

Article



## Molecular Cloning, Characterization, and Functional Analysis of Acetyl-CoA *C*-Acetyltransferase and Mevalonate Kinase Genes Involved in Terpene Trilactone Biosynthesis from *Ginkgo biloba*

# Qiangwen Chen, Jiaping Yan, Xiangxiang Meng, Feng Xu \*, Weiwei Zhang, Yongling Liao and Jinwang Qu

College of Horticulture and Gardening, Yangtze University, Jingzhou 434025, Hubei, China; chenqwx@foxmail.com (Q.C.); xiaolinyingxue@163.com (J.Y.); 15927855989@163.com (X.M.); zww8312@163.com (W.Z.); liaoyongling@yeah.net (Y.L.); qujinwang@163.com (J.Q.) \* Correspondence: xufeng198@126.com; Tel.: +86-71-6806-6264; Fax: +86-71-6806-6262

Academic Editor: Derek J. McPhee Received: 20 November 2016; Accepted: 27 December 2016; Published: 2 January 2017

Abstract: Ginkgolides and bilobalide, collectively termed terpene trilactones (TTLs), are terpenoids that form the main active substance of Ginkgo biloba. Terpenoids in the mevalonate (MVA) biosynthetic pathway include acetyl-CoA C-acetyltransferase (AACT) and mevalonate kinase (MVK) as core enzymes. In this study, two full-length (cDNAs) encoding AACT (GbAACT, GenBank Accession No. KX904942) and MVK (GbMVK, GenBank Accession No. KX904944) were cloned from G. biloba. The deduced GbAACT and GbMVK proteins contain 404 and 396 amino acids with the corresponding open-reading frame (ORF) sizes of 1215 bp and 1194 bp, respectively. Tissue expression pattern analysis revealed that GbAACT was highly expressed in ginkgo fruits and leaves, and GbMVK was highly expressed in leaves and roots. The functional complementation of GbAACT in AACT-deficient Saccharomyces cerevisiae strain  $\Delta$ erg10 and GbMVK in MVK-deficient strain  $\Delta$ erg12 confirmed that GbAACT mediated the conversion of mevalonate acetyl-CoA to acetoacetyl-CoA and GbMVK mediated the conversion of mevalonate to mevalonate phosphate. This observation indicated that GbAACT and GbMVK are functional genes in the cytosolic mevalonate (MVA) biosynthesis pathway. After G. biloba seedlings were treated with methyl jasmonate and salicylic acid, the expression levels of GbAACT and GbMVK increased, and TTL production was enhanced. The cloning, characterization, expression and functional analysis of *GbAACT* and *GbMVK* will be helpful to understand more about the role of these two genes involved in TTL biosynthesis.

Keywords: Ginkgo biloba; GbAACT; GbMVK; SA; MeJA; functional complementation

## 1. Introduction

*Gingko biloba* L., which dates back to more than 200 million years, is the only surviving member of ginkgophyta in gymnosperm family and considered a "living fossil" [1]. Ginkgo leaf extracts are used to treat cardiovascular and cerebrovascular diseases because of their highly specific and potent platelet-activating factor receptor antagonists [2]. *G. biloba* contains terpene trilactones (TTLs), such as diterpenoid ginkgolides and sesquiterpenoid bilobalide, as the main bioactive components [3]. TTLs in *G. biloba* possess unique biological properties and promote high activities of anti-platelet-activation factors [4]. TTLs also function as a selective glycine receptor and participate as the main bioactive substance in *G. biloba* [5]. However, there are many difficulties with respect to supply of ginkgo leaves and chemical synthesis is far from of being applicable for commercial-scale production. Different biotechnological strategies to improve TTL production have been used, including screening and

selection of in vitro ginkgo cultures, cell differentiation levels of these cultures, and optimization of culture conditions, feeding the elicitation strategies [6]. However, until now, the yields obtained from cell cultures have been low and new strategies such as the use of key genes for increasing TTL production by genetic engineering are imperative.

Until recently, the terpenoid origin of the plant was considered to have two involved precursors—isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl pyrophosphate (DMAPP). IPP and DMAPP are synthesized via two indenpendent isoprenoid pathways: the cytosolic mevalonate (MVA) pathway and the plastidial methylerythritol 4-phosphate (MEP) pathway (Figure 1). In the MVA pathway, three acetyl-CoA molecules are converted into isopentenyl diphosphate (IPP), and this process is catalyzed by seven enzymes. In the MEP pathway, IPP is similarly produced, and an IPP isomer called dimethylallyl pyrophosphate is formed from pyruvate and D-glyceraldehyde-3-phosphate via eight enzymatic reactions. Generally, the main MVA-derived isoprenoid end-products are sesquiterpenoids and sterols, whereas the monoterpenoids and diterpenoids are derived from the MEP pathway [7]. However, an IPP crosstalk exists between plastids and cytosol, and this observation suggests that the MVA pathway also contributes to TTL biosynthesis [8,9]. A few of genes involved in the MVA and MEP pathway have been cloned and identified in G. biloba, such as mevalonate diphosphate decarboxylase (MVD) [10,11], 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) [12], 1-deoxy-D-xylulose-5-phosphate synthase (DXS) [13], 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR) [14], 4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase (CMK) [15], 2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase (MECS) [16,17], 1-hydroxy-2-methyl-2-(E)butenyl-4-diphosphate synthase (HDS) [18], and 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase (IDS) [19]. TTL content in G. biloba could be enhanced by upregulating the transcript level of some of these genes [11,15,19,20]. Thus, cloning and characterization the genes involved in TTL biosynthesis could provide good candidates for metabolic engineering to increase TTL production in G. biloba.

In the MVA pathway, acetyl-CoA C-acetyltransferase (AACT) is the first enzyme that catalyzes the conversion of acetyl-CoA into acetoacetyl-CoA. Mevalonate kinase (MVK) is the fourth enzyme that catalyzes the conversion of MVA into mevalonate-5-phosphate, which undergoes an enzymatic reaction catalyzed by phospho-mevalonate kinase to generate mevalonate-5-diphosphate. Mevalonate-5-diphosphate is then transformed into isopentenyl diphosphate, as catalyzed by MVD. AACT and MVK genes have been described in plants; while reports on these genes are few, what we do know is that AACT is found in Elaeis guineensis Jacq [21], Ganoderma lucidum [22], and *Bacopa monnieri* [23], and *MVK* is found in *Eucommia ulmoides* [24] and *Hevea brasiliensis* [25]. However, cloning and characterization of AACT and MVK from G. biloba have not been reported in the literature. In this study, two novel cDNAs of AACT and MVK were cloned and characterized from G. biloba. Yeast complementation assays were conducted to identify the function of these genes. The expression patterns of *GbAACT* and *GbMVK* in various tissues, including roots, stems, leaves, fruits, male and female flowers were also examined to describe TTL synthesis. In addition, our previous work showed that the transcripts of HMGR [26] and MVD [11] genes involved in the MVA pathway are positively responsive to methyl jasmonate (MeJA) and salicylic acid (SA) treatments in ginkgo. Therefore, the expression profiles of *GbAACT* and *GbMVK* as well as TTL contents under the induction by MeJA and SA were also investigated, which will facilitate future work to map and regulate these important steps involved in TTL biosynthetic pathway at the level of molecular genetics.



**Figure 1.** The biosynthetic pathway of ginkgolides and bilobalide in *Ginkgo biloba*. MEP: methylerythritol 4-phosphate; MVA: mevalonate; AACT: acetyl-CoA *C*-acetyltransferase; HMGS: 3-Hydroxy-3-methylglutaryl-CoA synthase; HMGR: 3-Hydroxy-3-methylglutaryl-CoA reductase; MVK: mevalonate kinase; PMK: Phosphomevalonate Kinase; MPDC: Diphospho-MVA decarboxylase; IPPI: Isopentenyl diphosphate isomerase; FPPS: Farnesyl diphosphate synthase; GGPPS: Geranylgeranyl diphosphate synthase.

#### 2. Results and Discussion

## 2.1. Isolation and Characterization of the cDNA of GbAACT and GbMVK

The full-length cDNA sequences of *GbAACT* and *GbMVK* were cloned through RT-PCR (Reverse transcription polymerase chain reaction). Using total RNA isolated from the young leaves of *G. biloba*, a 2028-bp and 2057-bp fragment was amplified through RT-PCR. After aligating these genes into pMD19-T vector and sequencing, the results of sequence analysis showed that *GbAACT* contained an open-reading frame (ORF) of 1215 bp and that *GbMVK* contained an ORF of 1194 bp. The *GbAACT* encoded a protein containing 404 amino acids with a 6.33 isoeletric point, and the calculated molecular weight was approximately 41.5 kDa. Meanwhile, *GbMVK* encoded a protein containing 397 amino acids with a 5.71 isoeletric point, and the calculated molecular weight was approximately 42 kDa. A BLASTn search of *GbAACT* and *GbMVK* with other plant species showed that *GbAACT* and *GbMVK* are highly homologous to *AACT* genes and *MVK* genes from other plant species. Therefore, these genes were designated as *GbAACT* (GenBank accession No. KX904942) and *GbMVK* (GenBank accession No. KX904944).

#### 2.2. Bioinformatics Analysis of the Deduced GbAACT and GbMVK Protein

A BLASTp search against nonredundant (nr) protein in GenBank database (https://blast.ncbi. nlm.nih.gov) showed that the polypeptide sequence of GbAACT had 78%–81% homology with AACTs of many other plant species, such as *Nelumbo nucifera* (NnAACT, Identity: 81%, Accession No. XP\_010252788.1), *Amborella trichopoda* (AtAACT, Identity: 80%, Accession No. XP\_011628271.1), *Hevea brasiliensis* (HbAACT, Identity: 81%, Accession No. AFJ74323.1), *Ricinus communis* (RcAACT, Identity: 81%, Accession No. ALD84318.1), and *Euphorbia helioscopia* (EhAACT, Identity: 78%, Accession No. ALC76524.1). The deduced amino acid sequence of GbMVK had a high degree of homology, from 61% to 63% with the MVK protein of other plant species, such as *Prunus mume* (PmMVK, Identity: 62%, Accession No. XP\_009349394.1), *Jatropha curcas* (JcMVK, Identity: 62%, Accession No. XP\_012089078.1), *Ziziphus jujuba* (ZjMVK, Identity: 62%, Accession No. XP\_012089078.1), *Ziziphus jujuba* (ZjMVK, Identity: 62%, Accession No. AIO11226.1).

Multiple alignments of GbAACT with AACTs from other plants indicated that the plant AACTs were the most similar (Figure 2). The online InterPro result showed that the function of GbAACT harbors the activity of thiolase II, and the structure of GbAACT monomer contains three domains, including thiolase-like domain (14–280), N-terminal (15–273), and C-terminal (282–402). Based on the differences in catalytic activities, the thiolases are of two types: thiolase I (acetyl-CoA *C*-acyltransferase) and thiolase II (acetyl-CoA *C*-acetyltransferase). Thiolase I is a degradative thiolase, and thiolase II is a synthetic thiolase [27]. Residues of two cystines, one histidine, and one asparagine are present in GbAACT, which are highly conserved in AACT among the thiolases from different sources, and are important for catalytic activity (Figure 2, marked with "asterisk") [28,29]. One highly conserved domain (NVHGGAVSIGHPIGCSG) at the C-terminal end is also present. The thiolase II active site GVAGVCNGGGGASA at the last position is specific for AACT [30]. This evidence confirms the similarity function of GbAACT to AACTs from other plants.

Sequence alignment using Vector NTI 11.5.1 showed that the predicted GbMVK shared a high level of identity with MVKs from other plant species, implying that GbMVK harbors the activity of mevalonate kinase. The structure of GbMVK monomer contains ribosomal protein S5 2-type (8–230), GHMP kinase N-terminal domain (140–220), and GHMP kinase C-terminal domain (236–390). Ile<sub>146</sub>–Ala<sub>157</sub> is an ATP-binding conserved site (Figure 3, marked with red box), and these findings indicated that GbMVK has a similar catalytic function to other plant MVKs.

