## **Online Resource - Supplementary Information**

## Engineering of thioesterase YciA from *Haemophilus influenzae* for production of carboxylic acids

Laura Pöschel<sup>1,2</sup>, Mónica Guevara-Martínez<sup>3</sup>, David Hörnström<sup>3</sup>, Antonius J.A. van Maris<sup>3</sup>, Markus Buchhaupt<sup>1\*</sup>

<sup>1</sup> DECHEMA-Forschungsinstitut, Microbial Biotechnology, Theodor-Heuss-Allee 25, 60486 Frankfurt am Main, Germany

<sup>2</sup> Faculty of Biological Sciences, Goethe University Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany

<sup>3</sup> Department of Industrial Biotechnology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, AlbaNova University Center, SE 10691 Stockholm, Sweden

\*Corresponding author,

Phone: +49-69-7564-629, e-mail address: markus.buchhaupt@dechema.de

**Fig. S1** Growth kinetics and time-dependent concentration of 2-methylsuccinic acid, 2-hydroxy-3-methylsuccinic acid and mesaconic acid in supernatant of *M. extorquens* AM1 harboring YciAHI variants

**Fig. S2** Growth kinetics and time-dependent concentration of 2-methylsuccinic acid and mesaconic acid in supernatant of *M. extorquens* AM1 harboring plasmids encoding YciAHI FLAG, YciAHI F35L FLAG or an empty pCM160 vector

**Fig. S3** Uncorrected (2*S*)-methylsuccinyl-CoA conversion tests with purified YciAHI and YciAHI F35L thioesterases

**Fig. S4** Growth kinetics and time-dependent concentration of 2-methylsuccinic acid, 2-hydroxy-3-methylsuccinic acid and mesaconic acid in supernatant of *M. extorquens* AM1 harboring YciAHI F35 variants

Fig. S5 Alignment of TE6 family members (α-helix domain)

Fig. S6 Alignment of TE4 family members (α-helix domain)

Fig. S7 Alignment of TE7 family members (presumed α-helix domain)

**Fig. S8** Alignment of TE8 family members (α-helix domain)

**Fig. S9** Alignment of TE11 family members (α-helix domain)

Fig. S10 Alignment of TE31 family members (α-helix domain)



**Fig. S1** Growth kinetics and time-dependent concentration of 2-methylsuccinic acid, 2-hydroxy-3-methylsuccinic acid and mesaconic acid in supernatant of *M. extorquens* AM1 harboring YciAHI variants (untagged). EMCP-derived carboxylic acid products with titers insufficient for quantification (< 5 mg/L) are not displayed. Error bars represent standard deviations from three independent replicates



**Fig. S2** Growth kinetics and time-dependent concentration of 2-methylsuccinic acid and mesaconic acid in supernatant of *M. extorquens* AM1 harboring plasmids encoding C-terminal FLAG-tagged YciAHI, C-terminal FLAG-tagged YciAHI F35L or an empty pCM160 vector. Red squares indicate the sampling time points for SDS-PAGE samples. Error bars represent standard deviations from three independent replicates



**Fig. S3** Uncorrected (2*S*)-methylsuccinyl-CoA conversion tests with purified N-terminal Histagged YciAHI and N-terminal His-tagged YciAHI F35L thioesterases. Values were determined with a DTNB assay at pH 7.5 and 25°C using 1-150  $\mu$ M of (2*S*)-methylsuccinyl-CoA as substrate (color coded concentrations) and 7.5  $\mu$ M of purified enzyme. Prior to addition of the protein, the assay mixture containing all other components was equilibrated to 25 °C. The formation of DTNB-derived 5-thio-2-nitrobenzoate was measured at 412 nm. All measurements were done in triplicates



**Fig. S4** Growth kinetics and time-dependent concentration of 2-methylsuccinic acid, 2-hydroxy-3-methylsuccinic acid and mesaconic acid in supernatant of *M. extorquens* AM1 harboring YciAHI F35 variants (untagged). EMCP-derived carboxylic acid products with titers insufficient for quantification (< 5 mg/L) are not displayed. Error bars represent standard deviations from three independent replicates

