

TROPICAL AGRICULTURAL SCIENCE

Journal homepage: http://www.pertanika.upm.edu.my/

In-silico and Phylogenetic Analysis of Acetate: Succinate COA-Transferase (ASCT) from *Angiostrongylus malaysiensis*

Quincie Sipin¹, Suey Yee Low¹, Wan Nur Ismah Wan Ahmad Kamil², Kiew-Lian Wan³, Mokrish Ajat⁴, Juriah Kamaludeen^{5,6}, Sharifah Salmah Syed-Hussain⁷, Nur Indah Ahmad¹ and Nor Azlina Abdul Aziz^{1*}

¹Department of Veterinary Pathology and Microbiology, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400 Serdang, Malaysia
 ²Department of Microbiology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 Serdang, Malaysia
 ³Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
 ⁴Department of Veterinary Preclinical Science, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400 Serdang, Malaysia
 ⁵Department of Animal Science and Fishery, Faculty of Agriculture Science and Forestry, Universiti Putra Malaysia Bintulu Sarawak Campus, 97008 Bintulu, Sarawak, Malaysia
 ⁶Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
 ⁷Department of Veterinary Clinical Studies, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400

'Department of Veterinary Clinical Studies, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400 Serdang, Malaysia

ABSTRACT

The zoonotic capability of *Angiostrongylus malaysiensis* was recently observed after several years of doubt. This parasite was found in a high burden of Malaysian rats, which is alarming. There is currently no effective treatment for human neuroangiostrongyliasis. Acetate: succinate

ARTICLE INFO

Article history: Received: 26 July 2024 Accepted: 01 October 2024 Published: 16 May 2025

DOI: https://doi.org/10.47836/pjtas.48.3.06

E-mail addresses:

quinciesipin@gmail.com (Quincie Sipin) eulynnlow96@gmail.com (Suey Yee Low) wn_ismah@upm.edu.my (Wan Nur Ismah Wan Ahmad Kamil) klwan@ukm.edu.my (Kiew-Lian Wan) mokrish@upm.edu.my (Mokrish Ajat) juriahk@upm.edu.my (Juriah Kamaludeen) ssalmah@upm.edu.my (Nur Indah Ahmad) azlinaaziz@upm.edu.my (Nor Azlina Abdul Aziz) *Corresponding author CoA-transferase (ASCT) enzyme catalyses acetate production in helminth parasites. ASCT was classified into three subfamilies within the family I CoA-transferases (IA, IB, and IC). Acetate is an essential metabolic end product of many parasites, making it an attractive drug target since it is absent in mammalian hosts. The current study describes the in-silico analysis conducted for the identification and phylogenetic characterisation of *A. malaysiensis* ASCT and genetic variations between subfamilies of ASCT. The *Am*ASCT was identified from the ongoing de novo transcriptome assembly and annotation

of adult *A. malaysiensis*. The analysis of *Am*ASCT physiochemical properties, multiple sequence alignment and phylogenetic relations with the ASCTs of other helminths are conducted using standard bioinformatic tools. Pairwise comparisons between subfamilies of ASCT have also been conducted in silico. *Am*ASCT has the conserved regions of the family I CoA-transferases and is clustered with subfamily IB of ASCT. From the pairwise analysis, subfamilies IB and IC were most closely related between the three subfamilies. *Am*ASCT was predicted to be overall hydrophilic and stable in a neutral to slightly alkaline environment within the parasite. The phylogenetic analysis confirmed that *Am*ASCT belongs to subfamily IB of ASCTs. Further study on the biochemical activity of ASCT in *A. malaysiensis* is required to determine its enzymatic function.

Keywords: Acetate: succinate CoA-transferase (ASCT), acetate production, Angiostrongylus malaysiensis

INTRODUCTION

Human neuroangiostrongyliasis, caused by the rat lungworm, is a food-borne parasitic zoonosis distributed worldwide and endemic to Asia and the Pacific Basin. However, it has been reported in new regions beyond its traditional endemic range (Eamsobhana, 2014). The prime causative agent of human neuroangiostrongyliasis is *Angiostrongylus cantonensis* (Wang et al., 2008). Recently, *A. malaysiensis*, formerly known as the Malaysian strain of *A. cantonensis*, was revealed to be zoonotic after several years of doubt (Watthanakulpanich et al., 2021), increasing the risk of human neuroangiostrongyliasis infection. The newly found zoonotic capability of *A. malaysiensis* is alarming, given its high prevalence in rats in Malaysia (Low et al., 2023). There is currently no effective treatment for human neuroangiostrongyliasis. Current treatment focuses on alleviating pain, inflammation, and intracranial pressure in infected individuals (Sawanyawisuth et al., 2008; Slom et al., 2002). Although anthelmintics have been extensively used, their effectiveness is debatable (Murphy et al., 2013), as they leave the disease without an efficient cure. A new, effective approach is needed for treating this disease.

The complex life cycle of parasites enables them to inhabit more than one host organism throughout their life cycle. Most parasites have a complex life cycle, including free-living and distinct stages inhabiting one or more host organisms. As a result, parasites are forced to modify their metabolism to survive in habitats where nutrients and oxygen are scarce. These parasites have evolved metabolic strategies that produce acetate from acetyl-CoA as an essential metabolic product (Tielen et al., 2002). Two pathways of acetate formation from acetyl-CoA have been found in parasites and are catalysed by either a cytosolic acetyl-CoA synthetase (ACS) or an organellar acetate: succinate CoA-transferase (ASCT). Furthermore, the ACS pathway has been reported only in parasitic protists (Reeves et al., 1977; Sánchez et al., 2000; Stechmann et al., 2008), while the ASCT pathway was previously found in both helminths and protists (Steinbüchel & Müller, 1986; Saz et al., 1996; van Hellemond et al., 1998).

To date, all acetate/succinate CoA-transferases (ASCTs) found in eukaryotes are family I CoA-transferases that are distinguished by the catalysis of CoA group transfer, which involves the presence of a glutamate residue in the enzyme's active site (Heider, 2001). Recently, the corresponding ASCT genes were identified and characterised in two protists, *Trypanosoma brucei* (Rivière et al., 2004) and *Trichomonas vaginalis* (van Grinsven et al., 2008), and in the helminth *Fasciola hepatica* (van Grinsven et al., 2009). The ASCT genes from the three parasites showed little sequence homology, classifying them into three different subfamilies of family I CoA-transferases: *T. brucei* in subfamily IA, *F. hepatica* in subfamily IB, and *T. vaginalis* in subfamily IC (Tielens et al., 2010).

Acetate is a prime candidate for developing new antiparasitic medications because it is a crucial byproduct of the energy metabolism of many parasites absent in their mammalian hosts. Identifying and characterising acetate-producing enzymes in rat lungworms is crucial for developing new and effective drugs to treat human neuroangiostrongyliasis. Here, we conducted a phylogenetic and physiochemical characterisation of the identified putative gene sequence of *A. malaysiensis* ASCT (AmASCT) obtained from our ongoing transcriptomic data analysis and further analysed the genetic variations among the three subfamilies of ASCTs.

METHODS

In-silico Analysis

Identification of Putative A. malaysiensis ASCT (AmASCT)

Putative AmASCT was identified through ongoing transcriptomic analysis of adult *A. malaysiensis*. In short, the raw RNA-seq reads obtained were subjected to quality control using Trimmomatic (Galaxy version 0.39) (Bolger et al., 2014) and de novo transcriptome assembly using Trinity (Grabherr et al., 2011). Coding regions were predicted using TransDecoder (Galaxy version 5.5.0) (Haas et al., 2013). Subsequently, the assembled transcripts and predicted coding regions were subjected to homology-based similarity search using BLASTx and BLASTp against the latest non-redundant protein sequence (nr) database in the DIAMOND alignment tool (Galaxy version 2.0.15) (Buchfink et al., 2014). Hmmscan (Galaxy version 3.4) (Finn et al., 2011) searches were conducted using the protein family database (Pfam-A) database as a query, and TMHMM software (Galaxy version 0.0.17) (Cock et al. 2013) was used for transmembrane domains detection. Trinotate (Galaxy version 3.2.2) (Grabherr et al., 2011) integrated the BLAST, hmmscan, and TMHMM results.

A BLASTp search (Camacho et al., 2009) using *F. hepatica*'s ASCT gene (accession no. ACF06126.1) (van Grinsven et al., 2009) as a query was conducted against the nr database of *A. cantonensis*. The GenBank accession number of the top BLASTp hit was retrieved

and searched against the Trinotate integrated annotation results previously obtained. The gene sequence of the putative AmASCT was retrieved from the Trinotate result for further analysis.

Validation of the Obtained AmASCT

Several bioinformatic tools and molecular techniques were employed to validate the identified *Am*ASCT. The initial bioinformatic validation involved utilising BLASTx search by using the obtained putative *Am*ASCT gene sequence as a query against the nr database of *F. hepatica* to ensure whether the putative *Am*ASCT is a homolog to the previously characterised ASCT of *F. hepatica* (accession no. ACF06126.1) (van Grinsven et al., 2009). In addition, ORFfinder in the National Center for Biotechnology Information (NCBI) websites (https://www.ncbi.nlm.nih.gov/orffinder/) was used to identify and analyse the open reading frame (ORF) of the *Am*ASCT gene, aiding in the prediction of coding regions and potential functional domains.

Following the bioinformatic validation, molecular techniques were employed to confirm the presence of the putative AmASCT gene at the transcript and deoxyribonucleic acid (DNA) levels. Polymerase chain reaction (PCR), agarose gel electrophoresis, and Sanger sequencing were conducted. Briefly, adult A. malaysiensis worms collected from Low et al. (2023) were used for total ribonucleic acid (RNA) extraction using Trizol reagent (Thermo Fisher Scientific, USA), followed by cDNA synthesis using RevertAidTM First Strand cDNA Synthesis Kit (Toyobo, Japan) according to the manufacturer's instructions. Subsequently, the cDNA was used in polymerase chain reaction (PCR) amplification using a set of primers (forward: 5'-ATGCTGTGTCGGCTCTCATCC-3' and reverse: 5'-TTAGTCCACTTCAAGGCATC-3'). PCR was carried out in a thermal cycler with the cycling conditions set as follows: 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 55°C for 30 sec and extension at 72°C for 1 min. The cycling condition was ended with a final extension at 72°C for 5 min. Next, 1.5% agarose mixed with 1x TAE buffer was prepared, and the PCR product was electrophoresed (400W/80V) for 60 minutes and visualised under a UV transilluminator. Each primer's PCR product and 10 µl (10 µM) were outsourced to Apical Scientific Sdn. Bhd. (Selangor, Malaysia) for purification and sanger sequencing.

Physiochemical Profiling of AmASCT

The putative *Am*ASCT's gene sequence was retrieved and subjected to physiochemical characterisation using ExPASy's ProtParam tool (Wilkins et al., 1999). The default configuration of the ProtParam tool was utilised to obtain the putative *Am*ASCT's molecular mass, theoretical isoelectric point (pI), amino acid composition, aliphatic index, and grand average of hydropathy (GRAVY).

Phylogenetic Analysis

Retrieval of ASCT Protein Sequences from Other Helminths

Before multiple sequence alignment and phylogenetic analysis, ASCT gene sequences from other helminth parasites were retrieved from the NCBI GenBank. BLASTp searches using previously characterised ASCTs of *T. brucei* (accession no. EAN79240) (Rivière et al., 2004), *F. hepatica* (accession no. ACF06126.1) (van Grinsven, 2009), and *T. vaginalis* (accession no. XP_001330176) (van Grinsven, 2008) were conducted against the nr database of helminth parasites. Sequences with an E value smaller than 20 (E^{-20}) were retrieved. If multiple parasites of the same species were found as a hit, only the sequence with the highest hit was retrieved.

Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignment was conducted using the putative *Am*ASCT obtained in this study, the previously characterised ASCT protein sequences of *T. brucei*, *F. hepatica*, and *T. vaginalis*, and the retrieved ASCT protein sequences from other helminth parasites. Multiple sequence alignment was carried out using ClustalW (Thompson et al., 1994), and all sequences without the conserved ExG and GxGGxxD motifs of the family I CoA-transferases were removed. The aligned protein sequences of ASCTs were used for phylogenetic tree construction through the neighbour-joining (NJ) method in Molecular Evolutionary Genetics Analysis version 11 (MEGA11) software (Tamura et al., 2021). Pairwise comparisons to infer genetic variation within and between the three subfamilies of the family I CoA-transferases were also performed using MEGA11 software.