NnAACT	MGSTSG <mark>SDLLK</mark> SRDVOIVGVARTP <mark>IGGLSGSLSTLS</mark> ATKLGSIAIEAALKRAGVEPRIVOEVFFGNVLSANLGOAPA	77
	MADAAAADDSTKPRDVOVTCVARTEMCGETCSLSSLSSLSATKLGSTATECALKRANVDPATVOEVERGNULSANLGOADA	79
3 + 3 3 000		
ALAACT	MAPAAVSDSLKARDVCIVCVARTEWGGLCCSLSSLFATKLGSTATEOALKKANVDPAHVOEVFFGNVLSANLGQAPA	//
Hbaact	MS <mark>PSSDSIN</mark> PRDVCIVGVARTPMGGFLGSLSS <mark>FS</mark> ATKLGSIAI <mark>Q</mark> AAIKRANVDPSLVQEVFFGNVLSANLGQAPA	75
RCAACT	MAPTDSLKPRDVOLVGVARTENGGFLGSLSSLSATKLGSLAIHOALERARVDPSLVEEVFFGNVLSANLGOAPA	74
MANACT	MADY SARAGEDS IN PROVIDE TARTEMONT FLOAT SEVENTIC OSTATEANT KRAMUR PALVERONUL SANU COADA	80
N D D OTT		50
MAAACT	.MAPASNPSSSSIKPRDVCIVEVARTPIGFICSLSSLSATOLGSVAIKAALRRAGVEPSIVQEVFFGNVLSANLGQAPA	79
EhAACT	MASSS <mark>SDPIKP</mark> RDVC <mark>IVCV</mark> ARTP <mark>ICSFLCSLSSIS</mark> ATKLCSIAIQCALKRANVDPSLVCEVFFCNVLSANLCQAPA	76
PnAACT	MA PAAATDOSEST KARDVOTVGAARTPMCGFLCTLSSLSATKLGSTATOSALKRANVD PALVOEVEEGNVLSANLGOA PA	80
NI-DOCT		50
NDAAC'I'	·····MAPAAADSIKPRDVCIVGVARTPMGGFLGSDSSLSATRUGSIATEGPLKRANVLFSIVEDVFFGNVLSANLGQAPA	/6
	*	
Chart		157
GDAACI	KQAALGAGI FDS VICI IIINK VCASGWAAIMIAAQI VQIGINGI VVAGGWESMSMAPI CHENARAGI KIGAGI FVDGWEM	157
NNAAC'I'	RQAALGAGI P <mark>NSVV</mark> CTTIINKVCASGMKATIIDAAQSIQLG <mark>INDV</mark> VVAGGMESMSNAPKYIAEARK <mark>G</mark> SRLGHDTIVDGMLKD	159
Ataact	RQAALGAGIP <mark>NSVI</mark> CTTINKVCASG <mark>MKAIMLAAQSIQLGVNDV</mark> VVAGGMESMSNAPFYLPELRKGNRLGHGTVVDGMLTD	157
HDAACT	ROAALGAGT P <mark>NSVTCTTTNKVCASCMKATMLAALTTOACTNDT</mark> VVACCMESMSNAPKYLAEAR <mark>RC</mark> SRLGHDTTTDCMLKD	155
Babber	ROADLCACT DNOVECTUTNIXY/CACOMMATTEL AA OTT CAC THE TWAACOMESMSNAPWYLARADNOODLCUDTTT DOME HD	1 5 4
KCAAC I	KOARDORGIPHO VICI TINKVCHSCH AATHDRAGIIQAGINDIVVAGGNESNSNAFKI LABARNOSKUGADIIIDGHLAD	104
MdAACT	RQAALGAGIP <mark>NSVI</mark> CTTVNKVC <mark>A</mark> SGLKAIMLAAQ <mark>S</mark> IQIG <mark>I</mark> NDVVVAGGMESMSNVPKYLAEA <mark>RK</mark> GSRLGHDSLVDGMLKD	160
MaAACT	RQAALGAGIP <mark>NSGI</mark> CTTINKVC <mark>S</mark> SGMKAIMIAS <mark>QSIEL</mark> GTNDIVVAGGMESMSNTPKYLTDVRNGSRLGHDTIVDGMVKD	159
Ehaact	ROAALGAGT PNSVICTTTNKVCSSCM KATMLAAOTTOLG <mark>VHDT</mark> VVAGGMESMSN <mark>A</mark> PKVLPBARNGSRLGH <mark>DTTT</mark> DCTLKD	156
DEALOT		1.00
PIIAACI	RQAALGAGI PINS VVCI II INK VCASGMAAIMLAAQSI QIGINDI VVAGGMESNISN VPAILAEARAGSRIGHDSLVDG LLAD	100
NbAACT	RQAALGAGIP <mark>NTVV</mark> CTTVNKVCASGLKAIMLAVQSIQLG <mark>INDV</mark> VVAGGMESMSNVPKYIAEARK <mark>G</mark> SRLGHDSLVDGMLKD	156
a)		000
GDAAC'I'	GIWDYNNDIGWEVEADLOADQHHISKDEQDNYAVQSBDQEISAQNSGABAWDIVPVDISGEKEKPSTIIEKDEGIAKPDA	237
NnAACT	GLWDVYNNFGMCVCABLCADQHAITREQQDTY <mark>AIQ</mark> SFERCIAAQNSCAFAWEIVPVEVSCCRCKPSTVVDKDECLCKFDP	239
Ataact	GIWDYNNEGMGMCAELCADOHSITREBODSFAIKSEERGTAACNEGAEAWETVEWEVSGCREKTNTUTEKERC	2.37
That		227
HDAAC I	GIWDVINDFGWCVCAPICADQHNIIRDEODSIAIRSFERGNSACNGGVFSWEIVPVDVSGGRGKSVMVVDRDEGFIRFDA	235
RCAACT	GLWDVYND <mark>F</mark> GMCVCAE <mark>L</mark> CADRHTITREEQDSYAIRSFERGISAQNDGLFSWEIVPVEVSGGRGRLSTIIDKDEGI <mark>G</mark> KFDA	234
Mdaact	GLWDVYNDYGMCYCAPICADOHSVTREDODNFAVOSFERGIAAOGCDAFSWEIVPVEVSGGRGRESTIVDKDEGIGKEDA	240
Manner	CI WOWND COME VO COL CADOX KI SPERODA VA TRESPEROVSAO KNOUDAWE IT PUDVE COROK DST WITH TRESPERO	230
FIGAACT		255
ENAACT	GLWDVYNDFGWGVCGEICADQHKITIREQQDAYAIQSFERGISAQNAGLFSWEIAPVEVSAGRERSSTIVDRDEGIGRFDA	236
PnAACT	GLWDVYNDYGMCVCAEICAEQHGVSREEQDNYAIQSFERGIAAQNSGAEAWEIVPVEVSGCRGKPSTIVDKDEGICKFDA	240
NHAACT	GUTDVYKDOGMEVCAETCAENHKTTREFODDYAVOSEERGTAAOFAGAEAWETVPVEVEGEREKESTTVDKDDGUGKEDG	236
1101 1101		200
Gbaact	AKURKUREAFK, ECTVTACNASTISDCAAAIVIMSCEKALEICVKMIAKVACYADAAOAPDHETTAPSIATPKSISKAG	316
MDAACT	WE BELDE EVED CONTROLOND CONTROLOND AND A LOND CONTROL OF CALL AND A CALLED A LOND A CALLED A LOND CONTROL AND A LOND CONTROL AND A CALLED A LOND CONTROL AND A CALLED A LOND CONTROL AND A LOND CONTROL AND A CALLED A LOND CONTROL AND A LOND CONTROL AND A LOND CONTROL AND A LOND CONTROL AND A LOND C	210
MIGACI	VRIARIARI MEDGGGVI AGRAGO ISDCEARI VI VSGERATGIGGVVI ARTI GIEDARGEFETTI FERDARIA	515
Ataac'i'	VKLRKLREAFK.ECCSVTAGNASTLSDCAAAMVLVSGEKALNLCLHVLAKTTCYADAAQAPELFTTAPALALEKAVSNAH	316
Hbaact	AKLRKLRFSFK.NCCSVTAGNASIISDCAAALVLVSGEKAIELCLOVJIARIRCYCDAAQAPELFTTAPALAIPKAISNAG	314
RCAACT	AKLERKLERSEKEKESSTAGNASSTSDCAAATVIVSCEMATKICIOVIARTKCYCDAAOAPETETTAPATATPKATSNAG	314
MJAACI		
MCCAACT	AKLKKLRPSFRESGGSVTAGNASSISDGAAATVIASGEKVLKLGLQVIAKISGYADAAQAPPILFTTAPALAIPKAISNAG	200
		320
MaAACT	AKLRKLRE <mark>SFK.KCCTVTAGNASIISDCAAA</mark> EVL <mark>VSGEKALQLGLQVIAKIKCFA</mark> DAAQAPELFTTAFALAIPKAISNAG	320 318
MaAACT EhAACT	AKLRKLRESFK.KCCTVTAGNAS <mark>I</mark> ISDGAAAFVIVSGEKALQLGLOVIAKIKCFADAAQAPBLFTAPALAIEKAISNAG AKLRKLRESFKKNCCSVTAGNAS <mark>SI</mark> SDGAAALVIVSGEKAMKLCLHVIARIRCYCDAAQAPBLFTTAPALAIETAISNAG	320 318 316
MaAACT EhAACT PpAACT	AKLEKLERSFE, KOCTUTAGNAST ISOCAAAFUUVSOBKALOLOLOUTAKIKO FADAAQAPED PTTAPATA IPKALSNAG AKLEKLERSFEKINGOVTAGNASSISDOBAATUUVSOBKAMKLOLEVITAR IROYODAAQAPED PTTAPATA IP TAISNAG AVLEKTERSFEKINGOVTAGNASSISDOBAATUUVSOBTAMU VSOBTAVUVSOBTAVA	320 318 316
MaAACT EhAACT PnAACT	AKLRKLRESEK, KGCTVTAGNASTISDGAAAEVIVSGEKALOLGLOVIAKIKG PADAAQAPBIFTTAPAIAIPKAISNAG AKLRKLRESEKKNGCSVTAGNASSISDGAAALVIVSGEKANKLGLEVIARIRGYGDAAQAPBIFTTAPAIAIPTAFAIAIP AKLRKLRESEKETGCTVTAGNASSISDGAAALVIVSGEKANKLGLEVIARIRGYGDAAQAPBIFTTAPAIAIPKAISAG	320 318 316 320
MaAACT EhAACT PnAACT NbAACT	AKLEKLERSFE, KGCTVTAGNASTI SDCAAAFVUVSGEKALCICLCUTAKIKCFPADAAQPBD FTTAPATA I FKAISNAG AKLEKLERSFEKINGCSVTAGNASSI SDCAAAIVUVSGEKAMKICLEVIAR IRCYCDAAQAPBI FTTAPALA I PTAISNAG AKLEKLERSFE <mark>R</mark> EGTVTAGNASSI SDCGAALVIVSGETAVKICLCULAKISCVCDAAQSPD FTTSPALA I FKAISSAG AKLEKLERSFE <mark>RE</mark> GTVTAGNASSI SDCGAALVIVSGEKAIKICLNVIGKI SCVADAAQAPBI FTTAPALA I FKAISSAG	320 318 316 320 316
MaAACT EhAACT PnAACT NbAACT	AKLEKLERSFK, KGCTVTAGNASTISDCAAAFVIVSCEKALCICLOQIIAKIKGFADAAQAPEIPTTAPAIIAIPKAISNAG AKLEKLERSFKENGCSVTAGNASSISDCAAATVVVSCEKAMKLCLEVIARIRGYGDAAQAPEIFTTAPAIAIPKAISNAG AKLEKLERSFKETGCTVTAGNASSISDCGAATVVSCETAVKLCLOVLAKISGYGDAAQSPEIFTTSPAIAIPKAISSAG AKLEKLERSFKETGCTVTAGNASSISDCGAATVVSCEKAIKLCLOVLAKISGYGDAAQAPEIFTTAPAIAIPKAISSAG	320 318 316 320 316
MaAACT EhAACT PnAACT NbAACT	AKLRKLRESFK. KGGTVTAGNASIISDGAAAFVIVSGEKALQLGLOVIAKIKGFADAAQAPBIPTTAPAIAIPKAISNAG AKLRKLRESFKKNGGSVTAGNASSISDGAALVIVSGEKAMKLGLEVIAKIRGYGDAAQAPBIPTTAPAIAIPIAIA AKLRKLRESFKETGGTVTAGNASSISDGGAALVIVSGETAVKLGLOVLAKISGYGDAAQSPBIFTISPAIAIPKAISSAG AKLRKLRESFKEKDGTVTAGNASSISDGAAALVIVSGEKAIKLGLNVIGKISGYADAAQAPBIPTTAPAIAIPKAIKSAS * * *	320 318 316 320 316
MaAACT EhAACT PnAACT NbAACT GbAACT	AKLERKLERSFE, KGCTVTAGNAST ISDCAAAFUVIVSGEKALCICLOUTAKIKGFADAAQDEE PTTABAITAIEKAISNAG AKLERKLERSFEKENGGVTAGNASSISDCAAUVIVSGEKAMKLCLUVIARIRGYGDAAQAPEIFTTABAIAIETAISNAG AKLERLERSFEETGTVTAGNASSISDCGAAUVIVSGETAVKLCLOULAKISGYCDAAQSPEIFTTSPAIAIEKAISSAG AKLERLESFEETGTVTAGNASSISDCGAAUVIVSGETAVKLCLOULAKISGYCDAAQSPEIFTTSPAIAIEKAISSAG AKLERLESFEETGTVTAGNASSISDCGAAUVIVSGETAVKLCLOULAKISGYCDAAQSPEIFTTSPAIAIEKAISSAG * * * * * * HEALATEVYETNEAESUNALANOKTIKTKEETTÄVKLCLOULAKISGYCDAACSPEIFTTSPAIAIEKAISSAG	320 318 316 320 316
MaAACT EhAACT PnAACT NbAACT GbAACT	AKLIRKLIRESFK, KGCTUTAGNAST ISDCAAAFUUVSGBKALQLGLQUTAKTKG FADAAQAPED FTTAPATALI PKALSINAG AKLIRKLIRESFKING SVTAGNASS ISDCAAATUV VSGBKAMKLGLHVIAR IRGYGDAAQAPED FTTAPATALI PKALSINAG AKLIRKLIRESFKIETIG TUTAGNASS ISDCAATUV VSGBTAVKLGLOVLAKTSGYGDAAQAPED FTTAPATALI PKALSISAG AKLIRKLIRESFKETIG TUTAGNASS ISDCAATUV VSGBTAVKLGLOVLAKTSGYGDAAQAPED FTTAPATALI PKALSISAG AKLIRKLIRESFKETIG TUTAGNASS ISDCAATUV VSGBKATKLGLINUIGKTSGYADAAQAPED FTTAPATALI PKALSISAG AKLIRKLIRESFKETIG TUTAGNASS ISDCAATUV VSGBKATKLGLINUIGKTSGYADAAQAPED FTTAPATALI PKALSISAG * * * * LEALOTD VYELINDAF SVVALANQKTILGLINA VHGGAVSLGHPLGCSG ARTIVTLLGVI REPNGKYGVAGVENGGGGA	320 318 316 320 316 396
MaAACT EhAACT PnAACT NbAACT GbAACT NnAACT	AKLEKLERSFE, KGCTVTAGNASTI SDCAAAFVUVSGEKALCICLCUTAKIKGFADAAQPED FTTABATA I EKAISNAG AKLEKLERSFEKENGGSVTAGNASTI SDCAAAFVUVSGEKAMKLCLEVIAR IRGYGDAAQAPEI FTTABATA I EKAISNAG AKLEKLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAMKLCUQUAKISGYGDAAQAPEI FTTABATA I EKAISSAG AKLEKLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAKLCUQUAKISGYGDAAQAPEI FTTABATA I EKAISSAG AKLEKLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLCUVAGKISGYGDAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLGUVIGKI SGYGDAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLGUVIGKI SGYGDAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLGUVIGKI SGYGDAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLGUVGIGKI SGYADAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLGUVIGKI SGYADAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETGTVTAGNASSI SDCGAAUVVSGEKAIKLGUVGIGKI SGYADAAQAPEI FTTABATA I EKAISSAG AKLERLERSFERETTABATA I BALAIKSAS	320 318 316 320 316 396 399
MaAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT	AKLEKLERSFR, KGCTVTAGNASTI SICGAAAFUVIVSGBKALGLGLQUIAKTKGFADAAQAPBI PTTABATA I BKAISNAG AKLEKLERSFRKINGGSVTAGNASSI SICGAAI VIVSGBKAMKLGLHVIAR I REVGDAAQAPBI PTTABATA I PTAISNAG AKLEKLERSFRKINGTVTAGNASSI SICGAAI VIVSGBTAVKLGLQVIAKTSGVCDAAQSPBI PTTABATA I PTAISNAG AKLEKLERSFRKINGTVTAGNASSI SICGAAI VIVSGBTAVKLGLQVIAKTSGVCDAAQSPBI PTTABATA I PTAISNAG AKLEKLERSFRKINGTVTAGNASSI SICGAAI VIVSGBKAI KLGLNVIGKI SGVADAAQAPBI PTTABATA I PTAISNAG AKLEKLERSFRKINGTVTAGNASSI SICGAAI VIVSGBKAI KLGLNVIGKI SGVADAAQAPBI PTTABATA I PTAISNAG * * * * EACOID VIELENSSVVALANQKTI GLEFETV NVHGGAVSLGHPI GCSG ARI I VITLIGVI KOKNGKYGVSVONGGGGA LEASOID VIELENBASVVALANQKTI GLEFETV VHGGAVSLGHPI GCSG ARI I VITLIGVI KOKNGKYGVSVONGGGGA	320 318 316 320 316 396 399 396
MaAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT	AKLEKLERSFE, KGCTVTAGNASTI SDCAAAFVUVSGEKALCICLCUTAKIKGFADAAQPED FTTAPATATEKATSNAG AKLEKLERSFEKENGGSVTAGNASSI SDCAAIVUVSGEKAMKIC LHVTARTRGYGDAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGCTVTAGNASSI SDCAAIVUVSGEKAMKIC LHVTARTRGYGDAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGCTVTAGNASSI SDCAAIVUVSGEKAVKICLQULAKISGYGDAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGCTVTAGNASSI SDCAAIVUVSGEKATKICGLOVIGKI SGYADAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGCTVTAGNASSI SDCAAIVUVSGEKATKICGLOVIGKI SGYADAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGCTVTAGNASSI SDCAAIVIVSGEKATKICGLOVIGKI SGYADAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGVTAGNASSI SDCAAIVIVSGEKATKICGLOVIGKI SGYADAAQAPED FTTAPATATEKATSNAG AKLEKLERSFEETGVTAGNASSI SDCAAIVIVSGEKATKICGLOVIGKI SGYADAAQAPED FTTAPATATEKATKSAS * * * LEALOTD YEINEAFSVVALANOKTIGLEETVXVHGGAVSLGHELGCSG ARTIVTLLGVIKKENGKYGVAGVCNGGGGA LEASOTD YEINEAFSVVALANOKULTINPEKVNHGGAVSLGHELGCSG ARTIVTLLGVIKKESGSKYGVAAVCNGGGGA LEASOTD YEINEAFSVVALANOKULTINPEKVNHGGAVSLGHELGCSG ARTIVTLLGVIKKSGSKYGVAAVCNGGGGA	320 318 316 320 316 396 399 396 394
MaAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT	AKLIRKLIRESFK, KGCTVTAGNAST ISDCAAAFUVVSGBKAILULULU (IAKLKGFADAAQAPE) PTIAPATA IPKAISNAG AKLIRKLIRESFKENGGSVTAGNASSI SDCAAAIVVSGBKAMKLCHVTARIREYGDAAQAPEI PTIAPATA IPKAISNAG AKLIRKLIRESFKENGGVTAGNASSI SDCGAALVVVSGBKAMKLCHVTARIREYGDAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESFKENGGTVTAGNASSI SDCGAALVVVSGBKA IKLCHVVTGKTSVKDAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAALVVVSGBKA IKLCHVVTGKTSVKDAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAALVVVSGBKA IKLCHVVTGKTSVKDAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAALVVVSGBKA IKLCHVVTGKTSVKDAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAATI VVSGBKA IKLCHVVTGKTSVADAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAATI VVSGBKA IKLCHVVTGKTSVADAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAATI VVSGBKA IKLCHVVTGKTSVADAAQAPEI PTIAPATA IPKAISSAG AKLIRKLIRESPKENDGTVTAGNASSI SDCGAATI VVSGBKA IKLCHVVTGKSVGVAAQAPEI PTIAPATA IPKAISSAG IPKASOTO VYETINEAFSVVALANOKTI GLINPEKV VHGGAVSLGHPLGCSG ARI LVTLLGVI KORSKGKVGVAAV CNGGGGA IPKASOTO VYETINEAFSVVALANOKTI GLINPEKV VHGGAVSLGHPLGCSG ARI LVTLLGVI KORSKGKVGVAAV CNGGGGA IPKASOTO VYETINEAFSVVALANOKTI GLINPEKV VHGGAVSLGHPLGCSG ARI LVTLLGVI KINGKSKGVASTCNGGGGA	320 318 316 320 316 396 399 396 394
MaAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT RcAACT	AKLERKLRESFK, KGCTVTAGNAST SDCAAAFUVVSGBKALQLGLQUTAKTKGFADAAQAPED FTTAPATATEKAISNAG AKLERKLRESFKENGGSVTAGNASSISDCAAIVVVSGBKAMKLGLAVTARTREYGDAAQAPED FTTAPATATETAISNAG AKLERKLRESFKENGTVTAGNASSISDCAAIVVVSGBTAVKLGLQVLAKTSGYCDAAQAPED FTTAPATATETAISNAG AKLERKLRESFKENGTVTAGNASSISDCAAIVVVSGBKATKLGLQVLAKTSGYCDAAQAPED FTTAPATATEKAISSAG AKLERKLRESFKENGTVTAGNASSISDCAAIVVVSGBKATKLGLQVLAKTSGYCDAAQAPED FTTAPATATEKAISSAG AKLERKLRESFKENGTVTAGNASSISDCAAIVVVSGBKATKLGLQVLAKTSGYCDAAQAPED FTTAPATATEKAISSAG * * * LEALQID VETNEAFSVVALANQKTLGLKPETV VHGGAVSLGHPLGCSG ARTIVTLLGVTRERNGKYGVAGVGNGGGGA LEASQID VETNEAFSVVALANQKTLGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTRERNGKYGVAGVCNGGGGA LEASQID VETNEAFSVVALANQKTLGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTRERNGKYGVAGVCNGGGGA LEASQID VETNEAFSVVALANQKTLGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTRENGKYGVAGVCNGGGGA LEASQID VETNEAFSVVALANQKTLGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTRENGKYGVAGVCNGGGGA	320 318 316 320 316 399 396 394 394
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT AtAACT HbAACT RcAACT MdAACT	AKLRKLRESFK, KGCTVTAGNAST ISDCAAAFUVIVSGKALCICLCUTAKIKGFADAAQPBE PTIABATAIEKAISNAG AKLRKLRESFKENGGVTAGNAST ISDCAAAFUVIVSGBKAMKICLEVITARIRGYGDAAQAPBE PTIABATAIEKAISNAG AKLRKLRESFKENGGTVTAGNASSI SDCGAAUVIVSGBKAMKICLEVITARIRGYGDAAQAPBE PTIABATAIEKAISNAG AKLRKLRESFKENGGTVTAGNASSI SDCGAAUVIVSGBKAKIKUCLAVITARIRGYGDAAQAPBE PTIABATAIEKAISNAG AKLRKLRESFKENGGTVTAGNASSI SDCGAAUVIVSGBKAIKUCLAVITARIRGYGDAAQAPBE PTIABATAIEKAISNAG AKLRKLRESFKENGGTVTAGNASSI SDCGAAUVIVSGBKAIKUCLAVITARIRGYGDAAQAPBE PTIABATAIEKAISSAG AKLRKLRESFKENGGTVTAGNASSI SDCGAAUVIVSGBKAIKUCLAVIGKI SGYADAAQAPBE PTIABATAIEKAI KAIKSAS KERKLRESFKENGGTVTAGNASSI SDCGAAUVIVSGBKAIKUCLAVIGKI SGYADAAQAPBE PTIABATAIEKAI LEALOID VEINEAFSVVALANOKIIGLEVE VHGGAVSLGHELCCSG ARIIVTILLGVIRENGKYGVAGVCNGGGGA LEASQID VEINEAFSVVALANOKIIGLNPEKV NHGGAVSLGHELGCSG ARIIVTILGVIRKNGKYGVASI CNGGGGA LEASQID VEINEAFSVVALANOKIGLNGEN NHGGAVSLGHELGCSG ARIIVTILGVIRKNGKYGVASI CNGGGGA	320 318 316 320 316 399 396 394 394 400
MAAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT MdAACT MaAACT	AKLERKLRESFK.KGGTVTAGNASTISDGAAAFUVSGBKALGLGLQUIAKIKGFADAQAPBEPTTAPATATEKAISNAG AKLERKLRESFKINGGSVTAGNASGISDGAAIVUVSGBKAMKLGLHVTARIREYGDAAQAPBEPTTAPATATEKAISNAG AKLERKLRESFKENGGTVTAGNASSISDGAAIVVSGBTAVKLGLQVLAKISGYDAAQAPBEPTTAPATATETAISNAG AKLERKLRESFKENGTVTAGNASSISDGAAIVVSGBTAVKLGLQVLAKISGYDAAQAPBEPTTAPATATEKAISSAG AKLERKLRESFKENGTVTAGNASSISDGAAIVVSGBTAVKLGLQVLAKISGYDAAQAPBEPTTAPATATEKAISSAG AKLERKLRESFKENGTVTAGNASSISDGAAIVVSGBTAVKLGLQVLAKISGYDAAQAPBEPTTAPATATEKAISSAG AKLERKLRESFKENGTVTAGNASSISDGAAFUVSGBTAVKLGLQVLAKISGYDAAQAPBEPTTAPATATEKAISSAG AKLERLESFKENGTVTAGNASSISDGAAFUVSGBTAVKLGLNVGGAVSLGAPELGTVGVGVQVGVGVGGGGA LEASQIDVYEINEAFSVVALANQKILGLNPEKVVHGGAVSLGHFLGCSGARITVTLLGVIKQKNGKYGUSSVCNGGGGA LEASQIDVYEINEAFSVVALANQKILGLNPEKVVHGGAVSLGHFLGCSGARITVTLLGVIKQKSGKYGVAAVCNGGGGA LEASQIDVYEINEAFSVVALANQKILGLNPEKVVHGGAVSLGHFLGCSGARITVTLLGVIKKNGKYGVASICNGGGGA LEASQIDVYEINEAFSVVALANQKILGLNPEKVAHGGAVSLGHFLGCSGARITVTLLGVIKKNGKYGVASICNGGGGA LEASQIDPYEINEAFSVVALANQKILGLNPEKVAHGGAVSLGHFLGCSGARITVTLLGVIKKNGKYGVGGICNGGGGA LEASQIDPYEINEAFSVVALANQKILGLNPEKVAHGGAVSLGHFLGCSGARITVTLLGVIKKNGKYGVGGICNGGGGA LEASQIDPYEINEAFSVVALANQKILGLNPEKVAHGGAVSLGHFLGCSGARITVTLLGVIKKNGKYGVGGICNGGGGA	320 318 316 320 316 399 396 394 394 400 398
MAAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT MAAACT MAAACT EbAACT	AKLRKLRESFF.KGGTVTAGNASI ISDGAAAFVIVSGFAICULGUUIAKIKGFADAAQAPBI PTIAPAIAIPKAISNAG AKLRKLRESFFKINGGSVTAGNASI ISDGAAIVIVSGBKAMKLGLHVIARIRGYGDAAQAPBI PTIAPAIAIPKAISNAG AKLRKLRESFKENGGTVTAGNASI ISDGAAIVIVSGBKAMKLGLHVIARIRGYGDAAQAPBI PTIAPAIAIPKAISNAG AKLRKLRESFKENGGTVTAGNASI ISDGAAIVIVSGBKAKLGLQULAKISGYGDAAQAPBI PTIAPAIAIPKAISSAG AKLRKLRESFKENGGTVTAGNASI ISDGAAIVIVSGBKAIKLGLNVIGKISGYADAAQAPBI PTIAPAIAIPKAISSAG AKLRKLRESFKENGGTVTAGNASI ISDGAAIVIVSGBKAIKLGLNVIGKISGYADAAQAPBI PTIAPAIAIPKAISSAG AKLRKLRESFKENGGTVTAGNASI ISDGAAIVIVSGBKAIKLGLNVIGKISGYADAAQAPBI PTIAPAIAIPKAIKSAS ***********************************	320 318 316 320 316 399 396 394 394 400 398 394
MAAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT McAACT McAACT EhAACT EhAACT	AKLRKLRESFK.KGGTVTAGNASTISDCAAAFUVSGBKALCICLCUTAKTKGFQDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGGSVTAGNASSISDCAAIVVSGBKAMKLCLHVTARIRGYGDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGGVTAGNASSISDCAAIVVSGBTAVKLCLUVLAKTSGVDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGGVTAGNASSISDCAAIVVSGBTAVKLCLUVLKSGVDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGGVTAGNASSISDCAAIVVSGBTAVKLCLUVLKSGVDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAIVVSGBTAVKLCLUVLGKISGVDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAIVVSGBTAVKLCLUVLGKISGVDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAIVVSGBTAVKLCLUVSGBKATKLCLUVTGKISGVADAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAIVVSGBTAVKLCLUVSGBKATKLCLUVTGKISGVADAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAIVVSGBTAVKLCLUVTGKISGVADAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAIVVSGBTAVKLCLUVTGKISGVADAAQAPBEPTTABATATEKAISNAG LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKORSKYGVAAVCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVASTCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVASTCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVAGTCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVAGTCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVAGTCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVAGTCNGGGGA LEASOTDVYEINEAFSVVALANQKTLGLNPEKVNHGGAVSLGHPLGCSGARITVTLLGVTKKNGKYGVAGTCNGGGGA	320 318 316 320 316 399 396 394 394 400 398 396
MAAACT EhAACT PnAACT NbAACT GbAACT NnAACT AtAACT HbAACT McAACT McAACT EhAACT PnAACT	AKLRKLRESFF, KGCTVTAGNAST ISDCAAAFUVIVSGBKALCICLCUTAKIKGPADAAQPED FTIAPATATEKAISNAG AKLRKLRESFFKINGGVTAGNAST ISDCAAIVIVSGBKAMKIC LHVTARTREYGDAAQAPET FTIAPATATEKAISNAG AKLRKLRESFFKENGGVTAGNAST ISDCAAIVIVSGBKAMKIC LHVTARTREYGDAAQAPET FTIAPATATEKAISNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAAIVIVSGBKAMKIC LHVTARTREYGDAAQAPET FTIAPATATEKAISNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAAIVIVSGBKATKICLQULAKISGVCDAAQAPET FTIAPATATEKAISNAG AKLRKLRESFKENGGTVTAGNASS ISDCAAIVIVSGBKATKICLQULAKISGVCDAAQAPET FTIAPATATEKAISNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAAIVIVSGBKATKICGLOVIGKISGVADAAQAPET FTIAPATATEKAISNAG AKLRKLRESFKENGGTVTAGNASS ISDCAAIVIVSGBKATKICGLOVIGKISGVADAAQAPET FTIAPATATEKAISNAG AKLRKLRESFKENGGTVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRGKNGKSGTSSVCNGGGGA LEASOTDY YETINEAFS VVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRGKSGKYGVAAVCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKVNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKNNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKNNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKIGLENPEKNNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKILGLENPEKNNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKILGLENPEKNNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVAGTCNGGGGA LEASOTDY YETINEAFSVVALANGKILGLENPEKNNHGGAVSLGHPLGCSG ARTIVTLLGVIRKNGKYGVGACCNGGGGA	320 318 316 320 316 399 396 394 394 400 398 396 400
MAAACT EhAACT PhAACT NbAACT NbAACT AtAACT HbAACT McAACT McAACT McAACT PhAACT PhAACT NbAACT	AKLRKLRESFK.KGGTVTAGNASTISDGAAAFUVSGBKALCIGLGUTAKIKGFADAAQPBEPTHABATATEKAISNAG AKLRKLRESFKENGGSVTAGNASSISDGAAHUVSGBKAMKLGLHVIARIRGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGVTAGNASSISDGAAHUVSGBKAMKLGLHVIARIRGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAHUVSGBTAVKIGLGUTAKISGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAHUVSGBTAVKIGLGUTAKISGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAHUVSGBTAVKIGLGUTAKISGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAHUVSGBTAVKIGLGUTAKISGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAHUVSGBTAVKIGLGUTAKISGYGDAAQAPBEPTHABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAHUVSGBTAVKIGLGUTAKISGYGVAAQAPBEPTHABATATEKAISNAG LEASQIDTYEINEAFSVVALANOKTIGLKPETV VHIGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVGATCNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVGATCNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVGATCNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVGATCNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVGATCNGGGGA LEASQIDTYEINEAFSVVALANOKTIGLNPEKNNHGGAVSLGHELGCSGARITVTLLGVIKKNKKYGVGATCNGGGGA	320 318 316 320 316 399 396 394 394 400 398 396 400 396
MAAACT EhAACT NhAACT NhAACT AtAACT HbAACT MAAACT MAAACT EhAACT PhAACT NbAACT	AKLRKLRESFF, KGCTVTAGNAST ISDCAAAFUVIVSGEKALCICLCUTAKIKGPADAAQPED FTTAPATATEKATSNAG AKLRKLRESFFKINGGSVTAGNAST ISDCAAAFUVIVSGEKAMKLCHVTARTRGYGDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFFKENGGSVTAGNAST ISDCAAAFUVIVSGEKAMKLCHVTARTRGYGDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAATUVIVSGEKAMKLCHVTARTRGYGDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAATUVIVSGEKATKLCHVTARTRGYGDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAATUVIVSGEKATKLCHVTAKTSGYCDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFFKENGGTVTAGNASS ISDCAATUVIVSGEKATKLCHVTAKTSGYCDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFKENGGTVTAGNASS ISDCAATUVIVSGEKATKLCHVTAKTSGYCDAAQAPET FTTAPATATEKATSNAG AKLRKLRESFKENGGTVTAGNASS ISDCAATUVIVSGEKATKLCHVTAKTSGYCDAAQAPET FTTAPATATEKATSNAG LEASOTD YETNEAAS VVALANOKTIGLKPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGLSSVCNGGGGA LEASOTD YETNEAAS VVALANOKTIGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGASTCNGGGGA LEASOTD YETNEAAS VVALANOKTIGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGASTCNGGGGA LEASOTD YETNEAAS VVALANOKTIGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGASTCNGGGGA LEASOTD YETNEAAS VVALANOKTIGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGASTCNGGGGA LEASOTD YETNEAAFSVVALANOKTIGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGAAGTCNGGGGA LEASOTD YETNEAAFSVVALANOKTIGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGAAGTCNGGGGA LEASOTD YETNEAAFSVVALANOKTLGLNPEKV VHGGAVSLGHPLGCSG ARTIVTLLGVTKOKNGKYGAAGTCNGGGGA LEASOTD YETNEAAFSVVALANOKTLGLNPEKV NHGGAVSLGHPLGCSG ARTIVTLLGVTROKNGKYGAAGTCNGGGGA LEASOTD YETNEAAFSVVALANOKTLGLNPEKV NHGGAVSLGHPLGCSG ARTIVTLLGVTROKNGKYGAAGTCNGGGGA LEASOTD YETNEAFSVVALANOKTLGLNPEKV NHGGAVSLGHPLGCSG ARTIVTLLGVTROKNGKYGAAGTCNGGGGA LEASOTD YETNEAFSVVALANOKTLGLNPEKV NHGGAVSLGHPLGCSG ARTIVTLLGVTROKNGKYGAGTCNGGGGA LEASOTD YETNEAFSVVALANOKTLGLNPEKV NHGGAVSLGHPLGCSG ARTIVTLLGVTROKNGKYGAGGGGA LEASOTD YETNEAFSVVALANOKTLGLNPEKV NHGGAVSLGHPLGCSG ARTIVTLLGVTROKNGKYGAGGGA	320 318 316 320 316 399 396 394 400 398 396 400 398
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT HbAACT MAAACT HbAACT EhAACT NbAACT	AKLRKLRESFK, KGCTVTAGNAST, ISDCAAAFVUVSGFKALCICLGUTAKIKGFADAAQPBE PTIAPATAIPKAISNAG AKLRKLRESFKRUGGVTAGNAST ISDCAAAFVUVSGFKAMKLCLHVIARIRGYGDAAQAPBI PTIAPATAIPKAISNAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVUVSGFAVKICLGUTAKIKGFADAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVUVSGFAVKICLGUTAKISGYCDAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVVVSGFAVKICLGUTAKISGYCDAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVVVSGFAVKICLGUTAKISGYCDAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVVVSGFAVKICLGUTAKISGYCDAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVVVSGFAVKICLGUTAKISGYCDAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVVVSGFAVKICLGUTAKISGYCDAAQPBI PTIAPATAIPKAISAG AKLRKLRESFKETGCTVTAGNAST ISDCAAIVVSGFAVINGGAVSIGHPIGCSG ARIIVTLLGVIRENGKYGVAGVCNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVGAICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVGAICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVGAICNGGGGA LEASQID YEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVGAICNGGGGA LEASQID FYEINEAFSVVALANOKIIGLNPEKVVHGGAVSIGHPIGCSG ARIIVTLLGVIRKNGKYGVGAICNGGGGA	320 318 316 320 316 399 396 394 400 398 396 400 396
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT AtAACT RcAACT HbAACT EhAACT PhAACT NbAACT	AKLRKLRESFK.KGGTVTAGNAST ISDCAAAFUVSGBKALCICLGUTAKTKGFQDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGGSVTAGNASSISDCAAJUVVSGBKANKLCILVITARIREYGDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGGTVTAGNASSISDCAAJUVVSGBAYKLCILQUTAKTREYGDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVITARIREYGDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVITARIREYGDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVISGENAKTSGVCDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVISGENSKISGVAADAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVISGENAKTSGVCDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVISGENAKTSGVCDAAQAPBEPTTABATATEKAISNAG AKLRKLRESFKENGTVTAGNASSISDCAAJUVVSGBAKKLCILVISGKAKKISGVAADAQAPBEPTTABATATEKAISNAG LEASOIDVYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDVYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELCCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELGCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELGCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHELGCSG ARIIVTILGVIRKNGKYGVASTCNGGGGA LEASOIDFYEINEAFAVVALANQKIIGLNPEKVNHGGAVSLGHELGCSG ARIIVTILGVIRKNGKYGVAGYCNGGGGA	320 318 316 320 316 399 396 394 400 398 394 400 398
MAAACT EhAACT PhAACT NbAACT GbAACT HbAACT HbAACT MGAACT MGAACT PhAACT NbAACT ShAACT ShAACT	AKLRKLRESFF.KGCTVTAGNASTI SDCAAAFVUVSGEKALCICLCUTAKIKGPADAAQPED FTIAPATATEKAISNAG AKLRKLRESFFKOGSVTAGNASTI SDCAAAFVUVSGEKAMKLCILVTARIRGYGDAAQPED FTIAPATATEKAISNAG AKLRKLRESFFKENGGVTAGNASTI SDCAAIVUVSGEKAMKLCILVTARIRGYGDAAQPED FTIAPATATEKAISNAG AKLRKLRESFFETGTVTAGNASTI SDCAAIVUVSGEKAKLCICUTAKIKGPADAAQPED FTIAPATATEKAISNAG AKLRKLRESFFETGTVTAGNASTI SDCAAIVUVSGEKAKLCICUTAKISGYCDAAQPED FTIAPATATEKAISNAG AKLRKLRESFFETGTVTAGNASTI SDCAAIVUVSGEKAIKLCINVIGKI SGYADAAQPED FTIAPATATEKAISNAG AKLRKLRESFFETGTVTAGNASTI SDCAAIVVVSGEKAIKLCINVIGKI SGYADAAQPED FTIAPATATEKAISNAG AKLRKLRESFFETGTVTAGNASTI SDCAAIVVSGEKAIKLCINVIGKI SGYADAAQAPED FTIAPATATEKAISNAG AKLRKLRESFFETGTVTAGNASTI SDCAAIVVSGEKAIKLCINVIGKI SGYADAAQAPED FTIAPATATEKAISNAG LEASQID YEINEAFSVVALANQKI I GLNEEKV VHGGAVSLGHEI GCSG ARI I VITLLGVI RERNGKYGVAGVCNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV VHGGAVSLGHEI GCSG ARI I VITLLGVI RERNGKYGVAG LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVAG LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVAG LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVAGI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANQKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA LEASQID YEINEAFSVVALANGKI I GLNEEKV NHGGAVSLGHEI GCSG ARI I VITLLGVI RHNGKYGVGAI CNGGGGA	320 318 316 320 316 399 396 394 400 398 396 400 396
MAAACT EhAACT NDAACT GDAACT NDAACT ALAACT HDAACT ALAACT HDAACT EhAACT NDAACT GDAACT NDAACT	AKLRKLRESFF. KGCTVTAGNAST ISDCAAAFUVSGBKALCICLGUTAKTKGFQDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFKUGGSVTAGNASG ISDCAAAFUVSGBKAMKLCLHVTARTREYGDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKAMKLCLHVTARTREYGDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKAKLCLUVTAKTREYGDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG AKLRKLRESFFETGCTVTAGNASG ISDCAAAFUVVSGBKATKLCCHVTGKTSGVCDAAQAPBE PTHPATATEKATSNAG LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNYHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVASTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNYHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVASTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNYHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVASTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRHKNGKYGVAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRKNGKYGAAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRKNGKYGAAGTCNGGGGA LEAASOTDYEINEAFSVVALANQKTIGLNPEKVNHGGAVSLGHEFLGCSG ARTIVTLLGVTRKNGKYGAAGTCNGGGGA	320 318 316 320 316 399 396 394 394 400 398 396 400 396 404 416
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT AtAACT MdAACT MdAACT NbAACT NbAACT SchaACT NhAACT NhAACT	AKLRKLRESFF.KGGTVTAGNASTISDGAAAFUVIVSGBKALCIGLGUTAKIKGPADAAQPED FTIAPATATEKATSNAG AKLRKLRESFFKENGGVTAGNASTISDGAAHUVIVSGBKAMKIG LHVITAKIRGYGDAAQPED FTIAPATATEKATSNAG AKLRKLRESFFETGGTVTAGNASTISDGAAHUVIVSGBKAMKIG LHVITAKIRGYGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASTISDGAAHUVIVSGBKAMKIG LHVITAKIRGYGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASSISDGAAHUVIVSGBKATKIGLQULAKISGVGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASSISDGAAHUVIVSGBKATKIGLQULAKISGVGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASSISDGAAHUVIVSGBKATKIGLQULAKISGVGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASSISDGAAHUVIVSGBKATKIGLQULAKISGVGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASSISDGAAHUVIVSGBKATKIGLQULAKISGVGDAAQPED FTIAPATATETATISNAG AKLRKLRESFFETGGTVTAGNASSISDGAAHUVIVSGBKATKIGLQULAKISGVGDAAQPED FTIAPATATETATISNAG LEASOTDVETINEAFSVVALANOKTIGLKPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRKNGKYGVAGCONGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVAGTCNGGGGA LEASOTDVETINEAFSVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRHNGKYGVGATCNGGGGA LEASOTDVETINEAFAVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRKNGKYGVGVCNGGGGA LEASOTDVETINEAFAVVALANOKTIGLNPEKVNHGGAVSLGHELGCSG ARTIVTILGVIRKNGKYGVGVCNGGGGA	320 318 316 320 316 399 396 394 400 398 396 400 396 400 404 416 404
