

## **Supporting Information**

# **Identification of the full 46 cytochrome P450 (*CYP*) complement and modulation of *CYP* expression in response to water-accommodated fractions (WAFs) of crude oil in the cyclopoid copepod *Paracyclopsina nana***

**Jeonghoon Han<sup>a,#</sup>, Eun-Ji Won<sup>a,#</sup>, Hui-Su Kim<sup>a</sup>, David R. Nelson<sup>b</sup>, Su-Jae Lee<sup>c</sup>, Heum Gi Park<sup>d</sup>, and Jae-Seong Lee<sup>a,\*</sup>**

<sup>a</sup>*Department of Biological Science, College of Science, Sungkyunkwan University, Suwon 440-746, South Korea*

<sup>b</sup>*Department of Microbiology, Immunology, and Biochemistry, University of Tennessee, Memphis, Tennessee 38163, United States*

<sup>c</sup>*Department of Life Sciences, College of Natural Sciences, Hanyang University, Seoul 133-791, South Korea*

<sup>d</sup>*Department of Marine Resource Development, College of Life Sciences, Gangneung-Wonju National University, Gangneung 210-702, South Korea*

---

<sup>#</sup> These authors contributed equally to this work.

\*Corresponding author: Department of Biological Science, College of Science, Sungkyunkwan University, Suwon 440-746, South Korea.

Tel.: 82 31 290 7011; E-mail: jslee2@skku.edu (J.-S. Lee)

## **Table of contents**

<b>Supplementary Table</b>	-----	<b>S3</b>
<b>Figure legend</b>	-----	<b>S5</b>
<b>Figure</b>	-----	<b>S7</b>