|   | α-helix   |
|---|---|
|   |   |
| Conconsus   | C CC M M VATARC   |
| D44996LAset (VeiAUI) OC-Upemenhilus influenzas OV-71401 DDB-1VL                       |   |
| O8WYK01 Acot12 OS=Homo saniens OX=9606 PDR=387K                                       | HGNTEGGOTMAWMETVATTSASRI CWAHPE 30  |
| A0A656FDE91 Acot OS=Yersinia pestis PEXU2 OX=547046 PDB=5DM5                          | NGDIFGGWLMSOMDIGGAIOAKEIAOGRVV 30   |
| A0A2U0QTN1  Acot OS=Campylobacter jejuni OX=197 PDB=3D6L                              | - GNIFGGWILSQIDLAGAIAARELSPERVV 29  |
| Q81EE4  Acot OS=Bacillus cereus ATCC 14579 OX=226900 PDB=1Y7U                         | TLF <mark>GG</mark> KILSEMDMVASISASRHSRKECV 28 Crystal  |
| A0A0H3K033  Acot OS=Staphylococcus aureus Mu50 OX=158878 PDB=4NCP                     | TMFGGTLMANIDEIAAITAMKHAGAQVV 28 structure   |
| Q53VX2  Acot OS=Thermus thermophilus HB8 OX=300852 PDB=2EIS                           |   |
| R8GWI0L Acot OS=Racillus cereus VD196 OX=1053243 PDB=7C73                             | TIEGGKILAETDSTASTAAARHSRKHCV 28   |
| E3D482  Acot OS=Neisseria meningitidis alpha710 OX=630588 PDB=5SZU                    | - GNVHGGELLLLLDQVAYSCASRYSGNYCV 29  |
| Q91V12  Acot7 OS=Mus musculus OX=10090 PDB=6VFY                                       | H <mark>G</mark> FVH <mark>GGVTM</mark> KLMDE <mark>VA</mark> GIVAARHCKTNIV 30                      |
| Q5REE9  Uncharacterized protein OS=Pongo abelii OX=9601                               | QGNTFGGQIMAWMENVATIAASRLCRAHPT 30   |
| H9FH06  Acot11 isoform 2 (Fragment) OS=Macaca mulatta OX=9544                         | QGNTFGGQIMAWMENVATIAASRLCRAHPT 30   |
| I2CW57LAcot11 isoform 2 OS=Macaca mulatta OX=9544                                     | OGNTEGGOTMAWMENVATTAASRICRAHPT 30   |
| A0A1L8GMJ5  Uncharacterized protein OS=Xenopus laevis OX=8355                         | QGNTFGGQIMAWMVNVATIAASRLCHTHPT 30   |
| U3DQ68  Acyl-CoA thioesterase 11 OS=Callithrix jacchus OX=9483                        | QGNTF <mark>GGQIM</mark> AWMENVATIAASRLCRAHPT 30  |
| Q8CAL6  Uncharacterized protein OS=Mus musculus OX=10090                              | QGNTFGGQIMAWMENVATIAASRLCHAHPT 30   |
| Q8VHQ9  Acot11 OS=Mus musculus OX=10090   | QGNTFGGQIMAWMENVATIAASRLCHAHPT 30   |
| OROTTOL ACOLITIOS=BOS (durus OX=9913)   | OGNTEGGOTMAWMENVATTAASRLCRAHPT 30   |
| A2AVR6  Acot11 OS=Mus musculus OX=10090   | OGNTEGGOIMAWMENVATIAASRLCHAHPT 30   |
| K7BZE0  Acot11 OS=Pan troglodytes OX=9598   | QGNTFGGQIMAWMENVATIAASRLCRAHPT 30   |
| A0A6F9D6B2  Acot11 OS=Phallusia mammillata OX=59560                                   | H <mark>GNTFGGQVM</mark> SW <mark>MENVA</mark> S <mark>ISAGRLC</mark> RGFAT 30                      |
| Q501V4  Acot11 OS=Danio rerio OX=7955   | QVNTFGGQIMAWMENVATIAASRLCNAHPT 30   |
| V9K912  Acot11 OS=Callorninchus milli OX=7868   | QGYTEGGQIMAWMEGVAAIAARRECKAPAT 30   |
| O6GM80  Acot12 OS=Xenopus Jaevis OX=10110   | OGNTEGGOIMAWMETVANISASRLCRTHAT 30   |
| B3DLA1  Acot12 OS=Xenopus tropicalis OX=8364  | QGNTFGGQIMAWMEIVANISASRLCHTHAT 30   |
| Q9DBK0  Acot12 OS=Mus musculus OX=10090   | H <mark>GNTFGGQIMAWMETVA</mark> T <mark>ISASR</mark> LCHGHPF 30                                     |
| B5DFR9  Acot12 OS=Xenopus tropicalis OX=8364  | QGNTFGGQIMAWMETVANISASRLCHTHAT 30   |
| A0A0G2KAU8  Acot OS=Rattus norvegicus OX=10116  | - GNTFGGQIMAWMETVATISASRLCHGHPF 29  |
| C4TFE3I Acot OS=Orvzias latipes OX=8090   | OVSTEGGOIMAWMENVATIAASRLCKAHPT 30   |
| A0A6F9D5H3  Acot OS=Phallusia mammillata OX=59560                                     | HGKVTGGVTMKLMDEVAGICAFRHCHSNVV 30   |
| A0A287A7V2  Acot7 OS=Sus scrofa OX=9823   | LNPGVTMKLMDEVAGIVAARHCKTNIV 27  |
| V9L6E3  Acot7 OS=Callorhinchus milii OX=7868  | HGFVQGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| Q64559  Acot OS=Rattus norvegicus OX=10116<br>A0A5G200C4  Acot7 OS=Sus scrofs OX=0922 | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| A0A6D2WMR2L Acot7 isoform 6 OS=Pap troplodytes OX=9598                                | HGEVHGGVTMKLMDEVAGTVAARHCKTNTV 30   |
| Q2I166 Acot OS=Ictalurus punctatus OX=7998  | GVTMKLMDEVAGIVAARHCKTNIV 24   |
| U3F9E4  Acot7 isoform 1 OS=Micrurus fulvius OX=8637                                   | H <mark>GFVHGGVTM</mark> KL <mark>MDEVA</mark> GIVAARHCKTNIV 30                                     |
| A0A287BB59  Acot7 OS=Sus scrofa OX=9823   | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| T1E6N9  Acot7 OS=Crotalus horridus OX=35024   | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| A6MIZ1  Acot (Fragment) OS=Callithrix jacchus OX=9669                                 | FVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| Q0V9M5  Acot OS=Xenopus tropicalis OX=8364  | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| Q66I92  Acot7 OS=Danio rerio OX=7955  | H <mark>GFVHGGVTM</mark> KL <mark>MDEVA</mark> GIV <mark>AARHC</mark> KTNIV 30                      |
| A0A5G2QIZ9  Acot7 OS=Sus scrofa OX=9823   | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| U6DTD5  Acot (BACH) OS=Neogale vison OX=452646  | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| ADA6D2XOC1L Acot7 isoform 1 OS=Pan tradidutes OX=9598                                 | HGEVHGGVTMKLMDEVAGTVAARHCKTNTV 30   |
| A0A5K1VFS3  Acot7 OS=Macaca mulatta OX=9544   | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| Q4R7D6  similar to BACH OS=Macaca fascicularis OX=9541                                | H <mark>GFVHGGVTM</mark> KLMDE <mark>VA</mark> GIVAARHCKTNIV 30                                     |
| Q148I4  Acot7 OS=Bos taurus OX=9913   | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| H9F992  Isoform hBACHa OS=Macaca mulatta OX=9544                                      | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| R4DUX01 Acot OS=Homo saniens OX=9606  | HGEVHGGVTMKLMDEVAGTVAARHCKTNIV 30   |
| K7AYE0  Acot7 OS=Pan troglodytes OX=9598  | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| A0A1L8FMH2  Uncharacterized protein OS=Xenopus laevis OX=8355                         | H <mark>GYVH<mark>GG</mark>VT<mark>M</mark>KL<mark>MDEVA</mark>G<mark>IVAFRHC</mark>KTNIV 30</mark> |
| Q5R8T1  Uncharacterized protein OS=Pongo abelii OX=9601                               | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| FIRIK3 ACOL7 OS=SUS SCIOLA OX=9823  |   |
| A0A0E77CS6LAcot7 OS=Crotalus adamanteus OX=8729                                       | HGEVHGGVTMKI MDEVAGTVAARHCKTNTV 30  |
| U3EQU8  isoform hBACHa OS=Callithrix jacchus OX=9483                                  | HGFVHGGVTMKLMDEVAGIVAARHCKTNIV 30   |
| U3CAA5  isoform hBACHd OS=Callithrix jacchus OX=9483                                  | H <mark>GFVHGGVTM</mark> KLMDE <mark>VA</mark> GIVAARHCKTNIV 30                                     |
| Q5RCZ6  Uncharacterized protein OS=Pongo abelii OX=9601                               | HGFVHGGVTMKLMDEVAGIVAARHCKTIIV 30   |
| V9L288  Acot OS=Callorhinchus milii OX=7868   | HUFVQUGVTMKLMDEVAGIVAARHCKTNIV 30   |

