NLTIIYDNNQITCDGSVDLTNTEDVNAKMRACGWDVIDVEDGCYDVVGIVAALEKARAS
SPTREMBL|H0GCH3_9SA      SSLAGHLQLG.NLITFYDSNSISIDGKTSYSFDEDVLKRYEAYGWEVMEVDKGDDDMESISSALEKAKLS
SPTREMBL|I1BXP2_RHI      ASLAGHLQLG.KLIAIYDDNHITIDGDTDVSFTEDVIKRFEAYGWHTQVIDNGDTDIEDFKKAVVEAQKV
SPTREMBL|I1RJN7_GIB      IQLAGHWKLN.NLVVMYDNNQITCDGSVDLCNTEDVNAKMRACGWDVIEVEDGCYDVEGIVKALIKAKAS
SPTREMBL|I2K2H5_DEK      CSLAGHLKLG.KLICFYDDNKVSIDGSTKQEFDEDVAKRLEAYGWEVIEVPNADTDTGAILAALERAKKT
SPTREMBL|J9N9S0_FUS      VSLAGHWRLN.NLCIIYDNNSITCDGTADVANTEDINAKMEATGWNVLDVFDGDSDVTAIVNALIAARSS
SPTREMBL|J9VL23_CRY      CSLAGHLKLG.NLVAIYDDNKITIDGDTAVSFTEDVEMRFKSYGWNVLHIEKGDDDLAAIENAITEAKKS
SPTREMBL|L0PF66_PNE      ASLAGHLKLK.NLICFYDYNGISIDGDIECTFTDDVAERFRSCNWHVIIVDDGDSDIDSIERAIAEAKIV
SPTREMBL|L2FB38_COL      ISLAGSMQLD.NLVLIYDNNQVTCDGPLDWINTEDVNAKMRASGWHVLEIADGNHDVSAVVSALRFAKSL
SPTREMBL|Q5KHG5_CRY      ISLAGHLGLD.NLILVYDNNAVTVDGRIDNCFTENTSKKLQAQGWNVIDVYDGSNDLAAILEGFDKAKQF
SPTREMBL|Q6BN17_DEB      ISLAGHLGLD.NLIVIYDNNQITCDGSVDLANSEDINAKFMAQKWHVLTVDDGSFDLRSILAAIEQAKSV
SPTREMBL|Q6BXS5_DEB      ISFAGHLGLN.NLTAIYDNNQISCDGSVDLTNTENIGEKFKACNWNVIEVENGCYDVIAIVNALESSKKS
Consensus                 SLAGHL L  NLI  YD N I  DG     FTEDV     A GW V  V  G  D   I  A   AK
```



```
                               360       370       380       390       400       410       420
1 [transketolase Sp      TDKPTLIKVTTTICFGSPNKSNSYSVHGSALGSKEVEATRQNLGW.PYEPFHVPEEVKKHWSRH....TP
Paecilomyces divari      KDKPSVIKLTTTIGFGSKLQG.TGGVHGNPLKADDCQGVKKLFGFDPEKSFVVPQEVYDLYGKH....AA
SPTREMBL|A5DSN8_LOD      KNKPVLIRLVTTIGYGSKAAG.TGGVHGAPLKADDLKQLKKSWGFDEESNFYIPDEVANYFAEK....VE
SPTREMBL|A8PWG5_MAL      KTKPTLVCVRTTIGFGSVNQG.KASTHGSPLKDDDIAQLKQKFGFNPDEKFVVPDDVRAAYKSY....AE
SPTREMBL|B6K5F9_SCH      KDKPTLINMKTTIGYGSMKQG.THSVHGSPLSAEDAAHVKREFGFNPEQSFYVAPEVYDLYKKH....VA
SPTREMBL|C0LQF6_PIC      TNKPTLIRLPTIIGYGSLQQG.THGVHGSPLKPDDIKQLKKKFGFDPEQNFVVPREVTESYAKH....VA
SPTREMBL|C7ZAM9_NEC      KHMPTFINIRSTIGYGTTTAG.TAKSHHGTYSEADAALYAETSDQ...ASHRVSKECRAYFNERA....D
SPTREMBL|D7UPI2_OGA      VKAPTLINCRTVIGADAAFEN.HHADHGNSLGADGVREVKKKLGMNPAQKFHVPKEIYQFFSNKI....T
SPTREMBL|D7UPI3_OGA      D.KPTLINIHTVIGLSTPWEN.TAAVHGADIGAANVLKFKETVGIEADKTFYIPDEMYKYFSDIK....P
SPTREMBL|D8Q788_SCH      RGKPSFIHIRTIIGFGSRKQN.TGAAHGAALGEDEVSYVKQLFGFDPAKRFVIPDEVYDYFRPVI....Q
SPTREMBL|F4P3U6_BAT      TDKPSLIKIRTTIGFGSVLQG.EEKVHGSPLGAKDIIQLKIKFGLDPEVHFHVSDEVYKHYHDA....SA
SPTREMBL|F5HCT5_CRY      TGKPVFINIKTIIGIGSINQN.SGKVHGAALGEDDVARVKTALGFDPKDKFIVPDAVYDYFKETK....I
SPTREMBL|G3BBY9_CAN      TDKPTFVKIKTTIGIGSEQAG.TAAVHGAPLKPDDVKHLKKAIGFDPEQSFVVPSEVTDYFKKS....VE
SPTREMBL|G7DV49_MIX      KDKPSFITLKTTIGFGSKLQG.THGVHGNPLKPDDTEAIKKKFGFDPAAKFAIPDATKTAYAKI....GE
SPTREMBL|G8C0M6_TET      KKQPTLIKLTTTIGFGSLNEG.TYNVHGSPLKPDDVKQLKKRFGFNSEESFVVPQEVYDFYKEHM...IE
SPTREMBL|G9MFC9_HYP      VDRPTFINVRTIIGLDSAVAG.QAEAHGAAFGADDVKRMKRSYGFNEDEYFVIGDQVRQFFLDLP....N
SPTREMBL|H0GCH3_9SA      KDKPTIIKVTTTIGFGSLQQG.TAGVHGSALKADDVKQLKKRWGFDPNKSFVVPQEVYDYYKETV...VE
SPTREMBL|I1BXP2_RHI      TDKPSLIKVKTTIGYGSLNQG.KEKVHGAPLTEDDIKQVKEKFGFDPEVKFHVDNDIYELYHKR....AK
SPTREMBL|I1RJN7_GIB      TDKPTFINIHTVIGVGSKVAG.DAKAHGVAFSPEGVKAVKEHFGMDTDKQYVVSDEVYTFFRDRK....A
SPTREMBL|I2K2H5_DEK      DDRPTCIKMHTTIGYGSQKQG.TAGVHGSPLKPEDVAQLKQKWGFDPAKSFDVPEEVYRYYRGV....AA
SPTREMBL|J9N9S0_FUS      .DKPTFINIRTTIGFGSAAAG.NAKTHGAALGVDDVANIKKSFGLDSNEHFHIPQDVYDLFADVR....V
SPTREMBL|J9VL23_CRY      KDAPTIINLKTTIGFGSLRAG.GHDVHGAPLKKDDIAQLKKKFGFNPEESFAVPKETSDLYHKV....AE
SPTREMBL|L0PF66_PNE      .DRPTMIRLRTTIGYGSKLQG.TCNVHGAPLKDDDAMLVKQKFGFDPKESFVISQDVYEHCRQA....SF
SPTREMBL|L2FB38_COL      KGKPVFINIRTVIGLGTNVAG.TFKAHHGAFDAESVALSKRLAGQDPTATHQVPQPSLDYFRERK....T
SPTREMBL|Q5KHG5_CRY      KGKPSLVNIRTVIGYSSRKAN.TGPAHGQALGDDEVAYVKTQLGFDPAQKFVIPDKAYEYFAECK....I
SPTREMBL|Q6BN17_DEB      KGSPILINIRTIIGVDTNVAN.NAKAHGAAYGVEEGRRLKALYGFDPDQFIEVPKLVYDFFREGNEGAIS
SPTREMBL|Q6BXS5_DEB      ENKPTLINVYTNIGIGSNVEG.KAAAHGAAFGETEVDRLHRVNGFRPEDKFFIPKDVYTFFEDLP....A
Consensus                  KPT I   T IG GS   G     HG  L  DD    K   GFDP   F VP  VY
```

Figure 1F continued



```
                             430       440       450       460       470       480       490
1 [transketolase Sp  EGASLEAEWNTKFAEYEKKYPEDATEFKSITTG.EFPAGWEKALP.TYTPETPGDATRNLSQQCLNALAK
Paecilomyces divari  EGAAKEQEWNNLLKKYGTEYKAEHADLVRRLAG.KLPEGWEKSLP.TYKPTDSAVASRKLSEAVLEKIHD
SPTREMBL|A5DSN8_LOD  SNKKEQQAWEKLFNEYKQKYPKEGADLQRRLEG.KFPDGWKDNIP.TFTAKDKALATRKLSEGVLDYLYG
SPTREMBL|A8PWG5_MAL  QGAAKRAEWEKMFEEYGRKYENEHKEICRRIAH.KLPEGWDQVLP.RYEPSDKPMATRKLSELAITKLAS
SPTREMBL|B6K5F9_SCH  RGVEAQQKFDELFAKYKTEYPELAAQFERVVEK.KLPEGWEKKLP.TYKPTDPAIASRKLSEIVLDSIAP
SPTREMBL|C0LQF6_PIC  DNQQVEVEWNKLLTAYTKEYPELGQELHRRLDG.KLPENWQKALP.TYTVDDKPVASRKLSEIVLTSIEK
SPTREMBL|C7ZAM9_NEC  QGNCDQDLWLDTLDRYCQDFPREGTLLRARIAG...EVNYGDVLSRIQFPQ.EPAATRSLNGVIFQKLMD
SPTREMBL|D7UPI2_OGA  EGDQLVADWKKLVDNYVKEYPELGKEFLARVSG.ELPADWKSSLP..VQDYAGDTPTRAAARGLVQAAGK
SPTREMBL|D7UPI3_OGA  KGQAYEAQWNQLITSYEAAYPELAAEFQIKIKG.ELPANWKDYIP..TSFPNADTPSRKSGGLVLNPIAQ
SPTREMBL|D8Q788_SCH  RGADDEAAWNAMFAQYAKAHPAEHAELTQRLTG.TLPDGWKNDLPPKSALPTSAQATRKSSGIAVQALVP
SPTREMBL|F4P3U6_BAT  RGAAGYNAWVKLMQEYAVAHPDLASELNRRLNN.ELPTGWKDLLP.KYTPADPAVATRKLSENVLNKIAL
SPTREMBL|F5HCT5_CRY  KGAQYEQEWNDLFKRYKESFPSEAAEFQRRLDG.KLEEAWENKFPSKDALPKDPKATRQSSGIALRSVIP
SPTREMBL|G3BBY9_CAN  SKDKQYVEWQKLFKDYSHHYPELAEEINRREDG.RLPPDWERFLP.SYKPSDKAAATRKLSETVLGALTP
SPTREMBL|G7DV49_MIX  RGAKLESDWTELLSKYTEKFSAEGKDLERRLAG.KLPDGWEKCLP.VYSQKDDAIASRKLSETVITKLAD
SPTREMBL|G8C0M6_TET  RGKKANEEWNSTLEEYMRKYPELGAELKRRIAG.ILPPDWESALP.TFKPEDGDLATRKLSASVINSLAK
SPTREMBL|G9MFC9_HYP  RGERIVSQWNDLVGKYANAHPLLAEEFQRRLDG.KLPSDWKSLVPSHKKLSDKPMATRASSGLMINPLAK
SPTREMBL|H0GCH3_9SA  PGQKLNEEWDRMFEEYKTKFPEKGKELQRRLNG.ELPEGWEKHLP.KFTPDDDALATRKTSQQVLTNMVQ
SPTREMBL|I1BXP2_RHI  EGAEYEAKWNELFEKYKKEYPKEAKEILRRFNQ.ELPDGWEKALP.RYTSSDADAATRKLSENVLNAIAG
SPTREMBL|I1RJN7_GIB  RGDRLVEDWNILVDCYAKEYPELHQEFIKRVEG.RFTEDWRSIIPAKESLPTTPTPSRKSAGIICNPLAA
SPTREMBL|I2K2H5_DEK  AGAEAESRWQDNFARYSKKYPELAAEYARRQAG.KLPEGWEXVLP.XYTAADPKVATRKLSGDAIAAIVA
SPTREMBL|J9N9S0_FUS  RGDAYEAEWLKTVQRYKKEDPVLGTEFGLRVAG.KMPDDWTKCIPGSSELPTEPTSSRKSAGILTNILGE
SPTREMBL|J9VL23_CRY  NGAKAEAEWQALFKSYSEKYPKEAAELQRRIEG.RLPDGWEKALP.TYTTSDAAVGSRKLSETTITKLVE
SPTREMBL|L0PF66_PNE  KGAKTREKWLKLYDDYTKQYPCLAKEFQRRVNK.QFPENLKQVLP.IYKSTDPPVATRKLSETVLTKLSS
SPTREMBL|L2FB38_COL  RGKELEEEWNSMLQKYASAYPEKSLEFHGARNG.DRGTEAPQMLSEIGAGQFEGLATREVNGVLLEKLWK
SPTREMBL|Q5KHG5_CRY  KGAKANEEWNEKFEAYSKAYPDLYKQLTDRMNGRFAPDGWETLLPAKKDLPQGDQPTRKSSGIAVQTLVP
SPTREMBL|Q6BN17_DEB  KGVFHQEQWEKKLEAYSKKYPQLYEEVVSRING.KLPTDWKESLP..HSLPTDATASRKASGLVFTPLAA
SPTREMBL|Q6BXS5_DEB  KGDQIVNDWLRLVKSYKESYPDLGTDFERRISG.KLPCDWESLIP..ASFPTNPTASRKSSGLVINPLAK
Consensus             G      W      Y   YP    E  RR  G  LP  W   LP          ATRK S      L

                             500       510       520       530       540       550       560
1 [transketolase Sp  VIPGLLGCSADLASSNMTLLKMFGDFQKDTP..........EERNVRFCVREHGKGAICNGICLHSPG.F
Paecilomyces divari  IVPELLSCSADLTGSNNTRWKNAVDFQPPEYGI.....GDWSGRYLRYCVREHAMAAIMNGLAAYGT..V
SPTREMBL|A5DSN8_LOD  VLPELIGCSADLTGSNLTRPEASTDFQPPSTNL.....GNYGGRYIRYCVREHGKGAIMNGIDTFGAH.S
SPTREMBL|A8PWG5_MAL  VLPEMLGCSADLTGSNLTRWSDATDFQNPSLEI.....GKYSGRYIRYCVREHAMGAIMNGLCAYGMH..
SPTREMBL|B6K5F9_SCH  VLPELVGCSADLTGSNLTRWKGAVDFQPKESNL.....GDYSGRYIRYCVREHAMGGIMNGMAAYGT..V
SPTREMBL|C0LQF6_PIC  ELPELVGCSADLTGSNLTRWPDAVDFQPKSTGL.....GDFSGRYFRFCVREHGMGAIINGISAYGAN.F
SPTREMBL|C7ZAM9_NEC  HVPNIIACGSDLWNSNQMGDHSHRIFDGSR..........REGRVIRYCIREHAMAAISNGLAAYSPNAI
SPTREMBL|D7UPI2_OGA  AIPNIMACCADLSVSVNLQWPGVTYFQDPSLRTNCGLTGDYSGRYLEYCIREHAMCAIANGMAAFNKGTF
SPTREMBL|D7UPI3_OGA  HLNQFLVCTADLSPSVNMIWPGKVDFQDPKKETACGLNGDYTGRYIHYCIREHAMCAIANGIAAYNKGTF
SPTREMBL|D8Q788_SCH  KYKHFIACSADLMESTFVNFKGQVEFQSPDTG.....FGDYSGRQLRYCIREFAMVGAANGMNAYQNGMI
SPTREMBL|F4P3U6_BAT  SIPELVGCSADLTGSNLTRWKGAVDFQPDVSGL.....GSYAGRYIRFCVREHGMAAICNGMNAYGG..I
SPTREMBL|F5HCT5_CRY  EDKSFLVCSADLCESTFVNWDGMVEFQNPKSG.....YGDYSGRQIRYCIREHAMVAAANGLAAWHKGAI
SPTREMBL|G3BBY9_CAN  VIPELIGCAADLTPSTLTRTQNAVDFQPPSTGL.....GDYSGRYIRFCVREHAMGAMMNGIAAFGSN.F
SPTREMBL|G7DV49_MIX  AVPEFVCCSADLTGSNLTRWKTATDFQPPSTGL.....GSYAGRYMRYCIREHGMGAIANGMAAYGGNLI
SPTREMBL|G8C0M6_TET  TLPELIGCSADLTPSNLTRWDDAIDFQPDDTKI.....GSYMGRYIRYCVREHGMSAILNGISAYGYN.L
SPTREMBL|G9MFC9_HYP  DINSFMVCTADLSPSVHMAWDGKQDFQHPDLRTTCGINGSYKGRYIHYCVREHAMCAISNGMAAFAPGTF
SPTREMBL|H0GCH3_9SA  VLPELIGCSADLTPSNLTRWEGAVDFQPPITQL.....GNYAGRYIRYCVREHGMGAIMNGISAFGAN.Y
SPTREMBL|I1BXP2_RHI  TLPELLGCSADLTGSNLTRWKDAVDFQPPSTGL.....GDYSGRYFRFCVREHGMFAFLNGLSAYGG..I
SPTREMBL|I1RJN7_GIB  KLKNFMVCTADLSPSVNMIWKDKVDFQHPELRTTCGINGDYSGRYIHWCIREHAMASISNGLAAYSKGTI
SPTREMBL|I2K2H5_DEK  KVPEVIGCSADLTPSNNTRWAAARDFQPASSGL.....GDYSGRYIRYCVREHAMGAIMNGLSAYGLN.F
SPTREMBL|J9N9S0_FUS  RIKSFLVCTADLTPSCHVAFNNKVDFQSPDLRTACGLNGNYSGRYIHYCIREHAMCAISNGLAAFNKGTF
SPTREMBL|J9VL23_CRY  VLPELVGCSADLTGSNLTRWKGAEDFQHPSTGL.....GSYAGRYFRFCVREHGMTAICNGIAAYGG..I
SPTREMBL|L0PF66_PNE  LLPELMGCSADLTASNLTRWPDAVDFQHPSTKI.....GDYTGRYIRYCVREHAMIGIMNGLAAYGG..I
SPTREMBL|L2FB38_COL  SCPTLFGCGADLVNSNKVPYLEADVCHPSVS.........YAGRYIRYCIREHAMAAISNGLAAYGPGTL
SPTREMBL|Q5KHG5_CRY  KNNTFIACSADLLESTFVNFDGQVEFQNPSSG.....LGDYTGRQIRFCIREFAMVGLGNGMAAYHKGMF
SPTREMBL|Q6BN17_DEB  KYPQFLVCTADLSPSVNLLWPHKKDFQNPEIKTDCGINGDYSGRYLHYCIREHAMCAISNGISAYSKGAF
SPTREMBL|Q6BXS5_DEB  EINNFIVCSADLSPSVNLVWEGKKDFQNPRIKVSCNINGDYSGRYIHYCIREHAMAGIANGIAAFNEGTF
Consensus              P    GSADL  S    W    DFQ P         G Y GRYIRYG REHAM AI NG  AY
```

Figure 1F continued

```
                              570       580       590       600       610       620       630
                               .         .         .         .         .         .         .
1 [transketolase Sp   VPYCATFFVFTDYMRGAMRISALSEAGVIYVMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGN
Paecilomyces divari   IPAGGTFLNFVSYAAGAVRLSALSRVRVIHVATHDSIGLGEDGPTHQPIETLAHFRALPNCMVWRPADGN
SPTREMBL|A5DSN8_LOD   RTYGGTFLNFVSYAAGALRLSALSHQGVIWVATHDSIGLGEDGPTHQPIETLAHFRAIPNLSVWRPADGN
SPTREMBL|A8PWG5_MAL   IPTAGTFLNFVTYGWGAVRLSSLSNFQVIWVATHDSIGLGEDGPTHQPIEVAAALRALPNMDFWRPADGN
SPTREMBL|B6K5F9_SCH   IPFGGTFLNFVSYAAGAVRLAALGQFRTIYVATHDSIGLGEDGPTHQPIETYAHFRAMPNINCWRPADGN
SPTREMBL|C0LQF6_PIC   KAYGGTFLNFVSYASGAVRLSALSGHPIIWVATHDSIGLGEDGPTHQPIETLAHLRALPNLMVWRPADGN
SPTREMBL|C7ZAM9_NEC   IPVTATFFMFYLYAAPGVRMGALSSLKVIHVATHDSIGEGQNGPTHQPVEVDSLYRAMPNLLYIRPADGE
SPTREMBL|D7UPI2_OGA   IPITSTFFMFYLYAAPAIRMAGLQELKTIHIGTHDSINEGENGPTHQPIESPSLFRAMLNVYYMRPVDSA
SPTREMBL|D7UPI3_OGA   IPVTSTFFMFYLYAAPAVRMGALMNLKVIHVGTHDSIGTGEDGPTHQPIALANFYRALPNLYYIRPADSL
SPTREMBL|D8Q788_SCH   IPICSSYLAFWQYAMSACRLGALMGVRFIGLGTHDSIGVGEDGPTHQSVALAAFFRALPNFNLIRPADAE
SPTREMBL|F4P3U6_BAT   IPFGATFFNFISYALGAVRLSALSKHQVIYIMTHDSIGLGEDGPTHQPIETLASLRALPNLITLRPADGN
SPTREMBL|F5HCT5_CRY   VPIMSSYFIFWLYAAPSLRMAALMKLRFIAIATHDSIGVGEDGPTHQPIAFPLFLRALPNFNYIRPADAE
SPTREMBL|G3BBY9_CAN   KVFGGTFLNFVSYAAGAVRLSAISQLPIIWVGTHDSVGLGEDGPTHQPIETLAHLRAIPNLSVWRPADGN
SPTREMBL|G7DV49_MIX   LPAISTFANFTSYAAGAIRLSALSGFRVIWVATHDSIGLGEDGPTHQPIEVAAHFRALPNMHVWRPADGN
SPTREMBL|G8C0M6_TET   KPFGGTFLNFVSYASGAVRLAALSGYPVIWVATHDSIGLGEDGPTHQPIETVTHFRALPNMQVWRPADGN
SPTREMBL|G9MFC9_HYP   IPVTSSFFMFYLYAAPAVRMGALQGLKVIHAATHDSIGMGEDGPTHQPIELAALFRSMPNLLYIRPADSE
SPTREMBL|H0GCH3_9SA   KPYGGTFLNFVSYAAGAVRLAALSGNPVIWVATHDSIGLGEDGPTHQPIETLAHLRAIPNMHVWRPADGN
SPTREMBL|I1BXP2_RHI   IPFGGTFLNFLTYGWGAARLSALSHCRAIYIMTHDSIGLGEDGPTHQPVETLALTRATPNMLTFRPADGN
SPTREMBL|I1RJN7_GIB   LPVTSSFFMFYIYAAPGVRMGALQNLQAIHIATHDSIGTGEDGPTHQPIALPALYRSMPNILYIRPCDTE
SPTREMBL|I2K2H5_DEK   RPFAGTFLNFVSYAAGAVRLSALSGHPVVWVATHDSIGLGEDGPTHQPIETLAHFRAMPNLMVWRPADGN
SPTREMBL|J9N9S0_FUS   IPMTSSFFMFYLYAAPAVRMAALQGLQQIHIATHDSIGTGEDGPTHQPIALPALYRAMPNTLYIRPCDSE
SPTREMBL|J9VL23_CRY   IPFTATFLNFVSYAAGAVRLSALSHLRVLNVATHDSIGLGEDGPTHQPVETAAWLRAVPNLAFWRPADGN
SPTREMBL|L0PF66_PNE   IPYGGTFLNFVFYAAGSLRLSSISHLRVIIVATHDSIGLGEDGPTHQPIEAASWLRAQPNMMFWRPADGN
SPTREMBL|L2FB38_COL   LPITATFFMFYLYAAPGVRMGALSHLPVIHIATHDSFAEGQNGPTHQPVEIDSLYRAMPNLAYIRPCDAE
SPTREMBL|Q5KHG5_CRY   IPIMSTFFMFWIYAAPAARMAALQHLRFIGVATHDSIGIGEDGPTHQPIALAAFYRALPNINLIRPADAE
SPTREMBL|Q6BN17_DEB   IPITSSFFMFYLYSAPAVRMGALQNLQVIHVATHDSIGTGEDGPTHQPIALASFYRSLPNCLYVRPADNE
SPTREMBL|Q6BXS5_DEB   IPITSSFFMFYLYASPAVRYGALSKLQVIHVATHDSIGTGEDGPTHQPIALAALYRAMPNINYIRPCDSE
Consensus              P   TF  F  YAA AVR  ALS   VI VATHDSIG GEDGPTHQPIE  A  RA PN    RPAD