MAAACT EhAACT NbAACT NbAACT MAACT HbAACT HbAACT HbAACT MAAACT NbAACT NbAACT NbAACT NhAACT NhAACT	AKLRKLRESFK.KGGTVTAGNASI ISDGAAAFVIVSGBKALCIGLGUITAKIKGPADAAQPBETTIABAIA IBKAISNAG AKLRKLRESFKRGGSVTAGNASI SDGAAIVVIVSGBKAMKICILVIARIRGYGDAAQAPBIFTIABAIA IBKAISNAG AKLRKLRESFKERGGTVTAGNASI SDGAAIVVIVSGBKAMKICILVIARIRGYGDAAQAPBIFTIABAIA IBKAISNAG AKLRKLRESFKERGTVTAGNASI SDGAAIVVIVSGBKAKILGILVIARIRGYGDAAQAPBIFTIABAIA IBKAISNAG AKLRKLRESFKERGTVTAGNASI SDGAAIVVIVSGBKAKILGILVIARIRGYGDAAQAPBIFTIABAIA IBKAISNAG AKLRKLRESFKERGTVTAGNASI SDGAAIVVIVSGBKAIKICGIVIGKI SGVADAAQAPBIFTIABAIA IBKAISNAG AKLRKLRESFKERGTVTAGNASI SDGAAIVVIVSGBKAIKICGIVIGKI SGVADAAQAPBIFTIABAIA IBKAISSAG AKLRKLRESFKERGTVTAGNASI SDGAAIVVIVSGBKAIKICGIVIGKI SGVADAAQAPBIFTIABAIA IBKAISSAG AKLRKLRESFKERGTVTAGNASI SDGAAIVVIVSGBKAIKICGIVIGKI SGVADAAQAPBIFTIABAIA IBKAIKSAS * * * * * * * * * LEALOIDVIEINEAFSVVALANOKIIGINEKV VHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGVCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGVCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGUCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGUCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGUCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGUCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGUCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVAGUCNGGGGA LEASOIDVIEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGAICNGGGGA LEASOIDFYEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGAICNGGGGA LEASOIDFYEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGAICNGGGGA LEASOIDFYEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGAICNGGGGA LEASOIDFYEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGAICNGGGGA LEASOIDFYEINEAFSVVALANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGCONGGGGA LEASOIDFYEINEAFSVVSIGANOKIIGINEKV NHGGAVSIGHEIGCSG ARIIVTILGVIRKNGKYGVGCONGGGGA LEASOIDFYEINEAFSVVSIGANOKIGGNEKT	320 318 316 320 316 399 396 394 400 398 394 400 398 396 400 404 416 404
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT AtAACT RCAACT MdAACT RCAACT NbAACT SbAACT NhAACT NhAACT NhAACT	AKLRKLRESFE, KGCTVTAGNAST ISDCAAAFVUVSGEKALCICLCUTAKIKGPADAAQPED FTTAPATAIEKAISNAG AKLRKLRESFEKINGGVTAGNAST ISDCAAAFVUVSGEKAMKIC LHVIARIRGYGDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFETGTUTAGNASS ISDCAAIVIVSGEKAMKIC LHVIARIRGYGDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFETGTUTAGNASS ISDCAAIVIVSGEKAKILGLAVIARIGYGDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFETGTUTAGNASS ISDCAAIVIVSGEKAIKICGUVIGKISGVADAAQPED FTTAPATAIEKAISNAG AKLRKLRESFETGTUTAGNASS ISDCAAIVIVSGEKAIKICGUVIGKISGVADAAQPED FTTAPATAIEKAISNAG AKLRKLRESFETGTUTAGNASS ISDCAAIVIVSGEKAIKICGUVIGKISGVADAAQPED FTTAPATAIEKAISNAG AKLRKLRESFETGTUTAGNASS ISDCAAIVIVSGEKAIKICGUVIGKISGVADAAQPED FTTAPATAIEKAISNAG AKLRKLRESFKEKDCTVTAGNASS ISDCAAIVIVSGEKAIKICGUVIGKISGVADAAQPED FTTAPATAIEKAIKSAS * * * * * * * LEALQID YEINEAFSVVALANOKULGINPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGVASUCNGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGAAICNGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGKINGGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGKINGKINGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKINGKIGVGGGA LEASQID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKIGAACONGGGGA LEASGID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKIGAACONGGGGA LEASGID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKIGAACONGGGGA LEASGID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKIGAACONGGGGA LEASGID YEINEAFSVVALANOKULGUNPEKVSVHGGAVSLGHELGCSG ARIIVILLGVIRGNGKIGAACONGGGGA	320 318 316 320 399 396 394 400 398 394 400 398 400 396 400 404 411
MAAACT EhAACT PhAACT NbAACT MbAACT HbAACT HbAACT HbAACT HbAACT MdAACT MAAACT NbAACT SbAACT NDAACT AtAACT AtAACT RcAACT	AKLRKLRESFF.KGCTVTAGNASTISDCAAAFVUVSGBKALCLGLQUIAKIKGPADAAQPBETTIABAIAIEKAISNAG AKLRKLRESFFKUGGVTAGNASTISDCAAAIVUVSGBKAMKLGLHVIARIRGYGDAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEGTVTAGNASTISDCAAAIVUVSGBKAMKLGLHVIARIRGYGDAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEGTVTAGNASTISDCAAAIVUVSGBKAKLGLGULAKISGYGDAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVUVSGBKAKLGLGULAKISGYGDAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVUVSGBKAIKLGINVIGKISGYADAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVUVSGBKAIKLGINVIGKISGYADAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVVSGBKAIKLGINVIGKISGYADAAQAPBIFTIABAIAIEKAISNAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVVSGBKAIKLGINVIGKISGYADAAQAPBIFTIABAIAIEKAISAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVVSGBKAIKLGINVIGKISGYADAAQAPBIFTIABAIAIEKAISAG AKLRKLRESFFKEKGTVTAGNASTISDCAAAIVVSGBKAIKLGINVIGKISGYADAAQAPBIFTIABAIAIEKAISAG LEALQIDVEINEAFSVVALANQKIIGLREKVVHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVAGVCNGGGGA LEASQIDVEINEAFSVVALANQKIIGLNEKKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVAGTCNGGGGA LEASQIDVEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGICNGGGGA LEASQIDVEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGICNGGGGA LEASQIDVEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGICNGGGGA LEASQIDVEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGICNGGGGA LEASQIDFYEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGICNGGGGA LEASQIDFYEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGICNGGGGA LEASQIDFYEINEAFSVVALANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGGCNGGGGA LEASQIDFYEINEAFSVVSLANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGGCNGGGGA LEASQIDFYEINEAFSVVSLANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGGCNGGGGA LEASQIDFYEINEAFSVVSLANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGVGCNGGGGA LEASOIDFYEINEAFSVVSLANQKIIGLNEEKINHGGAVSLGHEIGCSG ARIIVTLLGVIRKNGKYGAAGYCNGGGGA LEASOIDFYEINEAFSVVSLANQKIGGNEGSGA SAVUEBIN	320 318 316 320 399 396 394 400 398 396 400 398 396 400 398 396 400 400 411 411
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT AtAACT HbAACT RCAACT NbAACT GbAACT AtAACT AtAACT AtAACT AtAACT AtAACT	AKLRKLRESFE, KGCTVTAGNAST ISDCAAAFVUVSGEKALCICLCUTAKIKGPADAAQPED FTTAPATAIEKAISNAG AKLRKLRESFEKINGGVTAGNAST ISDCAAAFVUVSGEKAMKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAMKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAMKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAIKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAIKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAIKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAIKLCILVITARIRGYCDAAQPED FTTAPATAIEKAISNAG AKLRKLRESFERTGTVTAGNASS ISDCAAIVIVSGEKAIKLCILVITAKISGYCDAAQPED FTTAPATAIEKAIKSAS * * * * * * * LEAIQIDYEINEAFSVVALANOKULGINPEKV VHGGAVSIGHELGCSG ARIIVTLLGVIKOKNGKYGISSVCNGGGGA LEASQIDYEINEAFSVVALANOKULGINPEKV VHGGAVSIGHELGCSG ARIIVTLLGVIKOKNGKYGVASICNGGGGA LEASQIDYEINEAFSVVALANOKULGINPEKV NHGGAVSIGHELGCSG ARIIVTLLGVIKOKNGKYGASICNGGGGA LEASQIDYEINEAFSVVALANOKULGINPEKV NHGGAVSIGHELGCSG ARIIVTLLGVIKOKNGKYGAGICNGGGGA LEASQIDPYEINEAFSVVALANOKULGINPEKV NHGGAVSIGHELGCSG ARIIVTLLGVIKOKNGKYGVAGICNGGGGA LEASQIDPYEINEAFSVVALANOKULGINPEKV NHGGAVSIGHELGCSG ARIIVTLLGVIKOKNGKYGVAGICNGGGGA LEASOIDPYEINEAFSVVGESL SALVIDEN SALVIDEN SALVIDEN SALVIDEN SALVIDEN	320 318 316 320 320 399 396 394 400 396 394 400 396 400 396 400 404 411 411 408
MAAACT EhAACT PhAACT NbAACT GbAACT NhAACT AtAACT RcAACT MdAACT NbAACT NbAACT SbAACT NhAACT AtAACT RcAACT MdAACT RcAACT MdAACT	AKLEKLERSFE, KGCTVTAGNAST ISDCAAAFVUVSGEKALCICLCUTAKIKGPADAAQPED FTIPATATEKATSNAG AKLEKLERSFEKINGGVTAGNAST ISDCAAAIVUVSGEKAMKIC LEVTARTREYGDAAQAPEI FTIPATATEKATSNAG AKLEKLERSFEETG TVTAGNAST ISDCAAIVUVSGEKAMKIC LEVTARTREYGDAAQAPEI FTIPATATETATEKATSNAG AKLEKLERSFEETG TVTAGNAST ISDCAAIVUVSGEKAMKIC LEVTARTREYGDAAQAPEI FTIPATATETATEKATSNAG AKLEKLERSFEETG TVTAGNAST ISDCAAIVUVSGEKAKIC LEVTARTREYGDAAQAPEI FTIPATATEKATSNAG AKLEKLERSFEETG TVTAGNAST ISDCAAIVUVSGEKAIKIC LEVTAKIG VOAAQAPEI FTIPATATEKATSNAG AKLEKLERSFEETG TVTAGNAST ISDCAAIVVSGEKAIKIC LEVTAKIG VOAAQAPEI FTIPATATEKATSNAG LEASOTD YETNEAFSVALANOKII GLEVENT VHGGAVSLGHEI GCSG ARTIVTLLGVI KERSKYG VAACONGGGGA LEASOTD YETNEAFSVALANOKII GLEVEKK VHGGAVSLGHEI GCSG ARTIVTLLGVI KORSKYG VAACONGGGGA LEASOTD YETNEAFSVALANOKII GLEVEKK VHGGAVSLGHEI GCSG ARTIVTLLGVI KORNGKYG VAGT CNGGGGA LEASOTD YETNEAFSVALANOKII GLEVEKK VHGGAVSLGHEI GCSG ARTIVTLLGVI KONGKYG VGGT CNGGGGA LEASOTD YETNEAFAVVALANOKII GLEVEKK VHGGAVSLGHEI GCSG ARTIVTLLGVI KONGKYG VGGT CNGGGGA LEASOTD YETNEAFAVVALANOKII GLEVEKK VHGGAVSLGHEI GCSG ARTIVTLLGVI KONGKYG VGGT CNGGGGA LEASOTD YETNEAFAVVGESI SALVIELM STINEFS	320 318 316 320 320 399 396 399 400 398 396 400 398 396 400 398 396 400 404 416 404 411 411
MAAACT EhAACT PhAACT NbAACT MbAACT HbAACT HbAACT MdAACT MdAACT MbAACT MbAACT MbAACT HbAACT MdAACT MdAACT MdAACT MaAACT	AKLRKLRESFK.KGGTVTAGNASTISDGAAAFUVSGBKALCIGLGUTAKIKGPADAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKMGGSVTAGNASSISDGAAIVIVSGBKAMKICLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAMKICLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAKILGLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAIKICGLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAIKICGLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAIKICGLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAIKICGLWITARIRGYGDAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAIKICGLWITAGTSGVAAAQAPBIPTTABATAIPKAISNAG AKLRKLRESFKENGGTVTAGNASSISDGAAIVIVSGBKAIKICGLWITAGTSGVAAAQAPBIPTTABATAIPKAISNAG LEASQIDTYEINEAFSVVALANOKIIGLKPETV VHGGAVSIGHPIGCSG ARIIVTILGVIRKKNKKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKKNKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKKNKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKKNKYGVASICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKKNKYGVGAICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKKNKYGVGAICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDTYEINEAFSVVALANOKIIGLNPEKVAHGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDFYEINEAFSVVALANOKIIGLNPEKVAYGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDFYEINEAFSVVALANOKIIGLNPEKVAYGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDFYEINEAFSVVALANOKIIGLNPEKVAYGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDFYEINEAFSVVALANOKIIGLNPEKVAYGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGVGAICNGGGGA LEASQIDFYEINEAFSVVALANOKIIGLNPEKVAYGGAVSIGHPIGCSG ARIIVTILGVIRKNKYGAXGYGANGGGGA SALVIELMSVXKVGRSI SALVIELMSVKKVGRSI	320 318 316 320 320 399 396 399 396 394 400 398 396 400 398 396 400 404 416 404 411 408 411
MAAACT EhAACT PNAACT NbAACT AtAACT MAAACT MAAACT RCAACT MAAACT NbAACT NbAACT SCAACT NDAACT NDAACT HDAACT HDAACT HDAACT HDAACT HDAACT HDAACT	AKLRKLRESFF.KGGTVTAGNASTISDGAAAFUVIVSGBKALCIGLGUTAKIKGPADAQ2PEIPTTAPATAIPKATAIPKATSNAG AKLRKLRESFFKUGGVTAGNASSISDGAAHUVIVSGBKAMKICILGUTAKIKGPADAQ2PEIPTTAPATAIPKATSNAG AKLRKLRESFFKUGGTVTAGNASSISDGAAHUVIVSGBKAMKICILGUTAKIKGPADAQ2PEIPTTAPATAIPKATSNAG AKLRKLRESFFKUGTVTAGNASSISDGAAHUVIVSGBKAKIKICILGUTAKIKGPADAQ2PEIPTTAPATAIPKATSNAG AKLRKLRESFFKUGTVTAGNASSISDGAAHUVIVSGBKAIKICILQUTAKIKGPADAQ2PEIPTTAPATAIPKATSNAG AKLRKLRESFFKUGTVTAGNASSISDGAAHUVIVSGBKAIKICILQUTAKISGVCDAAQ2PEIPTTAPATAIPKAISAG AKLRKLRESFFKUGTVTAGNASSISDGAAHUVIVSGBKAIKICILQUTAKISGVCDAAQ2PEIPTTAPATAIPKAISAG AKLRKLRESFFKUGTVTAGNASSISDGAAHUVIVSGBKAIKICILQUTAKISGVCDAAQ2PEIPTTAPATAIPKAIKSAS * * * * * * LEALOTDIVEINEAFSVVALANOKUIGILPEKVNVHGGAVSIGHPIGCSGARIIVTILGVIRERNGKYGVAGCONGGGGA LEASOTDIVEINEAFSVVALANOKUIGILPEKVNVHGGAVSIGHPIGCSGARIIVTILGVIRERNGKYGVAGCONGGGGA LEASOTDIVEINEAFSVVALANOKUIGILPEKVNVHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVAGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGILPEKVNVHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVAGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGILPEKVNVHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVAGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGILPEKVNVHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVAGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGILPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVAGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRHKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKVNHGGAVSIGHPIGCSGARIIVTILGVIRKNGKYGVGGICNGGGGA LEASOTDIVEINEAFSVVALANOKUIGINPEKNNHGGAVSIGHPIGCSGARIIVTILGVIRKNGKYGANGKYGAAGYCNGGGGA SALVELNINSKKYGF	320 318 316 320 320 399 396 400 398 396 400 398 396 400 398 396 400 411 411 411 411 411 415 413
MAAACT EhAACT PhAACT NbAACT MbAACT HbAACT HbAACT MdAACT MdAACT MdAACT NbAACT MbAACT MbAACT MbAACT MbAACT MdAACT MdAACT MdAACT HbAACT	AKLRKLRESFK.KGGTVTAGNASTISDGAAAFVIVSGFKALCICLGUTAKIKGPADAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGGVTAGNASSISDGAAIVIVSGFKAMKLGLHVIARIRGYGDAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAMKLGLHVIARIRGYGDAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAKLGLUVIARIRGYGDAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAKLGLUVIGKISGYADAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAKLGLUVIGKISGYADAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAKLGLUVISGFKAKLGUVIGKISGYADAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAKLGUVIGKISGYADAAQAPEDTTAPATAIPKAISNAG AKLRKLRESFKENGTVTAGNASSISDGAAIVIVSGFKAKLGUVIGGTGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNYHGGAVSLGHFLGCSG ARIIVTLLGVIRENGKYGVAGVCNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVASICNGGGGA LEASQIDYPEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVGATCNGGGGA LEASQIDPYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVGATCNGGGGA LEASQIDPYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVGATCNGGGGA LEASQIDPYEINEAFSVVALANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVGATCNGGGGA LEASQIDPYEINEAFSVVSLANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVGATCNGGGGA LEASQIDPYEINEAFSVVSLANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGVGACNGGGGA LEASQIDPYEINEAFSVVSLANQKIIGLNPEKVNHGGAVSLGHFLGCSG ARIIVTLLGVIRKNGKYGACNGGGGA LEASQIDPYEINEAFSVVSLANQKIIGLNPEKS SALVUELMYPLMYSLANQKIGGNSS SALVUELMYPLMYSLANQKIGSS SALVUELMYPLMYSLANQKIGSS SALVUELMYPLMYSLANQKIGSS SALVUELMYPLMUENG	320 318 316 320 316 399 396 394 400 398 394 400 398 396 400 396 400 404 416 404 411 408 411 408 415 413 408
MAAACT EhAACT PNAACT NbAACT GbAACT NDAACT AtAACT RCAACT MdAACT RCAACT NDAACT SCAACT NDAACT HDAACT HDAACT HDAACT HDAACT HDAACT HDAACT HDAACT HDAACT HDAACT HDAACT	AKLRKLRESFE, KGCTVTAGNAST ISDCAAAFUVIVSGEKALCICLGUTIKKKGPADAQ2PEIPTTPALAIPATAIPKATSNAG AKLRKLRESFEKUGGVTAGNAST ISDCAAAFUVIVSGEKAMKIC LHVITAR IRGYGDAAQ2PEIPTTPALAIPTAIPKAIPKAIPKA AKLRKLRESFEKUGTVTAGNAST ISDCAAAFUVIVSGEKAMKIC LHVITAR IRGYGDAAQ2PEIPTTPALAIPTAIPKAIPKAIPKA AKLRKLRESFEKUGTVTAGNAST ISDCAAFUVIVSGEKAKKICGUNIGKI SGYADAAQ2PEIPTTPALAIPKAIPKAIPKA AKLRKLRESFEKUGTVTAGNAST ISDCAAFUVIVSGEKATKICGUNIGKI SGYADAAQ2PEIPTTPALAIPKAIPKAIPKA AKLRKLRESFEKUGTVTAGNAST ISDCAAFUVIVSGEKATKICGUNIGKI SGYADAAQ2PEIPTTPALAIPKAIPKAIPKA AKLRKLRESFEKUGTVTAGNAST ISDCAAFUVIVSGEKATKICGUNIGKI SGYADAAQ2PEIPTTPALAIPKAIPKAIPKAIPKA AKLRKLRESFKEKUGTVTAGNAST ISDCAAFUVIVSGEKATKICGUNIGKI SGYADAAQ2PEIPTTPALAIPKAIPKAIPKAIPKAIPKAIPKAIPKAIPKAIPKAIPK	3200 3188 316 320 316 399 396 399 396 394 394 400 398 396 400 398 396 400 404 416 404 411 408 415 413 408