**Fig. S5** Alignment of TE6 family members (α-helix domain). Sequences were acquired from ThYme (<u>https://thyme.engr.unr.edu/v2.0/</u>). Alignment was done in COBALT (Papadopoulos and Agarwala 2007) and visualized with SnapGene (GSL Biotech LLC, San Diego, Canada). Consensus with a threshold of 80% is shown. Amino acids corresponding to consensus are highlighted in green

|  | α-helix   |
|--|---|
| Consensus<br>F1KX25] Acyl-CoA thioesterase 2 OS=Ascaris suum OX=6253 PE=2 SV=1<br>A5U5V6] Acyl-CoA thioesterase II OS=Mycobacterium tuberculosis H37Ra OX=419947 GN=tesB2<br>A0A656FGG1] Acyl-CoA thioesterase II OS=Wycobacterium marinum OX=1781 GN=tesB2<br>B0M0L9] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C17C3.1<br>B0M0L9] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C17C3.1<br>P0A62C]TESB_ECOLI Acyl-CoA thioesterase 2 OS=Mycobacterium avium 104 OX=243243 GN=MAV_2540<br>H2XYT6] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C17C3.1<br>P0A62C]TESB_ECOLI Acyl-CoA thioesterase 2 OS=Escherichia coli K-12 OX=83333 GN=tesB<br>Q73274] Uncharacterized protein OS=Mycobacterium avium subsp. paratuberculosis H37Ra OX=419947 GN=tesB1<br>Q57V11] Acetyl CoA thioesterase I OS=Mycobacterium avium SUBSP. paratuberculosis H37Ra OX=419947 GN=tesB1<br>Q57V11] Acetyl CoA thioesterase I IS ervicobacterium avice IOS=Arabidopsis thaliana OX=3702 GN=C17E1<br>A0A178W7C0] Cyclic nucleotide-binding domain-containing protein OS=6239 GN=C17C3.1<br>Q9110[S] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C17C3.1<br>Q9110[S] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C17C3.1<br>Q9110[S] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C12C3.1<br>Q9110[S] Uncharacterized Protein OS=Caenorhabditis elegans OX=6230 GN=C12C3.1<br>Q9110[S] Uncharacterized Protein OS=Caenor | a-helix   VFG6QGQALAAA-KTVVHS-H-   VYG6LLFAQSLAAAEETVPAQM-RVHAMHS   29<-FG6HVA6QSLVSAVRTVDPKY-NVHSLH628   VFG6QVV6QAIYAAKQTVPAER-TVHSFHS27   IGSQVAQQALWAAQRTVDPRER-VHSFHS27   IGSQVAQQALWAAQRTVDPER-AHSMHS27   IGSQVAQQALWAAQRTVDEQF-KPHSMHS27   IGGQUV6QALWAAQRTVDEQF-KPHSMHS27   IGGQUV6QALWAAQRTVDEQF-KPHSMHS27   IGGGUV6QALWAAQRTVDEQF-KPHSMHS27   VF6GQVV6QALWAAKETVPEER-LVHSFHS29   IGGGHAQSFVASSRTLTRHHLPPSAFSV29   VF6GQFF0QALAAASKTVDFHK-VHSHS27   VF6GQFF0QALAAASKTVDFK-VHSLHS29   VF6GQFF0QALAAASKTVDFHK-VHSHS27   VF6GQFF0QALAAASKTVDFK-VHSLHS29   VF6GQFF0QALAAASKTVDFK-VHSLH29   - GGLIFSQALAAASKTVDFK-VHSLH29   - GGQFF0QALAAASKTVDFK-VVHSLH29   - GGLIFSQALAAASKTVDFK-VVHSLH29   - GGUV6QALAAASKTVDFK-VVHSLH29   - GGUV6QALAAASKTVDFK-VVHSLH29   - GGQFV6QALAAASKTVDFK-VVHSLH29   - GGQFV6QALAAASKTVDFK-VVHSLH29   - GGQFV6QALAAASKTVDFK-VVHSLH29   - GGQFV6QALAAASKTVDFK-VVHSLH29   - VF6QQLV6QALAAASKTVDFK-VVHSLH29   - VF6QQLV6QALAAASKTVDFK-VVHSLH29 |
| 298151 J. Acyl-CoA thioseterase II OS-Yersinia pestis OX=632 GN=tesB<br>E4MWI3] mRNA, clone: RTFL01-11-K11 OS=Eutrema halophilum OX=98038<br>Q95Q68] Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C37H5.13   | VFGGQVVGQAIYAAKQTVPAER-TVHSFHS 29<br>VFGGQFMGQALAAASKTVDFLK-IVHSLHS 29<br>VYGGQVIGQALSAATATVEVGF-VPNSLHS 29   |
| Q95Q68J Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=C37H5.13<br>B3H5Z2J Acyl-CoA thioesterase family protein OS=Arabidopsis thaliana OX=3702 GN=At4g00520<br>B9FBX3J Os04g0558400 protein OS=Oryza sativa Japonica Group OX=39947 GN=0504g0558400<br>OSGYW7J Acyl-CoA thioesterase OS=arabidopsis thaliana OX=3702 GN=At101710/T1N6 21  | VYGGQVIGQALSAATATVEVGF-VPNSLHS 29<br>VYGGQLVGQALAAASKTVETMK-IVHNFHC 29<br>VFGGQLIGQALAAASKTVDCLK-AVHSLHA 29<br>VFGGQFVGQALAAASKTVDCLK-VVHSLHS 29  |
|  |   |

**Fig. S6** Alignment of TE4 family members (α-helix domain). Sequences were acquired from ThYme (<u>https://thyme.engr.unr.edu/v2.0/</u>). Alignment was done in COBALT (Papadopoulos and Agarwala 2007) and visualized with SnapGene (GSL Biotech LLC, San Diego, Canada). Consensus with a threshold of 80% is shown. Amino acids corresponding to consensus are highlighted in green