RESULTS AND DISCUSSION

Identification of the A. malaysiensis ASCT Gene Sequence

The ASCT sequence of *A. malaysiensis* was identified from our ongoing transcriptomic data analysis of adult *A. malaysiensis*. A BLASTP search conducted against the nr database of A. cantonensis using the previously characterised ASCT of *F. hepatica* (accession no. ACF06126.1) resulted in one blast hit (accession no. KAE9415559.1) that could be the putative ASCT of *A. cantonensis*. The previously characterised ASCT of *F. hepatica* was used as a query since it is a helminth parasite that could have genetic information similar to that of *A. malaysiensis*. In contrast, the other previously characterised parasites, *T. brucei* and *T. vaginalis*, are parasitic protists. Moreover, the nr database of *A. cantonensis* was searched against this gene due to the minimal protein and nucleotide sequences of *A. malaysiensis* being available in public databases. In addition, the reference genome of *A. malaysiensis* has yet to be sequenced. In contrast, the reference genome of *A. cantonensis*

is available in the NCBI GenBank (accession no. MQTX01000177.1, Xu et al., 2019). In addition, *A. cantonensis* is closely related to *A. malaysiensis*, and their morphological and genetic information is highly similar (Chan et al., 2020; Eamsobhana et al., 2015).

The gene annotation data from the transcriptomic data analysis of A. malaysiensis were searched for the gene homolog to the obtained putative A. cantonensis ASCT (accession no: KAE9415559.1). However, no similarities were found. For this reason, the putative A. cantonensis ASCT (accession no: KAE9415559.1) was further used as a query protein in a BLASTP search against the nr database to find other homologues. The top five hits of the BLASTP search (accession no. KAJ1347515.1, EYC13181.1, KIH58849.1, EYC13183.1, and EYC13182.1) were used further to search the transcriptomic gene annotation data of A. malaysiensis. A gene was found to be the homolog of the accession no. EYC13181.1, which could be the putative A. malaysiensis ASCT (AmASCT). The complete nucleotide length and the open reading frame (ORF) amino acid sequence length of the putative AmASCT gene were 1413 base pairs (bp) and 470 amino acids (aa), respectively (Figure 1). A BLASTP search conducted using the putative AmASCT protein sequence as a query revealed high similarity with a few other helminth parasites, with A. cantonensis (accession no. KAE9415559.1) exhibiting the highest similarity, followed by Parelaphostrongylus tenuis (KAJ1347515.1) and Ancylostoma ceylanicum (EYC13181). However, these homolog protein hits were described as hypothetical proteins because the gene has never been characterised in any of these helminths.

The putative AmASCTs were annotated as acetyl-CoA hydrolases or acetyl-CoA transferase N-terminal domains according to the protein family database (Pfam) annotation from the transcriptomic gene annotation data. Acetyl-CoA hydrolase (ACH) is one of the four known enzymes catalysing acetate formation from acetyl-CoA. The activity of ACH has been described in plants and animals (Hunt & Alexson, 2008; Zeiher & Randall, 1990). Although ACH activity has been reported in Ascaris suum parasites (de Mata et al., 1997), the complementary gene of this enzyme in this nematode parasite has yet to be identified. The CoA-transferase enzyme was previously confused with ACH in Saccharomyces cerevisiae (Lee et al., 1989, 1990); this enzyme was subsequently recharacterised as a CoA-transferase due to its minor hydrolase activity and was found to catalyse the transfer of the CoA moiety from succinyl-CoA to acetate, which is a particular characteristic of CoA-transferases (Fleck & Brock, 2009). One of the apparent characteristics of CoAtransferases that is absent in CoA-hydrolases is the glutamate active site for the formation of CoA-ester intermediates via a ping-pong bibi mechanism (Heider, 2001; White & Jencks, 1976). When the glutamate active site was replaced, CoA-transferase was found to be converted into a CoA-hydrolase (Mack & Buckel, 1997).