**Figure 2.** Amino acid sequence multiple alignments of GbAACT with other plant AACTs from *Nelumbo nucifera* (NnAACT, Accestion No. XP\_010267976.1), *Amborella trichopoda* (AtAACT, Accestion No. XP\_011628272.1), *Hevea brasiliensis* (HbAACT, Accestion No. AFJ74323.1), *Malus domestica* (MdAACT, Accestion No. XP\_008365553.1), *Morus alba* (MaAACT, Accestion No. ALD84318.1), *Euphorbia helioscopia* (EhAACT, Accestion No. ALC76524.1), *Panax notoginseng* (PnAACT, Accestion No. AIK21787.1), *Nicotiana benthamiana* (NbAACT, Accestion No. BAR94039.1). Dark blue: identity = 100%; red: 75% ≤ identity < 100%; light blue: 50% ≤ identity < 75%. The active sites of amino acid residues are marked with asterisk, Conserved domain (NVHGGAVSIGHPIGCSG) is marked with red frame, thiolase II active sites (GVAGVCNGGGGASA) are underlined.

GbMVK	MVHKIABVTGRAEGKLIIICGBHAVVHGSTALAASLGLYTQVRIQNSISTAAEPDGHUTIQUTDMNMHETWPIKRIEAAFQ	80
PmMVK	MEVKARAPGKIILLSGEHAVVHGSTAVASSIDLYTYVSLRFPTPSDNDDAURLQUKDVGLEESWPIGRIKEALS	73
PbMVK	MEVKARAPGKIILLSGEHAVVHGSTAVASSIDLYTYVSLRFPTPSDNDDAURLQUKDIGLESSWPIGRIKEVLS	73
JcMVK	MEVKARAPGKIILSGEHAVVHGSTAVAS <mark>SIN</mark> LYTNVTLSFPTSENDGTLKLQLKDMALESSWPIGKIKEALP	72
ZjMVK	MEVKARAPGKIILSGEHAVVHGSTAVAASIGLYTYVTLRFPTSSENDDTLKLQLEDLSLE <mark>B</mark> SWPIGKLKEALP	73
MnMVK	MEVKARAPGKIILSGEHAVVHGSTAVAASIDLYTYVTLRFPTPSDDDDDNLKLQLKDMALE <mark>I</mark> SWPIGRIKEELP	73
HbMVK	MEVKARAPGKIILSGEHAVVHGSTAVAASINLYTYVTLSFATAENDDSLKLQLKDMALE <mark>B</mark> SWPIGRIREALS	72
GmMVK	MEVKSRAPGKIIITGEHAVVHGSTAVAS <mark>SID</mark> IYTYVSLHFSTPSDNEDS <mark>LK</mark> OKIOETALEFSWPITRIRAAFP	73
	ATP-binding	
GbMVK	STATED VDPLSVKSCSPDTMOCTA ATVDEOR PEATTCVA ACVSATURIVISITER KPATVTVASTTER GSGLGSSAALC	160
PmMVK	EKDTENSS., VPTSCSTESTRSLAALVOPLKTPEAKTGLGAEVTAETWLYSSTOCYKPATVVTTSHLETGSGLGSSAALC	151
PbMVK	ELDIENSS. VPTSCSTELLKSLAALVDELKTPEAKIGLGAEVVAETWLVASTOCEKPATVVTSPLEIGSGLGSSAALC	151
JCMVK	NI GADSAS TPTSCSTESVKSTSAT VEPONT DEAKTALASCI ANTI MUVISTOCERDATIVVVTSI UBLOSCI CSSAAFO	150
ZIMVK	DI GEDESS TETEOSTEA I SST SVI VEDONT DEAKTOLSSC I SAFLWILVSST LEEKDATIVVVNSH USUGSC GSSAALG	151
MnMVK	GIGCEDSS LATSCSTEAAKVLAALVEDONTDEAKTTLASEVSAELWLYTSTOCYKPATTVVTSFLETGSGLGSSAAFC	151
HOMVK	NI GADDES TETSCSMEST KTTSALWEDEN TOPAK TALTSEV SALUMLYTSTOCEK DAT WWYTSTU DI GSGLGSSALEG	150
GmMVK	ESTACLSS TPNSCSVENAKATAALWEDINTEPAKI, GLASEVSADIWLYSSTOCFKPATVVVTSFILDIGSGLGSSASEC	151
0		101
GbMVK	WAGSCAMLAFSGALDLDVLOGDMLSLKDKKOEIVNKWAYEGDKUIHCKPSGIDNWYSTFCNLVKDKSCHLVHIPNVFPIK	240
PmMVK	VALSAALLSFSD <mark>SVNLD</mark> LSHOGWLTFGE <mark>SE</mark> LELLNKWAFEGERIIHGKPSGIDNTVSTYGNMIKF <mark>K</mark> SGSLTRIKT <mark>NM</mark> TLK	231
PbMVK	VALSCALLSFSDCV <mark>SLD</mark> LCNKEWTTFCE <mark>SE</mark> LELLNKWAFEGEK <mark>I</mark> HCKPSGIDNTVSTYCNMIKF <mark>R</mark> SGSLTRIKS <mark>NM</mark> TLK	231
JcMVK	VALSA2 FLALSDSVNVDKMHOGWLIFGESELELINKWAFEGEKITHCKPSGIDNTVSTYGNMIKERSCNLTRIKSNMLLK	230
Zimvk	VALSAAU SFSDSVNLUTSHOGWLLFGPNELELINKWAFEGPKILHGRPSGIDNTVSTYGNMIKERSCNLTRIKSKVPLI	231
MnMVK	VALSAAU AFSDLVKIDLSHOGWVMFAPDELDLINKWAFEGPKITHCRPSGIDNTVSTYGNLIKERSGSLTRIKSNLPLK	231
HDMVK	VAL SAAT DA FSD SVNVDTKHOGWSTFGESDLELINKWALEGEKITHGKPSGTDNTVSTVGNMTKEKSCNMTRTKSNMPLK	230
GmMVK	VALAAAUTAYTDSVSIJULKHOGWLSFGPKDLELVNKWAFEGPKITHCKPSGIDNTVSAYGNIISFKSGNLTHMKSSVPLK	231
GbMVK	MLTINTKVGRNTKALVAGVSERALRHPQSMMAVFTAVNAISEBMAETIQAPFSDDIS <mark>ISEQECKME</mark> ELMEMNQCLLQCMG	320
PmMVK	ML <mark>I</mark> TNT <mark>K</mark> VGRNTKALVAGV <mark>SERTL</mark> RH <mark>PEAMASVFNAVDS</mark> ISTE <b>LAA</b> T <mark>IRSPAPDDLSLTEKE</mark> SK <mark>IE</mark> ELME <mark>MNQGLLQC</mark> MG	311
PbMVK	ML <mark>I</mark> TNT <mark>K</mark> VGRNTKALVAGV <mark>SERTL</mark> RH <mark>PEAMASVFNAVDSISTEVANTIQSPAPDELSITEKEAKIE</mark> ELME <mark>MNQGLLQC</mark> MG	311
JcMVK	ML <mark>I</mark> TNT <mark>R</mark> VGRNTKALVAGV <mark>SERTL</mark> RH <mark>PNAM<mark>SF</mark>VFNAVDSISNELATI<mark>IQSPAPDELSITEKEAKLE</mark>ELMEMNQC<mark>LLQC</mark>MG</mark>	310
ZjMVK	ML <mark>I</mark> TNT <mark>K</mark> VGRNTKALVAGV <mark>SERAI</mark> RH <mark>PNAMSL</mark> VFNAVDSISEBLSII <mark>IQSDAPDDLSLTEKBE</mark> KIAELMEMNQGLLQCMG	311
MnMVK	ML <mark>ITNTK</mark> VGRNTKALVAGV <mark>SERTL</mark> RH <mark>ANAMSF</mark> VFNAVDSISEBLASTIQSPIGDDLSVPQKE <mark>B</mark> KIEELMEMNQGLLQCMG	311
HbMVK	ML <mark>V</mark> TNT <mark>R</mark> VGRNTKALVAGV <mark>SERTLRHPNAMSF</mark> VFNAVDSISNELANTIOSPAPDDVS <mark>ITEKE</mark> EKLEELMEMNOGLLOCMG	310
GmMVK	ML <mark>ITNTK</mark> VGRNTKALVAGV <mark>G</mark> ERMLRH <mark>PDIMAF</mark> VF <mark>SAVDSISNBITSTILKSPTPDELS<mark>VTEKEE</mark>KIE</mark> ELME <mark>V</mark> NQCMLQSMG	311
GbMVK	VSHSAIEAVLOTTS <mark>RYKL</mark> OSKLTGAGGGGCVLTLLEKLLS <mark>GKVVDKVKAD</mark> LES <mark>OGFOF</mark> LATIGGEGV <mark>OVCF</mark> QC <mark>SS</mark>	396
PmMVK	VSHASIETVLRTILKYKLASKLTGAGGGGCVLTLLPTLLSGTVVDKVTSELESCGFHCLTAALGCNGVEICFGDSS	387
PbMVK	VSHASIE <mark>TVLRTILK</mark> YKL <mark>A</mark> SKLTGAGGGGCVLTLLP <mark>TLLSGTVVD</mark> KVTÄELESCGFHC <mark>LTAALGCNGVEIRFAG</mark> PL	387
JcMVK	VSHASIETVLRTILKYKLASKLTGAGGGGCVLTLLPTLLSGTVVDKVIABLESCGFCCLIAGIGCNGVEICFGGSS	386
ZjMVK	VSHASIETVURTILKYKLISKLTGAGGGGCVLTLLPTLLSETVVDKVVSELESCGFOCFTAGIGGNGVEICFSGSS	387
MnMVK	VSHASIETVLRTTLKYKLASKLTGAGGGGCVLTLLPNLLSGTLADRVTSDLESSGFOCFTAGIGGNGVEIYLGGSS	387
HDMVK	VSEAS DETVURITILKYKLASKLTGAGGGGCVLTHLETILLSGTVVDKALAEDESCGFOCLLAGIGGNGVEFCFGGSS	386
GmMVK	VSHATTETULETTLKYKLASKLTGAGGGGCVLTLLETLLSGTWDKVVARLESCGEOCFIAGIGGGWEISFGVSS	387