| Consensus   | IC-PRNFGGFLMR-A-EL                 |
|---|------------------------------------|
| K7GPJ4  Acyl-CoA thioesterase 9 OS=Sus scrofa OX=9823 GN=ACOT9  | SLDICHPQERNIFNRIFGGFLMRRAYELGW 30  |
| T1PG11  Thiosterase OS=Musca domestica OX=7370  | PENRNAHNKVFGGFLMRNASEISW 24        |
| G3MLS1  Uncharacterized protein OS=Amblyomma maculatum OX=34609<br>G6A170211LTE2_0S=Htmukis_humkis_OX=2466  | ICQPESENLYNKVFGGFLMRNAFELAW 27     |
| AGA17932.1] TEZ OS-Huminus lupuus OX-3400<br>069MT11 0S0900517700 protein OS-Grvza sativa Japonica Group OX-39947 GN=0S0900517700                                   | ICOPOORNIHGRIEGGELMHRAFELAF 27     |
| 04V4C3 IP10924p (Fragment) OS-Drosophila melanogaster OX=727 GN=CG1638  | PEHRNQNNTVFGGFLMRSALEISW 24        |
| A0A090XCA2  Putative acyl-coa thioesterase ixodes scapularis acyl-coa thioesterase OS=Ixodes ricinus OX=34613   | ICQPEFENLYNKVFGGFLMRSAYELAW 27     |
| A0A5G2QT21  Acyl-CoA thioesterase 9 OS=Sus scrofa OX=9823 GN=ACOT9  | SLDICHPQERNIFNRIFGGFLMRRAYELGW 30  |
| A0A131XJS1  Putative acyl-coa thioesterase (Fragment) OS=Hyalomma excavatum OX=257692   | ICQPEFENLYNKVFGGFLMRNAFELAW 27     |
| H2QYE01 ACYI-COA Thioesterase 9 OS=Pan trogloodytes OX=9598 GN=ACU19  | SLEICHPUERNIFNRIFGGFLMRKAYELAW 30  |
| 09V9W4I GH08048p OS=Drosophila melanogaster QX=2227 GN=Dmel\CG1774  | PENRNAONTIEGGYLMROAVEISE 24        |
| O16725  Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=CELE_T07D3.9   | EICFPEYQNMYGKIFGGFLMRKALELAH 28    |
| F1LA44  Acyl-coenzyme A thioesterase 9 (Fragment) OS=Ascaris suum OX=6253   | ICFPSQRSVYNKVFGGYIMRIAFELAW 27     |
| F1R961  Acyl-CoA thioesterase 9, tandem duplicate 1 OS=Danio rerio OX=7955 GN=acot9.1   | -LEICHPQERNIFNRIFGGFLMRKAYELGW 29  |
| 12M333 Acyl-coenzyme A thioesterase 9,mitochondrial OS=Hydra Vulgans OX=6087 GN=AC019   | NLIICHPOSRNWYNKIFGGFLMRVAYELAW 30  |
| A0A762F1241 (thate cress) hypothetical protein OS=4 abidonsis thaliana OX=3702 GN=AT9943 1 OCUS21625  | ICOPOORNIHGRIEGGELMHRAFELAF 27     |
| A0A0K8TMD0  Putative acyl-coa thioesterase (Fragment) OS=Tabanus bromius OX=304241  | HPENRNAHNTVFGGFLMRNALEVSW 25       |
| A0A1R7T3J1  Thioesterase/thiol ester dehydrase-isomerase OS=Arabidopsis thaliana OX=3702 GN=At5g48370   | ICQPQQRNIHGRIFGGFLMHRAFELAF 27     |
| Q3SWX2 ACOT9_BOVIN Acyl-coenzyme A thioesterase 9, mitochondrial OS=Bos taurus OX=9913 GN=ACOT9   | SLDICHPQERNIFNRIFGGFLMRKAYELGW 30  |
| MIECBS1 Acyl-CoA thioesterase 9 (Fragment) OS=Mustela putorius furo OX=9669   | SLEVCHPQERNIFNRIFGGFLMRKAYELGW 30  |
| A0A5G9R7V11 AvvI-CoA thioesterase 9, milochonomiane 05-Prianusa manimala 0X-59500 Giv-AcoL9<br>A0A5G9R7V11 AvvI-CoA thioesterase 9, 0S-Sus scrofa 0X-9823 Giv-AcoL9 | SIDTCHPOERNTENRTEGGELMRRAVELGW 30  |
| J3JXX9] Uncharacterized protein OS=Dendroctonus ponderosae OX=77166 GN=D910_10473   | LAQPEDRNLHNTVFGGFIMRHATELSW 27     |
| Q7T175  Acyl-CoA thioesterase 9, tandem duplicate 1 (Fragment) OS=Danio rerio OX=7955 GN=acot9.1  | -LEICHPQERNIFNRIFGGFLMRKAYELGW 29  |
| A9NV70  Uncharacterized protein OS=Picea sitchensis OX=3332   | ICQPQQRNLHGRIFGGFLMHRASELAF 27     |
| F1L8V2  Acyl-coenzyme A thioesterase 9 OS=Ascaris suum OX=6253  | LCFPVDRNVYNKIFGGYLMRLAFELAW 27     |
| AUAAAAVD321 Uncharacteenzee g OSE-pan tronolodvas OXE-9823 GN-ACO19   | SIETCHPOERNTENRTEGGEIMRKAVELGW 30  |
| A0A5G20HL2LAcvI-CoA thioesterase 9 OS=Sus scrofa 0X=9823 GN=AC0T9   | SLDICHPOERNIENRIEGGELMRRAYELGW 30  |
| Q503Q9] Acyl-CoA thioesterase 9 OS=Danio rerio OX=7955 GN=acot9.1   | -LEICHPQERNIFNRIFGGFLMRKAYELGW 29  |
| A0A654G9F2  Uncharacterized protein OS=Arabidopsis thaliana OX=3702 GN=AN1_LOCUS25073   | ICQPQQRNIHGRIFGGFLMHRAFELAF 27     |