                              640       650       660       670       680       690       700
                               .         .         .         .         .         .         .
1 [transketolase Sp   ETAGSYKVAVENRK..TPSILALSRQK.LPNLPGTSIEGVEKGGYTITDNSSGN..KPDVILIGTGSELE
Paecilomyces divari   ETSAAYYSALTSKH..TPSILALTRQN.LPQLENSTIEGALKGGYIALDAPN.....ATVTLVSTGSEVS
SPTREMBL|A5DSN8_LOD   EVSAAYIAAIEAVD..HPSVIALTRQN.VPHLENSSIEQALKGGYTVWPVEDG...KPDLIIVSTGSEVS
SPTREMBL|A8PWG5_MAL   EVSAAYKVAIESYK..TPSVVSLTRQN.LPQLEGSSMEKAEKGGYVLLEDK.....EAAITLVSTGSEVS
SPTREMBL|B6K5F9_SCH   ETSAAYHSAISLAQ..TPSILALTRQN.LPQLNGSSIEKALKGGYVLVENKD.....ANVTLIGTGSEVS
SPTREMBL|C0LQF6_PIC   ETSAAYLRAIESKH..TPSIIALTRQN.LPQLEHSSIEKAAKGGYTVYPVEN.....PDIILVASGSEVS
SPTREMBL|C7ZAM9_NEC   EIIGAWSSALEVQG..QPVIISLARDPPEITIPNTSRTKESKGGYVIVD.....HPKPVVTLVSCGSELQ
SPTREMBL|D7UPI2_OGA   EVLGLFEKAIECP...YTSMLSLSRNEVLQYPGLSSPEKAKRGGYILED.....VENADVQLIGACAEME
SPTREMBL|D7UPI3_OGA   ETAGAYEVAIEAEG..YSSIISESRQNLVQYPENSKRDAVKFGAYVFDDF.NIPDAKKDLIIIGVGSEMC
SPTREMBL|D8Q788_SCH   EVLGAWEVALDAK..HTPTLMTLSRQG.VPLLEHSDRAKVHKGAYPIFANYDESSADPELVLIATGSEVW
SPTREMBL|F4P3U6_BAT   ETSGAYVAALENRH..RPSVLIFTRQN.LPQLEGSTVEKTLKGAYVLSESAN.....AKITLVGTGSEVS
SPTREMBL|F5HCT5_CRY   EVIGAWILGLRDA..DHPSLLSLTRQP.VPLLSGTDRNKVQYGGYVVYG..DENSIP.DITLVATGSEVA
SPTREMBL|G3BBY9_CAN   EVSAAYKAAISSVS..TPHVLALTRQG.VPHLEGSSIEKASKGGYVLVEAQD.....PDIILVSSGSEVS
SPTREMBL|G7DV49_MIX   ETSAAYLMAMMNAT..TPSMLLLTRQN.LPQLENSTIAHATKGGYVVHDTDG....ETDLTYVSTGSEVS
SPTREMBL|G8C0M6_TET   EVSAAYKVSIESKA..TPSIICLTRQN.VPQLEGSSIEKAAKGGYAISEYHGNTTPKLDLILVSSGSEVS
SPTREMBL|G9MFC9_HYP   ETAGAWIAAIETQN..ASSIISTSRHKVPQLKQTR.RDGVAKGAYVIQE.....DDDAQLTLIGVGAELS
SPTREMBL|H0GCH3_9SA   ETSAAYYSAIKSGR..TPSVVALSRQN.LPQLEHSSFEKALKGGYVIHDVE.....NPDIILVSTGSEVS
SPTREMBL|I1BXP2_RHI   EVSGAYLVAIQNKT..RPSVIALSRQG.VPQLENSSIEAVLKGGYILRTDEK.....PEVIFVATGTEVG
SPTREMBL|I1RJN7_GIB   ETAGAFITALEAKD..TPSIISLSRQNLEQYPQHSSREGTQKGAYVFIE.....DEDADVTLIGVGAEMV
SPTREMBL|I2K2H5_DEK   EVSAAYKTALESXX..TPSIVCLTRQG.LPELAGSSVEKAARGGYVVREASASX...AGLVLVATGSEVS
SPTREMBL|J9N9S0_FUS   EVAGAFIAAIQATE..TPTIISLSRQNLTQFPQHSSREGVSVGAYVFVEA.D..DDDFDVTLIGVGSEMG
SPTREMBL|J9VL23_CRY   ETSAAYLVAILSQH..TPSVFALSRQN.LPQLANSSIEKATKGGYVVEEVE.....NADVTLVSTGSEVT
SPTREMBL|L0PF66_PNE   ECSAAYYVAITSLN..MPSTIALSRQN.LPHLENTSIDGALKGGYVLRENSS.....ADLTIVSTGSEVS
SPTREMBL|L2FB38_COL   ETIGSWMLALNNRE..GPTMLSLGRDP.TGAVPSTDRFKVAKGAYVIED.....VAGAKLTLASCGTNLH
SPTREMBL|Q5KHG5_CRY   ECMGMWLLALSDQSANTPSIFALSRQP.VPLLSGTDRTKVAQGAYIVYG..DDRSPD..MTIIATGAEVA
SPTREMBL|Q6BN17_DEB   EVAGAWELAIETTN..KPTIISLSRQNLKQYPGITDRNKVKFGAYVLKEF.DSSSDSQKLQIISVGAESQ
SPTREMBL|Q6BXS5_DEB   ETAGAWEIAIKSQG..MPTIISLSRQNLPQYPGMSNRSLVKKGAYIFQER.....IDAVLNIVGVGAELK
Consensus             E   A   A        PS   L RQ   P L  S      KGGYV               L   GSEV
```

Figure 1F continued

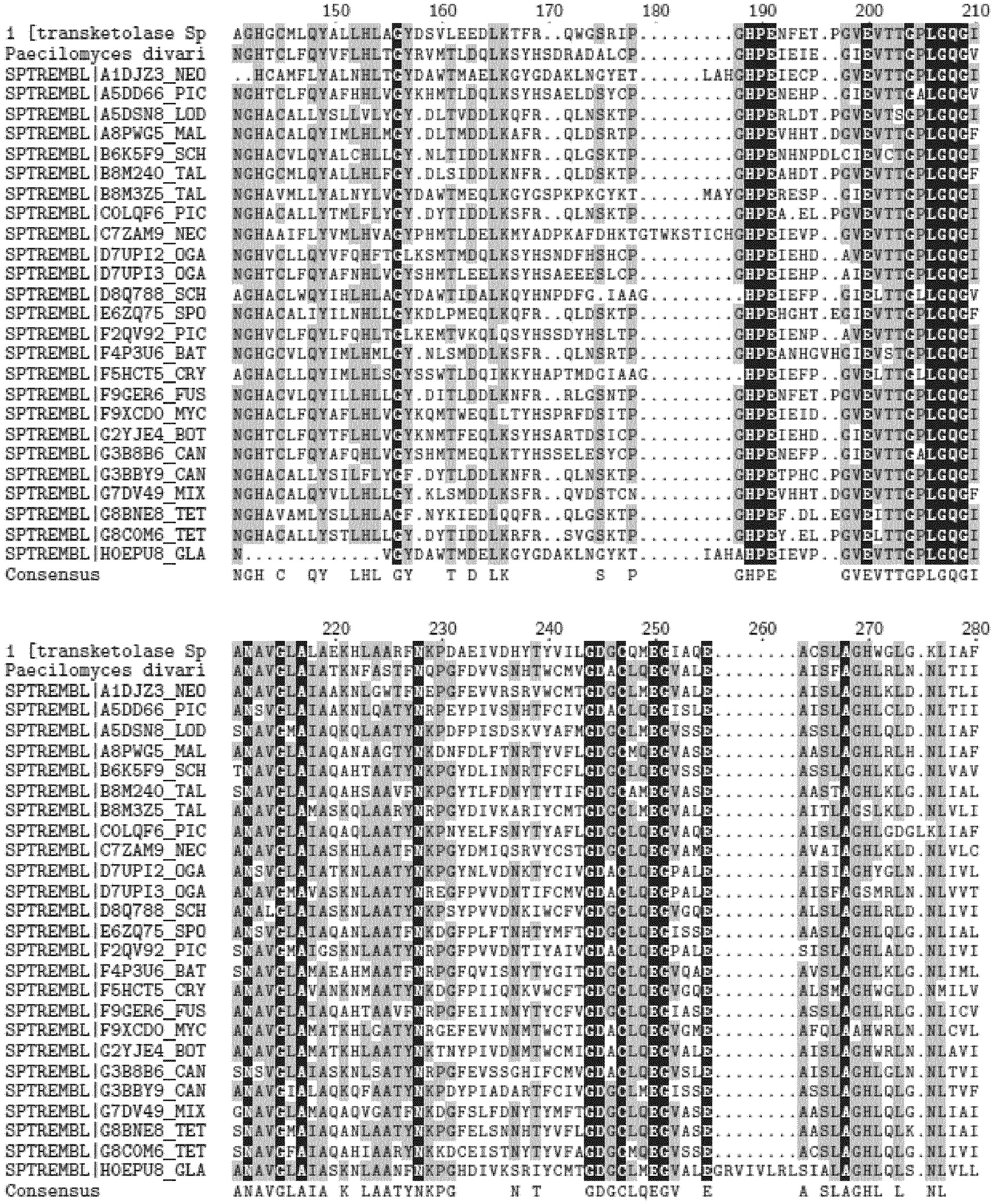
```
                              710       720       730       740       750       760       770
1 [transketolase Sp   IAAKAGDELRK.EGKAVRVVSFVSWELFEKQSDEYKESVLPS..DVTARVSIEAGSTFGWHKIVGSKGKA
Paecilomyces divari   ICIEAATYLKEKHNVAARVVSIPCFEVFDVQDKEYRLKVFP...DGIPVLSVEALSTMGWERYS...HEQ
SPTREMBL|A5DSN8_LOD   ISIDAAKLLAKEKNQKVAVVSIPDFHTFDQQPIDYRLSILP...DGVPIMSVEVLSTFGWSKYS...HEH
SPTREMBL|A8PWG5_MAL   ICVDAVAKLAE.KGIKARVVSIPCFRVFDLQPYDYRISVLP...SGHPILSVEAYSTFGWGVYS...HVH
SPTREMBL|B6K5F9_SCH   LCVEAAKVLEAEHGLKARIVSMPCCEIFDEQAQEYRLSVLP...DGIPIMSVEVYSTVGWKRYA...HEA
SPTREMBL|C0LQF6_PIC   IAIDGAKKLGT.EGVKASVVSIPDFFTFDSQARSYQLSVLP...DGVPIMSVEVMSTFGWSKYS...HEQ
SPTREMBL|C7ZAM9_NEC   FAVGAAAKLAEQ.GIGTRVVSMPCLRLFEAQSKDYQ..................................
SPTREMBL|D7UPI2_OGA   FAYKAAKILGRK.GLKVRVLSIPCTRLFDEHSLGSRRSVLRKDGSQVPPVIVDGHVAFGWERYS...TAS
SPTREMBL|D7UPI3_OGA   FAMGSAAILRSQ.GYNVRVVSFPCQRLFEQQSAEYRHSVLMR.QQMIPTVVIEAYAPNGWERYA...TAG
SPTREMBL|D8Q788_SCH   RAVEAAKKLAPL..R.TRVVSMPIQSVFDRQPAEYRRATLAT..GKALVVAIEPWASYGWARYA...HAS
SPTREMBL|F4P3U6_BAT   LAVDTAAMLAK.EGIQARVVSMPSWELFEDQSHEYRASVFL...DNIPTLSMEAMTTLGWSKYA...HAS
SPTREMBL|F5HCT5_CRY   RAVDTAELLKDK..YSVRVVSMPHTGRFDAQPLEYRRSVIPS..TKSLVVSIEPYASFGWAKYA...HAG
SPTREMBL|G3BBY9_CAN   IAVEAAKLLGK.NGTRARVVSMPSFFAFDQQSEKYRFSVLP...DGVPILSIEVMSPFGWSNYS...HVQ
SPTREMBL|G7DV49_MIX   ICMEAAKILMD.KGVKVRVVSLPCFEIFAAQPKDYRMSILP...DGHPVLSVEAYSTVSWGALA...HEH
SPTREMBL|G8C0M6_TET   LCCKSAEKYSKDHKVNIRVVSLPDFYTFGKQPFEYRISVLP...DEVPILSVEVYATLSMCKYS...HEQ
SPTREMBL|G9MFC9_HYP   HVLDAAAVLIETKGIRCRVVSFPCQRLFEGQSLEYKRNTLRR.HHGIPAVVVEPYTPNGWERYA...DAG
SPTREMBL|H0GCH3_9SA   ISIDAAKKLYDTKKIKARVVSLPDFYTFDRQSEEYRFSVLP...DGVPIMSFEVLATSSWGKYA...HQS
SPTREMBL|I1BXP2_RHI   IAVEASKKLNE.KKIPNRVVSMPCTEIFDEQSIEYKKQVFP...VGIPVISIEALGTFGWDRYS...HAH
SPTREMBL|I1RJN7_GIB   FAVKTREVLKNKFGIRARIVSFPCQRLFNSQPLEYKRQVLQY.RSNAPRVVIEAYAVNGWERYA...DAG
SPTREMBL|I2K2H5_DEK   LAVDAAGLLEA.EGIPTRVVSLPDFFTFDHQPAEYRKAVLP...DGVPILSVEVLSTFGWGKYS...HVH
SPTREMBL|J9N9S0_FUS   LTMQAKDVLLKEHGIKSRVVSFPCPRLFEQQSRQYKQSVLKP.RSGKATVVIEAYAANGWERYA...DAS
SPTREMBL|J9VL23_CRY   LCLQALEQLKS.KGIKARLVSLPCFEVFNNQPKDYKLSVLP...SGAPILSVEAYSTFGWGTYS...HDH
SPTREMBL|L0PF66_PNE   LCVDAIKVFEEKYNLRARLVSMPCFEVFDQQDFDYKLSVLK...DGIPVLGVEVYSTLGWKNYS...HEQ
SPTREMBL|L2FB38_COL   YAVEAAKTLSTC.GVPTRVVSCPSFTHFDQQSKQYRRSVFPT..DGTPIVSVEEYVATTWARYV...TAS
SPTREMBL|Q5KHG5_CRY   RAIESAKLLKSV..KKVRVVSMPSQKHFDIQTAEYKELVLKS..STALVVAIEAWASYGWARYA...HAS
SPTREMBL|Q6BN17_DEB   FAIDAAEILIES.NINVKIISFPCQRLFECQSTEYKRSVL...DPQIVTVAIEAYASNGWERYA...NAG
SPTREMBL|Q6BXS5_DEB   FALETANILNEK.GIKTRVISFPSHNLFEMQPDEYKRSILKR.GK.IPTVVIEAFAPNGWERYA...TAG
Consensus                 AA  L    G   RVVS P    F  Q   Y  SVL       P    E     GW  Y    H

                              780       790       800       810       820       830       840
1 [transketolase Sp   IGIDKFGASAPAGKIYQEYGITVEAVVEAAKSVC....................................
Paecilomyces divari   FGLNRFGASGPYKEVYKKFEFTPEGISKRAIATIDFYK.GHK.VRSPVNRAFQQLL..............
SPTREMBL|A5DSN8_LOD   FGLNRFGASGKAADLYKFFGFTPEGIAERAEKTIQFYK.GKV.VGSPVHTAFEELK..............
SPTREMBL|A8PWG5_MAL   HGINRFGVSAPYQKAYEKFQLTDHDIADKAVKVVEYYKSMGHIHSPIAVQAAFELQK.............
SPTREMBL|B6K5F9_SCH   FGIDSFGDSAPYQQLYAKYEMTPEGVAVRAKMTVDAWK.PVQNIRSPIHHAF..................
SPTREMBL|C0LQF6_PIC   FGINRFGISGPGPEIYKFFEFTAEGVADRASKVVQFYK.GKD.FYRHLISFESVMDVRLIAP........
SPTREMBL|C7ZAM9_NEC   ................EHMSLDS..............GHASA..RLLLG.....................
SPTREMBL|D7UPI2_OGA   YCMNTYGKSLPPDVIYEYFGYNPNTIAKKVEAYVKACRAD....PLLLHDYVDLKEKPKHDKVNKL....
SPTREMBL|D7UPI3_OGA   INMKTFGKSLPGMVCYDFFGYNKEKIATKVDAYLKQIKDT....PSVVYEFQDLN...............
SPTREMBL|D8Q788_SCH   MSMHTFGLSAPQETLYELFGYGIDTIVDKVSKFVESKRVDGEIVLPGVGEFEELLLGYAKKHAGPELGLH
SPTREMBL|F4P3U6_BAT   CGIDTFGASAPYLLLYKKFGLVPEAVSEKAKKVIEYYK..TRTPESKMSPALF.................
SPTREMBL|F5HCT5_CRY   AHMTGFGHSAPYSVLFEHFGFGPKTLAEKIRAWAEPKRNGDGWDIPGVGEFEELLVNSGNGH........
SPTREMBL|G3BBY9_CAN   FGVSRFGFSAKYTDIYDYLEFTPEGISSRANKTIDFYK.GSK.LLSPLSKPFYPQSFDLFTHGR......
SPTREMBL|G7DV49_MIX   YGMSTFGASGPYKDVYKKFGFTGENIASVGEKVIAFWK..GKTILSPQNTCLVAI...............
SPTREMBL|G8C0M6_TET   FGLDRFGMSGKSKEVYEYFEFTPDGVASRIAKVLSFYS.DKT.IRSPVRKALA.................
SPTREMBL|G9MFC9_HYP   ISVRRFGHSLPGPEAYKYFGFDVDSIVSKLSSYLTMIHQD....EILKQEFIDL................
SPTREMBL|H0GCH3_9SA   FGLDEFGRSGKGPEIYKLFDFTADGVASRAEKTINYYK.GKQ.LLSPMGRAF..................
SPTREMBL|I1BXP2_RHI   IGLHTFGSCGKVGDLYKHFGITADHAVEKAQKVIEHFK..KLGHVPEVNVQL..................
SPTREMBL|I1RJN7_GIB   YSMSTFGKSLPGKDAYKYFGFDETVIAPEVARLVEEVKEEGI..ESLRGEFRDLN...NVTHEH......
SPTREMBL|I2K2H5_DEK   LGLDRFGASGKGPEIFKRFGFTAEGIAQKGKQTVQFFK.GKQ.VISPLSAPDFAVRQ.............
SPTREMBL|J9N9S0_FUS   ISMRRFGKSLPSKATYEYFGFQPHNVATKIRGLVEEVKRDGI..EILRGDFRDLNGSLGVGFEH......
SPTREMBL|J9VL23_CRY   FGLKAWGASGPYDKVYAKFDITPEGIARRAEKVVDFYKKRGQPVFSPLISALDDISE.............
SPTREMBL|L0PF66_PNE   FGLDTFGASAPYKDLYKKYGFTVEEIALKAKETVDFWK.DVKPLRSPIRRAF..................
SPTREMBL|L2FB38_COL   ISMVSYGYSASNPSNYDRFELDSKGIVKRVRRYLDDLGEENA..RMAGWRQI..................
SPTREMBL|Q5KHG5_CRY   LSMQTFGHSAPQQQLYEHFGFSPKAMAEKIDEWAARWQTKGG..RPGLGDFEELLLRYTQH.........
SPTREMBL|Q6BN17_DEB   FHLNEFGISLPGKNAYEHFGFNGAYIASKIQKYLDDLQKD....DIMKFEYQELNITKNHHSSS......
SPTREMBL|Q6BXS5_DEB   INMKTYGKSLPGAAAYEYFGFNSTVIASKVEKYFSRWNQE....HEIRYEFLDLN...............
Consensus                  FG S P    Y  FG      A
```

X non conserved
X ≥ 50% conserved
X all match

Figure 1G