**Figure 3.** Multiple alignments of GbMVK with MVK proteins from *Prunus mume* (PmMVK, Accestion No. XP\_008246488.1), *Pyrus* × *bretschneideri* (PbMVK, Accestion No. XP\_009349394.1), *Jatropha curcas* (JcMVK, Accestion No. XP\_012089078.1), *Ziziphus jujuba* (ZjMVK, Accestion No. XP\_015875214.1), *Morus notabilis* (MnMVK, Accestion No. XP\_010105842.1), *Hevea brasiliensis* (HbMVK, Accestion No. AIO11226.1), *Glycine max* (GmMVK, Accestion No. NP\_001276217.1). Dark blue: identity = 100%; red: 75%  $\leq$  identity < 100%; light blue: 50%  $\leq$  identity < 75%.

### 2.3. Molecular Evolution Analysis

To investigate the evolutionary relations among deduced GbAACT with other AACTs, and among GbMVK with other MVKs from angiosperm, gymnosperm, fungus, and bacteria, two phylogenetic trees were constructed using the neighbor-joining (NJ) method with *p*-distance. Apparently, the bootstrap value is much high for each interior branch with a high identity (>90%). Therefore, the constructed phylogenetic trees of AACTs and MVKs are reliable. As shown in Figure 4A, AACTs from different species seemed to evolve into different groups, with fungus as an ancient group. Then the bacteria group containing AACTs diverged from *Sphingobacterium* and *Mucilaginibacter paludis*. The plant AACT group diverged later than the fungus and bacteria. Among the plant group, the GbAACT diverged a little earlier than AACT from other angiosperm plant species in the phylogenetic tree, which coincided with the evolutional position of *G. biloba* as the most ancient among gymnosperm plant species. As shown in Figure 4B, MVKs from different species evolved vertically from a common ancestor. The MVKs from different species evolved into various groups, with bacteria as an ancient group, and then the fungus group containing MVKs diverged from *Fusarium fujikuroi* 

and *Candida dubliniensis*. The plant species divided into angiosperm and gymnosperm species, and the GbMVK diverged earlier than MVK from other plant species. These results correspond with the fact that *G. biloba* is the most ancient among gymnosperm plant species. Given that both the AACTs and MVKs can be found in fungus, bacteria, and higher plants, the bioinformatics analysis also indicated that GbAACT was a plant AACT protein with AACT activity and that GbMVK was a plant MVK protein with GHMP kinase activity. Both *GbAACT* and *GbMVK* can also be inferred to be a class of highly conservative ancient genes.



**Figure 4.** Phylogenetic analysis of the amino acid sequences of AACTs (**A**) and MVKs (**B**). The neighbor-joining phylogenetic trees were constructed using the bootstrap method on MEGA 6.01 and the number of Bootstrap replications was 1000. The following are the protein sequences used in these trees: (**A**) AACT: *Populus trichocarpa* (XP\_002308755.1), *Nelumbo nucifera* (XP\_010252788.1), *Elaeis guineensis* (XP\_010936771.1), *Phoenix dactylifera* (XP\_008811276.1), *Schizosaccharomyces cryophilus* (XP\_013022052.1), *Cordyceps militaris* (XP\_006672862.1), *Mucilaginibacter paludis* (WP\_040627967.1), *Sphingobacterium* sp. 21 (WP\_013665750.1); (**B**) MVK: *Hevea brasiliensis* (ALR72881.1), *Prunus mume* (XP\_008246488.1), *Elaeis guineensis* (XP\_010906879.1), *Phoenix dactylifera* (XP\_008217536.1), *Pyrococcus yayanosii* (WP\_013905018.1), *Halorubrum litoreum* (WP\_008367863.1).

### 2.4. Functional Complementation of GbAACT and GbMVK in Saccharomyces Cerevisiae

The ergosterol synthesized from MVA pathway is essential for yeast survival [31,32]. A disruption of the MVA pathway genes is lethal in yeast [33,34]. To determine the function of *GbAACT* and *GbMVK*, two ergosterol auxotrophic strains of *Saccharomyces cerevisiae* that lacked the *AACT* or the *MVK* allele, named YPL028W ( $\Delta ERG10$ ) and YMR208W ( $\Delta ERG12$ ), respectively, were used for experiment. The pYES2 vectors, containing a yeast galactose-dependent promoter, were used as carrier for target genes in this study. The disrupted strains that harbored empty pYES2 could not grow on either the YPG expression medium or the YPD non-expression medium. Two expression vectors, namely, pYES2-GbAACT and pYES2-GbMVK, were constructed and transformed into strains YPL028W and YMR208W, respectively. YPL028W harbored pYES2-GbAACT, and YMR208W harbored pYES2-GbMVK, which grew well on the YPG medium. However, neither the YPL028W harbored with pYES2-GbAACT nor the YMR208W harbored with pYES2-GbMVK grew on the YPD medium (Figure 5). This proves that the transformed *GbAACT* can fix the functional loss of the *AACT* knockout

yeast and that the transformed *GbMVK* can compensate the functional lack of the *MVK* knockout yeast. These results confirmed that GbAACT and GbMVK have AACT and MVK activity respectively.



**Figure 5.** Functional complementation for the growth of haploid disrupted strains  $\Delta erg10$  complemented with *GbAACT* (**A**) and  $\Delta erg12$  complemented with *GbMVK* (**B**). The strains were grown on YPD and YPG at 30 °C for 3 days with the exception of  $\Delta erg10$  with *GbAACT* and  $\Delta erg12$  with *GbMVK*.

#### 2.5. Transcript Level of the Gene Expression Pattern of GbAACT and GbMVK in Different Tissues of G. biloba

To present, reports corresponding to the *AACT* and *MVK* of *G. biloba* have been inexistent. To determine the expression patterns of *GbAACT* and *GbMVK* genes among different tissues in *G. biloba*, total RNA was extracted from roots, stems, leaves, female flowers, male flowers, and fruit, and cDNA was synthesized as mentioned above. qRT-PCR (Real-time quantitative reverse transcription PCR) was performed for *GbAACT* and *GbMVK* (Figure 6).

qRT-PCR revealed that the expression profile of *GbAACT* is distributed throughout all ginkgo tissues. Among different tissues, *GbAACT* had highest expression in fruit, followed by the leaf and the male flower. The *GbAACT* expression in the roots was significantly lower than that in the other tissues (Figure 6A). *AACT* genes were reported to be tissue-specific genes in other plants [27]. In *Bacopa monnieri*, the *BmAACT* gene was highly expressed in the root, followed by the stem and leaf [23], whereas results of a recent study in *Isodon rubescens* corresponded with this study in that *AACT* gene is more significantly expressed in leaves and flowers than in roots and stems [35]. Our data revealed that the transcript of *GbAACT* was detected in all ginkgo organs, including the root, stem, leaves, flowers, indeed overlapped with those of *GbLPS* (Levopimaradiene synthase) [20], *GbIDS* (1-Hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate reductase) [19] and *GbMVD* (Mevalonate diphosphate decarboxylase) [11] in the roots and flowers. This result verifies the roots as the preferential site of TTL biosynthesis. However, one unexpected observation of this study is that the transcript level of *GbAACT* was higher in aerial tissues than roots. The possibility is that *GbAACT* in aerial parts of ginkgo is involved in the biosynthesis of yet to be identified terpenoids.

As shown in Figure 6B, *GbMVK* was highly expressed in leaves, roots, and stems, and the highest *GbMVK* expression was detected in roots and leaves. The expression levels of *GbMVK* in floral organs and fruits were significantly low. A similar expression pattern of the *MVK* gene was found in *Panax notoginseng*, though in a lesser degree, because *PnMVK* is significantly expressed in roots, flowers, and leaves but is seldom expressed in stems [36]. The expression profile is consistent with the TTL distribution from earlier reports [37], showing that higher contents of TTLs were found in the leaves and roots than other tissues of ginkgo. A correlation exists between the transcript level of *GbMVK* and the content of TTLs, thereby suggesting that *GbMVK* plays an important role in the production of TTLs in ginkgo.