Amoomomoooo	momonmo	20000000	AMCACCA	CCACT	manac	COM	CACACO	0020	00
AIGCIGIGICGGC 10	2	CGCTTG	30	ICCAC:	40	SGCT	50	GCAG	CG
MLCR	LSS	RL	I S	т т	S T	A	DR	A	A
TCGTCATTTCATO	SATAAGCG	CGAGATT	GCGTACC	CACTO	CAAGG	FAAG	GAACCC	AAAG	TG
SSFH	DKB	ET	A V	P T.	100 G	K	E P	K	v
TGCAATCTGAAGG	ACGCCTT	CAAACAG	ATCAAAI	CAGG	GACGA	CATA	TTTGTA	CACG	GA
130	1.	40	150		160		170		
CNLK	DAF	KQ	IK	SG	DD	I	FV	HO	G
ATTGCGGCGACAC	CTACACC	ACTTCTC	210	TGTGC	CGAGTA	TGTA	230	AACG	AT
IAAT	PTP	LL	KG	LC	EY	v	KA	NI	D
CTCAAAGGAATCA	AACTTCA	FCATCTT	CATCTGO	AAGG	GAGAC	GCCT	TGGACT	GCTG	AG
250	2.0	60	270		280	-	2.90	1.20	-
	K L H	H L	H L	E G	E T	P	W T	A	E
310	3:	20	330	. I IAC.	340	Inne	350	nnno	
D V R S	RIR	S N	SL	F T	GA	N	LR	K Z	A
GTTAATGAAGGCA	TCGCTGA	CTTCAAC	TCATGCI	TCCT	CAGGA	AGTG	CCAATG	CTGT	TT
VNE G		F N	390 S C	F T.	400 E	v	410 P M	т. 1	F
CGTAAGGGTGCAA	TCAAGCT	CAACGCC	GCTTTAA	TTCA	GTCTC	TCCT	CCAGAT	GCAA	AT
430	4.	40	450		460		470		
RKGA	IKL	N A	AL	I H	V S	P	P D	AI	N
GGATACTGCTCT	TAGGTAC	ATCTGTG	GACACCI	CAAGO	SGCGGGG	FGTA	GCAAAC	GCTG	AT
GYCS	LGT	s v	DT	SR	A G	v	AN	AI	D
CATGTTATAGCA	TGTCAAA	TAAGCAT	ATGCCAC	GGACT	TTTGG	TGAT	AGTCTT	ATTC	AC
550	5	60	570		580	-	590		
H V I A	M S N	K H	M P	RT	F G	D	S L	II	H
610	6	20	630	CALL	640	ICAL	650	CAIG.	11
SSHI	DVL	VE	DH	T F	PL	н	ER	н	v
GGTAGAGACAGT	GAAGAGGA	GAAGAAG	ATTGGAG	CAATA	ATTGC	FGAA	AATTTG	GTGG	AT
670	61	80	690		700	-	710		-
AATGGTGCTACAC	TACAGAT	GGTATT	GGGGGCCG	TACCA	GATGC	GCA	CTGTCA	GCTT	TG
730									
150	7.	40	750		760		770		
NGAT	LQM	40 G I	750 G A	V P	760 D A	A	770 L S	A	L
N G A T AAACATCACAAAG	L Q M GATCTCGG	G I FATTCAC	750 G A ACGGAGA	V P ATGTTO	760 D A	A	770 L S GTTTTA 830	A I GATC	L TG
N G A T AAACATCACAAAG 790 K H H K	L Q M SATCTCGG 80 D L G	40 GI FATTCAC 00 IH	750 G A ACGGAGA 810 T E	V P ATGTTC M F	760 D A CTCGGA 820 S D	A TGGT G	770 L S GTTTTA 830 V L	A D	L TG L
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG	L Q M SATCTCGG 80 D L G SCCATTAC	40 GI FATTCAC D0 I H AAATAGC	750 G A ACGGAGA 810 T E AAAAAGG	V P ATGTTC M F STGCT#	760 D A CTCGGA 820 S D CATCC	A TGGT G GGGA	770 L S GTTTTA 830 V L AAACTT	A D D TTGG	L TG L TT
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850	L Q M SATCTCGG 80 D L G SCCATTACI 81	40 GI FATTCAC DO IH AAATAGC 60	750 G A ACGGAGA 810 T E AAAAAGG 870	V P ATGTT(M F STGCT#	760 D A TCGGA 820 S D CATCC 880	A TGGT G GGGA	770 L S GTTTTA 830 V L AAACTT 890	A D GATC: D D TTGG:	L TG L TT
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N	L Q M SATCTCGG D L G SCCATTACI A I T	40 GI FATTCAC 00 IH AAATAGC 60 NS	750 G A ACGGAGA 810 T E AAAAAGG 870 K K	V P ATGTTC M F STGCT# V L	760 D A CTCGGA 820 S D CATCC 880 H P	A TGGT G GGGA	770 L S GTTTTA 830 V L AAACTT 890 K L	A D GATC: D D TTGG: L V	L TG L TT V
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATC 910	L Q M SATCTCGG D L G SCCATTACI A I T GGTTCCACI 92	40 G I FATTCAC 00 I H AAATAGC 60 N S AAAGTTG 20	750 G A EACGGAGA 810 T E EAAAAAGG 870 K K ETACAATT 930	V P ATGTTC M F STGCT# V L	760 D A CTCGGA 820 S D CATCC 880 H P GCACGA 940	A G GGGA G CAAT	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950	A D GATC: D D TTGG: L V ATCG	L TG L TT V AG
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y	L Q M SATCTCGG 80 D L G SCCATTACI 81 A I T GGTTCCACI 92 G S T	40 GI FATTCAC 00 IH AAATAGC 60 NS AAAGTTG 20 KL	750 G A ACCGGAGA 810 T E CAAAAAGG 870 K K TACAATT 930 Y N	V P ATGTTC M F STGCT# V L TTTTTC F L	760 D A CTCGGA 820 S D CATCC 880 H P CACGA 940 H D	A IGGT G GGGA G CAAT N	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L	A D D D D D D D D D D D D D D D D D D D	L TG L TT V AG
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATG 910 S F V Y TTTGGTGATGTT	L Q M SATCTCGG 80 D L G SCCATTAC 83 A I T GGTTCCAC 93 G S T CAGTGGGT	40 G I FATTCAC 00 I H AAATAGC 60 N S AAAGTTG 20 K L GAACGAT	750 G A ACCGGAGA 810 T E AAAAAGG 870 K K TACAATI 930 Y N CCTGCTG	V P ATGTTC M F GTGCT# V L TTTTTC F L STTAT#	760 D A CTCGGA 820 S D CATCC 880 H P GCACGA 940 H D ACGACA	A FGGT(G GGGA G CAAT(N GAAT(770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAAG	A D GATC: D D TTGG: L V ATCG I D GTTAC	L TG L TT V AG E CC
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATG 910 S F V Y TTTGGTGATGTTC 970 F G P V	L Q M SATCTCGC D L G CCATTACI 8 A I T GTTCCACI 9 G S T CAGTGGGT 9 9 0 V V	40 G I FATTCAC 00 I H AAAATAGC 60 N S AAAGTTG 20 K L SAACGAT 80 N P	750 G A RACGGAGA 810 T E RAAAAAGG 870 K K TACAATT 930 Y N CCTGCTG 990 P 2	V P ATGTTC M F STGCT# V L TTTTTC F L STTAT#	760 D A TCGGA 820 S D CATCC 880 H P CACGA 940 H D CGACA 1000 R	A IGGT G GGGA G CAAT N GAAT	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAAG 1010 P K	A D D D D D D D D D D D D D D D D D D D	L TG L TT V AG E CC
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATCCTC	L Q M SATCTCGG B L G SCCATTACI SCCATTACI SGTTCCACI 93 G S T CAGTGGGT 94 Q W V	40 G I FATTCAC 00 I H AAAATAGC 60 N S AAAGATG 20 K L SAACGAT 80 N D AATAGAT	750 G A CACGGAGA 810 T E CAAAAAGG 870 K K TACAATT 930 Y N CCTGCTG 990 P A CTCACAG	V P ATGTTC M F STGCT# V L TTTTTC F L STTAT# V I	760 D A TCGGA 820 S D ACATCC 880 H P GCACGA 940 H D ACGACA 1000 R Q ACTTGT	A FGGT G GGGA G CAAT N GAAT N	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAAG 1010 P K GACTCT	A D D D D D D D D D D D D D D D D D D D	L TG L TT AG CC T GA
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATTCTC 1030	L Q M E Q M GATCTCGG: 8 CCATTACI 8 CCATTACI 8 CATCCACI 9 CAGTGGGTT 2 CAGTGGGTT 9 0 0 0 0 0 0 0 0 0 0 0 0 0	40 G I FATTCAC 00 I H AAAATAGC 60 N S AAAAGTTG 20 K L SAACGAT 80 N D AATAGAT 040	750 G A ACGGAGA 810 T E AAAAAGG 870 K K TACAATT 930 Y N CCTGCTG 990 P A CCTACAG 1050	V P ATGTTC M F STGCT# V L STTTTC F L STTAT# V I SGGCA#	760 D A TCGGA 820 S D ACATCC 880 H P 5CACGA 940 H D ACGACA 1000 R Q AGTGT 1060	A FGGT G GGGA G CAAT GAAT N FTCT	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTG 950 P L CCAAG 1010 P K GACTCT 1070	A D D D D D D D D D D D D D D D D D D D	L TG L TT AG CC T GA
N G A T ARACATCACARAG 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATG 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATTCTC 1030 A I N S	L Q M E Q M GATCTCGG: 8 0 L G 5 CCATTACI 9 G S T CAGTGGGGT 9 Q W V 5 CTGGGG 0 1 0 1 1 A V E	40 G I TATTCAC 00 I H AAATAGC 60 N S AAAGTTG 20 K L SAACGAT 80 N D AATAGAT 040 I D	750 G A ACGACA 810 T E AAAAAG 870 K K TACAATT 930 X N CCTGCTG 990 P A CCTGCTG 1050 L T	V P ATGTTC M F STGCT# V L TTTTTC F L STTAT# V I SGGCA# G Q	760 D A TCGGA 820 S D CATCC 880 H P CCACGA 940 H D CCGACA 1000 R Q CGTGTT 1060 V V	A GGGGA GGGA CAAT SAAT SAAT	770 L S GTTTA 830 V L AAACTT 890 K L CCTTG 950 P L CCAAAG 1010 P K GACTCT 1070 D S	A D D D D D D D D D D D D D D D D D D D	L TG L TT AG E CC T GA
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATTCTC 1030 A I N S AAGCGTTTTCTGT	L Q M E Q M GATCTCGG: 8 0 L G 5 CCATTACI 9 0 0 0 0 0 0 0 0 0 0 0 0 0	40 G I FATTCAC 00 I H AAAATAGC 60 N S AAAGTTG 20 K L SAACGAT 80 N D AATAGAT 040 I D FGGTGGA	750 G A ACGGA27 810 T E ARAAAGO 870 K K TACAATT 930 Y N CCTGCTG 990 P A CTCACAC 1050 L T CCAGTGG	V P ATGTTC M F STGCTH V L TTTTTC F L STTATH V I SGGCAH G Q SACTT	760 DA TCGGAC 820 SD ACATCC 880 H P CACGACA 1000 R Q AGTGT 1060 V V CATTCC	A FGGT G GGGA CAAT CAAT S CAAT S FTCT S FGGT	770 L S GTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAAG 1010 P K GACTCT 1070 D S GCGGCA	A D D D D D D D D D D D D D D D D D D D	L TG L TT AG CC T GA GC
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATTCTC 1030 A I N S AAGCGTTTTCTGT 1090 K R F L	L Q M GATCTCGG: 8 0 L G CCATTAC: 8 3 3 G S T CAGTGGGTG 9 9 0 2 W V 5 CTGGGATT 1 1 A V E CTGGATT 1 5 G F	40 G I FATTCAC 00 I H AAATAGC 60 N S AAAGTTG 20 K L SAACGAT 80 N D AATAGAT 040 I D FGGTGGA G G G	750 G A ACGGAC 810 T E AAAAAG 870 K K TACAATI 930 Y N CCTGCTG 990 P A CCTCACAG 10500 L T CCAGTGG 10500 L T CCAGTGG 10100 Q V	V P MTGTTC M F STGCTH V L TTTTTC F L STTATH V I SGGCAH G Q SACTT D F	760 D A TCGGAC 820 S D ACATCC 880 H P CACGACA 1000 R Q AGTGT 1060 V V CATTCG 1120 I R	A FGGT G GGGA G CAAT S CAAT N FTCT S FGGT G	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A A	A D GATC: D D TTGG: L V ATCGJ GTTAC GTTAC V C ATTAC	L TG L TT AG CC T GA GC S
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTCT 970 F G D V GCCATTAATTCTC 1030 A I N S AAGCGTTTTCTGT 1090 K R F L GATGATGGCCAAG	L Q M BATCTCGG 8 D L G SCCATTACI 8 A I T GGTTCCACI 9 C S T CAGTGGGT 9 0 W V SCTGTGGAI 1 (1) A V E CTGGATT) 1 (A V E S G F S G F S G AAAGCC	40 G I FATTCAC 00 I H AAATAGC 60 N S K L GAACGAT 80 N D AATAGAT 040 I D FGGTGGA 100 G G GATCAT	750 G A ACGACE 810 T E AAAAAGO 870 K K TACAATT 930 Y N CCTGACG 990 P A CCTCACAG 1050 L T CAAGTGC 1110 Q V GCCTTGC	V P ATGTTC M F STGCTA V L TTTTTC F L STTATA V I SGGCAA G Q SACTT D F CCCTCA	760 D A TCGGA 820 S D CATCC 880 H P CCACGA 940 H D CCACGA 1000 R Q CGTGT 1060 V V CTTCGT 1120 I R KATCC	A FGGT G GGGA CAAT CAAT S FGGT G GAAG	770 L S GTTTA 830 V L AACTT 890 K L CCTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A A AAACGG	A D D D D D D D D D D D D D D D D D D D	L TG L TT AG E CC T GA GC S GC
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATG 910 S F V Y TTTGGTGATGTTCTG 970 F G D V GCCATTAATTCTG 1030 A I N S AAGCGTTTTCTGT 1090 K R F L GATGATGGCCAAC 1150	L Q M BATCTCGG 8 BL G CCATTACI 8 CCATTACI 8 C S T CAGTGCGCT 9 Q W V CCTGGGATT 1 CTGGGATT 1 S G F CCGGATC 1 S G F CCGGATCCO 1 1	40 G I FATTCAC 00 I H ANATAGC 60 N S ANAGTTG 20 K L SAACGAT 80 N D ANATAGAT 040 I D FGGTGGA 100 G G SATCAT 160	750 G A ACGGAGA ACGGAGA 810 T E AAAAAG 870 K K T TACAAT 930 Y N CCTGCCC CCTCACA 1050 L T CCAACGC 1110 Q V GCCTTCC CCT170	V P ATGTTC M F STGCTA V L TTTTTC F L STTATA V I SGGCAA G Q SACTT D F CCCTCA	760 D A TCGGA 820 S D CATCC 880 H P CCACGA 940 H D CCACGA 1000 R Q CATTGT 1060 V V CATTCGT 1120 I R LATCTC	A TGGT G GGGAI G GCAAT N N TTCT S S TGGT G G GAAG	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A AA AAAGGG 1190	A D D D D D D D D D D D D D D D D D D D	L TG L TT AG E CCC T GA G GC S GC
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATTCTC 1030 A I N S AAGCGTTTCTTG 1090 K R F L GATGATGGCCAAA 1150 D D G Q	L Q M BATCTCGG: 8 BL G CCATTAC: 8 A I T CATTCCAC: 9: G S T CAGTGGGT 0: 10 10 10 10 10 10 10 10 10 10	40 G I FATTCAC 00 I H ANATAGC 60 N S ANAGTTG 20 K L SCACCAT 80 N D AATAGAT 040 I D TGGTGGA G G SATCATT 160 I I	750 G A ACGGAGA ACGGAGA 810 T E AAAAAGC 870 K K TACAAT 930 Y N CCTGCTGC 1050 L T CCTGCTG 1110 Q V GCCTTGC 1170 A L	V P MTGTTC M F FTGCTJ V L TTTTTC F L STTATJ V I SGGCAJ G Q SACTTT D F SCCTCJ P S	760 D A 820 S D ACATCC 880 H P CACGACA 1000 R Q ACGACA 1060 V V CATCC 1120 I R ATCCC 1120 I S CACCA	A TGGT GGGAJ G CAAT N SAAT SAAT SAAT G G G SAAG K K	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAGG 1010 P K GACTCT 1070 D S GCGCA 1130 A A AAACT 1130 CCGCAG 1190 K G CCTTG	A D GATC: D D TTGG: L ATCGI GTTAC V GTCGC V CATTAC I S CAGAC	L TG L TT AG E CCC T GA G G G C S G C S S S
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 770 F G D V GCCATTAATTCTC 1030 A I N S AAGCGTTTTCTG 1090 K R F L GATGATGGCCAAC 1150 D D G Q AAGATTGTTCCTT 1210	L Q M E Q M SATCTCGG: 8 8 CCATTACI 8 CCATTACI 8 CATGGGT CAGTGGGT 9 9 Q W V 9 Q W V 9 0 1 1 S G F CTGGATT 1 S G F CGARAGCCO 1 1 CG K P CATACIANACON 1 1 CA K P CATACIANACON 1 1 CA K P CATACIANACON 1 1 CA K P CATACIANACON 1 1 CA K P CATACIANACON 1 CA K P CATACIANACON 1 CA K P CA	40 G I FATTCAC 00 I H ANATAGC 60 N S ANAGATG 80 N D AATAGAT 040 I D FGGTGGA 100 G G SATCATT 160 I I CGAGGGT 220	750 G A ACGGAGA ACGGAGA 810 T E ARAAAGG 870 K K TACAATT 930 Y N CCTGCTG 1050 L T CCAAGTGG 1110 Q V GCCTGCTGC 1170 A L GCTGGTG	V P TGTTC M F TGCTA V L TTTTC F L SGGCAA V I SGGCAA O F SACTT D F SCCTCA P S STCGTT	760 D A TTCGGA 820 S D CATCCC 880 H P CACACC CACACA 940 H D CGACAA 1000 R Q CGACAA 1000 V V CATCCC 1120 I R CATCCC 1180 I S CACACACC 1240	A TGGT GGGAJ G CAAT N SAAT SAAT G G GAGA K CAGA	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAGG 1010 P K GACTCT 1070 D S GCGCA 1130 A A AAACGG 1190 K G TCTCAC 1250	A 1 D 1 TTGGT L V ATCGJ GTTAG V 2 GTCGG V 0 GTCGG V 0 Q 2 GTACJ	L TT V AG E CC T GA G GC S GC S AT