```
                              10        20        30        40        50        60        70
1 [transketolase Sp  MAASSSLSTLSHHQTLLSHPKTHLPTTPASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASAVVRAAAVE
Paecilomyces divari  ...................................................MAGNGPNKKTELVSSTVTI
SPTREMBL|A1DJZ3_NEO  ..........................................MSP..................VTSRLIE
SPTREMBL|A5DD66_PIC  .............................................................MTTDD....
SPTREMBL|A5DSN8_LOD  ......................................................................
SPTREMBL|A8PWG5_MAL  ......................................................................
SPTREMBL|B6K5F9_SCH  ......................................................................
SPTREMBL|B8M240_TAL  ......................................................................
SPTREMBL|B8M3Z5_TAL  ......................MSDKLLQQQGLGQDDRENSQFSE..................KGRRIND
SPTREMBL|C0LQF6_PIC  ......................................................................
SPTREMBL|C7ZAM9_NEC  ..........................................MVP.................ANSRKLTA
SPTREMBL|D7UPI2_OGA  .............................................................MRIARAVST
SPTREMBL|D7UPI3_OGA  .............................................................MTVSESSAS
SPTREMBL|D8Q788_SCH  ..........................................MSSVVEAG....ALSGLSADPKVPAKGA
SPTREMBL|E6ZQ75_SPO  ......................................................................
SPTREMBL|F2QV92_PIC  ............................................................MARIPKAVSY
SPTREMBL|F4P3U6_BAT  ......................................................................
SPTREMBL|F5HCT5_CRY  ..........................................MSPVAVHNKHGDSASHGTALSKNPHVKS
SPTREMBL|F9GER6_FUS  ......................................MYAIIDPEEIKEYYIVPWSCTGETRPGKPVPL
SPTREMBL|F9XCD0_MYC  ...................................................................MEE
SPTREMBL|G2YJE4_BOT  ...................................MAPSAVETPSPTDLKMQLSPNSGVPDELVHVKFEG
SPTREMBL|G3B8B6_CAN  .............................................................MTVPESLQN
SPTREMBL|G3BBY9_CAN  ......................................................................
SPTREMBL|G7DV49_MIX  .....................................MTPALAPLELLRNDSTEGDASIASAPYRKTRQP
SPTREMBL|G8BNE8_TET  ......................................................................
SPTREMBL|G8C0M6_TET  ......................................................................
SPTREMBL|H0EPU8_GLA  ......................MAPGLVETGLLAVNHDDNKSKSHNLKADVAPAESPAQTTWETKKEVTG
Consensus

                              80        90       100       110       120       130       140
1 [transketolase Sp  ALESTDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEIMRYNPKNPYWFNRDRFVLS
Paecilomyces divari  LNKE......HDLVLKTFRLLIADLCRQFNGGHPGGAMGMAAIGIALWKYVMRYAPHTPDFFNRDRFVLS
SPTREMBL|A1DJZ3_NEO  ........EDDALAVLTLRNLIFDICNQNGGGHGGSAIGMAAIGVALWKYIMRYNPSNPQWFDRDRMTVG
SPTREMBL|A5DD66_PIC  ..........YDKILVTFRILIADLVQQYNGGHPGGAMGMAAIGVALWKHVLRISPSNPNFFNRDRFVLS
SPTREMBL|A5DSN8_LOD  ...MAPS..IDELSVSTIRGLSVDQVNAAQSGHPGAPLGLAPAAHALWRKMR.FNPQDPKWINRDRFVLS
SPTREMBL|A8PWG5_MAL  ..MVFTPSEIDLKSIATIRTLAADVVANAKSGHPGAPMGLAPLAHLLWSRIMVADPKESHWINRDRFVLS
SPTREMBL|B6K5F9_SCH  ..MVSTYTDLDTKAVNTIRTLAVDTTCKSKSGHPGAPMGMAPAAHVLFSRIMNFNPQKPRWINRDRFVLS
SPTREMBL|B8M240_TAL  ....MAFTEVDQLAVNTIRVLAVDGTAKANSGHPGAPMGLAPTAHVLFHKFMNFNPKNPKWANRDRFVLS
SPTREMBL|B8M3Z5_TAL  TDYS....EEDTLAVLSLRNLLFDMCNQNGGGHGGSAIGMAAIGVALWKYIMRHNPSNPEWFDRDRFVLS
SPTREMBL|C0LQF6_PIC  ...MS.....DLLAINTIRLLAVDTVAKANSGHPGAPMGLAPAAHTLFKQMR.FNPRNPAWINRDRFVLS
SPTREMBL|C7ZAM9_NEC  ..........DDIAVRDIRKLVIDCCRQNGGGHGGSAIGMAPLGVALWRHTLRFSTKNPAWFDRDRFVLS
SPTREMBL|D7UPI2_OGA  SDYE......HDQIIKYGRALVLDIVQQYDGGHPGSAMGGMALGIALWKYTMKYAPNDPTYFNRDRFVLS
SPTREMBL|D7UPI3_OGA  LIQAKQDQTKFDFMLKYYRNLIVDLVHNYKGGHGGGPNGMAAIGFALYKYVMKYNPENPSYFNRDRFILS
SPTREMBL|D8Q788_SCH  IDPE.....IEKLAVNTVRVLGADLTQAFKGGHPGTVMGAAAIGVALWRYQMRYNPKNPDWFNRDRFVLS
SPTREMBL|E6ZQ75_SPO  ..MSFTKAASDEKAIATIRVLAADVVAKSNSGHPGAPMGMAPMAHTLFSRFMNINPAHSKWINRDRFVLS
SPTREMBL|F2QV92_PIC  NDDI......HDLVIKTFRCYVLDLVEQYGGGHPGSAMGMVAIGIALWKYQMKYAPNDPDYFNRDRFVLS
SPTREMBL|F4P3U6_BAT  .....MPSTMDILAINTIRTLAADVVQKANSGHPGAPMGCAPMAHVLFSNFICQDPQHPDWISRDRFVLS
SPTREMBL|F5HCT5_CRY  SEVD.....TEKLVINAIRCLAADLCQQYKGGHPGTVMGAAAIATALWKYSMRYNPANPDWINRDRFVLS
SPTREMBL|F9GER6_FUS  NSHTATMNHSDETSINTIRLLAADATFRSNSGHPGAPMGLAPAAHVLFTKFLRCSPADPNWLNRDRFVLS
SPTREMBL|F9XCD0_MYC  FEKT.......DLVLKTFRCLIADLCHQFGCGHPGSAMGMAGIGVALWKYVMKYSPSNPDFFNRDRFVLS
SPTREMBL|G2YJE4_BOT  AELN.......SHVLKTFRCLIADLCEQFKGGHPGSAMGMAAIGVALWKYVMKYTPENPKFFNRDRFVLS
SPTREMBL|G3B8B6_CAN  RSQT......HQYVLKIFRVLIADLVQQFNGGHPGGAMDMAAIGIALFKYVMKYSPSNPDYFNRDRFVLS
SPTREMBL|G3BBY9_CAN  ...MTKD..IDQLSINTIKTLAVDSVSAANSGHPGAPLALATLAHAIFQKTK.FTPNHPKWPNRDRFVLS
SPTREMBL|G7DV49_MIX  VKMTFEAHADDWTCINTIRTLSLDVVNKANSGHPGAPMGMAPVAHILWSRFMRCNPKSSHWPNRDRFVLS
SPTREMBL|G8BNE8_TET  ...MSHFTEIDNLAVNTLRLLSVDTVDAAQSGHPGAPLGMSPAAHVVWTKFLNLNPKNPDFCNRDRFVLS
SPTREMBL|G8C0M6_TET  ...MSKFNDIDRLCISTVRLLAVDAVQKANSGHPGAPLGLAPAMHVLCRDFMNVNHENVDWPNRDRFVLS
SPTREMBL|H0EPU8_GLA  WIPE....TDDDFAILAVRNLVLDICMQNGGGHGGSALGMAAIGVALWKHVMRYNPANSEWFDRDRFVLS
Consensus                            T R L  D       GHPG  MG A     L    M   P  P W NRDRFVLS
```

Figure 1G continued



```
                              150       160       170       180       190       200       210
1 [transketolase Sp    AGHGCMLQYALLHLAGYDSVLEEDLKTFR..QWGSRIP.........GHPENFET.PGVEVTTGPLGQGI
Paecilomyces divari    NGHTCLFQYVFLHLTGYRVMTLDQLKSYHSDRADALCP.........GHPEIEIE..GIEVTTGPLGQGV
SPTREMBL|A1DJZ3_NEO    ..HCAMFLYALNHLTGYDAWTMAELKGYGDAKLNGYET......LAHGHPEIECP..GVEVTTGPLGQGI
SPTREMBL|A5DD66_PIC    NGHTCLFQYAFHHLVGYKHMTLDQLKSYHSAELDSYCP.........GHPENEHP..GIEVTTGALGQGV
SPTREMBL|A5DSN8_LOD    NGHACALLYSLLVLYGY.DLTVDDLKQFR..QLNSKTP.........GHPERLDT.PGVEVTSGPLGQGI
SPTREMBL|A8PWG5_MAL    NGHACALQYIMLHLMGY.DLTMDDLKAFR..QLDSRTP.........GHPEVHHT.DGVEVTTGPLGQGF
SPTREMBL|B6K5F9_SCH    NGHACVLQYALCHLLGY.NLTIDDLKNFR..QLGSKTP.........GHPENHNPDLCIEVCTGPLGQGI
SPTREMBL|B8M240_TAL    NGHGCMLQYALLHLFGY.DLSIDDLKNFR..QLDSKTP.........GHPEAHDT.PGVEVTTGPLGQGF
SPTREMBL|B8M3Z5_TAL    NGHAVMLLYALNYLVGYDAWTMEQLKGYGSPKPKGYKT......MAYGHPERESP..GIEVTTGPLGQGI
SPTREMBL|C0LQF6_PIC    NGHACALLYTMLFLYGY.DYTIDDLKSFR..QLNSKTP.........GHPEA.EL.PGVEVTTGPLGQGI
SPTREMBL|C7ZAM9_NEC    NGHAAIFLYVMLHVAGYPHMTLDELKMYADPKAFDHKTGTWKSTICHGHPEIEVP..GVEVTTGPLGQGI
SPTREMBL|D7UPI2_OGA    NGHVCLLQYVFQHFTGLKSMTMDQLKSYHSNDFHSHCP.........GHPEIEHD..AVEVTTGPLGQGI
SPTREMBL|D7UPI3_OGA    NGHTCLFQYAFNHLVGYSHMTLEELKSYHSAEEESLCP.........GHPEIEHP..AIEVTTGPLGQGI
SPTREMBL|D8Q788_SCH    AGHACLWQYIHLHLAGYDAWTIDALKQYHNPDFG.IAAG.........HPEIEFP..GIELTTGLLGQGV
SPTREMBL|E6ZQ75_SPO    NGHACALIYILNHLLGYKDLPMEQLKQFR..QLDSKTP.........GHPEHGHT.EGIEVTTGPLGQGF
SPTREMBL|F2QV92_PIC    NGHVCLFQYLFQHLTGLKEMTVKQLQSYHSSDYHSLTP.........GHPEIENP..AVEVTTGPLGQGI
SPTREMBL|F4P3U6_BAT    NGHGCVLQYIMLHMLGY.NLSMDDLKSFR..QLNSRTP.........GHPEANHGVHGIEVSTGPLGQGI
SPTREMBL|F5HCT5_CRY    AGHACLLQYIMLHLSGYSSWTLDQIKKYHAPTMDGIAAG.........HPEIEFP..GVELTTGLLGQGI
SPTREMBL|F9GER6_FUS    NGHACVLQYILLHLLGY.DITLDDLKNFR..RLGSNTP.........GHPENFET.PGVEVTTGPLGQGI
SPTREMBL|F9XCD0_MYC    NGHACLFQYAFLHLVGYKQMTWEQLLTYHSPRFDSITP.........GHPEIEID..GVEVTTGPLGQGI
SPTREMBL|G2YJE4_BOT    NGHTCLFQYTFLHLVGYKNMTFEQLKSYHSARTDSICP.........GHPEIEHD..GIEVTTGPLGQGI
SPTREMBL|G3B8B6_CAN    NGHTCLFQYAFQHLVGYSHMTMEQLKTYHSSELESYCP.........GHPENEFP..GIEVTTGALGQGI
SPTREMBL|G3BBY9_CAN    NGHACALLYSILFLYGF.DYTLDDLKNFR..QLNSKTP.........GHPETPHC.PGVEVTTGPLGQGI
SPTREMBL|G7DV49_MIX    NGHACALQYVLLHLLGY.KLSMDDLKSFR..QVDSTCN.........GHPEVHHT.DGVEVTTGPLGQGF
SPTREMBL|G8BNE8_TET    NGHAVAMLYSLLHLAGF.NYKIEDLQQFR..QLGSKTP.........GHPEF.DL.EGVEITTGPLGQGI
SPTREMBL|G8C0M6_TET    NGHACALLYSTLHLLGY.DYTIDDLKRFR..SVGSKTP.........GHPEY.EL.PGVELTTGPLGQGI
SPTREMBL|H0EPU8_GLA    N.............VGYDAWTMDELKGYGDAKLNGYKT......IAHAHPEIEVP..GVEVTTGPLGQGI
Consensus              NGH C  QY  LHL GY   T D LK        S  P         GHPE      GVEVTTGPLGQGI

                              220       230       240       250       260       270       280
1 [transketolase Sp    ANAVGLALAEKHLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQE........ACSLAGHWGLG.KLIAF
Paecilomyces divari    ANAVGLAIATKNFASTFNQPGFDVVSNHTWCMVGDACLQEGVALE........AISFAGHLRLN.NLTII
SPTREMBL|A1DJZ3_NEO    ANAVGLAIAAKNLGWTFNEPGFEVVRSRVWCMTGDGCLMEGVALE........AISLAGHLKLD.NLTLI
SPTREMBL|A5DD66_PIC    ANSVGLAIAAKNLQATYNRPEYPIVSNHTFCIVGDACLQEGISLE........AISLAGHLCLD.NLTII
SPTREMBL|A5DSN8_LOD    SNAVGMAIAQKQLAATYNKPDFPISDSKVYAFMGDGCLMEGVSSE........ASSLAGHLQLD.NLIAF
SPTREMBL|A8PWG5_MAL    ANAVGLAIAQANAAGTYNKDNFDLFTNRTYVFLGDGCMQEGVASE........AASLAGHLRLH.NLIAF
SPTREMBL|B6K5F9_SCH    TNAVGLAIAQAHTAATYNKPGYDLINNRTFCFLGDGCLQEGVSSE........ASSLAGHLKLG.NLVAV
SPTREMBL|B8M240_TAL    SNAVGLAIAQAHSAAVFNKPGYTLFDNYTYTIFGDGCAMEGVASE........AASTAGHLKLG.NLIAL
SPTREMBL|B8M3Z5_TAL    ANAVGLAMASKQLAARYNRPGYDIVKARIYCMTGDGCLMEGVALE........AITLAGSLKLD.NLVLI
SPTREMBL|C0LQF6_PIC    ANAVGLAIAQAQLAATYNKPNYELFSNYTYAFLGDGCLQEGVAQE........AISLAGHLGDGLKLIAF
SPTREMBL|C7ZAM9_NEC    ANAVGLAIASKHLAATFNKPGYDMIQSRVYCSTGDGCLQEGVAME........AVAIAGHLKLD.NLVLC
SPTREMBL|D7UPI2_OGA    ANSVGLAIATKNLAATYNKPGYNLVDNKTYCIVGDACLQEGPALE........AISIAGHYGLN.NLIVL
SPTREMBL|D7UPI3_OGA    ANAVGMAVASKNLAATYNREGFPVVDNTIFCMVGDACLQEGPALE........AISFAGSMRLN.NLVVT
SPTREMBL|D8Q788_SCH    ANALGLAIASKNLAATYNKPSYPVVDNKIWCFVGDGCLQEGVGQE........ALSLAGHLRLD.NLIVI
SPTREMBL|E6ZQ75_SPO    ANSVGLAIAQANLAATFNKDGFPLFTNHTYMFTGDGCLQEGISSE........AASLAGHLQLG.NLIAL
SPTREMBL|F2QV92_PIC    SNAVGMAIGSKNLAATYNRPGFPVVDNTIYAIVGDACLQEGPALE........SISLAGHLALD.NLIVI
SPTREMBL|F4P3U6_BAT    SNAVGLAMAEAHMAATFNRPGFQVISNYTYGITGDGCLQEGVQAE........AVSLAGHLKLG.NLIML
SPTREMBL|F5HCT5_CRY    ANAVGLAVANKNMAATYNKDGFPIIQNKVWCFTGDGCLQEGVGQE........ALSMAGHWGLD.NMILV
SPTREMBL|F9GER6_FUS    ANAVGLAIAQAHTAAVFNRPGFEIINNYTYCFVGDGCLQEGIASE........ASSLAGHLRLG.NLICV
SPTREMBL|F9XCD0_MYC    ANAVGLAMATKHLGATYNRGEFEVVNNMTWCTIGDACLQEGVGME........AFQLAAHWRLN.NLCVL
SPTREMBL|G2YJE4_BOT    ANAVGLAMATKHLAATYNKTNYPIVDNMTWCMIGDACLQEGVALE........AISLAGHWRLN.NLAVI
SPTREMBL|G3B8B6_CAN    SNSVGLAIASKNLSATYNRPGFEVSSGHIFCMVGDACLQEGVSLE........AISLAGHLGLN.NLTVI
SPTREMBL|G3BBY9_CAN    ANAVGIALAQKQFAATYNKPDYPIADARTFCIVGDGCLMEGISSE........ASSLAGHLQLG.NLTVF
SPTREMBL|G7DV49_MIX    GNAVGLAMAQAQVGATFNKDGFSLFDNYTYMFTGDGCLQEGVASE........AASLAGHLQLG.NLIAI
SPTREMBL|G8BNE8_TET    SNAVGMAIAQANLAATYNKPGFELSNNHTYVFLGDGCLQEGVASE........AASLAGHLQLK.NLIAI
SPTREMBL|G8C0M6_TET    SNAVGFAIAQAHIAARYNKKDCEISTNYTYVFAGDGCMQEGVSSE........ACSLAGHLQLG.NLIVI
SPTREMBL|H0EPU8_GLA    ANAVGLAIASKNLAANFNKPGHDIVKSRIYCMTGDGCLMEGVALEGRVIVLRLSIALAGHLQLS.NLVLL
Consensus              ANAVGLAIA K LAATYNKPG     N T    GDGCLQEGV  E        A SLAGHL L  NL
```

Figure 1G continued

```
                              290       300       310       320       330       340       350
1 [transketolase Sp   YDDNHISIDGDTAIAFTESVDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFG
Paecilomyces divari   YDNNQITCDGSVDLTNTEDINAKMRACGWDVIDIEDGCFDVSGIIAALNIARESRDKPTFVNVRTVIGVE
SPTREMBL|A1DJZ3_NEO   YDNNGVTCDGPLAWINTEDVNPKMRASGWHVLDVLDGSHNVQSIREALQHAKNLQGKPVFINIRTVIGVG
SPTREMBL|A5DD66_PIC   YDNNQITCDGSVDLTNTENINEKFRACNWNVIEIMDGSEDVSAIVTALEASRKSD.KPTLVNVQTKIGYG
SPTREMBL|A5DSN8_LOD   WDDNRISIDGDTACAFTENVPERYRAYGWHILEVPDANTDIDAISKAIDEAKTIKNKPVLIRLVTTIGYG
SPTREMBL|A8PWG5_MAL   FDDNHVSIDGDTKCSFTEDVPKRFESYGWNVLVVDNGDSDLQALYDAVVKAWDEKTKPTLVCVRTTIGFG
SPTREMBL|B6K5F9_SCH   WDNNKITIDGSTDVSFDEDVEKRYEAYGWQVLRVADGDSDYASIEQALLDAVACKDKPTLINMKTTIGYG
SPTREMBL|B8M240_TAL   YDDNHISIDGDTKCAFTEDVLKRFEAYGWHTLEVKDGDHDLQGIEEAIRKAQQVKDKPSVIKITTTIGFG
SPTREMBL|B8M3Z5_TAL   YDNNHITCDGPLEWTNTEDINAKMRASGWDVIDVQDGNYDVQAIVAALNHASSLRGKPVMINIRTVIGVD
SPTREMBL|C0LQF6_PIC   WDDNQISIDGDTNVSFTEDVPAKFRAQGWEVLSVKDGNDDLEGIAAALAKAKTT.NKPTLIRLPTIIGYG
SPTREMBL|C7ZAM9_NEC   YDNN.............QDTNSKMRSMGWNVIDVFDGDNSVESIVNALNLGKATKHMPTFINIRSTIGYG
SPTREMBL|D7UPI2_OGA   YDNNQVCADGSVDIANTEDISAKFKACNWNVIEVANASEDVATIVKALEYAQNEVKAPTLINCRTVIGAD
SPTREMBL|D7UPI3_OGA   YDNNQISCDGSVDITNTEDINAKFIACNWNVIDVENGSMDICAIVQALEDAKLSD.KPTLINIHTVIGLS
SPTREMBL|D8Q788_SCH   YDDNSVTVDGSIATCFTDDTSAKVKAMGFNVLEVYDGTNDLTAIIKAMDEAKATRGKPSFIHIRTIIGFG
SPTREMBL|E6ZQ75_SPO   YDDNHISIDGDTACSFTEDVEQRFLSYGWHVSHVLDGDKDLEGLYKAISEAKKVTDKPSIIRVKTTIGFG
SPTREMBL|F2QV92_PIC   YDNNQVCCDGSVDVNNTEDISAKFRAQNWNVIDIVDGSRDVATIVKAIDWAKAETERPTLINVRTEIGQD
SPTREMBL|F4P3U6_BAT   YDDNHITIDGDIEVGFTEDVVQRFESYGWHTQVLGDGDNDVAGIIAAIQKAKAVTDKPSLIKIRTTIGFG
SPTREMBL|F5HCT5_CRY   YDNNSVTVDGNIDICFTDDTSAKLKSLGWHVLEVEDGSNDLAAIVDAFEQAQKLTGKPVFINIKTIIGIG
SPTREMBL|F9GER6_FUS   YDDNRITIDGDISCTFTEDVVKRYESYGWHVITVEDGNSSLSAIEAAIQESKSEKSRPSLIKLRTTIGFG
SPTREMBL|F9XCD0_MYC   YDRNEITCDGSVDLCSSDDVNAKMRAVGFEVLEVADGNNDVEALVKAMIQARASKDKPTFINIRTTIGVG
SPTREMBL|G2YJE4_BOT   YDNNQVTCDGSVDLCNNEDVNAKMIACGWNVIDVHDGNYDIEAIVRALMAARNNKDKPTFINIRTVIGIG
SPTREMBL|G3B8B6_CAN   YDNNQITCDGSVDLTNTENIKKKFEACNWQVLEVQDGSNDVDGIIEALETAKKSEGNPTLIIAYTQIGID
SPTREMBL|G3BBY9_CAN   WDSNKISIDGSTDLAFTEDVPERYRAYGWHVIDVDEENPSLSTLSKAIDEASKVTDKPTFVKIKTTIGIG
SPTREMBL|G7DV49_MIX   YDSNKITIDGDTAVSFTEDVTKRMEAYGWHCQTVEQGDDDLEGMYSAILAAQQVKDKPSFITLKTTIGFG
SPTREMBL|G8BNE8_TET   YDHNHITIDGSTDVSFTEDVLKRFEAYGWEVLTVENGDSDLDAIYNAIAQAKQSNK.PTLIQNKTTIGYG
SPTREMBL|G8C0M6_TET   YDDNSITIDGKTNLSFTEEVAVRYRGYGWEVIEVKDGNDNLDAISQAIKLAKSNKKQPTLIKLTTTIGFG
SPTREMBL|H0EPU8_GLA   YDNNQVTCDGPLDWINTEDVNAKMVASGWEVINVADGRYDVRAIVSALDLAKNSSSKPVFINIRTVIGVD
Consensus             YD N I  DG      TEDV     A GW V  V DG  D   I  A   A     KP  I   T IG G