The biosynthesis organ of the TTL in *G. biloba* was not clearly reported, and long-distance transport was reported to be involved in the translocation of various compounds [38]. Therefore, considering the results of the present study and given that both *GbAACT* and *GbMVK* are highly expressed in the leaf and *GbMVK* is highly expressed in the root, we infer that the root and leaf could be most the important tissues for forming TTLs. To further determine which specific part of plant is involved in TTL biosynthesis, an isotopic tracer technique might be helpful.



**Figure 6.** Gene expression Patterns of *GbAACT* (**A**) and *GbMVK* (**B**) in the various organs of *G. biloba* using qRT-PCR. The gene expression level of *GbAACT* and *GbMVK* in the root was set to 1, and those of *GbAACT* and *GbMVK* in other tissues were accordingly accounted and presented as the relative fold changes, respectively. R: root, S: stem, L: leaf, Ff: female flower, Mf: male flower, F: fruit. Data from qRT-PCR were shown as the mean  $\pm$  SD (standard deviation) of three replicated assays. Means with different letters are significantly different at *p* < 0.05 by one-way ANOVA, with Tukey's honestly significant difference test.

## 2.6. Transcript Level of GbAACT and GbMVK and TTL Content Changes in Ginkgo biloba under the Induction of MeJA and SA Elicitors

To understand the expression pattern of *GbAACT* and *GbMVK*, *GbAACT* and *GbMVK* transcript levels and TTL content were measured after seedlings at 4–5-leaf stage were treated with the elicitors MeJA and SA. qRT-PCR experiments clearly showed that SA and MeJA induction significantly increased *GbAACT* and *GbMVK* expression, and the TTL content was significantly enhanced after SA and MeJA treatments (Figure 7). The expression level of *GbAACT* continuously increased after MeJA treatment and peaked at 64 h with an 3.73-fold increase compared with that of control seedlings. The SA-treated seedlings showed the highest level of *GbAACT* expression at 32 h, and then slowly decreased (Figure 7A). With regard to the *GbMVK* gene, the expression level quckly increased after the seedlings were treated with MeJA for 8 h, and expression level of *GbMVK* gene peaked at 24 h after MeJA treatment. The *GbMVK* expression slightly changed from 16 h to 32 h, and the *GbMVK* expression decreased from 32 h to 96 h after treatment. Under SA treatment, *GbMVK* was highly expressed 8 h after the treatment, and *GbMVK* expression peaked until 96 h, with a 3.34-fold increase compared with that of the control (Figure 7B). TTL increased by 5.8% 8 h after SA treatment, and TTL content continuously increased from 0 h to 64 h. A similar increase was observed among MeJA-treated seedlings (Figure 7C).

A

Relative expression level of GbAACT

В

4.0

3.5

4.0

35

■CK

SA 3.0 2.5 2.0 1.5 1.0 0.5 0.0

0

■CK

MeJA ■SA 3.0 2.5 2.0 1.5

a a

MeJA





Figure 7. *GbAACT* and *GbMVK* expression level (A,B) and total terpene trilactone (TTL) content (C) changes in ginkgo by salicylic acid (SA) and methyl jasmonate (MeJA). The expression levels were normalized to the house-keeping gene GAPDH (glyceraldehyde-3-phospate dehydrogenase). Data from qRT-PCR were analyzed as expression ratios relative to the level of control (CK), and are shown as the mean  $\pm$  SD of triplicate assays. Means with different letters from each time of post-treatments are significantly different at p < 0.05 by one-way ANOVA, with Tukey's honestly significant difference test.

With the regulation of SA and MeJA, plants were capable of adjusting to both abiotic and biotic stress [39]. SA and MeJA interact antagonistically against each other to induce transcription of defense-related genes by producing certain compunds and proteins [40–43]. Therefore, challenging ginkgo with SA and MeJA would activate defense-related genes in plants concomitantly increasing secondary metabolites. Indeed, the exogenous application of SA and MeJA has been performed to improve the biosynthesis of ginkgolide A and B and bilobalide in cell cultures of G. biloba [44]. In addition, some genes involved in TTL biosynthesis are upregulated after SA and MeJA treatments are administered in ginkgo. The transcript levels of GbDXS (1-deoxy-D-xylulose 5-phosphate synthase) [13], GbCMK2 (4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase) [15], GbHMGR [26], GbIDS2 [45], and GbMVD [11] are induced by SA and MeJA and are positively involved in TTL biosynthesis in G. biloba. Similar to these reports, our data showed that the expression of GbAACT and *GbMVK* was upregulated and the TTL content was increased after SA and MeJA treatments, implying that the transcription of these genes responsible for TTL biosynthesis was enhanced. However, TTLs comprise a wide range of compounds and an integrated reaction of a cluster of genes related to TTL biosynthesis. The SA or MeJA-induced increase in TTL production in the present study may be

arrtributed to an integrated effect on multiple genes related to TTL biosynthesis. Hence, it is interesting to unveil the expression profiling of other genes involved in TTL biosynthesis under SA or MeJA treatments in ginkgo, which will provide more insights into the regulatory role of SA and MeJA in TTL biosynthesis.

## 3. Experimental Section

## 3.1. Plant Materials

Plant samples were obtained from 18-year-old trees of *G. biloba* growing on the ginkgo garden of Yangtze University, Hubei, China (around N30.35, E112.14). The roots, stems, leaves, female and male flowers were collected in April 2014, and the fruits were collected in June 2014. Primer synthesis and DNA sequencing were performed by Sangon Biotechnology Company (Shanghai, China). Agarose Gel DNA Extraction Kit Ver. 4.0, pMD19-T vector kit, first-strand cDNA synthesis kit, dNTPs, RNasin, and *Taq* DNA polymerase were purchased from Takara Company (Dalian, China). A PrimeScript RT reagent kit with a gDNA (genomic DNA) Eraser and AceQ qPCR SYBR Green Master Mix were purchased from Vazyme Bio Inc. (Nanjing, China).

The *G. biloba* seedlings at 4–5-leaf stage were grown in a  $24 \pm 1$  °C incubator with 100 µmol·m<sup>-2</sup>·s<sup>-1</sup> light and a 16/8 h photoperiod. The seedlings were sprayed with 1 mM MeJA or 0.5 mM SA solution containing 0.1% (v/v) ethanol. Seedlings treated with 0.1% ethanol solution were used as control. The seedlings were harvested at 0, 8, 16, 24, 32, 64, and 96 h after treatment.

## 3.2. Cloning of the Full-Length cDNA of GbAACT and GbMVK

The RNA of *G. biloba* was extracted with TaKaRa MiniBEST Plant RNA Extraction Kit (Takara Bio Inc., Dalian, China) according to the manufacturer's instruction. The extracted RNA was purified and used as templates to constructed RNA-seq library according to the manufacturer's protocol. The transcriptome sequencing libraries were generated using a using a NEB Next<sup>®</sup> UltraTM RNA Library Prep Kit for Illumina<sup>®</sup> (NEB, Ipswich, UK). Sequencing run was performed at Biomarker Co., Beijing, China using Illumina Hiseq 2500 platform. In our previous work, we assigned the functional annotations of 35,113 unigenes [46]. The specific cDNA isolations were obtained through the PrimeScript first-strand cDNA synthesis kit. Primers were designed on the basis of the initial data of *AACT* and *MVK* unigenes in the transcriptome (GenBank accession numbers SRR3985386 and SRR3989536 for male and female strobilus, respectively). The sequences of all primers were shown Table 1. The PCR products were purified and ligated into pMD19-T vector (Takara Bio Inc., Dalian, China). The recombined plasmids were transformed into *E. coli* DH5 $\alpha$  competent cells (Vazyme Bio Inc., Nanjing, China) and sequenced.

Usage	Primer Name	Primer Sequence 5'-3'
	GbAACT-F	ATGGGTTCTACATCAGGCTCAGATTTGTT
ODE DOD	GbAACT-R	TTACATGAGCTCCAAGACAACTGCAG
ORF-PCK	GbMVK-F	ATATTGGGAGTTCAGAATGGTGC
	GbMVK-R	ACATTTATGAGCTTGAGCATTGGAA
	GbAACT-qF	CATTGGGGGTCTGAGTGGTTC
	GbAACT-qR	CATTGTTGCTTTCATTCCCGA
aRT-PCR	GbMVK-qF	GGCATTAGTTGCTGGAGTTTCTGAG
qKI-I CK	GbMVK-qR	GGTGCCTGAATAATCTCTGCCATCT
	GbGAPDHF	CTGGCGTAGAGTATGTGGTTGAAT
	GbGAPDHR	CACGCCAACAACGAACATG
	GbAACT-yup	CGGGATCCATGGGTTCTACATCAGGCTCAG
Verst DCD	GbAACT-yud	CGGAATTCTTACATGAGCTCCAAGACAACTG
reast-PCK	GbMVK-yup	CGGGATCCATATTGGGAGTTCAGAATGGTGC
	GbMVK-yud	CGGAATTCACATTTATGAGCTTGAGCATTGGAA

Table 1. Primers used in this study. ORF: open-reading frame.

#### 3.3. Bioinformatics Analysis and Molecular Evolution Analyses

Sequences were assembled and their ORF were analyzed with program Vector NTI 11.5.1 (Invitrogen, Paisley, UK,) [47]. Sequence comparison and amino acid translation were performed with DNAMAN 8.0 (Lynnon Biosoft, Quebec, QC, Canada). Protein isoelectric point (pI) and molecular weight were calculated via http://web.expasy.org/compute\_pi/. The protein-conserved domain was analyzed via InterProScan [48] and http://www.ncbi.nlm.nih.gov. A phylogenetic tree of GbAACT and GbMVK from *G. biloba* and other plants was constructed through NJ method with CLUSTALX 2.0 (Conway Institute UCD Dublin, Dublin, Ireland) and MEGA 6.06 (Biodesign Institute, Tempe, AZ, USA) [49].

#### 3.4. Construction of Expression Plasmids for Yeast Complementation

The coding regions of the *GbAACT* and *GbMVK* were amplified with two pairs of primers: GbAACT-yup and GbAACT-yud, GbMVK-yup and GbMVK-yud, respectively (Table 1). The forward primers contained the BamHI restriction site, and the reverse primers contained the EcoRI restriction site. The amplified products and pYES2 vector (Invitrogen, Carlsbad, CA, USA) were digested with BamHI and EcoRI, and then ligated into recombined vector pYES2-GbAACT and pYES2-GbMVK. PCR and sequencing were performed on positive clones. The constructed pYES2-*GbAACT* and pYES2-*GbMVK* plasmids were extracted and transformed into YPL028W ( $\Delta ERG10$ ) and YMR208W ( $\Delta ERG12$ ) with the Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA). The wild-type S. cerevisiae strain YSC1021 is a diploid yeast, and *S. cerevisiae* strain Δ*ERG10* lacking the *AACT* allele is the haploid yeast. The *S. cerevisiae* strain  $\Delta ERG12$  lacking the *MVK* allele is also a haploid yeast. The transformants were spotted on SC (-Ura) medium (6.7% yeast nitrogen base without amino acid, 2% galactose). The transformed diploid cells were induced to sporulate and subsequently formed haploid cells containing pYES2-*GbAACT* and pYES2-*GbMVK*. To further observe their growth condition, the diploid S. cerevisiae strain YSC1021 and transformed haploid strain YPL028W and YMR208W were separately grown on YPD (1% yeast extract, 2% Bacto Peptone, 2% glucose) and YPG (1% yeast extract, 2% Bacto Peptone, 2% galactose).

### 3.5. GbAACT and GbMVK Tissue-Specific Analysis

For quantification transcripts of GbAACT and GbMVK genes from G. biloba in different tissues, qRT-PCR was performed with primers designed (Table 1) using Primer premier 5.0 (Premier Biosoft International, Palo Alto, CA, USA). The PrimeScript RT reagent Kit was used with 100 ng of each total RNA to synthesize single-strand cDNA. The reaction was performed on MiniOpticon (Bio-Rad Laboratories, Inc., Alfred Nobel Drive Hercules, CA, USA) using SYBR Green detection with a reaction mixture (25 µL) containing 2× SYBR<sup>®</sup> Premix Ex Taq II (Tli RNaseH Plus, Takara Bio Inc., Dalian, China), 0.4  $\mu$ M each of forward and reverse primers, and 2 ng/ $\mu$ L of template cDNA. PCR amplification was performed under the following conditions: 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s and 60 °C for 30 s. All experiments were performed in triplicate, and the mean value was analyzed. Master mix without template was treated as negative control in each reaction. A glyceraldehyde-3-phospate dehydrogenase house-keeping gene (GAPDH) was used as internal control for normalization of all the reactions [50]. The primers used for the normalization (GbGAPDHF and GbGAPDHR) and expression analysis of GbAACT (GbAACT-qF and GbAACT-qR) and GbMVK (GbMVK-qF and GbMVK-qR) are shown in Table 1. All of the reactions were run in triplicate and repeated thrice. Gene expression analysis  $2^{-\Delta\Delta Ct}$  method [51] was used to normalize the relative gene expression of the transcript in different tissue types.

#### 3.6. Determination of TTL Contents in Ginkgo under Induction of MeJA and SA Elicitor

*G. biloba* seedlings were sampled and freeze-dried. Ginkgolide A, ginkgolide B, ginkgolide C, and bilobalide were extracted and quantified by gas chromatography with wide bore capillary column

method [52] with minor modification. In detail, an appropriate amount of purified methanolic fraction was evaporated to dryness and trimethylsilylated by adding 100  $\mu$ L of a silylating agent (Trisil BSA formula D, Sigma-Aldrich, Darmstadt, Germany). This mixture was vortex and heated for 2 h at 80 °C. Analysis by Gas chromatography-flame ionization detection (GC-FID) was performed with a Shimadzu GC-14B (Shimadzu, Japan) equipped with a 30 m × 0.53 mm × 1.5  $\mu$ m TC-1 capillary column (HP-5, Agilent, San Francisco, CA, USA). The temperatures of the column, injector, and detector were maintained at 290 °C, 320 °C, and 320 °C, respectively. He was used as the carrier gas at a flow rate of 3.1 mL/min. TTL content was calculated as the sum of ginkgolide A, ginkgolide B, ginkgolide C, and bilobalide contents by dry weight percentages. Each sample was evaluated in triplicate, and data were represented as means  $\pm$  SD (n = 3).

## 3.7. Statistical Analysis

Data were analyzed with using the statistical software SPSS 11.0 for Windows (SPSS Inc., Chicago, IL, USA). Comparisons between multiple treatment groups were performed using one-way ANOVA, with Tukey's honestly significant difference test. p < 0.05 was considered to be statistically significant.

#### 4. Conclusions

The genes *GbAACT* and *GbMVK* encoding AACT and MVK of the MVA pathway and the enzymes catalyzing the first and fourth steps in this pathway, respectively, were cloned and characterized. Bioinformatics analysis revealed that the deduced GbAACT and GbMVK harbored a highly similar identity to AACTs and MVKs of other plants. Functional complementation experiments in yeast demonstrated that *GbAACT* and *GbMVK* encode functional AACT and MVK, respectively. The transcript levels of *GbAACT* and *GbMVK* significantly increased in response to MeJA and SA treatments, and this increase corresponded to the increase in the TTL content after SA or MeJA treatment was performed. qRT-PCR analysis showed that *GbAACT* and *GbMVK* are tissue-specific genes. The transcript level of *GbAACT* is highly expressed in fruits, leaves, and floral organs, whereas the transcript level of *GbMVK* is highly expressed in leaves, roots, and stems. For further unveiling of the function of *GbAACT* and *GbMVK*, a plant expression vector containing the *GbAACT* or *GbMVK* has been constructed and a study of the genetic transformations of ginkgo is underway in order to investigate the potential role in enhancing TTL accumulation by genetic engineering.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China (31370680) and Natural Science Foundation of Hubei Province of China (2013CFA039).