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATCTC 1030 A I N S AAGCGTTTCTCTG 1090 K R F L GATGATGGCCAAA 1150 D D G Q AAGATTGTCCCC 1210 K I V P	I Q I Q BATCTCGG: 8: SCCATTAC: 8: G SCTATCCAC: G S G S C 9: Q W SCTGTGGGT 9: Q W SCTGTGGATT: 1: S G S G S G G X Q W SCTGTGGATT: 1: S G G K Q W S G G K Q X S G G K C 1: Y N	40 G I TATTCAC 00 I H HAAATAGC 60 N S AAAGTTG 20 K L SAACGAT 80 N D AATAGAT 100 G G GATCATT 160 I I CGAGGGT 220 E G	750 G A ACGGACR 810 T E ARAAAGG 870 K K TTACAATT 930 Y N CCTGCTG 990 P A CCTCACAG 1050 L T CCAGGGG 1110 Q V GCCTTGC 1110 Q V GCCTTGC 1110 A C CCGGGGG 1250 A G	V P MTGTTC M F STGCTA V L F L STTATA V I GGCAA G Q SACTT D F SCCTCA P S STCGT V V	760 D A TCGGA 820 S D CATCCC 880 H P CACACC 940 H D CACACA 1000 R Q CACTGT 1120 I R NATCTC 1120 I R NATCTC 1180 I S CACACAC I S CACACACA I S CACACACACA I S CACACACACA I S CACACACACA I S CACACACACACACACACACACACACACACACACACACA	A IGGTO G GGGAI G G GAATO N TTCTO G G SAAGO K CAGA	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A A AAACGG 1190 K G TCTCAC 1250 S H	A D GGATC' D D TTGG' I D GTTGG' V C GTTGG' V C GTTGG' V C GTTGG' V C GTTGG' V C GTTGG' V C	L TT V AG E CCC T GA G G G C S G C S S AT H
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 1030 A I N S AAGCGTTTTCTGT 1090 K R F L GATGATGGCCAAC 1150 D D G Q AAGATTGTTCCCT 1210 K I V P TATGTGGTTACG	I Q M I Q M SATCTCGG 81 D L G SGTCCACI 91 G S T CAGTGGGGT 91 Q W V SGTGTGGGAI 11 S G T SGTGTGGGAI 11 S G F SGAAAGCCC 11 1 G K P TATATCAAC 1 1 Y I N SAGTACGG 1 1	40 G I TATTCAC 00 I H AAAATAGC 60 N S AAAGTTG 20 K L SAACGAT 80 N D AATAGAT 040 I D FGGTGGA 100 G G SATCATT 160 I I CGAGGGT 220 E G FATTGCA	750 G A ACGGACG 810 T E ARARAGO 870 K K TTACAATT 930 Y N CCTGCTG 990 P A CTCACAC 1050 L T CCAGTGGTG 1110 Q V GCCTGGTG 11230 A C CCTGGTG 1230 A C CCTGGTGTG	V P MTGTTC M F STGCTA V L F L STTATA V I SGGCAA G Q SACTT D F SCCTCA P S STCGTT V V	760 D A TCGGA 820 S D CATCC 880 H P 940 H D CGACGA 1000 R Q NGTTGT 1120 I R NATCTC 1120 I R NATCTC 1120 I S CACGACA 1120 I R NATCTC 1240 T S SACAGACA	A IGGTO GGGGAI G GCAATO N SAATO SAATO SAATO SAATO GGTO GGAAGA K CAGA	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTGG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A A AAAGGG 1190 K G CTCACC 1250 S H AGCCAA	A 1 GATC' D 1 TTGG' I 1 GTTAC GTCGC V 0 GTCGC V 0 GTCGC Q 2 GTACI V 1 CCGTGC	L TT V AG E CC T GA GC S GC S S AT H CA
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATG 910 S F V Y TTTGGTGATGTTCTG 1030 A I N S AAGCGTTTTCTG 1090 K R F L GATGATGGCCAAG 1150 D G Q AAGATTGTTCCC 1210 K I V P TATGTGGTTACGG 1270 V V V 1270	L Q M BATCTCGG: 8 0 L G CCATTACI 8 A I T GGTTCCACI 9 0 S T CAGTGGGT(9 0 V CTGTGGAT 1 1 1 1 5 G F SGAAGCC(0 1 1 2 5 G K P 1 2 5 GAAAGCC(0 1 1 2 5 G K P 1 2 5 GAAAGCC(0 1 1 2 5 G K P 1 2 5 GAAAGCC(0 1 1 2 5 G K P 5 GAAGCC(0 1 1 2 5 G K P 5 GAAAGCC(0 1 1 2 5 G K P 5 GAAGCC(0 1 1 2 5 G K P 5 GAAAGCC(0 1 1 2 5 G K P 5 GAACGC(0 1 1 2 5 G K P 5 GAACGC(0 1 1 2 5 G K P 5 G 5 G 5 G 5 G 5 G 5 G 5 G 5 G	40 G I FATTCAC 00 I H AAATAGC 60 N S K L GAACGAT 80 N D AATAGAT 040 I D TGGTGGA 100 G G GAGTGGAT 11 CGAGGGT 220 E G FATTGCA 280	750 G A ACGGAC 810 T E AAAAAG 870 K K TACAATI 930 Y N CCTGCTG 990 P A CTCACAG 1050 L T CCAGTGC 1110 Q V GCCTGCC 1170 A L GCTGGTG 1230 A G CCAGTTGT 1290	V P MTGTTC M F STGCTH V L TTTTTC F L STTATH V I SGGCAN G Q SACTT D F SCCTCH P S STCGTT V V V SGGGGN	760 D A TCGGA S D ACATCC 880 H P CACGACA 1000 R Q ACGACA 1000 R Q V V CATTCC 1120 I R I S CACGACA 1200 T T S ACATC 1240 T S ACATCC	A GGGT G CAAT N SAAT SAAT S SAAT G G G SAAG K CAGA R R TATG	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A A AAAGGCA L250 S H AGGCAA 1310	A D GATC: D D TTGG: L T GTTGG GTCG Q GTCG Q GTCG Q Q GTACI V D CAGAC	L TT V AG E CC T GA GC S GC S S C C S AT H CA
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATG 850 I A C N TCGTTCGTTTATG 910 S F V Y TTTGGTGATGTTCC 970 F G D V GCCATTAATTCTG 1030 A I N S AAGCGTTTCTGT 1090 K R F L GATGATGGCCAAG 1150 D D G Q AAGATTGTTCCT 1210 K I V P TATGGTGATGCTACCG 1270 Y V V T TATGATGTGATGATGA	L Q M BATCTCGG' 8 (A CCATTAC) 8 (A CCATTAC) 9 (C CCATTAC) 9 (C CCATTAC) 9 (C 0 (C 1 (C CCATCAC) 1 (C CCAC) 1 (C	40 G I FATTCAC 00 I H ANATAGC 60 N S ANAGTTG 20 K L SANAGAT 80 N D ANTAGAT 040 I D FGGTGGA 10 G G SATCATT 160 I I CGAGGGT 220 E G IATTCCA 280 I A TCATCCCA	750 G A ACGGAAGA ACGGAGA 810 T E AAAAAGG 870 K K TACAAT 930 Y N CCTGCCTG CCTGCCT 1100 Q V GCCTTGCC 1170 A L GCCTGCCT 1230 A G CCAGTGC 1290 Q L	V P MTGTTC M F TGCTA V L TTTTC F L STTATA V I G Q G Q G Q SACTT D F CCCTCA P S STCGTT V V V V V GGGGGA W G	760 D A TCGGAX 820 S D ACATCC 880 H P CACGACA 1000 R Q ACGACA 1000 R Q V V V TCGACA 1120 I R ATCTC 1180 I S CACGACA 1240 T S ACACACA 1300 K N X	A FGGTI G GGGA CAATO N TTCTO S FTGGTO G G CAGAC R R R TATG M M	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A AAACG 1190 K G TCTCAC 1250 S H AGCCAA 1310 R Q AAACT	A D D D D D D D D D D D D D D D D D D D	L TT V AG CC T GA GC S GC S C S AT H CA A TT
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTCC 1030 A I N S AAGCGTTTATCTC 1030 K R F L GATGATCGCCAAA 1150 D D G Q AAGATTGTTCCCC 1210 K I V P TATGTGGTTACGC 1270 X V V T TATGGATGATCGATA	I Q A I SATCTCGG: 8: BACCATTAC: 8: SGTTCCAC: 9: G S T SGTTCCAC: 9: G S T AATGTGGGT 9: Q W V SCTGTGGAT 1: S G F SGAAGCCC 1: 1: G K P C W V 1: S G F 5: G K P 1: S G F 5: GARAGCCO: 1: 1: 1: Y I N 5: 3: J 1: 1: 1: 1: E Y G 1: 1:	40 G I FATTCAC 00 I H ANAGTGC 60 N S ANAGTGC 70 N D ANTGCACAT 100 G G SATCATT 160 I D FGGTGGA 100 G G SATCATT 160 I I CGAGGGT 220 E G FATTGCA 280 I A FCATCCG	750 G A ACGGAGA ACGGAGA 810 T E AAAAGG 870 K K TACAATT 930 Y N CCTGCTG 1050 L T CCACGCT 1110 Q V GCCTTGCC 1170 A L GCTGGTG 1230 A G CCAGTGT 1220 Q L TCGCAGC 1350	V P TGTTC M F TGCTA V L TTTTTC F L SGGCAA G Q SACTT D F SGGCAA D F STCGT V V V CGGGGAA W G CGTGAA	760 D A TTCGGA 820 S D CATCCC 880 H P CACGACA 1000 R Q CGACGA 1000 R Q CGACGA 1000 R Q 1060 V V CATCCC 1200 I R ATCCC I S CACGACA 1300 K N TTTCCT 1360	A GGGGA GGGA GGGA N N TTCT GGAA CAGA R R CAGA R R CAGA M GGAG	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAGG 1010 P K GACTCT 1070 D S GCGCA 1130 A A AAACG S H AGGCAA 1310 R Q AAATCA 1370	A D GATC: D D TTGGT ATCGI GTTAG GTTAG V C GTTAG GTTAG V C GTTAG V C GTTAG Q S GTACI V D CCAGAG V D CCAGAG R D CCAGAG R D CCAGAG C C CCAGAG C C C C	L TT V AG CC T GA GC S GC S C S AT H CA A TT
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTCC 070 F G D V GCCATTAATTCTC 1030 A I N S AAGCGTTTCCTG 1090 K R F L GATGATGGCCAAC 1150 D D G Q AAGATTGTTCCCT 1210 X I V P TATGTGGTTACGC 1270 Y V V T TATGAATTGAATTGAATA	L Q M BATCTCGG: 8 8 CCATTACI 8 4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5	40 G I FATTCAC 00 I H ANATAGC 60 N S ANAGATG 20 K L SAACGAT 80 N D AATAGAT 040 I D FGGTGGA 100 G G SATCATT 160 I I CGAGGGT 220 E G FATTCA 240 I A FCATCCG 340 H P	750 G A ACGGACG ACGGACG 810 T E AAAAAG 870 K K TACAATT 930 P A CCTGCTG 1050 P A CCTGCTG 1050 L T CCAAGTGG 1110 Q V GCCTGCTG 11200 A G CCAGTGGT 12200 A G CCAGTGGT 1290 Q L TCGCCAGC 3350 S Q	V P TGTTC M F TGCTA V L TTTTTC F L SGGCAA G Q SACTT D F STCGT V V V GGGGGA W G CGTGAA R E	760 D A TTCGGA 820 S D CATCCC 880 H P CACACA 940 H D CGACA 1000 R Q CGACA 1000 R Q CGACA CGACA 1000 CGACA CGACA 1000 CGACA CGACA 1000 CGACA	A GGGA GGGA G GGAAT N N TTCT S TGGT G GGAG R R R TATG M GGAG E	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAGG 1010 P K GACTCT 1070 D S GCGGCA 1130 X G TCTCAC 1250 S H AGCCA 1310 R Q AAACT 1370 K S	A D GATC: D D TTGG: L T ATCGI L T GTTGGI V C GTCGI V C C GTCGI V C C GTCGI V C C GTCGI V C C GTCGI V C C GTCGI V C C GTCGI V C C GTCGI V C C GTCGI V C C C C C C C C C C C C C C C C C C C	L TT V A G C C C T G A G C C C T G A G C C S S C C S S C C S S C T T T T T T T
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTATC 910 S F V Y TTTGGTGATGTTC 970 F G D V GCCATTAATCCG 1030 A I N S AAGCGTTTTCTG 090 K R F L GATGATGGCCAAG 1150 D D G Q AAGATTGTCCCC 1210 X I V P TATGTGGTTACG 1210 X I V T TATGAATGATAA 1330 Y E L I GAAAGGTTGAAGG	L Q M BATCTCGG: 8 8 CCATTACI 8 CCATTACI 8 CCATTACI 9 CATGGGGT CAGTGGGT 9 0 0 1 CAGTGGGT 1 1 S G F CTGGATT 1 1 S G F CTGGATACCI 1 1 S G F CAGTGGATACCI 1 1 S G F CAGTACCCI 1 1 S G F CAGTGGATACCI 1 1 S G F CAGTACCCI 1 1 S G F CAGTACCCI 1 1 S G F CAGTACCCI 1 1 S G F CAGTACCCI 1 1 S G F CAGTACCI 1 1 S G F CAGTACCI 1 S G F CAGTACCI CAGTACCI	40 G I FATTCAC 00 I H ANATAGC 60 N S ANAGATG 20 K L SAACGAT 80 N D AATAGAT 040 I D FGGTGGA 100 G G SATCATI 160 I I CCAGGGT 220 E G FATTGCA 280 I A FCATCCG 340 H P FTCACGGG	750 G A ACGGA26 810 T E ARAAAG 870 K K TTACAATT 930 Y N CCTGCTC 990 P A CCTCACAG 1050 L T CCAAGTGG 1110 Q V GCCTTGC 1110 A L GCTGGTC 11200 A G CCAGTGT 1290 Q L TCGCAGC 1350 S Q GCCTAA	V P TGTTC M F TGCTA V L TTTTTC F L SGGCAA G Q SACTT D F SCCTCA P S STCGTT V V CGGGGA R E	760 D A TCGGA 820 S D CATCC 880 H P CACACA 940 H D CGACA 1000 R Q CGACA 1000 V V CATTCC 1120 I R CACACA 1120 I R CACACA 1120 I S CACACA 1120 I S CACACA 1120 I S CACACA 1120 I S CACACA I I S CACACA I S CACACAC I S CACACACI I S CACACI I S CACAC	A TGGT G GGGA G G CAAT N N TTCT G G SAAC K CAGA R R TATG M GGAG E	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 K G TCTCAC 1250 S H AGCCAA 1310 R Q AATCA 1370 K S	A D GATC: D D TTGG: L T GTCGI V C GTCGI V C GTCGI V C GTCGI V D CCGTGI R D GCTT: A D	L TT V A G C C C T G A G C C C T G A G C C S G C S S C C A T T T T T T T T T T T T T T T T T
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTC 1030 A I N S AAGCGTTTTCTGT 1090 K R F L GATGATGGCCAAA 1150 D D G Q AAGATTGTTCCCT 1210 X V Y TATGTGGTTACGC 1270 Y V V T TATGAATTGATAAAA 1330 Y E L I GAAGGTTGAAGC 1390 E R F F	I Q M SATCTCGG: 8/3 B A I GGTCCAC: 9/3 GSTTCCAC: 9/3 GSTTCCAC: 9/3 GSTTCCAC: 9/3 GSTTCCAC: 9/3 GSTTCCAC: 9/3 GSTGTGGGA: 9/3 A V SGTGTGGGA: 1/1 SGTCTGGGAT: 1/1 SGTGTGGGAAGCCC 1/1 SGTATACCAC: 1/1 SGTATACCAC: 1/1 SGTATACCAC: 1/1 SGTATACCC: 1/1 SGTATACCC: 1/1 SGTAGATACCC: 1/1 SGTGAGATACCC: 1/1 SGTGAGATACCC: 1/1 SGTGATACCC: 1/1	40 G I FATTCAC 00 I H AAATAGAT 00 N S AAAGTTG 20 K L EAACGAT 80 N D AATAGAT 040 I D FGGTGGA 100 G G EATCATT 160 I CGAGGGT 220 C FATTGCA 280 A FCATCCG 340 H P FTCAGTG 400 S V	750 G A ACGGACG 810 T E ARAAAGG 870 K K TTACAATT 930 Y N CCTGCTG 990 P A CCTCACAG 1050 L T CCAGGTG 1050 L T CCAGGTG 1110 Q V GCCTGCTG 1230 A G CCTGGTG 1230 A C CCGGTG 1230 A S C CCGGTG 1230 A S C CCGGTG 1230 A C CCGGTG 1230 A C CCGGTG 1230 A C CCGGTG 1230 A C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG 1410 C CCGGTG C C CCGGTG C C CCGGTG C C CCGGTG C C CCGGTG C C CCGGTG C C CCGGTG C C CCGGTG C C CCGGTG C C CCGG C C C C	V P MTGTTO M F STGCTA V L F L STTATA V I SGGCAA D F SCCTCA P S STCGTTA V V SGGGGAA W G SGTGAA R E	760 D A TCGGA 820 S D CATCC 880 H P CCACGA 940 H D CCACGA 940 R Q CACTGA 1000 V V CATTGA 1120 I R 1120 I R 1120	A TGGT G GGAA G CAAT N TTCT S TGGT G SAAG K CAGA R R TATG M GGAG E	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAAG 1010 P K GACTCT 1070 D S GCGGCA 1130 A A AAAGGG 1190 K G CCTCAC 1250 S H AGGCCAA 1310 R Q AAATCA 1370 K S	A) GATC' D) TTGG' ATCG' I) GTCG' V (GTCG' V (GTCG' V) GTCG' V (GTCG' V) CCAGA(CCAGA(CCAGA(R) A)	L TG L TT V AG E CC T GA G GC S GC S S C S S C S S C T T C A G C T T G A G C T T T T T T T T T T T T T T T T T T
N G A T AAACATCACAAAA 790 K H H K ATTGCCTGCAATC 850 I A C N TCGTTCGTTTATC 910 S F V Y TTTGGTGATGTTCTC 1030 A I N S AAGCGTTTTCTGT 1090 K R F L GATGATGGCCAAA 1150 D D G Q AAGATTGTTCCCT 1210 K I V P TATGTGGTTACGC 1270 Y V V T TATGGAATTGATTGATTA 1330 Y E L I GAAGGTTGAACG 1390 E R L K	L Q M BATCTCGG: 8 8 8 CATTACI 8 8 1 T GGTTCCACI 9 3 G S T CAGTGGGT(9 0 W V CTGGGATT 1 1 1 CTGGATT 1 1 1 5 G F T 1 1 2 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4	40 G I FATTCAC 00 I H AAATAGC 60 N S AAAGTTG 20 K L GAACGAT 80 N D AATAGAT 040 I D TGGTGGA 10 G G GAACGAT 10 G G GAACGAT 11 CGAGGGT 220 E G FATTGCA 280 I A TCATCGG 340 H P TTCACGG 340 K V	750 G A ACGGAAGA ACGGAGA 810 T E AAAAAGG 870 K K TTACAATT 930 Y N CCTGCCC 990 P A CCTGCCC 1050 L T CCCACCC 1110 Q V GCCTTGC CACC 1230 A G CCCAGTTC 1290 Q L TCGCAGC 1350 S Q GACTAA 1410 D *	V P TGTTC TGTTY V L TTTTT GGCA G Q GGCA G Q GGCA CTTY V S GGCA V V V GGCGA R E	760 D A TCGGAX 820 S D ACATCC 880 H P CACGACA 1000 R Q ACTGTC 1120 I R ACTCC 1120 I S CACGACA 1200 K V V CATTCC 1240 T S ACATCC 1240 T S ACATCC 1300 K N TTTCT 1360 F L	A FGGT G GGGA G CAAT N FTCT N FTCT G GAAG K CAGA R R R R R R R R R R R R R	770 L S GTTTTA 830 V L AAACTT 890 K L CCTTTG 950 P L CCAAG 1010 P K GACTCT 1070 D S A AAAC 1130 A A AAACTT 1070 D S H AAACTT 1070 CCAAG 1190 K G CCGGCA 1190 K G CCGCCA 1190 K G CCGCCA 1190 K G CCGCCA 1190 K G CCGCCA 1190 K G CCGCCA 1190 K S H ACCTCT 1250 S H ACCCC 1310 R Q AAACT 1370 K S	A D GATC: D D TTGG: ATCGI GTTAG GTTCG GTTCG Q GTTCG Q GTACI Q CGTGG Q CGTGG Q CGTGG A D D CGTGG A D D D D D D D D D D D D D D D D D D D	L TG L TT V AG E CC T GA G GC S GC S GC S S C S AT H CA A TT F