                              360       370       380       390       400       410       420
1 [transketolase Sp   SPNKSNSYSVHGSALGSKEVEATRQNLGW.PYEPFHVPEEVKKHWSRHTPEG.ASLEAEWNTKFAEYEKK
Paecilomyces divari   TAVAG.DAVAHGVALGSKNVADLKTTYGFDPNEHFVISGAVRQFFEDIPGRGEALVR.GWSNLVDEYSRN
SPTREMBL|A1DJZ3_NEO   TSTAG.TAMAHHGAFDTESVAASKIRAGQDPSSTHSPSERGLAFFRERKAAGETLEE.EWNDLISQYAKK
SPTREMBL|A5DD66_PIC   SNIEG.TAAAHGAAFGQDEVDRLHSVNGFPVDKKIVVPQYMYDFFK....KEENEEK.QWNELLTKYQAA
SPTREMBL|A5DSN8_LOD   SKAAG.TGGVHGAPLKADDLKQLKKSWGFDEESNFYIPDEVANYFAEKVESN.KKEQQAWEKLFNEYKQK
SPTREMBL|A8PWG5_MAL   SVNQG.KASTHGSPLKDDDIAQLKQKFGFNPDEKFVVPDDVRAAYKSYAEQG.AAKRAEWEKMFEEYGRK
SPTREMBL|B6K5F9_SCH   SMKQG.THSVHGSPLSAEDAAHVKREFGFNPEQSFYVAPEVYDLYKKHVARG.VEAQQKFDELFAKYKTE
SPTREMBL|B8M240_TAL   SKLQG.TGGVHGNPLKADDAENVKKAFGFDPSQSFVVPQQVYDLYGKTAAAG.AAKEQEWNNLLQKYASE
SPTREMBL|B8M3Z5_TAL   TESAG.TAKAHHSLIDVESIRLSKIRAGQDPSSTHSVPEQALVFFRERKTFGERLEA.NWNILIQKYETA
SPTREMBL|C0LQF6_PIC   SLQQG.THGVHGSPLKPDDIKQLKKKFGFDPEQNFVVPREVTESYAKHVADN.QQVEVEWNKLLTAYTKE
SPTREMBL|C7ZAM9_NEC   TTTAG.TAKSHHGTY...SEADAALYAETSDQASHRVSKECRAYFNERADQGNCDQD.LWLDTLDRYCQD
SPTREMBL|D7UPI2_OGA   AAFEN.HHADHGNSLGADGVREVKKKLGMNPAQKFHVPKEIYQFFSNKITEGDQLVA.DWKKLVDNYVKE
SPTREMBL|D7UPI3_OGA   TPWEN.TAAVHGADIGAANVLKFKETVGIEADKTFYIPDEMYKYFSDIKPKGQAYEA.QWNQLITSYEAA
SPTREMBL|D8Q788_SCH   SRKQN.TGAAHGAALGEDEVSYVKQLFGFDPAKRFVIPDEVYDYFRPVIQRGADDEA.AWNAMFAQYAKA
SPTREMBL|E6ZQ75_SPO   SKLQG.LAATHGAPLKADDIEGVKAKFGMDPKQSFTVPDDVYAAYKEIADRG.AKADKEWTELFKAYGDK
SPTREMBL|F2QV92_PIC   SAFGN.HHAAHGSALGEEGIRELKTKYGFNPAQKFWFPKEVYDFFAEKPAKGDELVK.NWKKLVDSYVKE
SPTREMBL|F4P3U6_BAT   SVLQG.EEKVHGSPLGAKDIIQLKIKFGLDPEVHFHVSDEVYKHYHDASARG.AAGYNAWVKLMQEYAVA
SPTREMBL|F5HCT5_CRY   SINQN.SGKVHGAALGEDDVARVKTALGFDPKDKFIVPDAVYDYFKETKIKGAQYEQ.EWNDLFKRYKES
SPTREMBL|F9GER6_FUS   SSQEG.THAVHGNPLKSDDIAQMKQKWGFNADESFVIPQDVYDIYHQQASSG.SKARDEWMTLMAAYADH
SPTREMBL|F9XCD0_MYC   STVEG.DAKSHGADLGEKDIIQLKKKFGMDPEKTFVVSDEVYNYFREAVPRGQQLEK.DYAAMIAKYTEA
SPTREMBL|G2YJE4_BOT   SQSAG.HAKAHGAAFGAEDVANIKRTFGMDETKHFQVSDEVYAFFRDSKPRGRQIET.DWNQLLKSYATA
SPTREMBL|G3B8B6_CAN   TSVEG.QAVAHGAAFGQSEVDNLHTLYGFDPETKFFVPETVYQFFADVVPRGKQLEK.QWTELVHDYSLQ
SPTREMBL|G3BBY9_CAN   SEQAG.TAAVHGAPLKPDDVKHLKKAIGFDPEQSFVVPSEVTDYFKKSVESK.DKQYVEWQKLFKDYSHH
SPTREMBL|G7DV49_MIX   SKLQG.THGVHGNPLKPDDTEAIKKKFGFDPAAKFAIPDATKTAYAKIGERG.AKLESDWTELLSKYTEK
SPTREMBL|G8BNE8_TET   SVNEG.THSVHGSPLKADDIAQLKTKFGFDPKKKFDVPQEVYDLYKTQLAQRGLESNKIWDQNFAQYKQK
SPTREMBL|G8C0M6_TET   SLNEG.TYNVHGSPLKPDDVKQLKKRFGFNSEESFVVPQEVYDFYKEHMIERGKKANEEWNSTLEEYMRK
SPTREMBL|H0EPU8_GLA   MASAG.TAKAHHGGFDQDSIAASKKLAGLLPSSRYEIPERPLQFLRESRDKGTAAEQ.SWDKSMDKYAEE
Consensus             S   G     HG  L        K   G  P   F VP  V          G       W  L   Y
```

Figure 1G continued

```
                              430       440       450       460       470       480       490
1 [transketolase Sp   YPEDATEFKSITTG.EFPAGWEKALP.TYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTL
Paecilomyces divari   FPELATEFRSRIKG.ELPLDWQQLIP..TEFPSKPTASRASSGLVFNTIAEKVKTFMVGTADLSPSVNMA
SPTREMBL|A1DJZ3_NEO   HPVKAAKFAARRSGNLDTDSITKILMRLDSTVFEGTATRESNGLILQKLWAECPALCGGGADLVTSNKFL
SPTREMBL|A5DD66_PIC   HPELAKEFLNRVNG.KLPSNWESLVP..TQFPSTPTASRKSSGLCVQNLAENIKSLFFGTADLSPSVNLS
SPTREMBL|A5DSN8_LOD   YPKEGADLQRRLEG.KFPDGWKDNIP.TFTAKDKALATRKLSEGVLDYLYGVLPELIGGSADLTGSNLTR
SPTREMBL|A8PWG5_MAL   YENEHKEICRRIAH.KLPEGWDQVLP.RYEPSDKPMATRKLSELAITKLASVLPEMLGGSADLTGSNLTR
SPTREMBL|B6K5F9_SCH   YPELAAQFERVVEK.KLPEGWEKKLP.TYKPTDPAIASRKLSEIVLDSIAPVLPELVGGSADLTGSNLTR
SPTREMBL|B8M240_TAL   YKSEHADLTRRLAG.KLPEGWEKNLP.TYKSTDAAIASRKLSETVLEKIYAAVPELLSGSADLTSSNNTR
SPTREMBL|B8M3Z5_TAL   HPDTGTKFRNQRAG.IDPTPCLKALSDMSSLGFDGKATRESNGAVMSKIWELHPGVFAGGADLVNSNKIP
SPTREMBL|C0LQF6_PIC   YPELGQELHRRLDG.KLPENWQKALP.TYTVDDKPVASRKLSEIVLTSIEKELPELVGGSADLTGSNLTR
SPTREMBL|C7ZAM9_NEC   FPREGTLLRARIAGEVNYGDVLSRIQFPQEP....AATRSLNGVIFQKLMDHVPNIIAGGSDLWNSNQMG
SPTREMBL|D7UPI2_OGA   YPELGKEFLARVSG.ELPADWKSSLP..VQDYAGDTPTRAAARGLVQAAGKAIPNIMAGCADLSVSVNLQ
SPTREMBL|D7UPI3_OGA   YPELAAEFQIKIKG.ELPANWKDYIP..TSFPNADTPSRKSGGLVLNPIAQHLNQFLVGTADLSPSVNMI
SPTREMBL|D8Q788_SCH   HPAEHAELTQRLTG.TLPDGWKNDLPPKSALPTSAQATRKSSGIAVQALVPKYKHFIAGSADLMESTFVN
SPTREMBL|E6ZQ75_SPO   YPNEHAEIARRIAG.KLPEGWEKALP.VYQPSDAAVASRKLSETVLGKVADAVPELLGGSADLTGSNLTR
SPTREMBL|F2QV92_PIC   YPREGQEFLSRVRG.ELPKNWRTYIP..QDKPTEPTATRTSAREIVRALGKNLPQVIAGSGDLSVSILLN
SPTREMBL|F4P3U6_BAT   HPDLASELNRRLNN.ELPTGWKDLLP.KYTPADPAVATRKLSENVLNKIALSIPELVGGSADLTGSNLTR
SPTREMBL|F5HCT5_CRY   FPSEAAEFQRRLDG.KLEEAWENKFPSKDALPKDPKATRQSSGIALRSVIPEDKSFLVGSADLCESTFVN
SPTREMBL|F9GER6_FUS   FPAEYADIQRRVKG.NMPQDWESALP.NFSPSGESVATRKLSETVLSRIHTTIPELIGGSADLTPSNLTR
SPTREMBL|F9XCD0_MYC   HPQLAAEFQKRRNG.EMIADWTKMVPKLEDFPTEATPTRKAAGQVLNPIAENISNLMVGTADLTPSVNMA
SPTREMBL|G2YJE4_BOT   HPVLYEEFMLRVAG.KLTDDWAKFIPAKADFPTSPTPSRKSSGLVCNPLAENIKSFMVGTADLSPSVNMM
SPTREMBL|G3B8B6_CAN   YPELALEFENRRLG.KLPPNWQSFIP..NSFPTTPTPSRKSSGLVFNEVAKNVNNFIVGTADLSPSVNLI
SPTREMBL|G3BBY9_CAN   YPELAEEINRREDG.RLPPDWERFLP.SYKPSDKAAATRKLSETVLGALTPVIPELIGGAADLTPSTLTR
SPTREMBL|G7DV49_MIX   FSAEGKDLERRLAG.KLPDGWEKCLP.VYSQKDDAIASRKLSETVITKLADAVPEFVCGSADLTGSNLTR
SPTREMBL|G8BNE8_TET   YPELGAEFQRRLNG.ELPANWDSTLP.VYTPKDSAIATRKLSEMLLDSIHSTLPEVLGGSADLTPSNLTR
SPTREMBL|G8C0M6_TET   YPELGAELKRRIAG.ILPPDWESALP.TFKPEDGDLATRKLSASVINSLAKTLPELIGGSADLTPSNLTR
SPTREMBL|H0EPU8_GLA   YPEEAAEFASRQKG..TQGDFEKVLNEATS........................................
Consensus              P    E   R  G  LP  W   LP           A RK S          P    G ADL  S

                              500       510       520       530       540       550       560
1 [transketolase Sp   LKMFGDFQKD..........TPEERNVRFGVREHGMGAICNGICLHS.PGFVPYCATFFVFTDYMRGAMR
Paecilomyces divari   FPGKVDFQNPELRTTCGINWNYSGRYIHYGIREHAMCAISNGIAAFAPNTFIPVTSSFFMFYLYAAPAVR
SPTREMBL|A1DJZ3_NEO   YSETDIFFPNT.......GAGCKGRYIRYGIREHAMAAISNGLAAYHPRTFLPVTATFLMFYLYAAPGVR
SPTREMBL|A5DD66_PIC   WKSKVDFQNPAIDTACGTNGDYSGRYLHYGIREHSMAGIANGIAAYNPGTFIPVTSSFFMFYLYAAPAVR
SPTREMBL|A5DSN8_LOD   PEASTDFQPP.....STNLGNYGGRYIRYGVREHGMGAIMNGIDTFG.AHSRTYGGTFLNFVSYAAGALR
SPTREMBL|A8PWG5_MAL   WSDATDFQNP.....SLEIGKYSGRYIRYGVREHAMGAIMNGLCAYG..MHIPTAGTFLNFVTYGWGAVR
SPTREMBL|B6K5F9_SCH   WKGAVDFQPK.....ESNLGDYSGRYIRYGVREHAMGGIMNGMAAYG..TVIPFGGTFLNFVSYAAGAVR
SPTREMBL|B8M240_TAL   WKAAVDFQPP.....EYGIGDWSGRYLRYGVREHAMGAIMNGLAAYG..TVIPAGGTFFNFVSYAAGAVR
SPTREMBL|B8M3Z5_TAL   YTESDIFHP.........SVSYSGRFVRYGIREHAMAGIANGLAAYNEGTFLPITATFFMFYLYAAPSVR
SPTREMBL|C0LQF6_PIC   WPDAVDFQPK.....STGLGDFSGRYFRFGVREHGMGAIINGISAYG.ANFKAYGGTFLNFVSYASGAVR
SPTREMBL|C7ZAM9_NEC   DHSHRIFDG..........SRREGRVIRYGIREHAMAAISNGLAAYSPNAIIPVTATFFMFYLYAAPGVR
SPTREMBL|D7UPI2_OGA   WPGVTYFQDPSLRTNCGLTGDYSGRYLEYGIREHAMCAIANGMAAFNKGTFIPITSTFFMFYLYAAPAIR
SPTREMBL|D7UPI3_OGA   WPGKVDFQDPKKETACGLNGDYTGRYIHYGIREHAMCAIANGIAAYNKGTFIPVTSTFFMFYLYAAPAVR
SPTREMBL|D8Q788_SCH   FKGQVEFQSPDTG.....FGDYSGRQLRYGIREFAMVGAANGMNAYQNGMIIPICSSYLAFWQYAMSACR
SPTREMBL|E6ZQ75_SPO   WGNAVDFQAP.....STQLGDYSGKYIRYGVREHAMGAIMNGLAAYG..MHIPTAGTFLNFVSYAIGAVR
SPTREMBL|F2QV92_PIC   WDGVKYFFNPKLQTFCGLGGDYSGRYIEFGIREHSMCAIANGLAAYNKGTFLPITSTFYMFYLYAAPALR
SPTREMBL|F4P3U6_BAT   WKGAVDFQPD.....VSGLGSYAGRYIRFGVREHGMAAICNGMNAYG..GIIPFGATFFNFISYALGAVR
SPTREMBL|F5HCT5_CRY   WDGMVEFQNPKSG.....YGDYSGRQIRYGIREHAMVAAANGLAAWHKGAIVPIMSSYFIFWLYAAPSLR
SPTREMBL|F9GER6_FUS   WSAATDLQSPDL...ESVPGNYSGRYIRYGVREHAMGAIMNGLAAYG..TLLPYAGTFLNFVSYAAGAVR
SPTREMBL|F9XCD0_MYC   WKGKVDFQNPDLQTACGLNGRYDGRYIHWGIREHAMASASNGMAAFNKGTILPITSSFFMFYIYAAPGLR
SPTREMBL|G2YJE4_BOT   WKGKVDFQHPGLKTSCGINGDYSGRYIHWGIREHAMASISNGLAAFNKGTIVPVTSSFFMFYIYAAPGIR
SPTREMBL|G3B8B6_CAN   WKGKTDFQNPRVKTSCGINGDYSGRYIHYGVREHAMAAISNGIAAYSPKTFIPVTSSFFMFYLYAAPAVR
SPTREMBL|G3BBY9_CAN   TQNAVDFQPP.....STGLGDYSGRYIRFGVREHAMGAMMNGIAAFG.SNFKVFGGTFLNFVSYAAGAVR
SPTREMBL|G7DV49_MIX   WKTATDFQPP.....STGLGSYAGRYMRYGIREHGMGAIANGMAAYGGNLILPAISTFANFTSYAAGAIR
SPTREMBL|G8BNE8_TET   WKEAKDFQPP.....SSGLGDYSGRYIRYGIREHAMGAILNGISAYG.VGYKPYSGTFLNFVAYAAGAVR
SPTREMBL|G8C0M6_TET   WDDAIDFQPD.....DTKIGSYMGRYIRYGVREHGMSAILNGISAYG.YNLKPFGGTFLNFVSYASGAVR
SPTREMBL|H0EPU8_GLA   .....................FKGRYIRNGIREHAMASIANGMAAYGPGTFLPITATFFMFFIYAAPGVR
Consensus             W     DFQ P        G YSGRYIRYG REHAM AI NG AAY      P   TF  F  YAA AVR
```

Figure 1G continued

```
                             570       580       590       600       610       620       630
1 [transketolase Sp   ISALSEAGVIYVMTHDSIGLGEDGPTHQPIEHLASFRAMPNILMLRPADGNETAGSYKVAVENRK...TP
Paecilomyces divari   MGALQHLQVIHAATHDSIGMGEDGPTHQPVELAALYRAMPNLLYIRPGDSEETAGAWIAAIEAKG...TP
SPTREMBL|A1DJZ3_NEO   MGALSHLPTIHIATHDSFAEGQNGPTHQPVELDSLFRAMPNLTYIRPCDAEETLGAWIYALTSANAKTSP
SPTREMBL|A5DD66_PIC   YGALSKLKVIHVATHDSVGLGEDGPTHQPIALATFYRALPQINYIRPADSEETAGAWIAALKAD....TP
SPTREMBL|A5DSN8_LOD   LSALSHQGVIWVATHDSIGLGEDGPTHQPIETLAHFRAIPNLSVWRPADGNEVSAAYIAAIEAVD...HP
SPTREMBL|A8PWG5_MAL   LSSLSNFQVIWVATHDSIGLGEDGPTHQPIEVAAALRALPNMDFWRPADGNEVSAAYKVAIESYK...TP
SPTREMBL|B6K5F9_SCH   LAALGQFRTIYVATHDSIGLGEDGPTHQPIETYAHFRAMPNINCWRPADGNETSAAYHSAISLAQ...TP
SPTREMBL|B8M240_TAL   LSALSQVRVIWIATHDSIGLGEDGPTHQPIETLAHFRALPNLMVWRPADGNETSAAYYSALTSKH...TP
SPTREMBL|B8M3Z5_TAL   MGALSRLPVIHIATHDSFAEGQNGPTHQPVELDSLYRAMPSFTYIRPCDGEEVIGAWKVALANR...DGP
SPTREMBL|C0LQF6_PIC   LSALSGHPIIWVATHDSIGLGEDGPTHQPIETLAHLRALPNLMVWRPADGNETSAAYLRAIESKH...TP
SPTREMBL|C7ZAM9_NEC   MGALSSLKVIHVATHDSIGEGQNGPTHQPVEVDSLYRAMPNLLYIRPADGEEIIGAWSSALEVQ...GQP
SPTREMBL|D7UPI2_OGA   MAGLQELKTIHIGTHDSINEGENGPTHQPIESPSLFRAMLNVYYMRPVDSAEVLGLFEKAIECP....YT
SPTREMBL|D7UPI3_OGA   MGALMNLKVIHVGTHDSIGTGEDGPTHQPIALANFYRALPNLYYIRPADSLETAGAYEVAIEAEG...YS
SPTREMBL|D8Q788_SCH   LGALMGVRFIGLGTHDSIGVGEDGPTHQSVALAAFFRALPNFNLIRPADAEEVLGAWEVALDAKH...TP
SPTREMBL|E6ZQ75_SPO   LSALSQHRVIWVATHDSIGLGEDGPTHQPIETAIALRAIPNLMFWRPADGNETSGAYLVALESKS...TP
SPTREMBL|F2QV92_PIC   MAALQELKAIHIATHDSIGAGEDGPTHQPIALSSLFRAMPNFYYMRPADATEVAALFEVAVELE....HS
SPTREMBL|F4P3U6_BAT   LSALSKHQVIYIMTHDSIGLGEDGPTHQPIETLASLRALPNLITLRPADGNETSGAYVAALENRH...RP
SPTREMBL|F5HCT5_CRY   MAALMKLRFIAIATHDSIGVGEDGPTHQPIAFPLFLRALPNFNYIRPADAEEVIGAWILGLRDAD...HP
SPTREMBL|F9GER6_FUS   LSALSRLRVIWVATHDSIALGQDGPTHQPIETLAHFRALPNCMVWRPADGNETSAAYRIALTSKN...TP
SPTREMBL|F9XCD0_MYC   MGSLQNLQTIHIATHDSIGTGEDGPTHQPIELAALYRAMPNFHYIRPADGEEVAGAVIAAIGAKD...TP
SPTREMBL|G2YJE4_BOT   MGALQGLQQIHIATHDSIGTGEDGPTHQPIALASLYRAMPNILYIRPADSEETAGAYIAALAATS...TP
SPTREMBL|G3B8B6_CAN   YGALSQLQVIHVATHDSIGIGEDGPTHHPIALASLYRSMPNMLYIRPCDSTEVAGAWEHAIEADN...TP
SPTREMBL|G3BBY9_CAN   LSAISQLPIIWVGTHDSVGLGEDGPTHQPIETLAHLRAIPNLSVWRPADGNEVSAAYKAAISSVS...TP
SPTREMBL|G7DV49_MIX   LSALSGFRVIWVATHDSIGLGEDGPTHQPIEVAAHFRALPNMHVWRPADGNETSAAYLMAMMNAT...TP
SPTREMBL|G8BNE8_TET   LSALSGHPIVWVATHDSIGVGEDGPTHQPIETLAHFRALPNLHVWRPADGNEVSAAYKVAIESKS...TP
SPTREMBL|G8C0M6_TET   LAALSGYPVIWVATHDSIGLGEDGPTHQPIETVTHFRALPNMQVWRPADGNEVSAAYKVSIESKA...TP
SPTREMBL|H0EPU8_GLA   MGALSNLQVIHIATHDSFAEGQNGPTHQPVELDSLYRAMPNLSYIRPCDAEELIGAWSYAMRSV...HNP
Consensus               ALS    I  ATHDSIG GEDGPTHQPIE     RA PN    RPAD  E   A   A        TP