**Table S1.** GenBank accession numbers of genes in this study

GenBank Accession number	ORF size (a.a.)	Clan	Matched gene	Matched species	E-value
<i>CYP 18E1</i> (KP899564)	543	2	CYP18A1 (XP393885)	<i>Apis mellifera</i>	2e-154
<i>CYP306A1</i> (KP899565)	489	2	CYP306A1 (XP006557887)	<i>Apis mellifera</i>	3e-112
<i>CYP 307F1</i> (KP899566)	532	2	CYP307E1 (AIL94134)	<i>Tigriopus japonicus</i>	8e-123
<i>CYP 3029B1</i> KP899567	492	2	CYP2J2 (XP587546)	<i>Bos taurus</i>	3e-79
<i>CYP 3033B1</i> (KP899568)	558	2	CYP2K1 (KVF81929)	<i>Struthio camelus australis</i>	1e-67
<i>CYP 3034B1</i> KP899569	509	2	CYP2J2 (KFO34696)	<i>Fukomys damarensis</i>	6e-73
<i>CYP 3036B1</i> (KP899570)	504	2	CYP2J6 (XP004068236)	<i>Oryzias latipes</i>	4e-77
<i>CYP 3036B2</i> (KP899571)	508	2	CYP2J6 (AIE17355)	<i>Cynoglossus semilaevis</i>	3e-75
<i>CYP 3037B1</i> (KP899572)	483	2	CYP2J6 (KDR19530)	<i>Zootermopsis nevadensis</i>	7e-96
<i>CYP 3038C1</i> (KP899573)	509	2	CYP2J2 (KDR22559)	<i>Zootermopsis nevadensis</i>	3e-56
<i>CYP 3041B1</i> (KP899574)	480	2	CYP2J6 (KDR19530)	<i>Zootermopsis nevadensis</i>	4e-89
<i>CYP 3075A1</i> (KP899575)	519	2	CYP2K4 (XP005420420)	<i>Geospiza fortis</i>	2e-79
<i>CYP 3075A2</i> (KP899576)	521	2	CYP2K1 (KVF81930)	<i>Struthio camelus australis</i>	4e-78
<i>CYP 3075A3</i> (KP899577)	519	2	CYP2C9 (XP009814712)	<i>Gavia stellata</i>	3e-86
<i>CYP 3075A4</i> (KP899578)	518	2	CYP2J6 (XP008293655)	<i>Stegastes partitus</i>	2e-86
<i>CYP 3075A5</i> (KP899579)	508	2	CYP2J2 (XP004896334)	<i>Heterocephalus glaber</i>	8e-75
<i>CYP 3075B1</i> (KP899580)	353	2	CYP2J6 (XP005795706)	<i>Xiphophorus maculatus</i>	1e-69
<i>CYP 3075B2</i> (KP899581)	509	2	CYP2J2 (XP006879803)	<i>Elephantulus edwardii</i>	2e-94
<i>CYP 3075C1</i> (KP899582)	522	2	CYP2J2 (XP005956793)	<i>Pantholops hodgsonii</i>	2e-84
<i>CYP 3076A1</i> (KP899583)	481	2	CYP2J2 (XP008137597)	<i>Eptesicus fuscus</i>	2e-65
<i>CYP 3076A2</i> (KP899584)	495	2	CYP2J2 (KDR22559)	<i>Zootermopsis nevadensis</i>	2e-65
<i>CYP 3082A1</i> (KP899585)	348	2	CYP2K1 (XP006137093)	<i>Pelodiscus sinensis</i>	3e43
<i>CYP 3083A1</i> (KP899586)	285	2	CYP2J6 (NP034138)	<i>Mus musculus</i>	5e-45
<i>CYP 3129A1</i> (KP899587)	507	2	CYP2J6 (XP004068238)	<i>Oryzias latipes</i>	9e-79
<i>CYP 3130A1</i> (KP899588)	504	2	CYP2J6 (XP005948149)	<i>Haplochromis burtoni</i>	4e-79
<i>CYP 3026D1</i> (KP899589)	528	3	CYP6A13 (XP966437)	<i>Tribolium castaneum</i>	3e-95
<i>CYP 3026D2</i> (KP899590)	529	3	CYP6A13 (KDR19800)	<i>Zootermopsis nevadensis</i>	1e-88
<i>CYP 3026D3</i> (KP899591)	528	3	CYP6A2 (XP969875)	<i>Tribolium castaneum</i>	2e-89
<i>CYP 3027D1</i> (KP899592)	481	3	CYP3A24 (NP001123376)	<i>Ovis aries</i>	3e-70
<i>CYP 3027E1</i> (KP899593)	489	3	CYP3A24 (ACO15001)	<i>Caligus clemensi</i>	6e-112
<i>CYP 3027F1</i> (KP899594)	485	3	CYP3A24 (ACO10681)	<i>Caligus rogercresseyi</i>	1e-84
<i>CYP 3027F2</i> (KP899595)	486	3	CYP3A24 (ACO15001)	<i>Caligus clemensi</i>	4e-102
<i>CYP 3079A1</i> (KP899596)	507	3	CYP3A78 (ADR00355)	<i>Phascolarctos cinereus</i>	4e-70
<i>CYP 3081A1</i> (KP899597)	528	3	CYP9B2 (XP001855257)	<i>Culex quinquefasciatus</i>	3e-60
<i>CYP 3081B1</i> (KP899598)	510	3	CYP9E2 (XP008197525)	<i>Tribolium castaneum</i>	3e-69
<i>CYP 3081C1</i> (KP899599)	514	3	CYP9E2 (001603857)	<i>Nasonia vitripennis</i>	3e-70
<i>CYP 44D2</i> (KP899600)	519	MT	CYP44 (KDR13959)	<i>Zootermopsis nevadensis</i>	1e-153
<i>CYP 302A1</i> (KP899601)	456	MT	CYP302A1 (AGT57842)	<i>Leptinotarsa decemlineata</i>	3e-116

**Table S1.** Continued.

GenBank Accession number	ORF size (a.a.)	Clan	Matched gene	Matched species	E-value
<i>CYP 314A1</i> (KP899602)	524	MT	CYP314A1 (AGI92301)	<i>Laodelphax striatella</i>	1e-103
<i>CYP 315A1</i> (KP899603)	428	MT	CYP315A1 (KDR16797)	<i>Zootermopsis nevadensis</i>	1e-70
<i>CYP 3020B1</i> (KP899604)	500	MT	CYP301A1 (XP001948959)	<i>Acyrthosiphon pisum</i>	1e-99
<i>CYP 3077A1</i> (KP899605)	515	MT	CYP49A1 (XP001946744)	<i>Acyrthosiphon pisum</i>	2e-75
<i>CYP 3078A1</i> (KP899606)	509	MT	CYP49A1 (XP006562546)	<i>Apis mellifera</i>	1e-65
<i>CYP 3078B1</i> (KP899607)	501	MT	CYP301A1 (KDR20887)	<i>Zootermopsis nevadensis</i>	1e-64
<i>CYP20A1</i> (KP899608)	473	20	CYP20A1 (XP003223588)	<i>Anolis carolinensis</i>	1e-65
<i>CYP 3080A1</i> (KP899609)	432	26	CYP26A1 (XP007890111)	<i>Callorhinus milii</i>	1e-66