| F1SQ00  Acyl-CoA thioesterase 9 OS=Sus scrofa OX=9823 GN=ACOT9  | SLDICHPQERNIFNRIFGGFLMRRAYELGW 30  |
| A0A023EV021 Putative acyl-coa thioesterase OS=Aedes albopictus OX=760   | HPEDRNAHNKVFGGFLMRNALELSW 25       |
| AGAD25G6A41 PGGGUE GAY-COG UNDESCERSE OS-ANIDAYONINA DISC OX-221400   | ICOPOORNIHGRIEGGELMRKAFELAF 27     |
| Q95TK5  LD44914p OS=Drosophila melanogaster OX=7227 GN=Dmel\CG1635  | PEHRNHHNRVFGGFLMRSALEISW 24        |
| V5HF46  Putative acyl-coa thioesterase 9 (Fragment) OS=Ixodes ricinus OX=34613  | ICQPEFENLYNKVFGGFLMRSAYELAW 27     |
| A0A2R8QMR4  Acyl-CoA thioesterase 9, tandem duplicate 1 OS=Danio rerio OX=7955 GN=acot9.1   | -LEICHPQERNIFNRIFGGFLMRKAYELGW 29  |
| F1R5C2  Acyl-coA thioesterase 9, tandem duplicate 1 OS=Danio reno OX=7955 GN=ac09.1   | - LEICHPQERNIFNRIFGGFLMRKAYELGW 29 |
| Vocaça a Antoniz Inica a dibesterase y micorionario 05-cerados capitata 0x-7215 (N-ACO19<br>AGA17034.11 TEA OS-Humitis lunitis (X=3486                              | TCOPOORNIYGRIEGGELMRRAVELAE 27     |
| Q9H2R8  CGI16-iso OS=Homo sapiens OX=9606   | SLEICHPQERNIFNRIFGGFLMRKAYELAW 30  |
| H7C5Q2  Acyl-coenzyme A thioesterase 9, mitochondrial (Fragment) OS=Homo sapiens OX=9606 GN=ACOT9   | SLEICHPQERNIFNRIFGGFLMRKAYELAW 30  |
| A0A131XKN3  Putative acyl-coa thioesterase (Fragment) OS=Hyalomma excavatum OX=257692   | ICQPEFENLYNKVFGGFLMRNAFELAW 27     |
| Q498L3  MGC114623 protein OS=Xenopus laevis OX=8355 GN=acot9.5  | - LEICHPQERNIFNRIFGGFLMRKAYELGW 29 |
| Variati Acylocenzyme a dnoesterase 9 mitochondrial (Frament) OS=Rhinicenhalus pulchellus OX=72859   | ICOPESENI YNKVEGGELMRNAFELAW 27    |
| Q32PS11 Acyl-CoA thioesterase 9, tandem duplicate 2 OS=Danio rerio OX=7955 GN=acot9.2   | -LEICHPQERNIFNRIFGGFLMRKAYELGR 29  |
| U6DQW3  Acyl-coenzyme A thioesterase 9, mitochondrial (Fragment) OS=Neogale vison OX=452646 GN=ACOT9  | SLEVCHPQERNIFNRIFGGFLMRKAYELGW 30  |
| K9IS74  Putative acyl-coa thioesterase (Fragment) OS=Desmoduls rotundus OX=9430   | SLEICHPQERNIFNRIFGGFLMRKAYELGW 30  |
| QSU2X8  Acyl-CoA thioesterase 9 OS=Rattus norvegicus OX=10116 GN=Acot9  | SLDICHPQERNIFNRIFGGFLMRKAYELAW 30  |
| B72002] Oncharacterized protein OS=Danio Teno OX=7935 GN=46053.1  | SIDICHPOERNVENRIEGGELMRKAVELAW 30  |
| OPROX4/ACOT9 MOUSE Acvi-coenzyme A thioesterase 9, mitochondrial OS-Mus musculus OX=10090 GN=Acot9  | SLDICHPQERNVFNRIFGGFLMRKAYELAW 30  |
| C6T9W5  Uncharacterized protein OS=Glycine max OX=3847  | ICQPQQRNIHGRIFGGFLMRRAFELAF 27     |
| Q94245  Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=CELE_F57F4.1   | EICFPEYQNMYGKIFGGFLMRKALELAY 28    |
| A0A6A7GC351 HotDog domain-containing protein OS=Hirondellea gigas OX=1518452  | SILMTQPQANNAQDKIFGGYLMRKAFELGR 30  |
| F2ECL1  Predicted protein (Fragment) US=Hordeum Vuigare subsp. Vuigare UX=112509<br>B1WB361 Aorde protein OS=Vapapus tronicalis OX=8364 GN=2019                     | ICQPQQRNLYGRIFGGFLMHRAFELAF 27     |
| Q23044  Uncharacterized protein OS=Caenorhabditis elegans OX=6239 GN=CELE_T22B7.7   | IAQPEHENPYGSVFGGFLVRKGLETAE 27     |
| Q8T3X0  AT27044p OS=Drosophila melanogaster OX=7227 GN=CG13771  | PENRNESNTIFGGFIIRKAIEISY 24        |
| K9K318  Mitochondrial acyl-coenzyme A thioesterase 9-like protein (Fragment) OS=Equus caballus OX=9796  | SLEICHPQERNIFNRIFGGFLMRKAYELGW 30  |
| A0A0F72E06  Acyl-coenzyme A thioesterase 9, mitochondrial-like OS=Crotalus adamanteus OX=8729   | - LEICHPQERNIFNRIFGGFLMRKAYELGW 29 |
| BUZKYI ACUCA PROLEIN OS-DANO TENO OXA-7535 GN-46059.1   | -LEIGHPUERNIPNKIPOGPUMKKAYELGW 29  |
| or toring help our differences putative op-realizage diandatula OA-3000 GN-11442332   | TONT QUALITOR FOUR UNAFELAF 2/     |