                             640       650       660       670       680       690       700
1 [transketolase Sp   SILALSRQK.LPNLPGTSIEGVEKGGYTITDNSSGN..KPDVILIGTGSELEIAAKAGDELRKE.GKAVR
Paecilomyces divari   SIISTSRHALPQLTQTKR.QGVLKGAYVLEE..VEGPEPADVTLIGVCAELSFALDVAKELGETKNIRTR
SPTREMBL|A1DJZ3_NEO   TMISVARDP.VSHVPNTDRLQVLRGAYVIDEV.MDA....DITLASCGSSLCHVVDAARILRG.DGVKCR
SPTREMBL|A5DD66_PIC   TIISLSRHNLEQQITTNR.QMVAKGAYLLQE..DED...AQITIIGVGAELTFAVGASKVLRD.KGIKSR
SPTREMBL|A5DSN8_LOD   SVIALTRQN.VPHLENSSIEQALKGGYTVWPVEDG...KPDLIIVSTGSEVSISIDAAKLLAKEKNQKVA
SPTREMBL|A8PWG5_MAL   SVVSLTRQN.LPQLEGSSMEKAEKGGYVLLEDK.E....AAITLVSTGSEVSICVDAVAKLAEK.GIKAR
SPTREMBL|B6K5F9_SCH   SILALTRQN.LPQLNGSSIEKALKGGYVLVENKD.....ANVTLIGTGSEVSLCVEAAKVLEAEHGLKAR
SPTREMBL|B8M240_TAL   SIFALTRQN.LPQLQNSSIENALKGGYVAVDAPD.....AAVIIVSTGSEVGIAIDAAQYLKDKHGVSAR
SPTREMBL|B8M3Z5_TAL   TMLSLGRDP.VGPVPATDRYKVALGAYIIQDL.PNP....KLVLVSSGTNLHYTVSAAKHMSA.AGTPTR
SPTREMBL|C0LQF6_PIC   SIIALTRQN.LPQLEHSSIEKAAKGGYTVYPVEN.....PDIILVASGSEVSIAIDGAKKLGT.EGVKAS
SPTREMBL|C7ZAM9_NEC   VIISLARDPPEITIPNTSRTKESKGGYVIVDH.PKP....VVTLVSCGSELQFAVGAAAKLAE.QGIGTR
SPTREMBL|D7UPI2_OGA   SMLSLSRNEVLQYPGLSSPEKAKRGGYILEDV.EN....ADVQLIGAGAEMEFAYKAAKILGR.KGLKVR
SPTREMBL|D7UPI3_OGA   SIISESRQNLVQYPENSKRDAVKFGAYVFDDF.NIPDAKKDLIIIGVGSEMCFAMGSAAILRS.QGYNVR
SPTREMBL|D8Q788_SCH   TLMTLSRQG.VPLLEHSDRAKVHKGAYPIFANYDESSADPELVLIATGSEVWRAVEAAKKLAP...LRTR
SPTREMBL|E6ZQ75_SPO   SVLALTRQN.LPQLEGSTIEKTSKGGYVLVENE.S....ADITLVASGSEVALCNEAAAKLADK.GIKAR
SPTREMBL|F2QV92_PIC   TLLSLSRHEVDQYPGKTSAQGAKRGGYVVEDC.EGK...PDVQLIGTGSELEFAIKTARLLRQQKGWKVR
SPTREMBL|F4P3U6_BAT   SVLIFTRQN.LPQLEGSTVEKTLKGAYVLSESAN.....AKITLVGTGSEVSLAVDTAAMLAK.EGIQAR
SPTREMBL|F5HCT5_CRY   SLLSLTRQP.VPLLSGTDRNKVQYGGYVVYG...DENSIPDITLVATGSEVARAVDTAELLKDK..YSVR
SPTREMBL|F9GER6_FUS   SILALSRQN.LPQLNESSIGAAAKGGYVLHEPTDP....ADITLVATGSEVSLCVEAAKLLREQYGRKAR
SPTREMBL|F9XCD0_MYC   SMISVSRQNVPQYPQYSSREGVQKGAYVWVE..EEN...ADVTLIGVGAEMKFAMGAREELT.KAGLKVR
SPTREMBL|G2YJE4_BOT   TIISLSRQDLEQYPQYSSREGVQKGAYVFIE..KEE...ADVTLIGVGAEMVFAVKTASLLESEFGLRVR
SPTREMBL|G3B8B6_CAN   TIISLSRHNLEQYPQYSRKDKVRLGAYKFIE..NAG...AVINIISTGSEMQFAVEASQVLAQ.KGIIAN
SPTREMBL|G3BBY9_CAN   HVLALTRQG.VPHLEGSSIEKASKGGYVLVEAQD.....PDIILVSSGSEVSIAVEAAKLLGK.NGTRAR
SPTREMBL|G7DV49_MIX   SMLLLTRQN.LPQLENSTIAHATKGGYVVHDTDGE....TDLTYVSTGSEVSICMEAAKILMDK.GVKVR
SPTREMBL|G8BNE8_TET   SIIALTRQN.LPQLEGTSIEKASKGGYVLQDVAN.....PDLIFVATGSEVSLSVDAAKILASQN.IKAR
SPTREMBL|G8C0M6_TET   SIICLTRQN.VPQLEGSSIEKAAKGGYAISEYHGNTTPKLDLILVSSGSEVSLCCKSAEKYSKDHKVNIR
SPTREMBL|H0EPU8_GLA   SMISVARDP.VGEVPNTSREGVQRGAYVIQEN.SDA....AITLVSCGSNLHYAVAAAKKLEQ.KGTKCR
Consensus             S   L RQ        S      KG YV            D  L   GSE   A  AA  L    G   R
```

Figure 1G continued

```
                              710       720       730       740       750       760       770
1 [transketolase Sp   VVSFVSWELFEKQSDEYKESVLPS..DVTARVSIEAGSTFGWHKIVGSKGKAIGIDKFGASAPAGKIYQE
Paecilomyces divari   VVSFPCQRLFEKQSLSYKRNTLQR.HRGIPAVVIEPYTPNGWERYAD...AAICMNRFGHSLPGKAVYKY
SPTREMBL|A1DJZ3_NEO   IVSCPSFDLLDRQEREYQEGVFRL..DGKPIVSVEEYVATTWARYVT...ASIGMSGFGYSASNKSNYAR
SPTREMBL|A5DD66_PIC   VVSFPSQRLFEQQSLNYKRSVLKR.GK.VPCVVIEAYAPNGWERYAT...AGFNMTTFGKSLPGSIVYDY
SPTREMBL|A5DSN8_LOD   VVSIPDFHTFDQQPIDYRLSILPD..GV.PIMSVEVLSTFGWSKYS...HEHFGLNRFGASGKAADLYKF
SPTREMBL|A8PWG5_MAL   VVSIPCFRVFDLQPYDYRISVLPS..GH.PILSVEAYSTFGWGVYS...HVHHGINRFGVSAPYQKAYEK
SPTREMBL|B6K5F9_SCH   IVSMPCCEIFDEQAQEYRLSVLPD..GI.PIMSVEVYSTVGWKRYA...HEAFGIDSFGDSAPYQQLYAK
SPTREMBL|B8M240_TAL   VVSIPCFEVFDTQSKEYRLSVLPD..GI.PILSVEAASTLGWERYS...HEQFGINRFGASGPYKKVYEK
SPTREMBL|B8M3Z5_TAL   VVSAPCLRLFDAQSQEYRDKILPR..DGTPVVSIEEYVATTWARYTT...ASIGMTGFGCSAAKECNYER
SPTREMBL|C0LQF6_PIC   VVSIPDFFTFDSQARSYQLSVLPD..GV.PIMSVEVMSTFGWSKYS...HEQFGINRFGISGPGPEIYKF
SPTREMBL|C7ZAM9_NEC   VVSMPCLRLFEAQSKDYQEHMSLD..SG................HAS...ARLLLG..............
SPTREMBL|D7UPI2_OGA   VLSIPCTRLFDEHSLGSRRSVLRKDGSQVPPVIVDGHVAFGWERYST...ASYCMNTYGKSLPPDVIYEY
SPTREMBL|D7UPI3_OGA   VVSFPCQRLFEQQSAEYRHSVLMR.QQMIPTVVIEAYAPNGWERYAT...AGINMKTFGKSLPGMVCYDF
SPTREMBL|D8Q788_SCH   VVSMPIQSVFDRQPAEYRRATLAT..GKALVVAIEPWASYGWARYAH...ASMSMHTFGLSAPQETLYEL
SPTREMBL|E6ZQ75_SPO   IVSIPCFSVFDAQPIEYRLSVLPS..GK.PILSVEAYSTLGWSKYS...HAQHGINTFGASAPAKDLFKK
SPTREMBL|F2QV92_PIC   VLSFPCQRLFDEQSITYRRSVLRR..GEVPTVVVEAYVAYGWERYAT...AGYTMNTFGKSLPVEDVYKY
SPTREMBL|F4P3U6_BAT   VVSMPSWELFEDQSHEYRASVFLD..NI.PTLSMEAMTTLGWSKYA...HASCGIDTFGASAPYLLLYKK
SPTREMBL|F5HCT5_CRY   VVSMPHTGRFDAQPLEYRRSVIPS..TKSLVVSIEPYASFGWAKYAH...AGAHMTGFGHSAPYSVLFEH
SPTREMBL|F9GER6_FUS   IVSLPCWEVFDAQSKRYKLSVLVS..GI.PILSVEMLSTMGWHRYS...HEQFGIDSFGASAPAEELFNK
SPTREMBL|F9XCD0_MYC   VVSFPCQRIFEAQSKEYKKSVLG....TAPRVVVEAYSPNGWERYAD...AAYSMHSFGQSLTGDYVHER
SPTREMBL|G2YJE4_BOT   VVSFPCQRLFDEQSLSYKRQVLRY.RSQAPRVVIEAYAVNGWERYAD...SGYSMSSFGHSLPGADAYKY
SPTREMBL|G3B8B6_CAN   VISFPCQRLFEKQPLEYKRSILNR.GK.VPSVVVEAYAVNGWERYAT...AGINMSTFGKSLPGKMAYEY
SPTREMBL|G3BBY9_CAN   VVSMPSFFAFDQQSEKYRFSVLPD..GV.PILSIEVMSPFGWSNYS...HVQFGVSRFGFSAKYTDIYDY
SPTREMBL|G7DV49_MIX   VVSLPCFEIFAAQPKDYRMSILPD..GH.PVLSVEAYSTVSWGALA...HEHYGMSTFGASGPYKDVYKK
SPTREMBL|G8BNE8_TET   VVSLPDWLSFDRQPEQYRKSVLPD..GV.PILSVEVLATNGWGKYA...HQHIGMTTFGASGKAPLVFKH
SPTREMBL|G8C0M6_TET   VVSLPDFYTFGKQPFEYRISVLPD..EV.PILSVEVYATLSMCKYS...HEQFGLDRFGMSGKSKEVYEY
SPTREMBL|H0EPU8_GLA   LVSAPCLSLFDQQDSEYRKSVFPL..DGKPIVSVEEYVATTWARYVT...ASIGMTSFGYSASNESNYER
Consensus             VVS P    FD Q   YR SVL       P  S E     GW  Y            FG S P    Y

                              780       790       800       810       820
1 [transketolase Sp   YGITVEAVVEAAKSVC....................................
Paecilomyces divari   FGFEINPMAAKISRYLSDVQS..DEILR..WEFLEL................
SPTREMBL|A1DJZ3_NEO   FELDAEGIVRRVKGYLAFLNG..RDARG..VGWQSL................
SPTREMBL|A5DD66_PIC   FGFNSAKIADRIERYIKEYE...DPDIK..FEFQDLLAV.............
SPTREMBL|A5DSN8_LOD   FGFTPEGIAERAEKTIQFY..KGKVVGSPVHTAFEELK..............
SPTREMBL|A8PWG5_MAL   FQLTDHDIADKAVKVVEYY.KSMGHIHSPIAVQAAFELQK............
SPTREMBL|B6K5F9_SCH   YEMTPEGVAVRAKMTVDAW.KPVQNIRSPIHHAF..................
SPTREMBL|B8M240_TAL   FEITPEGISKRALATIDFY..KGVKPRSPINRAFQQLI..............
SPTREMBL|B8M3Z5_TAL   FGLDTDGIIKRLTLYLADLKG..ANARG..VGWRQL................
SPTREMBL|C0LQF6_PIC   FEFTAEGVADRASKVVQFY..KGKDFYRHLISFESVMDVRLIAP........
SPTREMBL|C7ZAM9_NEC   ....................................................
SPTREMBL|D7UPI2_OGA   FGYNPNTIAKKVEAYVKACRA..DPLLL..HDYVDLKEKPKHDKVNKL....
SPTREMBL|D7UPI3_OGA   FGYNKEKIATKVDAYLKQIKD..TPSVV..YEFQDLN...............
SPTREMBL|D8Q788_SCH   FGYGIDTIVDKVSKFVESKRVDGEIVLPGVGEFEELLLGYAKKHAGPELGLH
SPTREMBL|E6ZQ75_SPO   FKMTGDDVAAKAEHVAKVF.QGQ.TLVSPVELEAKLMQY.............
SPTREMBL|F2QV92_PIC   FGYTPEKIGERVVQYVNSIKA..SPQIL..YEFHDLKGKPKHDKL.......
SPTREMBL|F4P3U6_BAT   FGLVPEAVSEKAKKVIEYY..KTRTPESKMSPALF.................
SPTREMBL|F5HCT5_CRY   FGFGPKTLAEKIRAWAEPKRNGDGWDIPGVGEFEELLVNSGNGH........
SPTREMBL|F9GER6_FUS   FEFTPHGVVKRAIATIEFWSKKSSNIDSPIIRAF..................
SPTREMBL|F9XCD0_MYC   FNYNNEKIARKVEALVEEVKKEGIEVLR..GDFRDLNFDYPRDIQFAAYHKH
SPTREMBL|G2YJE4_BOT   FGFDEKVIAPKVFELVQDVRKNGIDTLR..GEFRDLNVRSH...........
SPTREMBL|G3B8B6_CAN   FGFNSERIAAKIEKYYEDWKQ..DESLR..FEFQDLN...............
SPTREMBL|G3BBY9_CAN   LEFTPEGISSRANKTIDFY..KGSKLLSPLSKPFYPQSFDLFTHGR......
SPTREMBL|G7DV49_MIX   FGFTGENIASVGEKVIAFW.KGK.TILSPQNTCLVAI...............
SPTREMBL|G8BNE8_TET   FGFTPEVVAEKAAKTVSFY..KGKTIVSPLSSPF..................
SPTREMBL|G8C0M6_TET   FEFTPDGVASRIAKVLSFY..SDKTIRSPVRKALA.................
SPTREMBL|H0EPU8_GLA   FGLDTAGIEKKVNKYLKSLNG..GSARE..AGWQQL................
Consensus             F
```

X non conserved
X ≥50% conserved
X all match

Figure 1H

```
                              10        20        30        40        50        60        70
1 [transketolase Sp   MAASSSLSTLSHHQTLLSHPKTHLPTTPASSLLVPTTSSKVNGVLLKSTSSSRRLRVGSASAVVRAAAVE
Paeciolomyces divar   .............................................MSSYKFPVD.LKQFKQLKLDPKS.P
SPTREMBL|A8NV31_COP   ...........................................MSHLVEKFPID.LNQFKKIKLDPAQ.S
SPTREMBL|D8Q4R0_SCH   ...........................................MSGTAAPFPID.LKQYKKIKLDPTN.P
SPTREMBL|E3JQR8_PUC   .............................................MSADKFPIPFVEKFHKVALDPSQ.P
SPTREMBL|F4RJP4_MEL   .MPAFSLVEIEAKMSFFSDFLNQVKTPSVASKQIYVSKVLIQITNFDQLDFDFQ.IKILNQVTLHPSQ.P
SPTREMBL|G7DVE0_MIX   .........................................MASHSDFSDFPVD.LSKFKKLSLDPKK.K
SPTREMBL|H0EEH2_GLA   ...........................................MATATADFGIDVFAKYKPISLDPTN.A
SPTREMBL|I4YAU9_WAL   ...........................................MTAVNAKFPVP.LESFKKLTLDYKN.P
SPTREMBL|I4YE20_WAL   ........................................MLRTVIIKMSFPVD.LSKFKQLKLDYTR.S
SPTREMBL|K1WGM0_MAR   ...........................................MASATAEFGIDVLARYKPIALDPAS.P
SPTREMBL|K5WES2_PHA   ...........................................MSGTDAKFPID.LKQLKQLKLDPSN.P
SPTREMBL|Q55T38_CRY   ..........MLSRFLYSEFPTFVVVFKNRNYHTVARMSVNKVQDWDFEKFPID.LKKYKPFPLDPTKDK
SPTREMBL|L1J2M8_GUI   ......................................MTEG........FEID.MSKFKEVKLDPWQHE
SPTREMBL|A0YNC4_LYN   .......................................MTATQA....TFPID.LGAYKPLSLDP.ANA
SPTREMBL|A0ZLZ6_NOD   ......................................................................
SPTREMBL|B4WLI3_9SY   .......................................MTATQAKPKSAFPID.LSAYKRISLDP.SNN
SPTREMBL|C9M6X8_9BA   ......................................................MSDLFGRRVNMRGFPY
SPTREMBL|D1Y3C7_9BA   ........................................................MLFDEAVKINSFQW
SPTREMBL|D2Z2G4_9BA   ................................................................MVKYDL
SPTREMBL|D3L2W3_9BA   .........................................................MRDTLLALAHFPK
SPTREMBL|D6THL7_9CH   ....................................................................MT
SPTREMBL|D8F7T0_9DE   ......................................................MTTIITMIPFEAFKRV
SPTREMBL|F4XPA7_9CY   .......................................................MTNLSLVPFREFERL
SPTREMBL|G2H8B8_9DE   ......................................................MQSSLAFVPADECRRI
SPTREMBL|G3J1Y8_9GA   ...........................................MN....TFPID.LGAYQHITLNA.SNT
SPTREMBL|G9PZY8_9BA   ..........................................................MKVTEEIRGFKS
Consensus