Figure 1. Full-length nucleotide sequence and ORF of the putative *Angiostrongylus malaysiensis* ASCT. The coding region comprises 1413 bp (green) of nucleotides and 470 aa (blue)

Validation of The Obtained Putative A. malaysiensis ASCT

Upon conducting the BLASTx search, the putative AmASCT exhibited the ASCT gene of *F. hepatica* as the top hit, indicating a significant alignment and similarity between the two sequences. This BLASTx result demonstrates a high degree of homology between the putative AmASCT and the previously characterised ASCT gene of *F. hepatica* (accession no. ACF06126.1). The positive amplification of the AmASCT gene was confirmed through both PCR and agarose gel electrophoresis. Upon subjecting the PCR product to electrophoresis and visualising it under a UV transilluminator, a distinct and well-defined band was observed at the expected size, indicative of successful amplification of the AmASCT gene and the presence of the gene in transcript and DNA levels. In addition, the obtained Sanger sequencing result of the amplified PCR product confirmed the identity of the putative AmASCT, providing a similar representation of the coding sequence of the putative AmASCT obtained from the de novo transcriptome annotation of adult *A. malaysiensis*. These bioinformatic and molecular approach results validate the reliability of the obtained putative AmASCT gene sequence from the de novo transcriptome annotation data.

Physiochemical Profiling of AmASCT

The physiochemical properties of AmASCT were analysed to understand and predict the enzyme function and behaviour under different conditions. The physiochemical profiling of AmASCT conducted using ExPASy's ProtParam tool revealed several crucial molecular characteristics of AmASCT (Table 1). The computational analysis predicted that AmASCT has a molecular weight of 51.33 kDA, suggesting that the enzyme has a moderate size. The theoretical isoelectric point (pI) of 7.74 indicates that the AmASCT is likely to be neutral or slightly alkaline, which suggests that AmASCT is most stable under such conditions within the parasite. The high aliphatic index (93.79) implies that AmASCT is enriched in hydrophobic amino acids and its adaptation and stability to functioning in hydrophobic nature of the enzyme, contributing to its solubility in aqueous environment. Despite containing enriched hydrophobic amino acids as indicated by the high aliphatic index, the overall hydrophilicity of AmASCT may be because of the distribution of hydrophilic residues and charged or polar amino acids presence on its surface.

Table 1Physiochemical profiling of AmASCT

No. of amino acids	Molecular weight (kDA)	Theoretical isoelectric point (pI)	Aliphatic index	GRAVY
470	51.33	7.74	93.79	- 0.162

Retrieval of ASCT from Helminth Parasites from GenBank

The retrieved amino acid sequences of helminth parasites were subjected to BLASTP searches against the non-redundant helminth protein database using *F. hepatica* (accession no. ACF06126.1), *T. vaginalis* (accession no. XP_001330176), and *T. brucei* (accession no. EAN79240), which represent the three subfamilies of the family I CoA-transferase, as tabulated in Table 1. BLASTP searches revealed 15, 22 and two amino acid sequences of subfamilies IA, IB and IC, respectively, of the family I CoA-transferase. The retrieved amino acid sequence of the helminth parasites in Table 2 shows that most helminth parasites

Table 2

Subfamily A	Subfar	nily B	Subfamily C
Ancylostoma ceylanicum	Fasciola hepatica	Litomosoides	Caenorhabditis
(EYC45313)	(ACF06126)	sigmodontis	remanei
Aphelenchoides bicaudatus	Fasciola gigantica	(VDK78337)	(EFO99043)
(KAI6175034)	(TPP58613)	Toxocara canis	Diploscapter
Bursaphelenchus	Heterobilharzia	(KHN81999)	pachys (PAV69794)
okinawaensis	americana	Bursaphelenchus	
(CAD5205821)	(CAH8437700)	xylophilus	
Nippostrongylus brasiliensis	Schistosoma mansoni	(CAD5213378)	
(VDL75362)	(XP 018648509)	Acanthocheilonema	
Caenorhabditis briggsae (XP_002629737)	Trichobilharzia regent (CAH8830533)	viteae (VBB34274) Brugia malayi (XP 042934641)	
(NP_496144)	(XP_024501800)	Pristionchus pacificus	
Meloidogyne enterolobii	Caenorhabditis remanei	(KAF8362703)	
(CAD2169927)	(EFP07577)	Mesocestoides corti	
Parelaphostrongylus tenuis	Parelaphostrongylus	(VDD82102)	
(KAJ1371690)	tenuis (KAJ1347515)	Taenia asiatica	
Pristionchus pacificus	Angiostrongylus	(VDR22328)	
(KAF8360361)	cantonensis	Hymenolepis	
Ditylenchus destructor	(KAE9415559)	microstoma	
(KA11702014)	Enterobius vermicularis	(CUU98198)	
Dracunculus medinensis	(VDD96370)	Echinococcus	
(VDN53078)	Steinernema	multilocularis	
Strongyloides ratti	carpocapsae	(CDS43154)	
(XP_024505624) Paragonimus westermani (KAF8572391)	(TMS35399)	Rodentolepis nana (VDO13213)	
Clonorchis sinensis (KAG5445983) Opisthorchis felineus (TGZ64319)			

Helminth parasites and their accession numbers of the family I CoA-transferase amino acid sequences with E values less than $1e^{-20}$ retrieved from the NCBI GenBank

have the CoA-transferase enzyme of subfamily IB, followed by subfamily IA and only two in subfamily IC. This finding is in agreement with that of Tielens et al. (2010), who reported that the subfamily IA is found only in trypanosomatid parasites and some metazoan parasites, while the subfamily IC of ASCT is usually possessed only by prokaryotes and fungi. On the other hand, most of the ASCT enzymes of helminth parasites were previously reported to be in subfamily IB of family I CoA-transferases (Tielens et al., 2010).

Multiple Sequence Alignment of Helminth ASCT

Figure 2 presents the multiple amino acid sequence alignments of the putative ASCT gene of *A. malaysiensis* obtained in this study with amino acid sequences of the ASCT gene of other helminth parasites retrieved from the BLASTP searches and the ASCTs of previously characterised parasites. The helminth parasites were grouped into subfamilies based on the query ASCT sequence. Note that although *T. brucei* and *T. vaginalis* are parasitic protists, their ASCT proteins were included in this study as references since their ASCT genes were previously characterised.