**Fig. S7** Alignment of TE7 family members (presumed  $\alpha$ -helix domain). Sequences were acquired from ThYme (<u>https://thyme.engr.unr.edu/v2.0/</u>). Alignment was done in COBALT (Papadopoulos and Agarwala 2007) and visualized with SnapGene (GSL Biotech LLC, San Diego, Canada). Consensus with a threshold of 80% is shown. Amino acids corresponding to consensus are highlighted in green

|  | α-he               | lix                     |            |    |
|--|--------------------|-------------------------|------------|----|
|  |                    |                         |            |    |
| Consensus  | L-GGA-LVD-         | A                       |            |    |
| Q18187  4HBT domain-containing protein OS=Caenorhabditis elegans OX=6239 GN=C25H3.3  | FNTLHGGCTSTLIDI    | FTTGALLL                | TKP AR     | 28 |
| T2M4G7  Acyl-coenzyme A thioesterase 13 (Fragment) OS=Hydra vulgaris OX=6087 GN=ACOT13   | - GTLHGGYTSFLADY   | TTSIALAA                | - IN - NKN | 27 |
| A0A069DNT8  Putative hgg motif-containing thioesterase (Fragment) OS=Panstrongylus megistus OX=65343   | - GSLHGGFTSTLVDI   | LSGVALCT                | HKY-YPP    | 28 |
| A0A1E1WZG8  Putative paai thioester (Fragment) OS=Amblyomma aureolatum OX=187763   | -NTLHGGMATALIDI    | CTCVLMST                | AYE-ERL    | 28 |
| Q9V225) GE012033D1 OS=Drosopnila metanogaster OX=7227 GN=Dmet(C616986  | LHGGYIMILVDL       | TATATTP                 | AVE TOK    | 24 |
| A0A023F2451 Putative had motif-containing thioesterase (Fragment) OS=Triatoma infestans OX=30076   | - GSLHGGLTSTLVDV   | SGVALCT                 | HKY-YPP    | 28 |
| Q4QPU9  IP04554p (Fragment) OS=Drosophila melanogaster OX=7227   | LHGGYIMTLVDL       | ITTYALMS                | - KP CH    | 24 |
| A0A0C9SAM8  Putative paai thioester OS=Amblyomma americanum OX=6943  | - NTLHGGMATTLIDI   | CTCVLMST                | AYE-EKM    | 28 |
| H2QSD9  Acyl-CoA thioesterase 13 OS=Pan troglodytes OX=9598 GN=ACOT13  | IGTLHGGLTATLVDN    | ISTMALLC                | - TE - RGA | 28 |
| QODER5  Acyl-CoA thioesterase 13 OS=Xenopus tropicalis OX=8364 GN=acot13   | - GTLHGGLTATLVDT   | VSTVALLH                | - TE - RGA | 27 |
| 135FW/14HB1 domain-containing protein OS=Lotus Japonicus OX=34305<br>ONDP131 And-consymp A thioasterize 13 OS=Homo sphips OX=9606 GN=4COT13                                      | TOTI HOGI TATI VDN | TSTMALLC                | TE-RGA     | 28 |
| A0A178W7W71 (thale cress) hypothetical protein OS=Arabidopsis thaliana OX=3702 GN=At1004290  | LHGGATATLVDL       | IGSAVIYT                | AGASH      | 25 |
| F6P1Y9  Acyl-CoA thioesterase 13 OS=Danio rerio OX=7955 GN=acot13  | - GTLHGGMTATLVDM   | ISTMAIMY                | - SE - RGA | 27 |
| A0A6F9D5W1  Acyl-coenzyme A thioesterase 13-like OS=Phallusia mammillata OX=59560 GN=Acot13  | - GNLHGGFTATLVDV   | STYALGT                 | - LD - DPR | 27 |
| A0A0F7ZE08  Acyl-coenzyme A thioesterase 13-like OS=Crotalus adamanteus OX=8729  | - GTLHGGLTATLVDI   | VSTAALMY                | - SE - RGA | 27 |
| A6QQ83  THEM2 protein OS=Bos taurus OX=9913 GN=THEM2   | MGTLHGGMIATLVDV    | ISSLALLC                | - TE - RGI | 28 |
| OSEIXS1 4HB1 domain-containing protein OS=Corectriella appendiculata OX=13/0023<br>ANAN121G141 Acv1-conpryme A thiosesterses 13 (fragment) OS=Bseudodiantomus poplasia OX=213370 | MGTLHGGESAHLVDT    | TTMALMS                 | - NP - FGR | 27 |
| G712N81 AcvI-CoA thioesterase, putative OS=Medicago truncatula OX=3880 GN=11419954   | SLHGGATAALVDV      | GSAAIPA                 | SGYLGRN    | 28 |
| W5UTE9  Acyl-coenzyme A thioesterase 13 OS=Ictalurus punctatus OX=7998 GN=Acot13   | - GTLHGGLTATLVDV   | ISTTAFLY                | - TE - RGA | 27 |
| F1LEU7  Acyl-coenzyme A thioesterase 13 OS=Ascaris suum OX=6253  | - GTLHGGCTATLVDI   | VTTT <mark>A</mark> LMA | - TE - RAH | 27 |
| A9ULW5  LOC100137631 protein OS=Xenopus laevis OX=8355 GN=acot13.S   | - GTLHGGLTATLVDT   | VSTVALLH                | - TE - RGA | 27 |
| TIPG86(Thioesterase OS=Musca domestica OX=7370 GN=101888219  | LGGLHGGFTATLVDM    | VTTYALMS                | - KP CH    | 27 |
| QSR833 AC/I-coenzyme A thioesterase 13 05=20ngo abelli 0X=9011 GN=AC0113   | IGILHGGLIAILVDN    | TTYALMT                 | - TE - RGA | 28 |
| G8G915I AcvI-coenzyme A thioesterase 13 OS=Epinephelus coioides OX=94232   | - GTLHGGMTATLIDD   | ISTLAIMY                | -SE-RGA    | 27 |
| I3S9K1  4HBT domain-containing protein OS=Lotus japonicus OX=34305   | SLHGGATATLVDV      | GSAAIPT                 | VGYSSAS    | 28 |
| Q1HPG9  Thioesterase superfamily member 2 OS=Bombyx mori OX=7091   | - GTLHGGFIAHLVDA   | ISTYALTT                | - NENVDT   | 28 |
| M1EB28  Acyl-CoA thioesterase 13 (Fragment) OS=Mustela putorius furo OX=9669   | LGTLHGGMTATLVDN    | ISTIALLC                | - TE - RGA | 28 |