**Author Contributions:** Q.C., J.Y. and X.M. performed the experiments and analyzed the data. Q.C. and F.X. drafted the manuscript. W.Z. performed the qRT-PCR analysis. Y.L. and J.Q. contributed the functional complementation. F.X. designed the experiments. All authors read and approved the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- 1. Carrier, D.J.; van Beek, T.A.; van der Heijden, R.; Verpoorte, R. Distribution of ginkgolides and terpenoid biosynthetic activity in *Ginkgo biloba*. *Phytochemistry* **1998**, *48*, 89–92. [CrossRef]
- Hosford, D.; Domingo, M.; Chabrier, P.; Braquet, P. Ginkgolides and platelet-activating factor binding sites. *Methods Enzymol.* 1990, 187, 433–446. [PubMed]
- Stromgaard, K.; Nakanishi, K. Chemistry and biology of terpene trilactones from *Ginkgo biloba*. *Angew. Chem.* 2004, 43, 1640–1658. [CrossRef] [PubMed]
- 4. Braquent, P. BN 52021 and related compounds: A new series of highly specific paf-acether receptor antagonists. *Prostaglandins* **1985**, *30*, 687. [CrossRef]
- 5. Jaracz, S.; Nakanishi, K.; Jensen, A.A.; Strømgaard, K. Ginkgolides and glycine receptors: A structure-activity relationship study. *Chemistry* **2004**, *10*, 1507–1518. [CrossRef] [PubMed]

- Sabater-Jara, A.B.; Souliman-Youssef, S.; Novo-Uzal, E.; Almagro, L.; Belchí-Navarro, S.; Pedreño, M.A. Biotechnological approaches to enhance the biosynthesis of ginkgolides and bilobalide in *Ginkgo biloba*. *Phytochem. Rev.* 2013, *12*, 191–205. [CrossRef]
- 7. Lange, B.M.; Croteau, R. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13172–13177. [CrossRef] [PubMed]
- 8. Lichtenthaler, H.K.; Rohmer, M.; Schwender, J. Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. *Physiol. Plant.* **1997**, *101*, 643–652. [CrossRef]
- 9. Schwarz, M.A.D. Ginkgolide biosynthesis. In *Comprehensive Natural Products Chemistry*; Cane, D., Ed.; Pergamon Press: Oxford, UK, 1999; pp. 367–400.
- Pang, Y.; Shen, G.A.; Bergès, T.; Cardier, H.; Wu, W.; Sun, X.; Tang, K. Molecular cloning, characterization and heterologous expression in *Saccharomyces cerevisiae* of a mevalonate diphosphate decarboxylase cDNA from *Ginkgo biloba*. *Physiol. Plant.* 2006, 127, 19–27. [CrossRef]
- 11. Liao, Y.; Xu, F.; Huang, X.; Zhang, W.; Cheng, H.; Wang, X.; Cheng, S.; Shen, Y. Characterization and transcriptional profiling of *Ginkgo biloba* mevalonate diphosphate decarboxylase gene (*GbMVD*) promoter towards light and exogenous hormone treatments. *Plant Mol. Biol. Rep.* **2016**, *34*, 566–581. [CrossRef]
- Shen, G.; Pang, Y.; Wu, W.; Liao, Z.; Zhao, L.; Sun, X.; Tang, K. Cloning and characterization of a root-specific expressing gene encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase from *Ginkgo biloba. Mol. Biol. Rep.* 2006, *33*, 117–127. [CrossRef] [PubMed]
- Gong, Y.F.; Liao, Z.H.; Guo, B.H.; Sun, X.F.; Tang, K.X. Molecular cloning and expression profile analysis of *Ginkgo biloba DXS* gene encoding 1-deoxy-D-xylulose 5-phosphate synthase, the first committed enzyme of the 2-C-methyl-D-erythritol 4-phosphate pathway. *Planta Med.* 2006, 72, 329–335. [CrossRef] [PubMed]
- Kim, S.M.; Kuzuyama, T.; Chang, Y.J.; Song, K.S.; Kim, S.U. Identification of class 2 1-deoxy-D-xylulose 5-phosphate synthase and 1-deoxy-D-xylulose 5-phosphate reductoisomerase genes from *Ginkgo biloba* and their transcription in embryo culture with respect to ginkgolide biosynthesis. *Planta Med.* 2006, 72, 234–240. [CrossRef] [PubMed]
- Kim, S.M.; Kim, Y.B.; Kuzuyama, T.; Kim, S.U. Two copies of 4-(cytidine 5'-diphospho)-2-C-methyl-Derythritol kinase (CMK) gene in *Ginkgo biloba*: Molecular cloning and functional characterization. *Planta* 2008, 228, 941–950. [CrossRef] [PubMed]
- Gao, S.; Lin, J.; Liu, X.; Deng, Z.; Li, Y.; Sun, X.; Tang, K. Molecular cloning, characterization and functional analysis of a 2*C*-methyl-D-erythritol 2,4-cyclodiphosphate synthase gene from *Ginkgo biloba*. *J. Biochem. Mol. Biol.* 2006, 39, 502–510. [PubMed]
- Kim, S.M.; Kuzuyama, T.; Chang, Y.J.; Kim, S.U. Cloning and characterization of 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MECS) gene from *Ginkgo biloba*. *Plant Cell Rep.* 2006, 25, 829–835. [CrossRef] [PubMed]
- Kim, S.M.; Kim, S.U. Characterization of 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate synthase (HDS) gene from *Ginkgo biloba*. *Mol. Biol. Rep.* 2010, *37*, 973–979. [CrossRef] [PubMed]
- Kim, S.M.; Kuzuyama, T.; Kobayashi, A.; Sando, T.; Chang, Y.J.; Kim, S.U. 1-Hydroxy-2-methyl-2-(*E*)-butenyl
  4-diphosphate reductase (IDS) is encoded by multicopy genes in gymnosperms *Ginkgo biloba* and *Pinus taeda*.
  *Planta* 2008, 227, 287–298. [CrossRef] [PubMed]
- Kim, J.H.; Lee, K.I.; Chang, Y.J.; Kim, S.U. Developmental pattern of *Ginkgo biloba* levopimaradiene synthase (GbLPS) as probed by promoter analysis in *Arabidopsis thaliana*. *Plant Cell Rep.* 2012, *31*, 1119–1127. [CrossRef] [PubMed]
- 21. Khadijah Hanim, A.R.; Samian, M.R. Cloning and promoter identification of acetoacetyl-CoA thiolase gene from oil palm, *Elaeis guineensis* Jacq. *Int. J. Environ. Sci. Dev.* **2014**, *5*, 362–366.
- 22. Fang, X.; Shi, L.; Ren, A.; Jiang, A.L.; Wu, F.L.; Zhao, M.W. The cloning, characterization and functional analysis of a gene encoding an acetyl-CoA acetyltransferase involved in triterpene biosynthesis in *Ganoderma lucidum*. *Mycoscience* **2013**, *54*, 100–105. [CrossRef]
- 23. Vishwakarma, R.K.; Singh, S.; Sonawane, P.D.; Srivastava, S.; Kumari, U.; Kumar, R.J.S.; Khan, B.M. Molecular cloning, biochemical characterization, and differential expression of an acetyl-CoA *C*-acetyltransferase gene (*AACT*) of Brahmi (*Bacopa monniera*). *Plant Mol. Biol. Rep.* **2013**, *31*, 547–557. [CrossRef]
- 24. Wuyuntana; Wang, L.; Ye, S.J. Identification and bioinformatics analysis of *EuMK* gene in *Eucommia ulmoides*. *Nonwood For. Res.* **2014**, *32*, 6–12.

- 25. Deng, X.M.; Shao-Hua, W.U.; Dai, X.M.; Tian, W.M. Expression analysis of MVA and MEP metabolic pathways genes in latex and suspension cells of *Hevea brasiliensis*. *Guihaia* **2016**, *36*, 449–455.
- 26. Liao, Y.L.; Xu, F.; Huang, X.H.; Zhang, W.W.; Cheng, H.; Li, L.L.; Cheng, S.Y.; Shen, Y.B. Promoter analysis and transcriptional profiling of *Ginkgo biloba* 3-hydroxy-3-methylglutaryl coenzyme a reductase (*GbHMGR*) gene in abiotic stress responses. *Not. Bot. Horti Agrobot.* **2015**, *43*, 25–34. [CrossRef]
- 27. Dyer, J.H.; Maina, A.; Gomez, I.D.; Cadet, M.; Oeljeklaus, S.; Schiedel, A.C. Cloning, expression and purification of an acetoacetyl CoA thiolase from sunflower cotyledon. *Int. J. Biol. Sci.* **2009**, *5*, 736–744. [CrossRef] [PubMed]
- 28. Anderson, V.E.; Bahnson, B.J.; Wlassics, I.D.; Walsh, C.T. The reaction of acetyldithio-CoA, a readily enolized analog of acetyl-CoA with thiolase from *Zoogloea ramigera*. J. Biol. Chem. **1990**, 265, 6255–6261. [PubMed]
- 29. Williams, S.F.; Palmer, M.A.; Peoples, O.P.; Walsh, C.T.; Sinskey, A.J.; Masamune, S. Biosynthetic thiolase from *Zoogloea ramigera*. Mutagenesis of the putative active-site base Cys-378 to Ser-378 changes the partitioning of the acetyl *S*-enzyme intermediate. *J. Biol. Chem.* **1992**, *267*, 16041–16043. [PubMed]
- Yang, S.Y.; Yang, X.Y.; Healy-Louie, G.; Schulz, H.; Elzinga, M. Nucleotide sequence of the fadA gene. Primary structure of 3-ketoacyl-coenzyme A thiolase from *Escherichia coli* and the structural organization of the fadAB operon. *J. Biol. Chem.* 1990, 266, 10424–10429.
- Sando, T.; Takaoka, C.; Mukai, Y.; Yamashita, A.; Hattori, M.; Ogasawara, N.; Fukusaki, E.; Kobayashi, A. Cloning and characterization of mevalonate pathway genes in a natural rubber producing plant, *Hevea brasiliensis. Biosci. Biotechnol. Biochem.* 2008, 72, 2049–2060. [CrossRef] [PubMed]
- 32. Servouse, M.; Karst, F. Regulation of early enzymes of ergosterol biosynthesis in *Saccharomyces cerevisiae*. *Biochem. J.* **1986**, 240, 541–547. [CrossRef] [PubMed]
- 33. Trocha, P.J.; Sprinson, D.B. Location and regulation of early enzymes of sterol biosynthesis in yeast. *Arch. Biochem. Biophys.* **1976**, 174, 45–51. [CrossRef]
- 34. Oulmouden, A.; Karst, F. Isolation of the ERG12 gene of *Saccharomyces cerevisiae* encoding mevalonate kinase. *Gene* **1990**, *88*, 253–257. [CrossRef]
- 35. Zhu, Y.H.; Su, X.H.; Dong, C.M.; Chen, S.Q.; Shao, Y.Y.; Zhang, F.B. Cloning and expression analysis of acetyl-CoA C-acetyltransferase gene in *Isodon rubeseens. J. Chin. Med. Mater.* **2016**, *39*, 37–41.
- 36. Guo, X.; Luo, H.M.; Chen, S.L. Cloning and analysis of mevalonate kinase (*PnMVK1*) gene in *Panax notoginseng. Acta Pharm. Sin.* **2012**, *47*, 1092–1097.
- Cartayrade, A.; Neau, E.; Sohier, C.; Balz, J.P.; Carde, J.P.; Walter, J. Ginkgolide and bilobalide biosynthesis in *Ginkgo biloba*. 1. Sites of synthesis, translocation and accumulation of ginkgolides and bilobalide. *Plant Physiol. Biochem.* **1997**, *35*, 859–868.
- 38. Erb, M.; Lenk, C.; Degenhardt, J.; Turlings, T.C. The underestimated role of roots in defense against leaf attackers. *Trends Plant Sci.* 2009, *14*, 653–659. [CrossRef] [PubMed]
- 39. Robertseilaniantz, A.; Grant, M.; Jones, J.D.G. Hormone crosstalk in plant disease and defense: More than just jasmonate-salicylate antagonism. *Annu. Rev. Phytopathol.* **2011**, *49*, 317–343. [CrossRef] [PubMed]
- Koornneef, A.; Leonreyes, A.; Ritsema, T.; Verhage, A.; Den Otter, F.C.; Van Loon, L.C.; Pieterse, C.M. Kinetics of salicylate-mediated suppression of jasmonate signaling reveal a role for redox modulation. *Plant Physiol.* 2008, 147, 1358–1368. [CrossRef] [PubMed]
- 41. Li, J.; Brader, G.; Palva, E.T. The WRKY70 transcription factor: A node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell* **2004**, *16*, 319–331. [CrossRef] [PubMed]
- Küpper, F.C.; Gaquerel, E.; Cosse, A.; Adas, F.; Peters, A.F.; Müller, D.G.; Kloareg, B.; Salaün, J.P.; Potin, P. Free fatty acids and methyl jasmonate trigger defense reactions in *Laminaria digitata*. *Plant Cell Physiol*. 2009, 50, 789–800. [CrossRef] [PubMed]
- 43. Beckers, G.J.; Spoel, S.H. Fine-tuning plant defence signalling: Salicylate versus jasmonate. *Plant Biol.* **2006**, *8*, 1–10. [CrossRef] [PubMed]
- 44. Kang, S.M.; Min, J.Y.; Kim, Y.D.; Kang, Y.M.; Park, D.J.; Jung, H.N.; Kim, S.W.; Choi, M.S. Effects of methyl jasmonate and salicylic acid on the production of bilobalide and ginkgolides in cell cultures of *Ginkgo biloba*. *In Vitro Cell. Dev. Biol. Plant* **2006**, *42*, 44–49. [CrossRef]
- 45. Kang, M.K.; Nargis, S.; Kim, S.M.; Kim, S.U. Distinct expression patterns of two Ginkgo biloba 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate reductase/isopentenyl diphosphate synthase (HDR/IDS) promoters in Arabidopsis model. *Plant Physiol. Biochem.* **2013**, *62*, 47–53. [CrossRef] [PubMed]

- 46. Liao, Y.L.; Shen, Y.B.; Chang, J.; Zhang, W.W.; Cheng, S.Y.; Xu, F. Isolation, expression, and promoter analysis of *GbWRKY2*: A novel transcription factor gene from *Ginkgo biloba*. *Int. J. Genom.* **2015**, 2015, 607185.
- 47. Lu, G.; Moriyama, E.N. Vector NTI, a balanced all-in-one sequence analysis suite. *Brief Bioinform.* **2004**, *5*, 378–388. [CrossRef] [PubMed]
- 48. Quevillon, E.; Silventoinen, V.; Pillai, S.; Harte, N.; Mulder, N.; Apweiler, R.; Lopez, R. InterProScan: Protein domains identifier. *Nucleic Acids Res.* **2005**, *33*, 116–120. [CrossRef] [PubMed]
- 49. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [PubMed]
- 50. Jansson, S.; Meyer-Gauen, G.; Cerff, R.; Martin, W. Nucleotide distribution in gymnosperm nuclear sequences suggests a model for GC-content change in land-plant nuclear genomes. *J. Mol. Evol.* **1994**, *39*, 34–46. [CrossRef] [PubMed]
- 51. Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 2008, *3*, 1101–1108. [CrossRef] [PubMed]
- 52. Liao, Y.L.; Xu, F.; Zhu, J.; Wang, Y.; Cheng, S.Y. Separation and dtermination of terpene trilactones by gas chromatography with wide bore capillary column. *Acta Agric. Boreali-Occident. Sin.* **2008**, *17*, 146–149.

**Sample Availability:** Samples of *Ginkgo biloba* and cDNAs of *GbAACT* and *GbMVK* are available from the authors.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).