**Fig. S1.** PAH composition of the water-accommodated fraction (WAF), (A) concentrations of each component in WAF and (B) the total concentrations of 16 PAHs and alkylated PAHs in WAF. Abbreviations: N0, Naphthalene; N1, C1-Naphthalene; N2, C2-Naphthalene; N3, C3-Naphthalene; N4, C4-Naphthalene; AC, Acenaphthylene; AE, Acenaphthene; F, Fluorene; F1, C1-Fluorene; F2, C2-Fluorene; F3, C3-Fluorene; D0, Dibenzothiophene; D1, C1-Dibenzothiophene; D2, C2-Dibenzothiophene; D3, C3-Dibenzothiophene; P0, Phenanthrene; P1, C1-Phenanthrene; P2, C2-Phenanthrene; P3, C3-Phenanthrene; P4, C4-Phenanthrene; AN, Anthracene; FL, Fluoranthene; PY, Pyrene; AA, Benz[a]anthracene; C0, Chrysene; C1, C1-Chrysene; C2, C2-Chrysene; C3, C3-Chrysene; BB, Benzo[b]fluoranthene; BK, Benzo[k]fluoranthene; BE, Benzo[e]pyrene; BA, Benzo[a]pyrene; PE, Perylene; IC, Indeno[1,2,3-cd]pyrene; DB, Dibenzo[a,h]anthracene; BZ, Benzo[ghi]perylene

**Fig. S2.** Phylogenetic tree of the different CYP2 clan members.

**Fig. S3.** Phylogenetic tree of the different CYP3 clan members.

**Fig. S4.** Phylogenetic tree of the different mitochondrial clan, clan 20, and clan 26 members.

**Fig. S5.** (A) Phylogenetic tree of the *PN-CYP* genes with three WAF-sensitive *CYP* genes of *T. japonicus* and (B) the modulations of three *CYP* genes in response to WAF exposure. The values of modulated mRNA expression by WAF exposure are from Han et al. (2014).

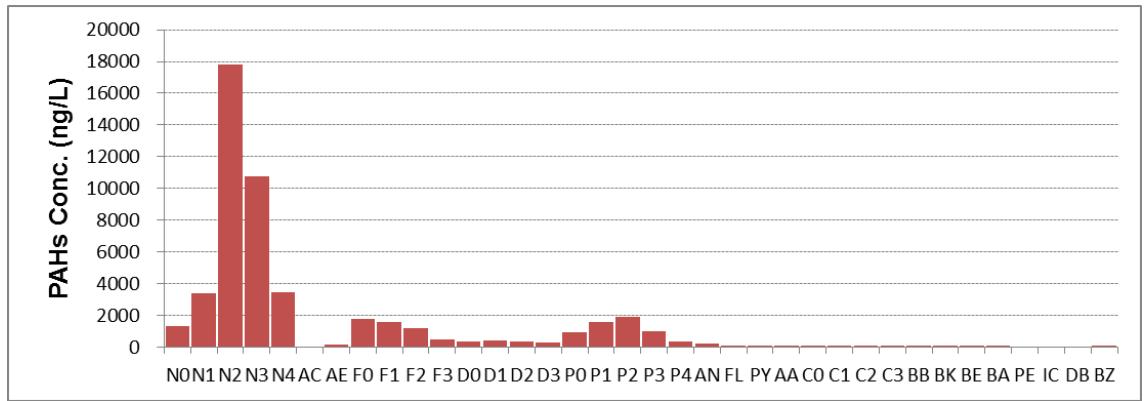
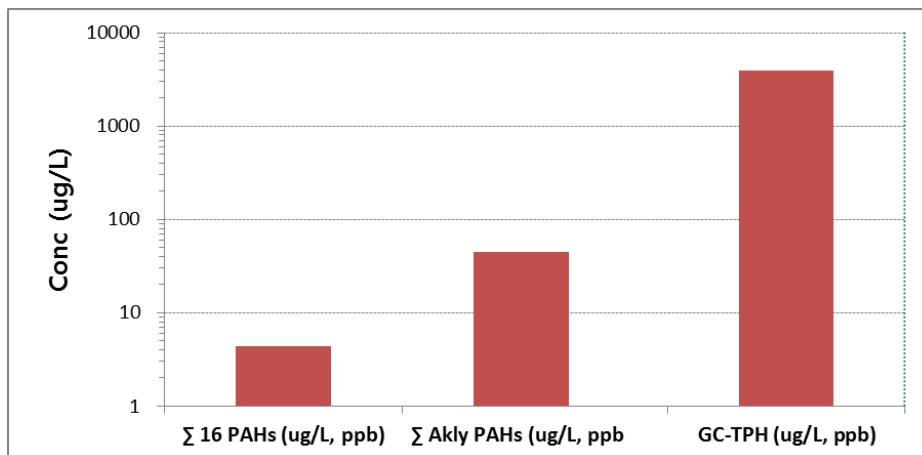
**Fig. S6.** Expression patterns of the four *PN-CYP* genes, (A) *CYP3027D1*, (B) *CYP3027E1*, (C) *CYP3027F1*, and (D) *CYP3027F2* in response to different concentrations (control, 20, 40, 60, and 80%) of WAF.