According to the multiple amino acid sequence alignments, all the ASCT enzymes of helminth parasites included in this study contained the conserved ExG and GxGGxxD motifs of the ASCTs in the family I CoA-transferases (Tielens et al., 2010). The subfamily

	ExG	GxGGxxD		
EAN 9240 Trypanosoma brucei (ASCT) EXC45313 Ancylostoma ceylanicum KAT617034 Apbelenchoides bicaudatus CAD5205821 Bursaphelenchus okinawaensis VDL75362 Nippostrongylus brasiliensis XP 002629737 Caenonhabditis briggsac NP 496144 Caenonhabditis elegans CAD2169927 Meloidogyne enterolobii KAJ1371609 Parelaphostcongylus tenuis KAF8360361 Pristionnus padifous KAT1702014 Dilylenchus destructor VDN53078 Dracunculus mediensis XP 024505624 Strongyloidos ratti KAF857191 Paragonius westermani KAG545983 Clonorchis sinensis TG264319 Ocisthorchis felineus	VNLGIGIPTLSSNYIPAGVNVULGENGLIGGPPTEDEKVDADW ANLGIGPTLCPNIPSGKVPLGENGVIGGPPTKKKERDADL VNLGIGPTLSNNTLSGVVULGENGVIGGPPTKKKERDADL VNLGIGPTLSNNTLSGVVULGENGVIGGPPTKKGERDADL ANLGIGPTLAPNIPSGIVVLGENGVIGGPPTKKGERDADL ANLGIGPTLAPNIPSGIVVLGENGLIGGPTKKGERDADL VNLGIGPTLAPNIPSGIVVLGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLPKGISVULGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLPKGIVLGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLPKGIVLGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLPKGIVLGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLPKGIVLGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLPKGIVVLGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLFKGVVLKGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLFKGVVLKGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLFKGVVLKGENGVIGGPTKKGERDADL VNLGIGPTLCPNVLFKGVVLKGENGVIGGPTKKGERDADL VNLGIGPTLGPTLSVVKKGVVLKGENGVIGGPTKKGERDADL VNLGIGPTLGCTFKGVVLKKSNGVGPTFKGERDADL	338 FM1PGKLVSGFGGALLVSC-GTKVVV 359 MM1PGKLVSGFGGALLVSC-GTKVVV 355 MM1PGKLVSGMGGALLVSAPGRVIV 355 MM1PGKMVSGMGCALLVSAPGRVIV 357 MM1PGKLVSGMGGALLVSAPGRVVV 359 MM1PGKLVSGMGGALLVSAPGRVVV 359 MM1PGKLVSGMGGALLVSAPGRVVV 359 MM1PGKLVSGMGGALLVSAPGRVVV 350 MM1PGKLVSGMGGALLVSAPGRVVV 354 MM1PGKLVSGMGGALLVSAPGRVVV 355 MM1PGKLVSGMGGALLVSAPGRVVV 356 MM1PGKKVSGMGGALLVSAPGRVVV 359 MMVPGKKVSGMGGALLVSAPGRVVV 359 MMVPGKKVSGMGGALLVSAPGRVVV 359 MMVPGKKVSGMGGALLVSAPGRVVV 350 MMVPGKKVSGMGGALLVSAPGRVVV 350 MMVPGKKVSGMGGALLVSAPGRVVV 351 MMVPGKLVSGMGGALLVSAPGRVVV 357 WVVPGSLISGMGGALLSVSAPGRVVC 357 WVPGSLISGMGCALLSVSAPGRVVV 357 WVPGSLISGMGCALLSVSAPGRVVV 357 WVPGSLISGMCCALLSVSAPGRVVV	414 435 431 433 433 435 437 432 430 432 435 415 435 435 434 433 433	A
RCE06115 Opisitoccils Teilmeus RCE06126 Fasciola hopatica (ASCT) TPP55613 Fasciola gigantica Ch48437700 Hotorobilharzia amoricana XP 016648509 Schistosoma mansoni CAF8830533 Tritobbilharzia regenti Angiostrongylus malaysiensis (TRIS STUDY) KAJ1347515 Parelaphostrongylus catononsis VD06370 Enterobius vermicularis TM55399 Steinernema carpocapsae XP 026401800 Strongyloides ratti EF01577 Caenorhabditis remanei VDK78373 Bursapholenchus xylophilus VAN1999 Toxoora canis CAD521378 Bursapholenchus xylophilus VAD82102 Haesocatoides conti VDR84018 Tationohus pacificus VDR9210 Facinonus pacificus Facin	TKIVPILKLGGGVVTTRAIVIVVI SVG LAVLFGKNLGGAILL TKIVPILKLGGVVTTRAIVIVVI SVG LAVLFGKNLGGAILL SKISCLKPGGVVTTRAIVIVVI SVG LAVLFGKNLGGAILL SKISCLKPGGVVTRAIVIVVI SVG LAVLFGKLGGAALL SKISCLKPGGVVTRAIVIVVI SVG LAVLFGKLGGAALL SKISCLKPGGVVTSRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGAGVVTSRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVVTSRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVVTSRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVVTSRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVVTSRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVTTRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVTTRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVTTRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGGVTTRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVI SVG LAVLGKGKNGGAYLLI SKIVPINEGGVTURGAVVINI SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGGUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAIVIVINVING SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAIVIVINVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAIVIVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAVGGVTINGAVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAVINGANVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAVINGANVIN SVG LAVLGKGKNGGAYLLI SKIVPINEGQUTTRAVINGANVINGANGANGAYLLI SKIVPINEGQUTTRAVINGANVINGANGANGANGANGAYLLI SKIVPINEGQUTTRAVINGANGANVINGANGANGANGANGAYLI	357 REVERSELT SCHOOL PERGANSLO 452 DSIGTTIYSCFCOQUE FLRGANVSLOC 553 DSIGTRIYSCFCOQUE FLRGANVSLOC 557 DSVGSRIVSCFCOQUE FLRGANSLOC 575 DSVGSREPLSCFCOQUE FLRGANSLODC 586 DSVGSREPLSCFCOQUE FLRGANSLODC 570 DSVGSREPLSCFCOQUE FLRGANSLODC 571 DSVGSREPLSCFCOQUE FLRGANSLODC 572 DSVGSREPLSCFCOQUE FLRGANSLODC 573 DSVGSREPLSCFCOQUE FLRGANSLODC 574 DSVGSREPLSCFCOQUE FLRGANSLODC 575 DSVGSREPLSCFCOQUE FLRGANSLODC 576 DSVGSREPLSCFCOQUE FLRGANSLODC 576 DSVGSREPLSCFCOQUE FLRGANSLODC 576 DSVGSREPLSCFCOQUE FLRGANSLODC 570 DSIGSREPLSCFCOQUE FLRGANSLODC 570 DSIGSREPLSCFCOQUE FLRGANSLODC 570 DSIGSREPLSCFCOQUE FLRGANSLODC 570 DSIGSREPLSCFCOQUE FLRGANSLODC 570 <td< td=""><td> 43.3 391 396 396 396 383 383 383 387 392 384 390 387 394 395 394 395 394 396 396 397 398 398</td><td>В</td></td<>	 43.3 391 396 396 396 383 383 383 387 392 384 390 387 394 395 394 395 394 396 396 397 398 398	В
EFO99043 Cacnorhabditis remanci PAV69794 Diploscapter pachys <mark>ΚΡ_001330176 Trichomonas vaginalis (λSCT)</mark>	GDISSVVPFASHIDHAEHDVDILVIEQG.ADLRGLAPRERARAVI GKISAIVPQASHVDHINQDVQVIVIEQG.ADLRGLSPRQRARAVI TGISCVVPMCTHVDHTEHDLDVIVIEQG.ADLRGLAPVERARIMI	217 HVCGIKMMN <mark>GIGGSCI</mark> FARNAHLAIFV 166 LVMGSRIQNGIGGSCIFARNAYVSIFM 467 LINGTKLVN <mark>GIGGSCI</mark> FLRNGYLSIMII	165 114 411	с

Figure 2. Amino acid multiple sequence alignments of the conserved ExG and GxGGxxD regions of family 1 CoA-transferases

IAs of ASCTs have a broader SENG motif since they are grouped with the mammalian succinyl-CoA:3-ketoacid CoA transferase (SCOT) enzyme (Rivière et al., 2004). This conserved SENG motif contains the glutamate residue active site of CoA-transferase (Rangarajan et al., 2005). On the other hand, ASCTs in subfamily IB and subfamily IC do not have the conserved SENG motif but instead possess a conserved ExG motif, which is believed to be a constituent of the more prominent SENG motif found within subfamily IA (Tielens et al., 2010). All the subfamilies of ASCT have the conserved GxGGxxD motif, an integral component of the oxyanion hole (Rangarajan et al., 2005; Tielens et al., 2010). The relative position of the GxGGxxD motif to the ExG motif was consistent between subfamily IB and IC. However, there was a notable difference in the expression of subfamily IA. In subfamily IA, the ExG motif was located proximal to the N-terminal region, whereas the GxGGxxD motif was situated toward the C-terminal region. The opposite trend occurs in subfamilies IB and IC.

Phylogenetic Analysis of Helminth ASCT

The phylogenetic tree constructed from the multiple amino acid alignment of ASCTs from helminth parasites is presented in Figure 3. The phylogenetic tree showed that the ASCTs of helminth parasites consisted of two clades: subfamilies IB and IC were in the same clade, while subfamily IA was in the other clade. Subfamilies IB and IC were divided into two distinct clusters within their clades. The phylogenetic tree in Figure 2 shows that subfamilies IB and IC are closer to each other than to subfamily IA. The putative *Am*ASCT obtained in this study clustered with the previously characterised ASCT of *F. hepatica* in subfamily IB.

The CoA-transferases responsible for the generation of acetate in parasites have been well-recognised as succinate-dependent enzymes and are thus referred to as acetate/ succinate CoA-transferases. The CoA moiety of acetyl-CoA was transferred to succinate, resulting in the production of acetate and succinyl-CoA. To date, all eukaryotic ASCTs found are family I CoA-transferases. These enzymes are distinguished by their ability to transfer CoA groups, which involve a glutamate residue located in the active region of the enzyme (Heider, 2001). The identification and characterisation of the ASCT gene have been conducted in the helminth parasite *F. hepatica* (van Grinsven et al., 2009) and two other protist parasites, *T. brucei* (Rivière et al., 2004) and *T. vaginalis* (van Grinsven et al., 2008). The ASCT genes from these three parasites were found to share little homology, thus characterising them into three different subfamilies of the family I CoA-transferases (Tielens et al., 2010).

The ASCT enzymes of *T. brucei* were classified into subfamily IA of the family I CoA-transferases. ASCT in subfamily IA was found to exhibit significant homology to SCOT. This enzyme plays a crucial role in using ketone bodies inside the mitochondria of the human brain and muscle (Fukao et al., 2004). Among the three subfamilies, only



Figure 3. Phylogenetic tree of the ASCT gene of helminth parasites

subfamily IA exhibited homology to enzymes present in mammals. ASCT in subfamily IA occurs in the aerobic mitochondria of trypanosomatids and other metazoans (Rivière et al., 2004). The ASCT enzyme of *F. hepatica* was previously classified into subfamily IB and occurs in the anaerobically functioning mitochondria of metazoan organisms. The ASCTs of parasitic helminths mainly belong to subfamily IB of the family I CoA-transferases (Tielens et al., 2010). ASCT of *T. vaginalis* was classified into the subfamily IC. Unlike those in subfamilies IA and IB, ASCT in subfamily IC did not occur in mitochondria but in anaerobic hydrogenosomes (van Grinsven et al., 2008). Hydrogenosomes are membrane-bound organelles that are involved in the production of hydrogen. These organelles are closely linked to mitochondria, although they develop independently in different protists (Cavalier-Smith & Chao, 1996).