| D3ZA93  Acyl-CoA thioesterase 13 OS=Rattus norvegicus OX=10116 GN=Acot13   | FGTLHGGLTATLVDS    | ISTMALMC                | - TE - RGA | 28 |
| P34419 PdLatve esterase F42F10.0 OS=Caelionabulus elegans 0A=023 GN=42F10.0  | LGTI HGGI TATI TDS | ISTEALLC                | - TE - RGA | 28 |
| A9NZ49  4HBT domain-containing protein OS=Picea sitchensis OX=3332   | LHGGAIAAFVDI       | IGSAAIFT                | TGAKS      | 25 |
| G0ZRI5  Thioesterase superfamily member 2 OS=Penaeus chinensis OX=139456   | - GTLHGGLTATLVDI   | STMALMT                 | - TE - KAV | 27 |
| Q9CQR4  Acyl-coenzyme A thioesterase 13 OS=Mus musculus OX=10090 GN=Acot13   | LGTLHGGLTATLVDS    | ISTMALMC                | - TE - RGA | 28 |
| FIRUE0  Acyl-CoA thioesterase 13 OS=Sus scrota OX=9823 GN=ACOT13   | MGTLHGGMTATLVDC    | STYALLC                 | - TE - RGA | 28 |
| FIOYO9 Acyl-CoA thioesterase 13 OS=Danio rerio OX=7955 GN=acot13   | - GTLHGGMTATLVDM   | ISTMATMY                | - SE-RGA   | 28 |
| FLBY51 Esterase OS-Ascaris suum OX=6253  | TLHGGOTAALVDM      | ITARAAGI                | -TI-KDR    | 26 |
| C6TBG1  4HBT domain-containing protein OS=Glycine max OX=3847  | SLHGGAIAALVDV      | AGSAAIPT                | VGYSAPN    | 28 |
| A0A2R8QNU6  Acyl-CoA thioesterase 13 OS=Danio rerio OX=7955 GN=acot13  | - GTLHGGMTATLVDM   | ISTMAIMY                | - SE - RGA | 27 |
| B5FXY0  Putative thioesterase superfamily member 2 OS=Taeniopygia guttata OX=59729   | - GTLHGGLTATLVDV   | STAALLY                 | - TE - RAV | 27 |
| AUAU23EF/3] Putative cpiju16/40 conserved protein OS=Addes albopictus OX=/160  | - GGLHGGFTATIVDV   | TSVALMS                 | - KD - NCL | 27 |
| TIF7281 4HBT domain-containing protein OS=Anopheles aguasalis OX=42839   | - GGLHGGYTATIVDV   | TTYALMT                 | -KE-NAV    | 20 |
| A0A023EGG4  Putative cpij016740 conserved protein (Fragment) OS=Aedes albopictus OX=7160   | - GGLHGGFTATIVDV   | TTYALMT                 | - KD - NCL | 27 |
| Q9VZZ6  CG16985 protein OS=Drosophila melanogaster OX=7227 GN=BcDNA:RE30174  | - GTLHGGLTATIVDN   | CTTYALMS                | - KG SH    | 26 |
| A0A0P6IXR1  4HBT domain-containing protein OS=Aedes aegypti OX=7159  | - GGLHGGFTATIVDV   | VTTYALMT                | - KE - NCL | 27 |
| A0A023EFU9  Putative cpij016/40 conserved protein (Fragment) OS=Aedes albopictus OX=/160   | - GGLHGGFTATIVDV   | STTALLN                 | - KD - NCL | 27 |
| R400b3 (middesterabe superiarinity memoer 2 05=Calionninchus mini 0x=7806 Giv=dot13<br>004SO11 At1o04200/E10P10 27 OS=Arabidonsis thaliana 0X=3702 Giv=At1o04290                 | LHGGATATIVD        | TGSAVIVI                | AGASH      | 25 |
| B9ELN1 Thioesterase superfamily member 2 OS=Salmo salar OX=8030 GN=THEM2   | - GTLHGGLTATLVDV   | ISTTAINY                | - TE - RGA | 27 |
| A9PCW5  4HBT domain-containing protein OS=Populus trichocarpa OX=3694 GN=POPTR_004G134066  | LHGGATATLVDL       | GSAAIFT                 | VGAPA      | 25 |
| Q8IG45  4HBT domain-containing protein OS=Caenorhabditis elegans OX=6239 GN=C25H3.14   | TLHGGCTAALIDC      | FTTG <mark>A</mark> LLL | TKE AR     | 26 |
| AOAOK8RL60  4HBT domain-containing protein OS=Ixodes ricinus OX=34613  | MGSLQGGMAATLVDV    | ISTYALLT                | L - RDV    | 27 |
| V5IGA8  Putative hgg motif-containing thioesterase OS=Ixodes ricinus OX=34613  | - GSLQGGMAATLVDI   | ISTYALLT                | L-RDV      | 26 |
| VSGWEDI Putative hos notificantialing protein 05-ixodes ricinus 0X=34613   | MGSLOGGMAATLVDV    | ISTYALLT                | L-RDV      | 27 |
| A0A0K8RL72  4HBT domain-containing protein OS=Ixodes ricinus OX=34613  | - GSLQGGMAATLVDV   | ISTYALLT                | L - RDV    | 26 |
| A0A023G703  4HBT domain-containing protein OS=Amblyomma triste OX=251400   | LGTMQGSFSASLVDV    | ISTYALLT                | L - R D V  | 27 |
| A0A131X9S4  4HBT domain-containing protein OS=Hyalomma excavatum OX=257692   | LGTLQGSFSASLVDV    | ISTYALLT                | L - R D V  | 27 |
| A0A1E1WXB2  4HBT domain-containing protein (Fragment) OS=Amblyomma aureolatum OX=187763  | LGTMQGSFSASLVDV    | ISTYALLT                | L-RDV      | 27 |
| ANAN23G9P51 4HBT domain-containing protein OS=Amblyomma triste OY=251400   | LGTMQGSFSASLVDV    | ISTYALLT                | L-RNV      | 27 |
| 07XV63I 0s04q0436100 protein OS=0ryza sativa Japonica Group 0X=39947 GN=0s04q0436100   | - GYLLSGVTATLADO   | LGSGVFLS                | SGIGT      | 27 |
| A0A8J8Y0M0  Acyl-coenzyme A thioesterase 13 OS=Zea mays OX=4577 GN=Acot13_1  | VGYLRSGVTATLADQ    | GSAVFFC                 | SGIPS      | 28 |
| B6UFQ3  Thioesterase superfamily member 2 OS=Zea mays OX=4577 GN=100286193   | - GYLRSGVTATLADQ   | LGSAVFFC                | SGLPS      | 27 |
| F2DNV6  Predicted protein OS=Hordeum vulgare subsp. vulgare OX=112509  | - GYLLSDVTASLADQ   | LGSAVFFS                | SG VGT     | 27 |