                              80        90       100       110       120       130       140
1 [transketolase Sp   ALESTDIDQLVEKSVNTIRFLAIDAVEKANSGHPGLPMGCAPMGHILYDEIMRYNPKNPY.WFNRDRFVL
Paeciolomyces divar   QLSAQQKNDLLHNINIFRDAIIAFTATGAARGLAGHTGGPFDTAPEVCILLAFMNANPD...GFVDALYD
SPTREMBL|A8NV31_COP   ELTEEQRSALKANTQLLRDAIVYFTATGAARGVSGHTGGAYDTVPEVCILLALFEGNKD...AYVPILFD
SPTREMBL|D8Q4R0_SCH   KLTDEQKQTLTANINLLRDAIVLFTATGAARGVSGHTGGAFDTVPEVCLILSLFDAAPE...KFVPIVFD
SPTREMBL|E3JQR8_PUC   KLSADQKSKLLENISILRDAIVFFTATGAARGVSGHTGGPFDTVPEVCLLLALFASDPDG.KKFDHTFWD
SPTREMBL|F4RJP4_MEL   KLTQEEKSKLLNNTSILRDSIVFFTDTGAARGVGGHAGGPFDTVREVVLLLASFASGSDS.KIFDHTVSD
SPTREMBL|G7DVE0_MIX   TLSEQERSDLLSNIKLFRDAIVFFTATGAARGVSGHTGGPYDTAPEVNILLAFFNHSPA...QFLPIFFD
SPTREMBL|H0EEH2_GLA   KLSDQQKEDLLSNISLLRDTIVYFTSSGAARGVSGHTGGPYDTIPELCILEAFFLHSSE...KFLPIVYD
SPTREMBL|I4YAU9_WAL   QLSSQQRADLLHNINIFRDAIVAFTATGSARGVSGHTGGPFDTAPEVCILLAFINANPE...KFVDTFFD
SPTREMBL|I4YE20_WAL   TLSEQERSDLLSNINIFRDAIVAFSATGNARGVSGHTGGPFDTAPEVCILLALINANPE...KFIDTFFD
SPTREMBL|K1WGM0_MAR   KLSAEQKADLLSNISLLLDTIVYFTASGAARG.........DTIPEVCILEAFFKHSSE...KYMPVIYD
SPTREMBL|K5WES2_PHA   QLTDEQLSTLKHNIQLMRDAIVLFTATGAARGVSGHTGGAYDTVPEVNMLLSLFNHS.N...KFVPILFD
SPTREMBL|Q55T38_CRY   KLSQEQKDGLIANISLLRDVIVFFTATGAARGLAGHTGGAFDTIPEVVILLSFLLADQDK.SKYVDILFD
SPTREMBL|L1J2M8_GUI   SIPADIKADILANIELCRDAVIFFTACGSASGYGGHTGGAFDTMPEVMLLRSFFNARPD...KFVPIYFD
SPTREMBL|A0YNC4_LYN   TLTNEQRETLKANIQFCRDLIVFFTATGAARGVGGHTGGPYDTVPEVMILDALFRGAAD...KFVPIFFD
SPTREMBL|A0ZLZ6_NOD   .MATQEQLHQWYELAQQLRVDSIRATTVAGSGHPTSSMSPADLMAVLLSKYFRYDFQNPD.NPNNDRLIF
SPTREMBL|B4WLI3_9SY   KLTDEQRDALQANIQLCREAIVFFTATGGARGVGGHTGGPYDTVPEVMVLDALFRGNPD...KFVPTFFD
SPTREMBL|C9M6X8_9BA   ESLPADVVADLENAARVCRALAVAMTTRANSGHPAGSLSSMKMYMAAYGVANLTPDNAAD..LDRDYIAI
SPTREMBL|D1Y3C7_9BA   EKLPQAEERQLNKAAQVCKGLAVAMVARANSGHPAGALSSMKMYMAAYGAANVSPQNCDS..LDRDFVVV
SPTREMBL|D2Z2G4_9BA   NDVKLN...EMEEAARRCSGLAVSMVARAKSGHPAGSLSSMKLYLAAYDVADLTPENCDG..TDRDYVVI
SPTREMBL|D3L2W3_9BA   DKLDKVDTEAFLIAAKKCRGWVLTMTTAANSGHPGGSMSSMEMYLVTYAVSNLRRERVEE..TDRDYVVI
SPTREMBL|D6THL7_9CH   TTQTQPQIQQWTNLAQQLRVDSIRCTTAAGSGHPTSSMSAADLMAVLMISYLRYDFDNPK.HPNNDHLIF
SPTREMBL|D8F7T0_9DE   QKAEMDTNRRLGLLADMCRANALAAVKQAGSGHLGSSFSAMDIVVRLYYEEMNALQLGPD.HPDRDVYFS
SPTREMBL|F4XPA7_9CY   KDLDITPETHTELFATYCRINTLYMIKQAGSGHIGSSFSSLDIVCWLFLNELRVRKTNSN..QPRDIYFS
SPTREMBL|G2H8B8_9DE   LDLPVGHLEKRQLFADVCRINTLFMITKAGSGHIGTSFSCLDILAALFINGLREGNDAATGYPYRDLFFS
SPTREMBL|G3J1Y8_9GA   TLTDEQRASLKANIQLCRDAIVFFTATGAARGVGGHTGGPYDTVPEVVIMDALFRGDAD...KYVPIFFD
SPTREMBL|G9PZY8_9BA   EGISAEGVAELRGAAREARVWAVVATAAAKSGHPAGALSSMDIYMTLLGAANVTPELADS..PERDRIVV
Consensus                         N          F  T A  G  G      D
```

Figure 1H continued

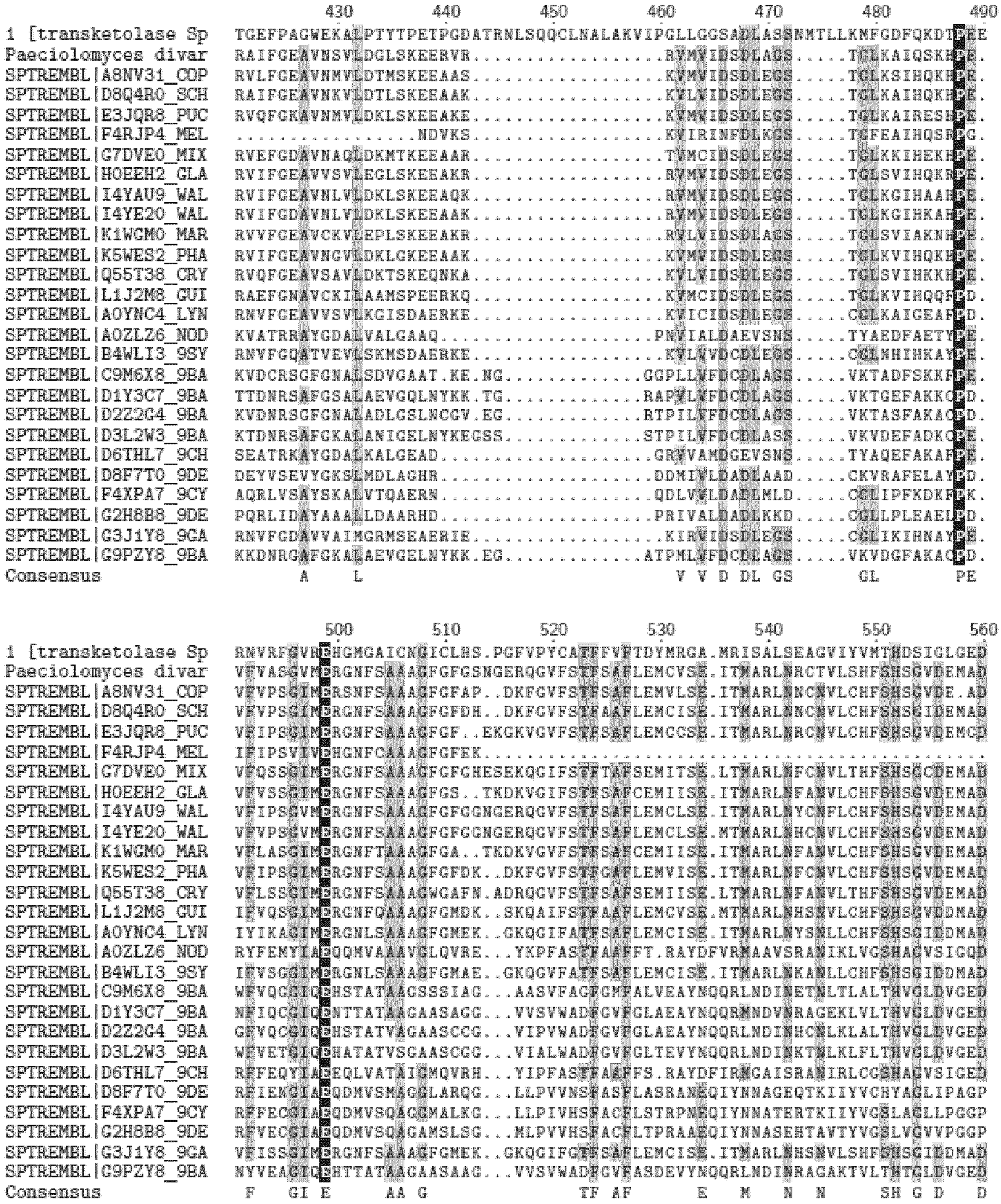
```
                             150       160       170       180       190       200       210
1 [transketolase Sp  SAGHGCMLQYALLHLAGYDSVLEEDLKTFRQWGSRIPGHPENFETPGVEVTTGPLGQGIANAVGLALAEK
Paeciolomyces divar  EAGHRVATQYLLAALDGKID..PDHLLNYRAADSKLPGHPELGLTPGIKFSSGRLGHMWAMCNGISMAHK
SPTREMBL|A8NV31_COP  EAGHRVATQYLLAAMEGSIP..AEHLLNYRAANSKLPGHPELGLTPGIKFSSGRLGHIWPLVNGVAMAHR
SPTREMBL|D8Q4R0_SCH  EAGHRVATQYLLAALEGSLP..AEHLLHYREANSKLPGHPELGLTPGVKFSSGRLGHVWPLVNGVALANK
SPTREMBL|E3JQR8_PUC  EAGHRVATQYLVAALEGHLP..AEQLMMYREAQSKLPGHPELGLTPGVKFSSGRLGHMWPLVNGVAMANR
SPTREMBL|F4RJP4_MEL  EAGHR..........................AQSKLPGHPQLGLTPGVKFSSVVVDWATCGLFSRVSHSP
SPTREMBL|G7DVE0_MIX  EAGHRVATQYLLSVLNGDMP..ASHLLNYRAASSKLPGHPELGLTPGIKFSSGRLGHMWPLVNGVALANS
SPTREMBL|H0EEH2_GLA  ESGHRVATQYLRAALHKHLP..FEHLLHYREANSKLPGHPELGLTPGIQFSSGRLGHIWSFVNGVCMANP
SPTREMBL|I4YAU9_WAL  EAGHRVATQYLLSTLDGHIS..PEHLLNYREAHSKLPGHPELGLSDGVKFSSGRLGHMWAMINGVAMANR
SPTREMBL|I4YE20_WAL  EAGHRVATQYLLSTLDGHID..PEHLLNYREANSKLPGHPELGLTPGVKFSSGRLGHMWAMINGVALANK
SPTREMBL|K1WGM0_MAR  ESGHRVTTQYL..............................LSLTPGVAFSSGRLGHLWSFVNGVALAHP
SPTREMBL|K5WES2_PHA  EAGHRVATQYLVSAIDGHIP..PEHLLHYREANSKLPGHPELGLTPGVKFSSGRLGHVWPWINGVALANR
SPTREMBL|Q55T38_CRY  EAGHRVATQYLLSALDGHIP..VEHLLHYREANSKLPGHPELGLTPGVKFSSGRLGHMWPLVNGVALAEK
SPTREMBL|L1J2M8_GUI  EAGHRVATQYLLSALDGHVT..FAHLKKYRVGHAGLPGHPELHCTPGVKFSSGRLGQMWPMVNGVGMANP
SPTREMBL|A0YNC4_LYN  EAGHRVATQYLMAALGGELP..AEKLLHYREAHSTLPGHPELGLTPGVKFSSGRLGHMWPYVNGVALANS
SPTREMBL|A0ZLZ6_NOD  SKGHAAPLLYSMYKAAGVIN..DEELRSLRRFGSRLEGHPTPILP.WVDVATGSLGQGLPIGVGIALAGK
SPTREMBL|B4WLI3_9SY  EAGHRVGTQYLMSTLNGDLP..AEQLMQYRAADSKLPGHPELGLTPGVKFSSGRLGHMWPFVNGVALANP
SPTREMBL|C9M6X8_9BA  SHGHTSPAAYATLAYYGFVDG.LDAVREFRRCGSLFQGHVERDVP.GIDWSSGDLGQGLSAGAGFALAQT
SPTREMBL|D1Y3C7_9BA  SHGHTSAAAYATLAYYGFVDA.FDAVNEFRRTGSRFQGHVERMVP.GIDWGSGCLGQGLSAGAGFALAQK
SPTREMBL|D2Z2G4_9BA  SHGHTSPAAYAVLAHNGFVDP.LEAVADFRRCGSAFQGHVERAIP.GIDWGSGNLGQGLSAGVGYGLALK
SPTREMBL|D3L2W3_9BA  SHGHTAPAAYSALAYLGFFDP.LELLGNFRRTGSPFQGHVERSVP.GIDWGSGNLGQGLAAGVGYALAQK
SPTREMBL|D6THL7_9CH  SKGHAAPLLYSMYKAAGAID..DAELMTLRKFGSRLEGHPTPALP.YVDVATGSLGQGLPIGVGVAMAGK
SPTREMBL|D8F7T0_9DE  SKGHDVPGLYAVLFSLGRLP..EDKLFNLRRFGG.LDGHPDIACP.GVEVNSGSLGMGISKGRGIAWAKQ
SPTREMBL|F4XPA7_9CY  SKGHDVPGFYSVLIGHGLLE..FERIHSLRRLDG.LPGHPDISIP.LLETNTGSLGMGVSKAKGMISANR
SPTREMBL|G2H8B8_9DE  SKGHDVPAVYACLIALGHLP..FDSIHTLRRLGG.LPGHPDRRTP.GLPCNTGSLGMGISKAKGFVHAAR
SPTREMBL|G3J1Y8_9GA  EAGHRAATQYLMSALNGDLP..AERLMEYRAAHSHLPGHPELGLTPGVKFSSGRLGHMWPYVNGVAIANP
SPTREMBL|G9PZY8_9BA  SHGHTSAGFYAALAEYGFFPA.HEMAANFRRAGSPYQGHVERDVP.GVDWGTGNLGQGLAAGVGFALAAR
Consensus              GH V   Y      G      E L   R   S LPGHPE  L  GV  SSG LG       G A A

                             220       230       240       250       260       270       280
1 [transketolase Sp  HLAARFNKPDAEIVDHYTYVILGDGCQMEGIAQEACSLAGHWGLGKLIAFYDDNHISIDGDTAIAFTES.
Paeciolomyces divar  D..............KNVLLLGSDGSQQEGNDAEAARIAVAKNLK.VKLFIDNNDVTIAGHPSIYLKGYE
SPTREMBL|A8NV31_COP  D..............KTVFCLGSDGSQQEGNDAEAARIAVAQHLN.VKLLIDDNDVTIAGHPSQYLKGFD
SPTREMBL|D8Q4R0_SCH  D..............KTVFCLGSDGSQQEGDDAEAARLAVAQGLN.VKLLIDDNDVTIAGHPSEYMKGYE
SPTREMBL|E3JQR8_PUC  G..............KTIFCLGSDGSQQEGNDAEAARLAVAQNLN.VKLLIDDNDVTIAGHPSDYMKGYE
SPTREMBL|F4RJP4_MEL  T..............ETVFCFCSDGSQHEGSDAEAARLARAQKLN.IKLLIDNNNVTISGHTSGYLKGYK
SPTREMBL|G7DVE0_MIX  D..............KNVILLGSDGSQQEGDDAEAARLAVAKGLN.VKLLIDDNDVTIAGHPSQYMPGYS
SPTREMBL|H0EEH2_GLA  D..............KTVFMLGSDGSQQEGDDAEAARLAVAQNLN.VKILIDDNDVTIAGHPSDYLKGYS
SPTREMBL|I4YAU9_WAL  S..............KNVILLGSDGSQQEGNDAEAARLAVAQNLN.VKVFVDDNDVTISGHPSQYLKGFN
SPTREMBL|I4YE20_WAL  D..............KNVIMLGSDGSQQEGDDAEAARLAVAQNLN.VKLFIDDNDVTISGHPSDYMKGYS
SPTREMBL|K1WGM0_MAR  E..............KTLFMLGSDGSQQEGDDAEAARLAVAQGLN.VKLLLDDNDVTIAGHPSEYLRGYS
SPTREMBL|K5WES2_PHA  D..............KTVFLLGSDGSQQEGNDAEAARLAVAQNLN.IKLIIDDNDVTIAGHPSEYLKGYD
SPTREMBL|Q55T38_CRY  N..............KAVFILGSDGSQQEGDDAEAARLAVAQGLN.VKLFVDDNDVTIAGHPSEYLKGYS
SPTREMBL|L1J2M8_GUI  D..............KIVWCLASDGAQQEGNDAEAARLAAARDIN.VKLIVDDNNVTIAGHPQDYLKGYN
SPTREMBL|A0YNC4_LYN  G..............KTVFCLGSDGSQQEGDDAEAARLAVSQQLN.VKLLIDDNDVTIAGHPSDYMPGYS
SPTREMBL|A0ZLZ6_NOD  YLDQLP.........YHTWVLLGDSETAEGSVWEAFDHAAHYTLDNLIAIIDVNRLGQRGQTELGWNTKA
SPTREMBL|B4WLI3_9SY  G..............KTTFLLGSDGAQQEGDDAEAARLAVAKGIN.VKLIIDDNDVTIAGHPSEYLGGYS
SPTREMBL|C9M6X8_9BA  ARKNK..........SRVYVLMGDGEQGKGQLAEARRLIAARNLP.VTALIDANDIQISGRVEDVMPIH.
SPTREMBL|D1Y3C7_9BA  ARGYD..........GRVYVLMGDGEQPKGQIAEARRLIAARKLTSVTALIDVNDIQISGRCGDVMPAH.
SPTREMBL|D2Z2G4_9BA  KRGLD..........RHVYVLMGDGEQPKGQLAEARRIAVAHGIKDITVLVDMNDIQISGKVEDVMPVN.
SPTREMBL|D3L2W3_9BA  KRRGE..........GHVFVLMGDGEQTKGQVSEARRIAAKEGLDNVTALVDFNHIQISGFVEDVMPVQ.
SPTREMBL|D6THL7_9CH  YLDKLP.........YHVWVVLGDSEMAEGSIWEACDKAGFYKLDHLIGILDCNRLGQRGETELGWNTAA
SPTREMBL|D8F7T0_9DE  YLGRG..........GHVFVMTGDGELQEGQNYEAMQSAVQQGVGNLTVIVDHNKVQSDREVARIVNLQD
SPTREMBL|F4XPA7_9CY  LMGRE..........GNVYILTGDGELQEGQFWESLQPAANLGFHEMTVIVDHNKIQSDTWVSQVSDLGD
SPTREMBL|G2H8B8_9DE  AQGLD..........RRCFVLTGDGELQEGQIWESLQGAANRGLHEITVIVDHNKIQSDIWVHRTSDLGD
SPTREMBL|G3J1Y8_9GA  S..............KVLFCLGSDGSQQEGNDAEAARLSVAQNLN.VKLIIDDNDVTIAGHPSKYLPGCS
SPTREMBL|G9PZY8_9BA  ARGSR..........AHTWVVMGDGEQVKGQVAEARRIAAKEKLTNITALVDWNDIQISGRLEEVMPVN.
Consensus                             V  L  DG QQEG  AEA RLA A  L  V    D NDV I G
```

Figure 1H continued

```
                              290       300       310       320       330       340       350
1 [transketolase Sp   VDLRFEALGWHVIWVKNGNTGYDEIRAAIKEAKTVTDKPTLIKVTTTIGFGSPNKSN.........SYSV
Paeciolomyces divar   IARTLEGHGLKVIRA..QGENLDSLYGAMCEIINYN.GPAAVVVDRKMAAGIEEIEG..........ETH
SPTREMBL|A8NV31_COP   VAKTLEGHGLKVVTV..DGENIDSLWGGVSSVLQYD.GPAAVVIKRVMAPGIPDIEG..........STH
SPTREMBL|D8Q4R0_SCH   LQKTLEGHGLKVVTV..QGEDLDSLWGGVHEILNYD.GPAAVVSKRLMAPGVDDIEG..........TPH
SPTREMBL|E3JQR8_PUC   VSKTLEGHGLKVFHA..EGENIDSLWSGLCSVISHQ.GPAALIAKRLMCPGIQGLEG..........SNH
SPTREMBL|F4RJP4_MEL   VGKTLEAHALKIVRA..EGE..................................................
SPTREMBL|G7DVE0_MIX   VAKTLAGHGLKVFDT..EGEDIDKLYASVCDLVAHD.GPTALIVKRVMCPGIDGLEG..........STH
SPTREMBL|H0EEH2_GLA   LASTLGGHGMKVFET..SGEDIDQLYAAICGAVSHN.GPAAVIAKRKMAPGIKGLEG..........TTH
SPTREMBL|I4YAU9_WAL   TAKTLEGHGLKVFRT..EGENIDKLYASMCEIISYD.GPAAVVASRKIAPEVEGIEG..........ESH
SPTREMBL|I4YE20_WAL   VQKTLEGHGLKVVRA..EGENIDSLFKAVCEVVAHD.GPAAVVASRKIAPEIEGIEG..........ESH
SPTREMBL|K1WGM0_MAR   LLKTLEGHGLKTIVA..QGEDIDELYAAVCTAVTTE.GPAAVVVKRKMAPGIKGLEG..........STH
SPTREMBL|K5WES2_PHA   IGKTLAGHGLKVVTV..QGEDTAALWSALVEVINYD.GPAAVIAKRLMAPGIADIEG..........TPH
SPTREMBL|Q55T38_CRY   VARTLEGHGLKVVEA..NGEDLDSLYSAIVEVMNHK.GPAAVVTHRPMAPKIKGIEG..........SPH
SPTREMBL|L1J2M8_GUI   VEKTLEGHGMTCAVC..DGEDLDALYKAMRAAVLKD.GPVAVVAKRKMCPGVEGVEG..........TCH
SPTREMBL|A0YNC4_LYN   VAQTLAGHGLKILEG..NGEDIDDLYRRICEAVNTD.GPVAVVNKRPMAVGIEGIEG..........SEH
SPTREMBL|A0ZLZ6_NOD   YAKRATAFGWQAIEI..DGHDLTDIDQAYSAAVTVNDRPTVIIARTKKGQGVASLED.........LGGW
SPTREMBL|B4WLI3_9SY   VQRTLEGHGIKTLVN..DGEDIDALYANICEAIATD.GPIAVICKRPMTVGIEGLEG..........STH
SPTREMBL|C9M6X8_9BA   LRQIWEADGWQVIEC..DGDNFPSLYAGLRQASLSD.RPTVLICRTTMGQGVSFMED.........RPDY
SPTREMBL|D1Y3C7_9BA   IKELWEADGWQTLLC..DGDSFAELYVALKAARAAG.RPTAILCRTTMGKGVSFMED.........TPDY
SPTREMBL|D2Z2G4_9BA   IEALWKADGWAVLKA..DGHDFRSIYCAMEKAQGMD.VPAVILCSTVMGKGVSFMED.........RPDY
SPTREMBL|D3L2W3_9BA   YKRLWEIDGWQVYEC..DGHDPSDIYSALKEAYEDD.GPSVVFCNTVMGKGVSFMEG.........IPDY
SPTREMBL|D6THL7_9CH   YAARARAFGWHAIEI..DGHDYEQINQAYAEALQTEGKPTLIIAKTKKGKGVSFLED.........KDGW
SPTREMBL|D8F7T0_9DE   LEKKLSAFGWEVARC..DGHDQHALARVFERFRHLKDIPKILIADTIKGRGVSFMEHPLALREGAGFYRW
SPTREMBL|F4XPA7_9CY   LERKFAAYGWYVERC..DGNDVKALGNAIAKLKAVTNQPKILIADTIKGKGISFMEH.TAIRPEDNLYHF
SPTREMBL|G2H8B8_9DE   LDAKFAAFGWHVQRC..NGHDPAALDTALAN.AATDPRPSVIIADTQKGAGAEAFLH.TAAHDSSELYRF
SPTREMBL|G3J1Y8_9GA   IAKTLAGHGLTVLEG..DGENIDDLYKRICTAINTD.GPVAVINHRPMCPGIIGLEG..........STN
SPTREMBL|G9PZY8_9BA   IPALWAADGWLVTEC..DGHDYEALYRAMKAAKGSP.VPTVILCRTTMGKGVSFMEN.........IPDY
Consensus                 L   G  V      G D   L             P        M  G    E
```