**Fig. S7.** Expression patterns of the four *PN-CYP* genes; *CYP3027D1*, *CYP3027E1*, *CYP3027F1*, and *CYP3027F2*) in response to a time course of exposure to alkylated

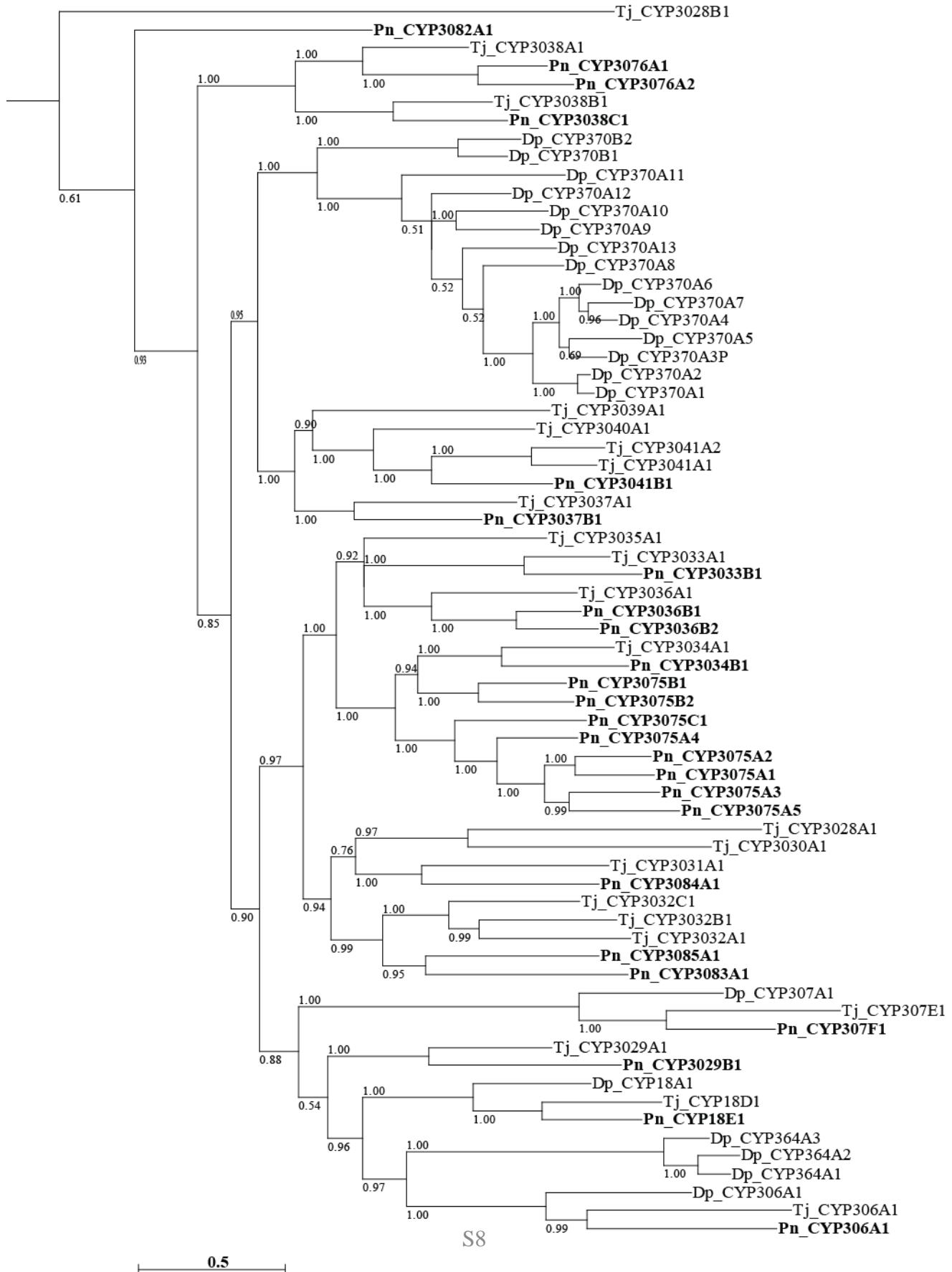
PAHs, (A) 2-ethyl-phenanthrene (0.1 mg/L) and (B) 1,8-dimethyl-9H-fluorene (0.2mg/L).