**Fig. S8** Alignment of TE8 family members (α-helix domain). Sequences were acquired from ThYme (<u>https://thyme.engr.unr.edu/v2.0/</u>). Alignment was done in COBALT (Papadopoulos and Agarwala 2007) and visualized with SnapGene (GSL Biotech LLC, San Diego, Canada). Consensus with a threshold of 80% is shown. Amino acids corresponding to consensus are highlighted in green

|  | ————————————————————————————————————— |   |
|--|---------------------------------------|---|
|  |                                       |   |
| Consensus  | LHGG-SES                              |   |
| B4XYA6  Azi13 OS=Streptomyces sahachiroi OX=285525 GN=azi13  | HGILHGGVHCAVVESVASAAADRWLGDR 28       | 8 |
| B6U112  4HBT domain-containing protein OS=Zea mays OX=4577   | VLNGGVSALVAESTASVGAFMASG 24           | 4 |
| AAF76242.1  FcbC2 OS=Arthrobacter sp. TM1 OX=75752   | - GLVHGGAYCALAEMLATEATVAVVHEK 27      | 7 |
| C6T4X8  4HBT domain-containing protein OS=Glycine max OX=3847 GN=100527631   | - KVLHGGVSALVAESLASIGAHMASG 25        | 5 |
| Q6AV57  Os03g0691400 protein OS=Oryza sativa Japonica Group OX=39947 GN=LOC_Os03g48480   | - KVLHGGVSALIAEGLASMGAHMASG 25        | 5 |
| Q9FI76 DNAT2_ARATH 1,4-dihydroxy-2-naphthoyl-CoA thioesterase 2 OS=Arabidopsis thaliana OX=3702 GN=DHNAT2  | - KVLHGGVSALIAEALASLGAGIASG 2         | 5 |
| A9NTP7  4HBT domain-containing protein OS=Picea sitchensis OX=3332   | VLHGGISAFLAESLGSLGAIIASG 24           | 4 |
| B7FMY5[ 4HBT domain-containing protein OS=Medicago truncatula OX=3880  | -KVLHGGVSAMISESLASIGAHIACG 25         | 5 |
| A9NKE01 4HBT domain-containing protein OS=Picea sitchensis OX=3332   | VLHGGISAFLAESLGSLGAFITSG 24           | 4 |
| 1Q41 B crystal structure of 4-hydroxybenzoyl CoA thioesth 4-hydroxyphenyl CoA OS=Arthrobacter sp. SU OX=/1255  | - GLVHGGAYCALAEMLATEATVAVVHEK 2       | 7 |
| AUADHZWYZ8 COMAZ Taminy protein US=Staphylococcus aureus subsp. aureus COL 0X=93062 GN=SACOL0947   | FGYLHGGASTALGETACSLGSANLIDIT 28       | 8 |
| AUA178U9Y4 DHNA12 OS=Arabidopsis trainana OX=3702 GN=At5g48950   | - KVLHGGVSALIAEALASLGAGIASG 25        | 5 |
| ASPOND 4HB1 domain-containing process of the statements of a 332   | VLHGGVSAFLAESLGSLGAYMASG 24           | 4 |
| P2E022 Predicted protein OS=hordedin Vulgare Subsp. Vulgare OA=112309  |                                       | 2 |
| CODECAL (AURT domain_constaining protein OS=7a3 mays OY=4577   |                                       | 5 |
| 00441614HBT_ARTSP_4_budgavbergroup_Cod_btioesterses_OS=atthrobarter.sp_OX=1667_GN=frbC   | - GLVHGGAVCALAEMLATEATVAVVHEK 2       | 7 |
| Q04410[4161_AKT39 4410[416][416][416][416][416][416][416][416]   | -KVLHGGVSSLTSEGLASMGAHTASG 25         | 5 |
| CA177823.11 4-chlorobenzovi CoA thioesterase OS=Arthrobacter sp. EG1 OX=405744   | - GLVHGGAYCALAEMLATEATVAVVHEK 2       | 7 |
| OCOPDS61 14-dihydroxy-2-naphthoyl-CoA thioesterase 1 OS=Zea mays OX=4577 GN=100383169  | -KVLHGGVSALVAEALASMGAHMASG 25         | 5 |
| COPCH8  4HBT domain-containing protein OS=Zea mays OX=4577   | LNGGVSALMAETTASIGGYVASG 23            | 3 |
| C6T483] 4HBT domain-containing protein OS=Glycine max OX=3847 GN=100527403   | - KVLHGGVSALIAESLASMGAHMASG 25        | 5 |
| F2DIJ2  Predicted protein OS=Hordeum vulgare subsp. vulgare OX=112509  | - KVLHGGVSALVSEGLASMGAHMASG 25        | 5 |
| 13SG93  4HBT domain-containing protein OS=Lotus japonicus OX=34305   | -KVLHGGVSAMIAESLASMGAHMASG 25         | 5 |
| F2D139  Predicted protein OS=Hordeum vulgare subsp. vulgare OX=112509  | - KVLHGGVSALIAEGLASMGAHMASG 25        | 5 |
| G7IQD5  Acyl-CoA thioesterase OS=Medicago truncatula OX=3880 GN=11441252   | -KVLHGGVSAMISESLASIGAHIACG 25         | 5 |
| P0A8Y9 ENTH_ECOL6 Proofreading thioesterase EntH OS=Escherichia coli CFT073 OX=199310 GN=entH  | FGLLHGGASAALAETLGSMAGFMMTRD- 27       | 7 |
| A0A2P1CQE3  Thioesterase 2 OS=Petunia x hybrida OX=4102  | -QVLHGGVSALIAESLASMGAHVACG 25         | 5 |
| P9WIM3 Y1847_MYCTU Putative esterase Rv1847 OS=Mycobacterium tuberculosis H37Rv OX=83332 GN=Rv1847   | - GVVHGGVYCAMIESIASMAAFAWLNSHGE 29    | 9 |
| Q9SX65 DNAT1_ARATH 1,4-dihydroxy-2-naphthoyl-CoA thioesterase 1 OS=Arabidopsis thaliana OX=3702 GN=DHNAT1  | -KVLHGGVSALIAESLASMGAHMASG 2          | 5 |
| B6TLJ7  Acyl-CoA thioesterase/ catalytic/ hydrolase, acting on ester bonds OS=Zea mays OX=4577   | -QVLHGGVSALVAEGLASMGAHMASG 25         | 5 |
| 135426  4HBT domain-containing protein OS=Medicago truncatula OX=3880  | VLHGGVSAVIAEALASIGAHVACG 24           | 4 |
| AUA2PICQU/  Inioesterase 1 05=Petunia x hybrida 0x=4102  | -KVLHGGISALIAESLASMGAHIASG 25         | 5 |
| BBLQG214HB1 domain-containing protein OS=Picea sitchensis OX=3322  | - KVLHGGVSSLISEGLASMGAHIASG 2         | 5 |
| Q913A4111016_PSEAE Putative esterase PA1018 OS=PSeudonionas deruginosa PA01 0X=208904 GN=PA1018  | CLUBORTYCAL AEMLATEATVAVV HEK 2       | B |
| CODET1114-dibydrowy-2-paphboyLCoA thiosetarase 1 OS=72a mays OY=4577 GN=100283760  | -OVI HOOVSALVAEGLASMGAHMASG           | 5 |
| D45083/11161 HAFIN Dirtative esterase HT 1161 OS Haemonbilus influenza Rd KW/20 OY=71421 GN=HI 1161  | EGVLHGGVSVALAETTGSLAGSLCLEE 2         | 7 |
| A0A6A7G8H0I Esterase OS=Hirondellea rigas OX=1518452   | EGVLHGGASVVLAETLGSVAGNMATDED 28       | 8 |
| P77781IMENI ECOLI 1.4-dihydroxy-2-nanhthoyl-CoA hydrolase OS=Escherichia coli K-12 OX=83333 GN=ment  | EGILHGGASVVLAESTGSVAGVLCTEGE 2        | 8 |
| F2CSG6I Predicted protein OS=Hordeum vulgare subsp. vulgare OX=112509  | -KVLHGGVSALVSEGLASMGAHMASG 2          | 5 |
| The second second standard second standard second standard second s |                                       | - |

**Fig. S9** Alignment of TE11 family members ( $\alpha$ -helix domain). Sequences were acquired from ThYme (<u>https://thyme.engr.unr.edu/v2.0/</u>). Alignment was done in COBALT (Papadopoulos and Agarwala 2007) and visualized with SnapGene (GSL Biotech LLC, San Diego, Canada). Consensus with a threshold of 80% is shown. Amino acids corresponding to consensus are highlighted in green



**Fig. S10** Alignment of TE31 family members (α-helix domain). Sequences were acquired from ThYme (https://thyme.engr.unr.edu/v2.0/). Alignment was done in COBALT (Papadopoulos and Agarwala 2007) and visualized with SnapGene (GSL Biotech LLC, San Diego, Canada). Consensus with a threshold of 80% is shown. Amino acids corresponding to consensus are highlighted in green

## **Reference list**

Papadopoulos JS, Agarwala R (2007) COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics 23:1073–1079. https://doi.org/10.1093/bioinformatics/btm076