```
                              360       370       380       390       400       410       420
1 [transketolase Sp   HGSALGSKEVEATRQNLGWPYEPFHVPEEVKKHWSRHTPEGASLEAEWNTKFAEYEKKYPEDATEFKSIT
Paeciolomyces divar   AHDVIPVDIARKYLTKRGYSKEQL.AFYD..QIKPGSNPHQYQGS..................TKEKGAN
SPTREMBL|A8NV31_COP   GHDVIPVNAAVKYFESRGLKSLTG..LYD..GIRPVPNPYLYIGS..................TKEVGAN
SPTREMBL|D8Q4R0_SCH   GHDVIPVKAALNYLKKRGHP.DYS.NILL..EIKPNANPYLFVGS..................TKESAAN
SPTREMBL|E3JQR8_PUC   GHDVIPVPKAIAYLKDRPYGEAAA.KILN..SLTPSSSPYLYIGC..................TKEKDAC
SPTREMBL|F4RJP4_MEL   ........................................KYTGC.........................
SPTREMBL|G7DVE0_MIX   GHDVIPVPKAIKYLKDRGYPDTVG.EILS..NIKPAPNPYLYTGV..................SKELGAC
SPTREMBL|H0EEH2_GLA   AHDVIPVKIAVEYLEERGYTQAAE.KIK...AITPNSLAYLYEGS..................TTEKSAN
SPTREMBL|I4YAU9_WAL   AHDVIPVAVAKKYLTKRGYSEDQL.SFYD..QIKPQKYQYEYTGV..................SADKGAN
SPTREMBL|I4YE20_WAL   AHDVIPVAVAKKYLTKRGYTEEQL.KFYD..DIKPKSVKYDYKGV..................TADKGAN
SPTREMBL|K1WGM0_MAR   ADDVIPVKKAVEYLQSKGYDQAVS.KLE...AITPNSLAYLYSGS..................GSEKFAN
SPTREMBL|K5WES2_PHA   GHDVVPVKAALKYFEARGFPAEFAGNILN..NIKPYALPYLYIGS..................TKELGAN
SPTREMBL|Q55T38_CRY   AHDAIKVEPAIEYLDARHPKCAAI...LR..AIQPSNYAELLSGS..................TKERGAC
SPTREMBL|L1J2M8_GUI   GHDAIAVAKAITYLEKKGYKEAVQWLQGPATKAEKDPRAGTYMGTQ.................GSSVGAN
SPTREMBL|A0YNC4_LYN   GHDVIPLDKAIKYLEAKGKTAAVDYLKS....IEKPKNTYTFMGS..................SDKLGSN
SPTREMBL|A0ZLZ6_NOD   HGKALESEDAKRAISELGG.ERQIIISVDTPDDQAQPASIGNAQP...............LQLPKYDKSK
SPTREMBL|B4WLI3_9SY   GHDVIPTDLAIKYLEKHGRQEAADYLKN....IEKPKHSYEFKGS..................SDEVGSN
SPTREMBL|C9M6X8_9BA   HGKAATGDLYRQAMDELKA.PDLVAMANSSSVR..TAGAKMDRPR..........PALNVGAPRVYKLGE
SPTREMBL|D1Y3C7_9BA   HGKAASGDLYRQAMKELGQ.PDLVALAAEHKKAKFSAARAVEPLE..........AVLDLGAAKVYGPAD
SPTREMBL|D2Z2G4_9BA   HGKAATGDLYRQAMEELGQ.PDILVEAASLGDQVHYRGRSVEVPA..........VSVDTGSPRTYGVEE
SPTREMBL|D3L2W3_9BA   HGKALNKEQYLRAMEELGEDPSILEEALRARRGPLPEGRRVSVSA..........PPIQTGHPFTYSRET
SPTREMBL|D6THL7_9CH   HGKPVPKDKEDEALQELHRPARSMTFSVQKPEN.LQPVSIPPAKA...............VDLPRYEVGS
SPTREMBL|D8F7T0_9DE   HAGAPPDDAFMAAHSEILARINSRMADEGLDAFALKEVPPPSKKELFY...TLQGEPESKIENQSVVKVT
SPTREMBL|F4XPA7_9CY   HSGAPDDEAYTTGVAELIAAANTQLEQVGASPLALEMLSYAKP...................AVAD...G
SPTREMBL|G2H8B8_9DE   HSGAPAPEARDAALAELADRIERSLRDVGAATPLLLHREPAAP...................VPAPPAAT
SPTREMBL|G3J1Y8_9GA   GHDVVSVKVAVDYLESRGRQAAADHLKG....IKAPKNDAVFLGS..................GDKWDAN
SPTREMBL|G9PZY8_9BA   HGKPAAGEKLELALKELGSDISVFNAALAERKGPLPKGRRITPEK..........PVLELGVPHTYTKED
Consensus                              G
```

Figure 1H continued



```
                             430       440       450       460       470       480       490
1 [transketolase Sp  TGEFPAGWEKALPTYTPETPGDATRNLSQQCLNALAKVIPGLLGGSADLASSNMTLLKMFGDFQKDTPEE
Paeciolomyces divar  RAIFGEAVNSVLDGLSKEERVR..................RVMVIDSDLAGS.....TGLKAIQSKHPE.
SPTREMBL|A8NV31_COP  RVLFGEAVNMVLDTMSKEEAAS..................KVMVIDSDLEGS.....TGLKSIHQKHPE.
SPTREMBL|D8Q4R0_SCH  RAIFGEAVNKVLDTLSKEEAAK..................KVLVIDSDLEGS.....TGLKAIHQKHPE.
SPTREMBL|E3JQR8_PUC  RVQFGKAVNMVLDKLSKEEAKE..................KVMVIDSDLEGS.....TGLKAIRESHPE.
SPTREMBL|F4RJP4_MEL  .................NDVKS..................KVIRINFDLKGS.....TGFEAIHQSRPG.
SPTREMBL|G7DVE0_MIX  RVEFGDAVNAQLDKMTKEEAAR..................TVMCIDSDLEGS.....TGLKKIHEKHPE.
SPTREMBL|H0EEH2_GLA  RVIFGEAVVSVLEGLSKEEAKR..................RVMVIDSDLEGS.....TGLSVIHQKRPE.
SPTREMBL|I4YAU9_WAL  RVIFGEAVNLVLDKLSKEEAQK..................RVMVIDSDLEGS.....TGLKGIHAAHPE.
SPTREMBL|I4YE20_WAL  RVIFGDAVNLVLDKLSKEEAAK..................RVMVIDSDLEGS.....TGLKGIHKAHPE.
SPTREMBL|K1WGM0_MAR  RVVFGEAVCKVLEPLSKEEAKR..................RVLVIDSDLAGS.....TGLSVIAKNHPE.
SPTREMBL|K5WES2_PHA  RVIFGEAVNGVLDKLGKEEAAK..................KVMVIDSDLEGS.....TGLKVIHQKHPE.
SPTREMBL|Q55T38_CRY  RVQFGEAVSAVLDKTSKEQNKA..................KVLVIDSDLEGS.....TGLSVIHKKHPE.
SPTREMBL|L1J2M8_GUI  RAEFGNAVCKILAAMSPEERKQ..................KVMCIDSDLEGS.....TGLKVIHQQFPD.
SPTREMBL|A0YNC4_LYN  RNVFGEAVVSVLKGISDAERKE..................KVICIDSDLEGS.....CGLKAIGEAFPD.
SPTREMBL|A0ZLZ6_NOD  KVATRRAYGDALVALGAAQ....................PNVIALDAEVSNS.....TYAEDFAETYPE.
SPTREMBL|B4WLI3_9SY  RNVFGQATVEVLSKMSDAERKE..................KVLVVDCDLEGS.....CGLNHIHKAYPE.
SPTREMBL|C9M6X8_9BA  KVDCRSGFGNALSDVGAAT.KE.NG.............GGPLLVFDCDLAGS.....VKTADFSKKFPE.
SPTREMBL|D1Y3C7_9BA  TTDNRSAFGSALAEVGQLNYKK.TG.............RAPVLVFDCDLAGS.....VKTGEFAKKCPD.
SPTREMBL|D2Z2G4_9BA  KVDNRSGFGNALADLGSLNCGV.EG.............RTPILVFDCDLAGS.....VKTASFAKACPD.
SPTREMBL|D3L2W3_9BA  KTDNRSAFGKALANIGELNYKEGSS.............STPILVFDCDLASS.....VKVDEFADKCPE.
SPTREMBL|D6THL7_9CH  SEATRKAYGDALKALGEAD....................GRVVAMDGEVSNS.....TYAQEFAKAFPE.
SPTREMBL|D8F7T0_9DE  DEYVSEVYGKSLMDLAGHR....................DDMIVLDADLAAD.....CKVRAFELAYPD.
SPTREMBL|F4XPA7_9CY  AQRLVSAYSKALVTQAERN....................QDLVVLDADLMLD.....CGLIPFKDKFPK.
SPTREMBL|G2H8B8_9DE  PQRLIDAYAAALLDAARHD....................PRIVALDADLKKD.....CGLLPLEAELPD.
SPTREMBL|G3J1Y8_9GA  RNVFGDAVVAIMGRMSEAERIE..................KIRVIDSDLEGS.....CGLIKIHNAYPE.
SPTREMBL|G9PZY8_9BA  KKDNRGAFGKALAEVGELNYKK.EG.............ATPMLVFDCDLAGS.....VKVDGFAKACPD.
Consensus                  A    L                             V V D DL GS      GL       PE

                             500       510       520       530       540       550       560
1 [transketolase Sp  RNVRFGVREHGMGAICNGICLHS.PGFVPYCATFFVFTDYMRGA.MRISALSEAGVIYVMTHDSIGLGED
Paeciolomyces divar  VFVASGVMERGNFSAAAGFGFGSNGERQGVFSTFSAFLEMCVSE.ITMARLNRCTVLSHFSHSGVDEMAD
SPTREMBL|A8NV31_COP  VFVPSGIMERSNFSAAAGFGFAP..DKFGVFSTFSAFLEMVLSE.ITMARLNNCNVLCHFSHSGVDE.AD
SPTREMBL|D8Q4R0_SCH  VFVPSGIMERGNFSAAAGFGFDH..DKFGVFSTFAAFLEMCISE.ITMARLNNCNVLCHFSHSGIDEMAD
SPTREMBL|E3JQR8_PUC  VFIPSGIMERGNFSAAAGFGF..EKGKVGVFSTFSAFLEMCCSE.ITMARLNRCNVLCHFSHSGVDEMCD
SPTREMBL|F4RJP4_MEL  IFIPSVIVEHGNFCAAAGFGFEK...............................................
SPTREMBL|G7DVE0_MIX  VFQSSGIMERGNFSAAAGFGFGHESEKQGIFSTFTAFSEMITSE.LTMARLNFCNVLTHFSHSGCDEMAD
SPTREMBL|H0EEH2_GLA  VFVSSGIMERGNFSAAAGFGS..TKDKVGIFSTFSAFCEMIISE.ITMARLNFANVLCHFSHSGVDEMAD
SPTREMBL|I4YAU9_WAL  VFIPSGVMERGNFSAAAGFGFGGNGERQGVFSTFSAFLEMCLSE.ITMARLNYCNFLCHFSHSGVDEMAD
SPTREMBL|I4YE20_WAL  VFVPSGVMERGNFSAAAGFGFGGNGERQGVFSTFSAFLEMCLSE.MTMARLNHCNVLCHFSHSGVDEMAD
SPTREMBL|K1WGM0_MAR  VFLASGIMERGNFTAAAGFGA..TKDKVGVFSTFSAFCEMIISE.ITMARLNFANVLCHFSHSGVDEMAD
SPTREMBL|K5WES2_PHA  VFIPSGIMERGNFSAAAGFGFDK..DKFGVFSTFGAFLEMVISE.ITMARLNFCNVLCHFSHSGVDEMAD
SPTREMBL|Q55T38_CRY  VFLSSGIMERGNFSAAAGWGAFN.ADRQGVFSTFSAFSEMIISE.LTMARLNFANVLTHFSHSGVDEMAD
SPTREMBL|L1J2M8_GUI  IFVQSGIMERGNFQAAAGFGMDK..SKQAIFSTFAAFLEMCVSE.MTMARLNHSNVLCHFSHSGVDDMAD
SPTREMBL|A0YNC4_LYN  IYIKAGIMERGNLSAAAGFGMEK..GKQGIFATFSAFLEMCISE.ITMARLNYSNLLCHFSHSGIDDMAD
SPTREMBL|A0ZLZ6_NOD  RYFEMYIAEQQMVAAAVGLQVRE...YKPFASTFAAFFT.RAYDFVRMAAVSRANIKLVGSHAGVSIGQD
SPTREMBL|B4WLI3_9SY  IFVSGGIMERGNLSAAAGFGMAE..GKQGVFATFSAFLEMCISE.ITMARLNKANLLCHFSHSGIDDMAD
SPTREMBL|C9M6X8_9BA  WFVQGGIQEHSTATAAGSSSIAG...AASVFAGFGMFALVEAYNQQRLNDINETNLTLALTHVGLDVGED
SPTREMBL|D1Y3C7_9BA  NFIQCGIQENTTATAAGAASAGG...VVSVWADFGVFGLAEAYNQQRMNDVNRAGEKLVLTHVGLDVGED
SPTREMBL|D2Z2G4_9BA  GFVQCGIQEHSTATVAGAASCCG...VIPVWADFGVFGLAEAYNQQRLNDINHCNLKLALTHVGLDVGED
SPTREMBL|D3L2W3_9BA  WFVETGIQEHATATVSGAASCGG...VIALWADFGVFGLTEVYNQQRLNDINKTNLKLFLTHVGLDVGED
SPTREMBL|D6THL7_9CH  RFFEQYIAEEQLVATAIGMQVRH...YIPFASTFAAFFS.RAYDFIRMGAISRANIRLCGSHAGVSIGED
SPTREMBL|D8F7T0_9DE  RFIENGIAEQDMVSMAGGLARQG...LLPVVNSFASFLASRANEQIYNNAGEQTKIIYVCHYAGLIPAGP
SPTREMBL|F4XPA7_9CY  RFFECGIAEQDMVSQAGGMALKG...LLPIVHSFACFLSTRPNEQIYNNATERTKIIYVGSLAGLLPGGP
SPTREMBL|G2H8B8_9DE  RFVECGIAEQDMVSQAGAMSLSG...MLPVVHSFACFLTPRAAEQIYNNASEHTAVTYVGSLVGVVPGGP
SPTREMBL|G3J1Y8_9GA  VFISSGIMERGNFSAAAGFGMEK..GKQGIFGTFSAFLEMCISE.ITMARLNHSNVLSHFSHSGIDDMAD
SPTREMBL|G9PZY8_9BA  NYVEAGIQEHTTATAAGAASAAG...VVSVWADFGVFASDEVYNQQRLNDINRAGAKTVLTHTGLDVGED
Consensus             F   GI E     AA G              TF AF      E   M   N  N     SH G D   D
```

Figure 1H continued

```
                            570       580       590       600       610       620       630
1 [transketolase Sp  GPTHQPIEHLASFRAMPN...ILMLRPADGNETAGSYKVAVENRKTPSILALSRQKLPNLPGTSIEGVEK
Paeciolomyces divar  NTCHFGLNLFFADNGLMDAESTSLYFPADGEQMKAVVNKVFWDKSMRFIFSTRSKVPYILKEGTDQKLYG
SPTREMBL|A8NV31_COP  NTCHFGINNMFADNGLSDVQS.FLYFPADPAQMVAIVKRVFFERGMRFVFSTRAKVPWILREDG.TRFFE
SPTREMBL|D8Q4R0_SCH  NTCHFGVNNLFADNGLADANS.WLYFPADAAQMIACVKRVFFERGVRFIFSTRSKVPYLLKEDG.SRMFG
SPTREMBL|E3JQR8_PUC  NTCHFGLNTFFADNGLEDGYPTRLYFPADVHQMTAIVNEVFWDMGLRFVFSTRSKVPNILKEGSQEHYFG
SPTREMBL|F4RJP4_MEL  ......................................................................
SPTREMBL|G7DVE0_MIX  NTCHFGINSFFLDNGLLDAHETRMYFPADPSQLRAVIDAVFKDRGIRFVFSTRSKVPWILREGTQEKLYE
SPTREMBL|H0EEH2_GLA  NTCHFGINTFFADNALIDGDVTRLYFPADATQMRACVAKVFWDQGIRLVLSTRAKVPLILKERSQEPFYA
SPTREMBL|I4YAU9_WAL  NTCHFGLNHFFADNGLLDTAQTNLYFPADGEQMKAVVQKVFFDKGPKFVFSTRAKVPYIYKEGTETKLFG
SPTREMBL|I4YE20_WAL  NTCHFGLNHFFADNGLEDTAPTSLYFPADGEQMKAIVQTIFWDQGIRIVFSTRSKVPYILKQGTQEQLYG
SPTREMBL|K1WGM0_MAR  NTCHFGLNTFFADNGLVDGDTTRLYFPADANQMRATVSKVFWDQGIRLVISTGAKVPLIFKEGTQEPYYG
SPTREMBL|K5WES2_PHA  NTCHFGINHFFADNGLADTES.YLYWPADGAQMIKIVERVFFERGLRFVFSTRSKVPYILKEDGQTRLFG
SPTREMBL|Q55T38_CRY  NTCHFGINQFFLDNGLEDGYETRLYFAADCSQMDAIVDRVFYDKGLRFVFSTRSKVPWILKEDG.SRFFD
SPTREMBL|L1J2M8_GUI  NSCHYGLNNFFADNGLDEHETTRLYFPADYHQLLKVVPAVFWDHGIRCIFTTRSKLPAIQDEQG.KEMFG
SPTREMBL|A0YNC4_LYN  NTCHFGLNNMFADNGLDDGYETRLYFPADANQMKACVESVFSDPGLRFIFSTRSKVPYILNTDG.SNFFG
SPTREMBL|A0ZLZ6_NOD  GASQMALEDLAAFRAVWS...STVLYPC.DANQTAKLVAQMSDRNGIVYLRTTREKT......PVIYEHG
SPTREMBL|B4WLI3_9SY  NTCHFGINNMFADNGLDDGYETRLYFPADAHQMKACVEKVFHDPGLRFIFSTRAKVPMLQDDSG.QTLYD
SPTREMBL|C9M6X8_9BA  GKTHQCIDYVALAANCFG...WKLIVPADANQTDRALRWAVQEGGNVCLAMGRSKLP.VLELD.GRPAFE
SPTREMBL|D1Y3C7_9BA  GMTHQCIDYVSLMSNFFN...WKLVVPADPNQTDRATRWALKTAGNVCLAMGRSKLP.ALCAN.GKPLLA
SPTREMBL|D2Z2G4_9BA  GMTHQCIDYVSLLANTFG...WKLVVPSDPNQADRATRWALAERGNVCLAMGRSTLH.PLSDGKGKPFFG
SPTREMBL|D3L2W3_9BA  GMTHQCIDYCGLLRNLFG...WKLTVPIDPNQTDRITRWALSEAGNICVAMGRSKVPILTREGEDTPFYG
SPTREMBL|D6THL7_9CH  GPSQMALEDLAMMRAVFG...STVLYPS.DPNQAAQLVAEMAKHDGIVYLRSTREKT......PVLYAPD
SPTREMBL|D8F7T0_9DE  GKSHQSIRDISLFGALPN...CEILEPCNGEETRMVVDYAVNEARENVMIRLIIGPS......PRKIMLP
SPTREMBL|F4XPA7_9CY  GHSHQSVRDISALGAIPN...MVMIEPANEAEVSLALDYCLHRTISSCYIRLVSIPC......QIPYQLP
SPTREMBL|G2H8B8_9DE  GHSHQSIRDIALMGGVPG...SIALEPCCEAEVAPALRACLAHPGPS.YLRLVSIPC......DIPYSLP
SPTREMBL|G3J1Y8_9GA  NTCHFGLNNLFADNGLDDGYATNLYFPADPLQMKACLEAIFFNPGLRFVFSTRSKVPTILNTQG.QDYFG
SPTREMBL|G9PZY8_9BA  GMTHQCIDYVGLFRNTFG...WNVVVPADPNQTDRATRWMLAEPGNVCLAMGRSVLPVILKE.DGTPFYG
Consensus               H           L       L  PAD  Q            G       RSK P

                            640       650       660       670       680       690       700
1 [transketolase Sp  GGYTITDNSS...GNKPDVILIGTG.SELEIAAKAGDELRK.EGKAVRVVSFVSWELFEKQSDEYKESVL
Paeciolomyces divar  DGYEFVPGKEEVIRKGSAGYVVSYG.DMLYRSLDAVERLRK.EGLDVGLINKPTLNIVDEDTIKVYGSTP
SPTREMBL|A8NV31_COP  GDYEFVPGKDEIIVEGKAGYVVSFG.EMLYRSWDAVLRCRQ.EGLDVGLVNKPSLNVVDEQTIRKIGSSP
SPTREMBL|D8Q4R0_SCH  DDYEFIPGKDEFMATGTAGYVVTYG.EMTYRAWDAVLRLRQ.EGLDVGLVIKSTLNIIDEQALTTVGKTP
SPTREMBL|E3JQR8_PUC  GDYKFKPGKDEFIRRGKIGTVICFG.EIVHRALDAVDRLRE.EGFDVGLVNKSTLNVIDEEAIAEYGKQK
SPTREMBL|F4RJP4_MEL  .......GKE...KMRKLDAVISFG.EIVHRALDAGDQLGI.EGFDVGLVNKSTLNVIDEKPWMNMDIRN
SPTREMBL|G7DVE0_MIX  GDYKFVPCKDEYVRTGKAGTIVTYG.EMTYRCLDAVERLRA.EGYDIALVNKVTMNVVDEDAIKEYGSSP
SPTREMBL|H0EEH2_GLA  DGYEFEPCKDEFIRVGTAGYVVAFG.EMVYRAIDAVDRLRG.EGVDVGLINKSTLNVVDEDAIRKYGSTP
SPTREMBL|I4YAU9_WAL  EGYDFVPGKEEVIRKGSAGYVISYG.DMLYRSLDAVERLRK.EGLDVGLINKPTLNVVDEETIKEYGKTG
SPTREMBL|I4YE20_WAL  GDYKFVPGKEEIVRTGSAGYVISFG.DMLYRSLDAVERLRE.QGIDVGLINKPTLNVVDEETLKVYGSSP
SPTREMBL|K1WGM0_MAR  AGYEFTPSKDDFIRTGKVGYVVCYG.EMVYRALDAVDRLRA.EGVDVGLVNKSTLNVVDEAATRIYGSGP
SPTREMBL|K5WES2_PHA  DNYQFVPGKDEVIAEGTAGYVVSFG.DMVYRSWDAVLRCRQ.EGLNVGLINKPTLNIVDEETIRKVGSSP
SPTREMBL|Q55T38_CRY  SDYKFVPGKDEVIRKGTKGYVVAYG.EILYRALDAVDRLRK.EGLDVGLINKSTLNVVDEDMIKEIGSTE
SPTREMBL|L1J2M8_GUI  EGYTFVPGKDDVIRKGTKGYIVSYG.DSLYRSLDAVERLRA.EGVDVGLINKCTLNVVDDECMQLIGKTG
SPTREMBL|A0YNC4_LYN  DGYKFVPGKDEVIREGTAGYIISFG.AALYRSLDAVERLKQ.EGIDVGLINKATLNVVDEDMMAKVGQSP
SPTREMBL|A0ZLZ6_NOD  EEFPIGGSKIVRSSDQDQAAVIGAG.ITVHEAIKAYDRLKD.EGITVRVIDAYSVKPIDVQTLHQAARDT
SPTREMBL|B4WLI3_9SY  GDYTFVPGKDDVIREGTDGYIVSFG.DGLYRALDAVERLKE.EGLNIGLINKSTLNVVDEQMLEKVGKSP
SPTREMBL|C9M6X8_9BA  KP..FVYGDVVRLREGKNGAILALG.AMANQALEAAALLAQ.KGIQVSVYSVSCPLVCCDETLKEAFETG
SPTREMBL|D1Y3C7_9BA  RD..FTYGEAVKVREGKDAAILALG.AMAGRAVQAAELLAE.KGVETAVYAVSCPLEIDERALTEAFQTG
SPTREMBL|D2Z2G4_9BA  GDETFRYGDAHLLRDGGDCAVLALG.AMASKAMEAADRLSD.KGVEVKVYAVSCPLEVDMDALKDACSTG
SPTREMBL|D3L2W3_9BA  ENYVFRYGAIDLIRDGRHASVFAMG.HMVYRAVQARELLEK.HGLSLRVYGVSCPLSIDEKAIEEAAELG
SPTREMBL|D6THL7_9CH  EKFTVGGSKVVKQSAQDKVTVVAAG.ITLHEALKAYEQLKG.ENIHARIIDAYSVKPIDEETILAAAEEA
SPTREMBL|D8F7T0_9DE  EGYRLEKGKGVSLAHGGDAVLFAYGPVMIHEALTTSELLAE.KGIGLRVINSPWLNRMDRNWMMKQVDAV
SPTREMBL|F4XPA7_9CY  EDYQLELGQGIALTEGQDAILFGYGPVLLPQAYQAAQLLAQNHGITLKVVNLPWLNFVDLDWLEQILQGY
SPTREMBL|G2H8B8_9DE  EDHRLEKGKGHVLLPGTQVAVIGYGPVLLPQAWHAAHRLAA.EGVSVTVIDLPWLNHVDPAWLAEELSGK
SPTREMBL|G3J1Y8_9GA  GDYKFVSGKDEVIREGTQGYIVSFG.EALYRALDAVERLKQ.EGIDVGLINKPTLNVIDEEMMVKIGNSP
SPTREMBL|G9PZY8_9BA  EGYEFRNGAIDLLRDGTDVAILAMG.HLAGRAVEARELLAK.EGISARVLHCATPLSMDENELFTLVGEL
Consensus              Y F  GK    R G        G     RA  A  RL   EG  V   N   LN  DE
```

Figure 1H continued