A previous study revealed the possibility that hydrogenosomes and mitochondria are evolutionarily related and that hydrogenosomes have undergone evolutionary changes due to adaptation to anaerobic environments (Martin & Muller, 1998; Roger et al., 1996). The exact evolutionary relationship between mitochondria and hydrogenosomes is currently disputed in the academic community. ASCT coupled with the succinyl-CoA synthase (SCS) cycle is the only catabolic pathway identified in mitochondria and hydrogenosomes. Nevertheless, based on the phylogenetic tree (Figure 3), the mitochondrial ASCT of T. brucei in subfamily IA lacks similarity with the hydrogenosomal ASCT of T. vaginalis in subfamily IC, as does the mitochondrial ASCT of F. hepatica in subfamily IB. Hence, if both mitochondria and hydrogenosomes are evolutionarily related, it may be inferred that the occurrence of ASCT enzyme activity within organelles derived from endosymbiotic events has separately undergone convergent evolution on at least two separate occasions. Additionally, based on the multiple sequence alignment previously discussed in this study (Figure 2), there was a significant difference in the relative location of the oxyanion hole GxGGxxD motif and the active site ExG motif between subfamily IA and subfamilies IB and IC. The positioning of the two conserved motifs is the same in subfamilies IB and IC but different in subfamily IA (Tielens et al., 2010). This relative location of both conserved regions may contribute to positioning the three subfamilies in the phylogenetic tree.

Pairwise analyses were also conducted to infer genetic variations between and within the three subfamilies of the family I CoA-transferases. The genetic variations within subfamilies ranged from 0.04% to 0.85% for subfamily IA (Table 3), 0.005% to 0.78% for subfamily IB (Table 4), and 0.56% to 0.63% for subfamily IC (Table 5). In subfamily IA, the highest genetic variation between helminths was observed between *Caenorhabditis briggsae* and *Opistorchis felineus* (0.760%), while the lowest was observed between *Clonorchis sinensis* and *Opisthorchis felineus* (0.039%). In subfamily IB, the highest genetic variation between helminths was observed between *Trichobilharzia regenti* and *P. pacificus* (0.781%), while the lowest variation was observed between *Fasciola*

Table 3 Pairwise	genetic 1	variation	in percen	tage (%) ;	for helmin	ıth parasi	ites* in su	bfamily L	Ч							
	T.b	A.c	A.b	B.0	N.b	C.b	C.e	M.e	P:t	P.p	D.d	D.m	S.r	P.w	C.s	0.f
T.b																
A.c	0.663															
A.b	0.648	0.341														
B.o	0.665	0.351	0.228													
N.b	0.664	0.112	0.341	0.355												
C.b	0.713	0.149	0.355	0.355	0.176											
C.e	0.720	0.141	0.342	0.348	0.174	0.061										
M.e	0.663	0.324	0.307	0.315	0.329	0.359	0.354									
P.t	0.671	0.172	0.345	0.361	0.163	0.183	0.187	0.345								
P.p	0.680	0.211	0.342	0.338	0.209	0.206	0.197	0.315	0.262							
D.d	0.696	0.314	0.318	0.341	0.325	0.332	0.324	0.268	0.321	0.277						
D.m	0.670	0.336	0.363	0.375	0.346	0.352	0.341	0.361	0.368	0.318	0.335					
S.r	0.706	0.351	0.411	0.410	0.378	0.353	0.335	0.396	0.368	0.366	0.408	0.395				
P.w	0.802	0.671	0.591	0.645	0.651	0.717	0.707	0.664	0.679	0.681	0.662	0.643	0.658			
C.s	0.827	0.682	0.623	0.669	0.683	0.747	0.709	0.663	0.670	0.691	0.678	0.638	0.684	0.289		
O.f	0.846	0.680	0.628	0.667	0.673	0.760	0.725	0.656	0.676	0.702	0.685	0.649	0.706	0.272	0.039	
Note. T.l Nippostr P.p = Pis Clonorch	o = Trypa ongylus t tionchus uis sinens	nosoma b rasiliensi pacificus, is, Of = C	rucei, $A.c$ is, $C.b = 0$ $D.d = D_i$ Disthore.	c = Ancylı Caenorha itylenchu: his felineı	ostoma ce _. Ibditis bri _š s destructa US	ylanicum, zgsae, C.c ər, D.m =	A.b = Ap e = Caeno Dracuno	helencho orhabditis ulus medi	ides bican elegans, inensis, S.	udatus, B. $M.e = Me$ $r = Stron$	o = Bursc eloidogyn gyloides 1	ıphelench e enterold atti, P.w	tus okinar obii, P.t = = Parago	waensis, i Parelaph nimus we	V.b = ostrongylu stermani,	is tenuis, C.s =

	asites* in subfamily IB
	or helminth parc
	vercentage (%) j
	c variation in p
Table 4	Pairwise geneti

E.m
H.m
T.a
M.c
P.p
B.m
$\mathbf{A}.\mathbf{V}$
B.x
T.c
L.s
C.r
S.r
S.c
E.v
A.c
n P.t
. A.I
n T.ı
a S.1
g H.
h F.
Ŀ.

R.n

_	
7	
Τ.	

	0
0.005	0000
F.8	7.7

- H.a 0.388 0.392
- S.m 0.378 0.375 0.170
- T.r 0.463 0.468 0.277 0.362
- A.m 0.597 0.597 0.680 0.637 0.689
- *P.t* 0.583 0.583 0.680 0.667 0.701 0.057
- A.c 0.597 0.596 0.681 0.636 0.690 0.005 0.062
- E.v 0.588 0.593 0.667 0.658 0.692 0.242 0.240 0.246
- S.c 0.611 0.617 0.683 0.653 0.705 0.185 0.169 0.189 0.252
- S.r 0.631 0.629 0.698 0.689 0.737 0.346 0.327 0.349 0.355 0.345
- C.r 0.632 0.629 0.669 0.645 0.671 0.174 0.175 0.177 0.281 0.201 0.322

Pertanika J. Trop. Agric. Sci. 48 (3): 747 - 766 (2025)

- L.s 0.645 0.644 0.680 0.657 0.714 0.303 0.317 0.302 0.266 0.333 0.381 0.306
- *T.c* 0.602 0.607 0.670 0.652 0.687 0.198 0.179 0.202 0.129 0.224 0.331 0.220 0.247
- $0.648\ 0.644\ 0.690\ 0.704\ 0.710\ 0.281\ 0.264\ 0.280\ 0.290\ 0.291\ 0.341\ 0.291\ 0.330\ 0.272$ B.X
- $0.642\ 0.641\ 0.680\ 0.640\ 0.720\ 0.307\ 0.326\ 0.304\ 0.277\ 0.337\ 0.386\ 0.326\ 0.113\ 0.271\ 0.352$ A.v
- $0.645\ 0.644\ 0.682\ 0.648\ 0.728\ 0.300\ 0.308\ 0.297\ 0.247\ 0.325\ 0.374\ 0.290\ 0.098\ 0.241\ 0.328\ 0.082$ B.m
- 0.696 0.702 0.725 0.714 0.781 0.316 0.318 0.321 0.295 0.291 0.353 0.275 0.374 0.291 0.401 0.372 0.348 $d \cdot d$
 - $0.296\ 0.295\ 0.441\ 0.465\ 0.505\ 0.614\ 0.594\ 0.610\ 0.622\ 0.602\ 0.632\ 0.629\ 0.657\ 0.633\ 0.661\ 0.640\ 0.646\ 0.691$ M.c
- $0.292\ 0.292\ 0.432\ 0.447\ 0.519\ 0.634\ 0.612\ 0.630\ 0.642\ 0.613\ 0.647\ 0.656\ 0.669\ 0.639\ 0.658\ 0.652\ 0.658\ 0.701\ 0.062$ T.a
- $0.291\ 0.290\ 0.444\ 0.461\ 0.518\ 0.613\ 0.594\ 0.610\ 0.614\ 0.608\ 0.629\ 0.629\ 0.644\ 0.661\ 0.635\ 0.641\ 0.691\ 0.693\ 0.078\ 0.081$ H.m
- $0.299\ 0.298\ 0.442\ 0.457\ 0.521\ 0.642\ 0.637\ 0.646\ 0.629\ 0.653\ 0.658\ 0.675\ 0.643\ 0.663\ 0.658\ 0.666\ 0.702\ 0.062\ 0.052\ 0.081$ E.m
- $0.295\ 0.294\ 0.463\ 0.465\ 0.522\ 0.620\ 0.602\ 0.615\ 0.623\ 0.612\ 0.631\ 0.639\ 0.652\ 0.636\ 0.671\ 0.640\ 0.647\ 0.695\ 0.070\ 0.073\ 0.028\ 0.070$ R.n

Note. Fh = Fasciola hepatica, F.g = Fasciola gigantica; H.a = Heterobilharzia americana; S.m = Schistosoma mansoni; T.r = Trichobilharzia regent; A.m.= Angiostrongylus malaysiensis; Pt = Parelaphostrongylus tenuis; A. c = Angiostrongylus cantonensis; E. v = Enterobius vermicularis; S. c = Steinernemaxylophilus; A.v = Acanthocheilonema viteae; B.m = Brugia malayi; Pp = Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; <math>H.m = xylophilus; A.v = Aconthocheilonema viteae; B.m = Brugia malayi; Pp = Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; <math>H.m = xylophilus; A.v = Aconthocheilonema viteae; B.m = Brugia malayi; Pp = Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; <math>H.m = xylophilus; A.v = Aconthocheilonema viteae; B.m = Brugia malayi; Pp = Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; H.m = xylophilus; A.v = Aconthocheilonema viteae; B.m = Brugia malayi; Pp = Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; H.m = Xylophilus; A.v = Aconthocheilonema viteae; B.m = Brugia; Pp = Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; H.m = Xylophilus; A.v = Aconthocheilonema viteae; B.m = Brugia; Pristionchus pacificus; M.c = Mesocestoides corti; Ta = Taenia asiatica; H.m = Xylophilus; A.v = Aconthocheilonema viteae; Pristionchus pacificus; Pristionchus; Pristionchus pacificus; Pristionchus pacicarpocapsae; Sr = Strongyloides ratti; Cr = Caenorhabditis remanei; L.s = Litomosoides sigmodontis; <math>Tc = Toxocara canis; B.x = BursaphelenchusHymenolepis microstoma; E.m = Echinococcus multilocularis; R.n = Rodentolepis nana

hepatica and *Fasciola* gigantica (0.005%). Similarly, low variation was observed between Angiostrongylus malaysiensis and Angiostrongylus cantonensis (0.005%). Although both C. briggsae and O. felineus are in subfamily IA, C. briggsae is a nematode helminth, while O. felineus is a trematode. The same is true for T. regenti and P. pacificus in subfamily IB; T. regenti is a trematode, while P. pacificus is a nematode. Moreover, Clonorchis sinensis and Opisthorchis felineus in subfamily IA are both trematodes. In subfamily IB, F. hepatica and F. gigantica are parasites of the same genus, similar to A. malaysiensis and A. cantonensis, which explains the slight variation observed between these

Table 5

Pairwise	genetic variation in percentage (%) for
helminth	parasites* in subfamily IC

	C.r	D.p	T.v
C.r			
D.p	0.555		
T.v	0.627	0.625	

Note. C.r = Caenorhabditis remanei; D.p = Diploscapter pachys; T.v = Trichomonas vaginalis

Table 6

Pairwise genetic variation between subfamilies IA, IB, and IC

	IA	IB	IC	
IA				
IB	3.828			
IC	3.651	1.933		

parasites. There was high genetic variation between parasites in the subfamily IC (Table 5). However, only two ASCT sequences of helminth parasites were found with the addition of the previously characterised ASCT of *T. vaginalis*.

A pairwise genetic comparison between the subfamilies was also conducted, where subfamilies IB and IC were found to be most closely related to each other, with a mean genetic variation between the two groups of 1.933%. Moreover, subfamily IA was found to have more significant genetic variation than both subfamily IB (3.828%) and subfamily IC (3.651) (Table 6). The three subfamilies of the family I CoA-transferases exhibit only remote genetic relationships. Notably, subfamilies 1B and 1C demonstrated a closer genetic affinity to each other than to subfamily 1A.

CONCLUSION

The putative *Am*ASCT identified in this study possessed the ExG motif, which contains the active site's conserved glutamate residue, and the conserved GxGGxxD motif, which is part of the oxyanion hole. Both motifs are characteristic of the family I CoA-transferases. The phylogenetic tree showed that the putative *Am*ASCT gene was clustered with the previously characterised subfamily IB gene of F. hepatica ASCT, suggesting that the putative *Am*ASCT gene is a subfamily IB gene of the family I CoA-transferase. Acetate, a significant byproduct of energy metabolism in many parasites but not in their mammalian hosts, presents an appealing opportunity to advance novel antiparasitic medications. Further biochemical characterisation of this enzyme in rat lungworms and other parasites

is crucial for further understanding its role in parasite energy metabolism for survival in mammalian hosts.