```
                              710       720       730       740       750       760       770
1 [transketolase Sp   PSDVTARVSIEAGSTFG..WHKIVGSKG............KAIGIDKFGASAPAGKIYQEYGITVEAVVE
Paeciolomyces divar   F..VVVVESIAQKTGLGSRLGSHLLERKLTP.......KFKAIGAVREG.SGGLYEQINAQGLGPNDIIA
SPTREMBL|A8NV31_COP   F..VLVVESFNQKTGLGSKFGTWLLERQLTP.......KYGYMGTTKEG.CGGLWEQIPYQNLDPQSIIS
SPTREMBL|D8Q4R0_SCH   F..VLVVESANQKTGLGSKFGTWLLERGLTP.......KYGYMGTTKEG.SGGLWEQVYHQGLDPQSIIQ
SPTREMBL|E3JQR8_PUC   F..VLVVESLNQKTGLGSKMGTWLLERGLTP.......RYGYMGTNKEG.CGGLSAQIFHQGLDPASIIR
SPTREMBL|F4RJP4_MEL   L..F..................................................................
SPTREMBL|G7DVE0_MIX   F..VLVVESANQKTGLGSKMGTWLLERGLKP.......TYGYMGTNKEG.CGGLSEQIFHQGLGPQDILL
SPTREMBL|H0EEH2_GLA   F..VLVVESLNQKTGLGSKMGTWLLERQLHP.......KYKYIGTTKEG.CGGLGEQIFHQGLDPQSVVK
SPTREMBL|I4YAU9_WAL   F..VFVVESISQKNGLGSRLGTQLLQRGFHT.......KYDFMGAHTEG.CGGVEFQIPHQGLDSPSIQA
SPTREMBL|I4YE20_WAL   F..VLVVESISQKNGLGSKLGTQLIQRGFTP.......KYKHIGTFHEG.CGGIADQIPYQKLHPADIEA
SPTREMBL|K1WGM0_MAR   F..VLVVESLNQKTGLGSKMGTWLLERQLTP.......KYKYIGTTKEG.CGGLGEQIFHQGLDAQSVMK
SPTREMBL|K5WES2_PHA   F..VLVVESLNQKTGLGLRFGTWLLERSLTP.......KYAYMGTNKEG.CGGLTEQIPHQGLDPQSIIL
SPTREMBL|Q55T38_CRY   F..VFVAESLNRKTGLGSKFGTWLLERDLRP.......RYNYIGTSKEG.CGGLGEQIGHQNLGSSDIAL
SPTREMBL|L1J2M8_GUI   F..VLVVEPLNRKTGLGVRFGTWLLERGLSP.......KYGYIGSTREG.CGGLWEHAYHQGYDADSILA
SPTREMBL|A0YNC4_LYN   M..VIVSEAFNRRTGLGSRFGTWLLERGFSP.......KFASLGTHHEG.CGGLWEQFPHQGIDPTGIIN
SPTREMBL|A0ZLZ6_NOD   EGNLVVVEDHWIEGGLGAAVLDAFAGTSNTPIYDGPQLQLIKLAVRDMPGSGTPEELLHAAKIDADAIVE
SPTREMBL|B4WLI3_9SY   L..VMVVESFNEKTGLGARYGTWLLERGYTP.......KFAHAGTHEEG.GGGLWEQFYHQGLDPESIMQ
SPTREMBL|C9M6X8_9BA   A..VLTLEDHNARTGMGSLWLARASELGLTG.......RALRLGVTHYGMSGPSSDVYAQMELDGQSVAE
SPTREMBL|D1Y3C7_9BA   T..VLTVEDHNVVSGMGSLWLARAEELGLHS.......VARRLGVHRYGDSGPSEEVYAAMGLSADKIAE
SPTREMBL|D2Z2G4_9BA   K..VIALEDHCARTGMGSIWSSAAAEAGLTG.......RWSFMGVKRYGDSGPSSDVYEAMGLSSEGLVA
SPTREMBL|D3L2W3_9BA   P..IFTFEDHHAGTGLGREIVSKVQSLGLHC.......RVKPLGVWRYGDSGPFDEVFEAMGLSSKALAE
SPTREMBL|D6THL7_9CH   GNKIVTVEDHWPEGGLGDAVTEVFTQR......DGPMPQVVKLAVQSMPGSGTPAELLEEAGISAHHIVQ
SPTREMBL|D8F7T0_9DE   D.HVFVLEDHAPVGGFADFLLNEMVGAGGLG.....SRSFVKFAVEGYPACGTPAEALRFHGLDGASLAN
SPTREMBL|F4XPA7_9CY   S.WVFTLDNHYVIGGQGDRIVSGLAELGLSG.....QIQVRQFGVFDIPKCGQNQEVLRAHGLDAESLAF
SPTREMBL|G2H8B8_9DE   T.LVVTLDNHLLDGGQGRRIAAAMAGAGIA.......VPVLHLGPEDFPPCGRNDEVLAKLGLDAGGIAR
SPTREMBL|G3J1Y8_9GA   M..VLIVESYNRKTGLGSRFGSWLLERGLTP.......KYAHIATHKEG.CGGLWEQFPHQGIDPAGIMT
SPTREMBL|G9PZY8_9BA   P..LLTCEDHHADTGIGAVAALAFARAGKAV.......RIKNLGVTRYGLSGSNGDVLADMGLDADGIAG
Consensus                V   E     TGLG        E G               G    G  G   E     GLD   I

                              780       790       800       810       820       830       840
1 [transketolase Sp   AAKSVC................................................................
Paeciolomyces divar   AVKEVSGK..............................................................
SPTREMBL|A8NV31_COP   KIKQLTQ...............................................................
SPTREMBL|D8Q4R0_SCH   KIKELNQGLSLNQGGTHTHH..................................................
SPTREMBL|E3JQR8_PUC   KVKVLAA...............................................................
SPTREMBL|F4RJP4_MEL   ......................................................................
SPTREMBL|G7DVE0_MIX   KAKSLSN.......ASPK....................................................
SPTREMBL|H0EEH2_GLA   AIRSLQ................................................................
SPTREMBL|I4YAU9_WAL   RIKKVAGRK.............................................................
SPTREMBL|I4YE20_WAL   EIRKLHA...............................................................
SPTREMBL|K1WGM0_MAR   AVRSLQPFRLHRTASTPKHEAAYHQHLTEVLYHSTTVTVQPYNRTTVQQITGKRCPVPAEPEPGASNEIF
SPTREMBL|K5WES2_PHA   KIRQLSQGSQRA..........................................................
SPTREMBL|Q55T38_CRY   KVKQMIK...............................................................
SPTREMBL|L1J2M8_GUI   KAKSMI................................................................
SPTREMBL|A0YNC4_LYN   KVKQLAK...............................................................
SPTREMBL|A0ZLZ6_NOD   AVKSQVRQPVGV..........................................................
SPTREMBL|B4WLI3_9SY   KVKKLAGK..............................................................
SPTREMBL|C9M6X8_9BA   RFARLIR...............................................................
SPTREMBL|D1Y3C7_9BA   SLEELKKVAR............................................................
SPTREMBL|D2Z2G4_9BA   KVERVLRG..............................................................
SPTREMBL|D3L2W3_9BA   CVLKELEK..............................................................
SPTREMBL|D6THL7_9CH   AVKALI................................................................
SPTREMBL|D8F7T0_9DE   RILFHMD...............................................................
SPTREMBL|F4XPA7_9CY   EINKTMM...............................................................
SPTREMBL|G2H8B8_9DE   AIRDTLQSIR............................................................
SPTREMBL|G3J1Y8_9GA   NVKDLLKANPDASAFPVL....................................................
SPTREMBL|G9PZY8_9BA   G.AKELLK..............................................................
Consensus
```

Figure 1H continued

```
                           850       860       870       880       890       900       910
1 [transketolase Sp   ......................................................................
Paeciolomyces divar   ......................................................................
SPTREMBL|A8NV31_COP   ......................................................................
SPTREMBL|D8Q4R0_SCH   ......................................................................
SPTREMBL|E3JQR8_PUC   ......................................................................
SPTREMBL|F4RJP4_MEL   ......................................................................
SPTREMBL|G7DVE0_MIX   ......................................................................
SPTREMBL|H0EEH2_GLA   ......................................................................
SPTREMBL|I4YAU9_WAL   ......................................................................
SPTREMBL|I4YE20_WAL   ......................................................................
SPTREMBL|K1WGM0_MAR   FGPAVRLVTCSKSSTKKGMAAAESQWEPTKEISGHDDQGVVAGDVEDLERDGHANFNRLGWKRLTVVLLV
SPTREMBL|K5WES2_PHA   ......................................................................
SPTREMBL|Q55T38_CRY   ......................................................................
SPTREMBL|L1J2M8_GUI   ......................................................................
SPTREMBL|A0YNC4_LYN   ......................................................................
SPTREMBL|A0ZLZ6_NOD   ......................................................................
SPTREMBL|B4WLI3_9SY   ......................................................................
SPTREMBL|C9M6X8_9BA   ......................................................................
SPTREMBL|D1Y3C7_9BA   ......................................................................
SPTREMBL|D2Z2G4_9BA   ......................................................................
SPTREMBL|D3L2W3_9BA   ......................................................................
SPTREMBL|D6THL7_9CH   ......................................................................
SPTREMBL|D8F7T0_9DE   ......................................................................
SPTREMBL|F4XPA7_9CY   ......................................................................
SPTREMBL|G2H8B8_9DE   ......................................................................
SPTREMBL|G3J1Y8_9GA   ......................................................................
SPTREMBL|G9PZY8_9BA   ......................................................................
Consensus

                           920       930       940       950       960       970       980
1 [transketolase Sp   ......................................................................
Paeciolomyces divar   ......................................................................
SPTREMBL|A8NV31_COP   ......................................................................
SPTREMBL|D8Q4R0_SCH   ......................................................................
SPTREMBL|E3JQR8_PUC   ......................................................................
SPTREMBL|F4RJP4_MEL   ......................................................................
SPTREMBL|G7DVE0_MIX   ......................................................................
SPTREMBL|H0EEH2_GLA   ......................................................................
SPTREMBL|I4YAU9_WAL   ......................................................................
SPTREMBL|I4YE20_WAL   ......................................................................
SPTREMBL|K1WGM0_MAR   EAIALGALSIPSAFATLGMVPGVITTVGLGFVAIYTSYIVGQVKLKFPHVAHYADAGRLIAGKFGYELIG
SPTREMBL|K5WES2_PHA   ......................................................................
SPTREMBL|Q55T38_CRY   ......................................................................
SPTREMBL|L1J2M8_GUI   ......................................................................
SPTREMBL|A0YNC4_LYN   ......................................................................
SPTREMBL|A0ZLZ6_NOD   ......................................................................
SPTREMBL|B4WLI3_9SY   ......................................................................
SPTREMBL|C9M6X8_9BA   ......................................................................
SPTREMBL|D1Y3C7_9BA   ......................................................................
SPTREMBL|D2Z2G4_9BA   ......................................................................
SPTREMBL|D3L2W3_9BA   ......................................................................
SPTREMBL|D6THL7_9CH   ......................................................................
SPTREMBL|D8F7T0_9DE   ......................................................................
SPTREMBL|F4XPA7_9CY   ......................................................................
SPTREMBL|G2H8B8_9DE   ......................................................................
SPTREMBL|G3J1Y8_9GA   ......................................................................
SPTREMBL|G9PZY8_9BA   ......................................................................
Consensus
```

Figure 1H continued

```
                           990      1000      1010      1020      1030      1040      1050
1 [transketolase Sp ......................................................................
Paeciolomyces divar ......................................................................
SPTREMBL|A8NV31_COP ......................................................................
SPTREMBL|D8Q4R0_SCH ......................................................................
SPTREMBL|E3JQR8_PUC ......................................................................
SPTREMBL|F4RJP4_MEL ......................................................................
SPTREMBL|G7DVE0_MIX ......................................................................
SPTREMBL|H0EEH2_GLA ......................................................................
SPTREMBL|I4YAU9_WAL ......................................................................
SPTREMBL|I4YE20_WAL ......................................................................
SPTREMBL|K1WGM0_MAR AMFSLQLIFLVGSHCLTGTIAFLKITESSTCSLVFGVISAVILLLLALPPSFADVAILGYIDFVSIIGAI
SPTREMBL|K5WES2_PHA ......................................................................
SPTREMBL|Q55T38_CRY ......................................................................
SPTREMBL|L1J2M8_GUI ......................................................................
SPTREMBL|A0YNC4_LYN ......................................................................
SPTREMBL|A0ZLZ6_NOD ......................................................................
SPTREMBL|B4WLI3_9SY ......................................................................
SPTREMBL|C9M6X8_9BA ......................................................................
SPTREMBL|D1Y3C7_9BA ......................................................................
SPTREMBL|D2Z2G4_9BA ......................................................................
SPTREMBL|D3L2W3_9BA ......................................................................
SPTREMBL|D6THL7_9CH ......................................................................
SPTREMBL|D8F7T0_9DE ......................................................................
SPTREMBL|F4XPA7_9CY ......................................................................
SPTREMBL|G2H8B8_9DE ......................................................................
SPTREMBL|G3J1Y8_9GA ......................................................................
SPTREMBL|G9PZY8_9BA ......................................................................
Consensus

                          1060      1070      1080      1090      1100      1110      1120
1 [transketolase Sp ......................................................................
Paeciolomyces divar ......................................................................
SPTREMBL|A8NV31_COP ......................................................................
SPTREMBL|D8Q4R0_SCH ......................................................................
SPTREMBL|E3JQR8_PUC ......................................................................
SPTREMBL|F4RJP4_MEL ......................................................................
SPTREMBL|G7DVE0_MIX ......................................................................
SPTREMBL|H0EEH2_GLA ......................................................................
SPTREMBL|I4YAU9_WAL ......................................................................
SPTREMBL|I4YE20_WAL ......................................................................
SPTREMBL|K1WGM0_MAR GITVIATGLTRGDMPAQADWSAWPKENTSFTDGFIAISNIVFAYSFAVCQFSFMDEMHTPKDFVKSIWAL
SPTREMBL|K5WES2_PHA ......................................................................
SPTREMBL|Q55T38_CRY ......................................................................
SPTREMBL|L1J2M8_GUI ......................................................................
SPTREMBL|A0YNC4_LYN ......................................................................
SPTREMBL|A0ZLZ6_NOD ......................................................................
SPTREMBL|B4WLI3_9SY ......................................................................
SPTREMBL|C9M6X8_9BA ......................................................................
SPTREMBL|D1Y3C7_9BA ......................................................................
SPTREMBL|D2Z2G4_9BA ......................................................................
SPTREMBL|D3L2W3_9BA ......................................................................
SPTREMBL|D6THL7_9CH ......................................................................
SPTREMBL|D8F7T0_9DE ......................................................................
SPTREMBL|F4XPA7_9CY ......................................................................
SPTREMBL|G2H8B8_9DE ......................................................................
SPTREMBL|G3J1Y8_9GA ......................................................................
SPTREMBL|G9PZY8_9BA ......................................................................
Consensus
```

Figure 1H continued

```
                            1130      1140      1150      1160      1170      1180      1190
1 [transketolase Sp  ......................................................................
Paeciolomyces divar  ......................................................................
SPTREMBL|A8NV31_COP  ......................................................................
SPTREMBL|D8Q4R0_SCH  ......................................................................
SPTREMBL|E3JQR8_PUC  ......................................................................
SPTREMBL|F4RJP4_MEL  ......................................................................
SPTREMBL|G7DVE0_MIX  ......................................................................
SPTREMBL|H0EEH2_GLA  ......................................................................
SPTREMBL|I4YAU9_WAL  ......................................................................
SPTREMBL|I4YE20_WAL  ......................................................................
SPTREMBL|K1WGM0_MAR  GLIEIVIYTVTGSLIYAFVGKDVESPALLSAGPLMAKIAFGVALPVIFISGSINGTVVGRYLHGRMYKDS
SPTREMBL|K5WES2_PHA  ......................................................................
SPTREMBL|Q55T38_CRY  ......................................................................
SPTREMBL|L1J2M8_GUI  ......................................................................
SPTREMBL|A0YNC4_LYN  ......................................................................
SPTREMBL|A0ZLZ6_NOD  ......................................................................
SPTREMBL|B4WLI3_9SY  ......................................................................
SPTREMBL|C9M6X8_9BA  ......................................................................
SPTREMBL|D1Y3C7_9BA  ......................................................................
SPTREMBL|D2Z2G4_9BA  ......................................................................
SPTREMBL|D3L2W3_9BA  ......................................................................
SPTREMBL|D6THL7_9CH  ......................................................................
SPTREMBL|D8F7T0_9DE  ......................................................................
SPTREMBL|F4XPA7_9CY  ......................................................................
SPTREMBL|G2H8B8_9DE  ......................................................................
SPTREMBL|G3J1Y8_9GA  ......................................................................
SPTREMBL|G9PZY8_9BA  ......................................................................
Consensus

                            1200      1210      1220      1230      1240      1250      1260
1 [transketolase Sp  ......................................................................
Paeciolomyces divar  ......................................................................
SPTREMBL|A8NV31_COP  ......................................................................
SPTREMBL|D8Q4R0_SCH  ......................................................................
SPTREMBL|E3JQR8_PUC  ......................................................................
SPTREMBL|F4RJP4_MEL  ......................................................................
SPTREMBL|G7DVE0_MIX  ......................................................................
SPTREMBL|H0EEH2_GLA  ......................................................................
SPTREMBL|I4YAU9_WAL  ......................................................................
SPTREMBL|I4YE20_WAL  ......................................................................
SPTREMBL|K1WGM0_MAR  IVRFINTKQGWLTWILFITVITVIAWVIAEAIPFFSDLLSISSALFISGFTFYFPAIFWFMLLKEGKWNA
SPTREMBL|K5WES2_PHA  ......................................................................
SPTREMBL|Q55T38_CRY  ......................................................................
SPTREMBL|L1J2M8_GUI  ......................................................................
SPTREMBL|A0YNC4_LYN  ......................................................................
SPTREMBL|A0ZLZ6_NOD  ......................................................................
SPTREMBL|B4WLI3_9SY  ......................................................................
SPTREMBL|C9M6X8_9BA  ......................................................................
SPTREMBL|D1Y3C7_9BA  ......................................................................
SPTREMBL|D2Z2G4_9BA  ......................................................................
SPTREMBL|D3L2W3_9BA  ......................................................................
SPTREMBL|D6THL7_9CH  ......................................................................
SPTREMBL|D8F7T0_9DE  ......................................................................
SPTREMBL|F4XPA7_9CY  ......................................................................
SPTREMBL|G2H8B8_9DE  ......................................................................
SPTREMBL|G3J1Y8_9GA  ......................................................................
SPTREMBL|G9PZY8_9BA  ......................................................................
Consensus
```

Figure 1H continued

```
                                1270      1280      1290      1300      1310
1 [transketolase Sp   ......................................................
Paeciolomyces divar   ......................................................
SPTREMBL|A8NV31_COP   ......................................................
SPTREMBL|D8Q4R0_SCH   ......................................................
SPTREMBL|E3JQR8_PUC   ......................................................
SPTREMBL|F4RJP4_MEL   ......................................................
SPTREMBL|G7DVE0_MIX   ......................................................
SPTREMBL|H0EEH2_GLA   ......................................................
SPTREMBL|I4YAU9_WAL   ......................................................
SPTREMBL|I4YE20_WAL   ......................................................
SPTREMBL|K1WGM0_MAR   TKMNILLSLVNVAVFLIGMIVLVGGTYSSIVDIKDQYSSGAVRGAFTCAPIGSE
SPTREMBL|K5WES2_PHA   ......................................................
SPTREMBL|Q55T38_CRY   ......................................................
SPTREMBL|L1J2M8_GUI   ......................................................
SPTREMBL|A0YNC4_LYN   ......................................................
SPTREMBL|A0ZLZ6_NOD   ......................................................
SPTREMBL|B4WLI3_9SY   ......................................................
SPTREMBL|C9M6X8_9BA   ......................................................
SPTREMBL|D1Y3C7_9BA   ......................................................
SPTREMBL|D2Z2G4_9BA   ......................................................
SPTREMBL|D3L2W3_9BA   ......................................................
SPTREMBL|D6THL7_9CH   ......................................................
SPTREMBL|D8F7T0_9DE   ......................................................
SPTREMBL|F4XPA7_9CY   ......................................................
SPTREMBL|G2H8B8_9DE   ......................................................
SPTREMBL|G3J1Y8_9GA   ......................................................
SPTREMBL|G9PZY8_9BA   ......................................................
Consensus
```

X non conserved
X ≥50% conserved
X all match