ACKNOWLEDGEMENTS

The authors thank Mr Mohd Noor Md Isa for assistance with transcriptomic data analysis. This study was financially supported by the Fundamental Research Grant Scheme (FRGS) Research Grant (FRGS/1/2019/STG03/UPM/02/16) from the Ministry of Higher Education Malaysia.

REFERENCES

- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Journal of Bioinformatics*, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170
- Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. Nature Methods, 12(1), 59–60. https://doi.org/10.1038/nmeth.3176
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10, 421. https://doi.org/10.1186/1471-2105-10-421
- Cavalier-Smith, T., & Chao, E. E. (1996). Molecular phylogeny of the free-living archezoan *Trepomonas agilis* and the nature of the first eukaryote. *Journal of Molecular Evolution*, 43(6), 551–562. https://doi.org/10.1007/bf02202103
- Chan, A. H. E., Chaisiri, K., Dusitsittipon, S., Morand, S., & Ribas, A. (2020). Mitochondrial ribosomal genes as novel genetic markers for discrimination of closely related species in the *Angiostrongylus cantonensis* lineage. *Acta Tropica*, 211, 105645. https://doi.org/10.1016/j.actatropica.2020.105645
- Cock, P. J. A., Grüning, B. A., Paszkiewicz, K., & Pritchard, L. (2013). Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology. *PeerJ*, 1, e167. https://doi.org/10.7717/peerj.167
- de Mata, Z. S., de Bruyn, B., Saz, H., & Lo, H. S. (1997). Acetyl-CoA hydrolase activity and function in Ascaris suum muscle mitochondria. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 116(3), 379–383. https://doi.org/10.1016/s0305-0491(96)00266-0
- Eamsobhana, P. (2014). Eosinophilic meningitis caused by *Angiostrongylus cantonensis* A neglected disease with escalating importance. *Tropical Biomedicine*, *31*(4), 569–578.
- Eamsobhana, P., Lim, P., & Yong, H. (2015). Phylogenetics and systematics of *Angiostrongylus* lungworms and related taxa (Nematoda: Metastrongyloidea) inferred from the nuclear small subunit (SSU) ribosomal DNA sequences. *Journal of Helminthology*, 89(3), 317–325. https://doi.org/10.1017/s0022149x14000108
- Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research*, 39, W29–W37. https://doi.org/10.1093/nar/gkr367
- Fleck, C. B., & Brock, M. (2009). Re-characterization of Saccharomyces cerevisiae Ach1p: Fungal CoAtransferases are involved in acetic acid detoxification. Fungal Genetics and Biology, 46(6-7), 473–485. https://doi.org/10.1016/j.fgb.2009.03.004

- Fukao, T., Lopaschuk, G. D., & Mitchell, G. A. (2004). Pathways and control of ketone body metabolism: On the fringe of lipid biochemistry. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 70(3), 243–251. https://doi.org/10.1016/j.plefa.2003.11.001
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K. ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644–652. https://doi. org/10.1038/nbt.1883
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N. ... Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084
- Heider, J. (2001). A new family of CoA-transferases. *FEBS Letters*, 509(3), 345–349. https://doi.org/10.1016/ S0014-5793(01)03178-7
- Hunt, M. C., & Alexson, S. E. (2008). Novel functions of acyl-CoA thioesterases and acyltransferases as auxiliary enzymes in peroxisomal lipid metabolism. *Progress in Lipid Research*, 47(6), 405-421. https:// doi.org/10.1016/j.plipres.2008.05.001
- Lee, F. J., Lin, L. W., & Smith, J. A. (1989). Purification and characterization of an acetyl-CoA hydrolase from Saccharomyces cerevisiae. European Journal of Biochemistry, 184(1), 21–28. https://doi. org/10.1111/j.1432-1033.1989.tb14985.x
- Lee, F. J., Lin, L. W., & Smith, J. A. (1990). A glucose-repressible gene encodes acetyl-CoA hydrolase from Saccharomyces cerevisiae. Journal of Biological Chemistry, 265(13), 7413–7418. https://doi.org/10.1016/j. jbc.1990.003
- Low, S. Y., Lau, S. F., Ahmad, N. I., Sharma, R. S. K., Rosli, M. Z., Mohd-Taib, F. S., Md. Ajat, M. M., Kamaludeen, J., Syed Hussain, S. S., Wan, K.-L., Salleh, A., & Abdul Aziz, N. A. (2023). A cross-sectional study of *Angiostrongylus malaysiensis* in rats and gastropod hosts from recreational parks in Kuala Lumpur, Malaysia: Detection, risk factors and pathology. *Zoonoses and Public Health*, 70(7), 636-646. https://doi.org/10.1111/zph.13072
- Mack, M., & Buckel, W. (1997). Conversion of glutaconate CoA-transferase from Acidaminococcus fermentans into an acyl-CoA hydrolase by site-directed mutagenesis. FEBS Letters, 405(2), 209–212. https://doi. org/10.1016/s0014-5793(97)00187-7
- Martin, W., & Müller, M. (1998). The hydrogen hypothesis for the first eukaryote. Nature, 392, 37–41. https:// doi.org/10.1038/32096
- Murphy, G. S., & Johnson, S. (2013). Clinical aspects of eosinophilic meningitis and meningoencephalitis caused by Angiostrongylus cantonensis, the rat lungworm. Hawaii Journal of Medicine and Public Health, 72, 35–40.
- Rangarajan, E. S., Li, Y., Ajamian, E., Iannuzzi, P., Kernaghan, S. D., Fraser, M. E., Cygler, M., & Matte, A. (2005). Crystallographic trapping of the glutamyl-CoA thioester intermediate of family I CoA transferases. *Journal of Biological Chemistry*, 280(52), 42919–42928. https://doi.org/10.1074/jbc.m510522200

- Reeves, R. E., Warren, L. G., Susskind, B., & Lo, H. S. (1977). An energy-conserving pyruvate-to-acetate pathway in *Entamoeba histolytica*. *Journal of Biological Chemistry*, 252, 726–731.
- Rivière, L., van Weelden, S. W. H., Glass, P., Vegh, P., Coustou, V., Biran, M., van Hellemond, J. J., Bringaud, F., Tielens, A. G. M., & Boshart, M. (2004). Acetyl CoA-transferase in procyclic *Trypanosoma brucei*. *Journal of Biological Chemistry*, 279(44), 45337–45346. https://doi.org/10.1074/jbc.m407513200
- Roger, A. J., Clark, C. G., & Doolittle, W. F. (1996). A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proceedings of the National Academy of Sciences of the United States of America*, 93(25), 14618–14622. https://doi.org/10.1073/pnas.93.25.14618
- Sánchez, L. B., Galperin, M. Y., Müller, M. (2000). Acetyl-CoA synthetase from the amitochondriate eukaryote *Giardia lamblia* belongs to the newly recognized superfamily of acyl-CoA synthetases (Nucleoside diphosphate-forming). *Journal of Biological Chemistry*, 275, 5794–5803. https://doi.org/10.1074/ jbc.275.8.5794
- Sawanyawisuth, K., & Sawanyawisuth, K. (2008). Treatment of angiostrongyliasis. Transactions of the Royal Society of Tropical Medicine and Hygiene, 102(10), 990-996. https://doi.org/10.1016/j.trstmh.2008.04.021
- Saz, H. J., deBruyn, B., & de Mata, Z. (1996). Acyl-CoA transferase activities in homogenates of Fasciola hepatica adults. Journal of Parasitology, 82(5), 694–696. https://doi.org/10.2307/3283876
- Slom, T. J., Cortese, M. M., Gerber, S. I., Jones, R. C., Holtz, T. H., Lopez, A. S., Zambrano, C. H., Sufit, R. L., Sakolvaree, Y., Chaicumpa, W., Herwaldt, B. L., & Johnson, S. (2002). An outbreak of eosinophilic meningitis caused by *Angiostrongylus cantonensis* in travelers returning from the Caribbean. *New England Journal of Medicine*, 346(9), 668-675. https://doi.org/10.1056/nejmoa012462
- Stechmann, A., Hamblin, K., Pérez-Brocal, V., Gaston, D., Richmond, G. S., van der Giezen, M., Clark, C. Graham, & Roger, A. J. (2008). Organelles in *Current Biology*, 18, 580–585. https://doi.org/10.1016/j. cub.2008.03.037
- Steinbüchel, A., Müller, M. (1986). Anaerobic pyruvate metabolism of *Tritrichomonas foetus* and *Trichomonas vaginalis* hydrogenosomes. *Molecular and Biochemical Parasitology*, 20(1), 57–65. https://doi.org/10.1016/0166-6851(86)90142-8
- Tamura, K., Stecher, G., Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution, 38(7), 3022-3027. https://doi.org/10.1093/molbev/msab120
- Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673-4680. https://doi.org/10.1093/nar/22.22.4673
- Tielens, A. G. M., Rotte, C., Van-Hellemond, J. J., Martin, W. (2002). Mitochondria as we don't know them. *Trends in Biochemical Sciences*, 27, 564–572. https://doi.org/10.1016/s0968-0004(02)02193-x
- Tielens, A. G. M., van Grinsven, K. W. A., Henze, K., van Hellemond, J. J., & Martin, W. (2010). Acetate formation in the energy metabolism of parasitic helminths and protists. *International Journal for Parasitology*, 40(4), 387–397. https://doi.org/10.1016/j.ijpara.2009.12.006
- van Grinsven, K. W. A., Rosnowsky, S., van Weelden, S. W. H., Pütz, S., van der Giezen, M., Martin, W., van Hellemond, J. J., Tielens, A. G. M., & Henze, K. (2008). Acetate CoA-transferase in the hydrogenosomes

of *Trichomonas vaginalis*: Identification and characterization. *Journal of Biological Chemistry*, 283(3), 1411–1418. https://doi.org/10.1074/jbc.m702528200

- van Grinsven, K. W. A., van Hellemond, J. J., Tielens, A. G. M. (2009). Acetate CoA-transferase in the anaerobic mitochondria of *Fasciola hepatica*. *Molecular and Biochemical Parasitology*, 164(1), 74–79. https://doi. org/10.1016/j.molbiopara.2008.11.008
- van Hellemond, J. J., Opperdoes, F. R., & Tielens, A. G. M. (1998). Trypanosomatidae produce acetate via a mitochondrial acetate CoA transferase. *Proceedings of the National Academy of Sciences of the United States of America*, 95(6), 3036–3041. https://doi.org/10.1073/pnas.95.6.3036
- Wang, Q.-P., Lai, D.-H., Zhu, X.-Q., Chen, X.-G., & Lun, Z.-R. (2008). Human angiostrongyliasis. *The Lancet Infectious Diseases*, 8, 621–630. https://doi.org/10.1016/s1473-3099(08)70229-9
- Watthanakulpanich, D., Jakkul, W., Chanapromma, C., Ketboonlue, T., Dekumyoy, P., Lv, Z., Chan, A. H. E., Thaenkham, U., & Chaisiri, K. (2021). Co-occurrence of *Angiostrongylus malaysiensis* and *Angiostrongylus cantonensis* DNA in cerebrospinal fluid: Evidence from human eosinophilic meningitis after ingestion of raw snail dish in Thailand. *Food and Waterborne Parasitology*, 24, e00128. https://doi. org/10.1016/j.fawpar.2021.e00128
- White, H., & Jencks, W. P. (1976). Mechanism and specificity of succinyl-CoA:3-ketoacid coenzyme A transferase. *Journal of Biological Chemistry*, 251(6), 1688–1699. https://doi.org/10.1016/S0021-9258(17)33704-3
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the *ExPASy server*. *Methods in Molecular Biology*, 112, 531–552. https://doi.org/10.1385/1-59259-584-7:531
- Xu, L., Xu, M., Sun, X., Xu, J., Zeng, X., Shan, D., Yuan, D., He, P., He, W., Yang, Y., Luo, S., Wei, J., Wu, X., Liu, Z., Xu, X., Dong, Z., Song, L., Zhang, B., Yu, Z. ... Wu, Z. (2019). The genetic basis of adaptive evolution in parasitic environment from the *Angiostrongylus cantonensis* genome. *PLoS Neglected Tropical Diseases*, 13(11), e0007846. https://doi.org/10.1371/journal.pntd.0007846
- Zeiher, C. A., & Randall, D. D. (1990). Identification and characterization of mitochondrial acetyl-coenzyme A hydrolase from *Pisum sativum* L. seedlings. *Plant Physiology*, 94(1), 20-27. https://doi.org/10.1104/ pp.94.1.20