**Fig. S8.** Complete cDNA sequences of (A) *AhR* and (B) *ARNT*.

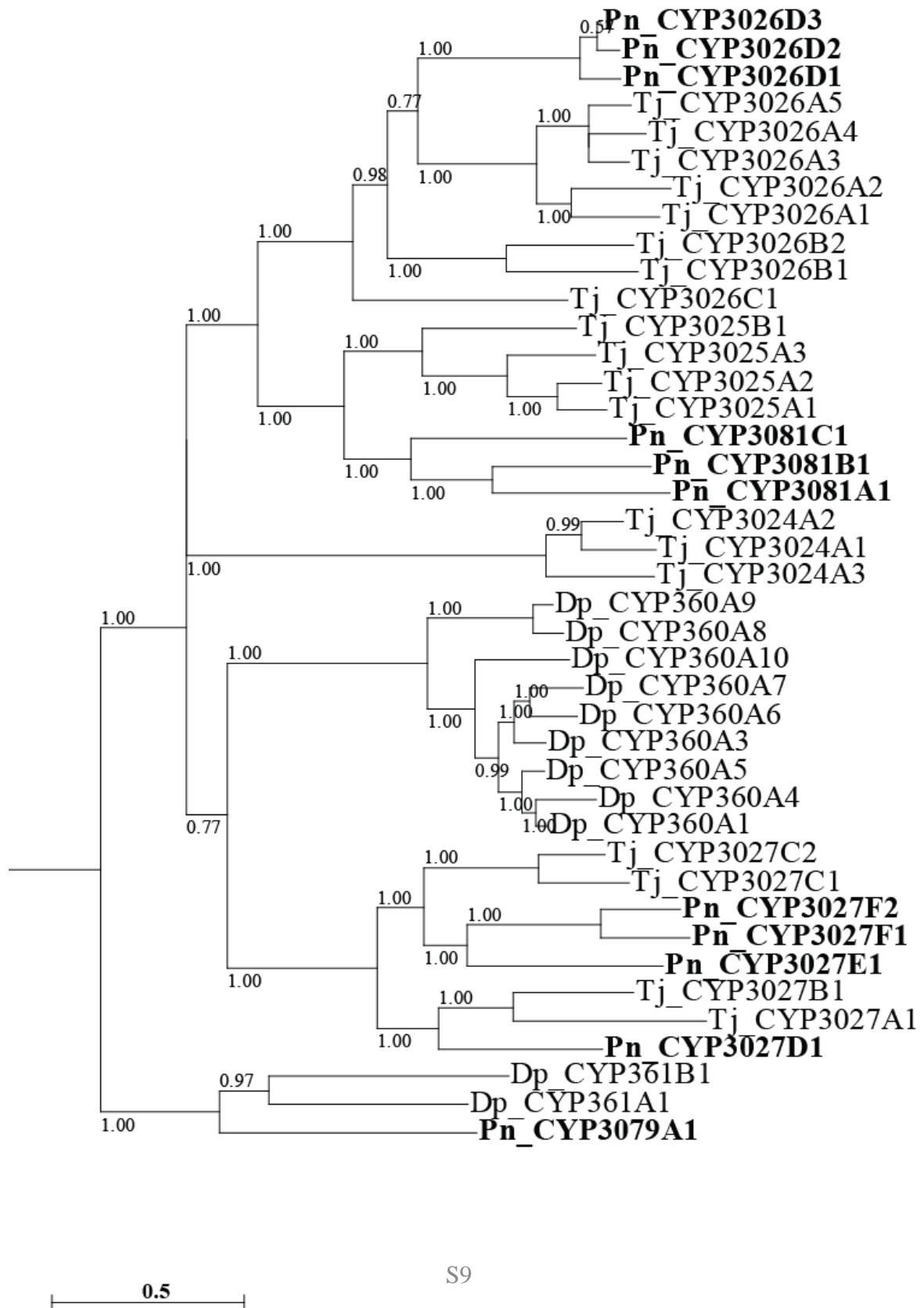
**Fig. S9.** Expression patterns of the (A) *AhR* and (B) *ARNT* genes from *T. japonicus* in response to WAF exposure.

**A)****B)**

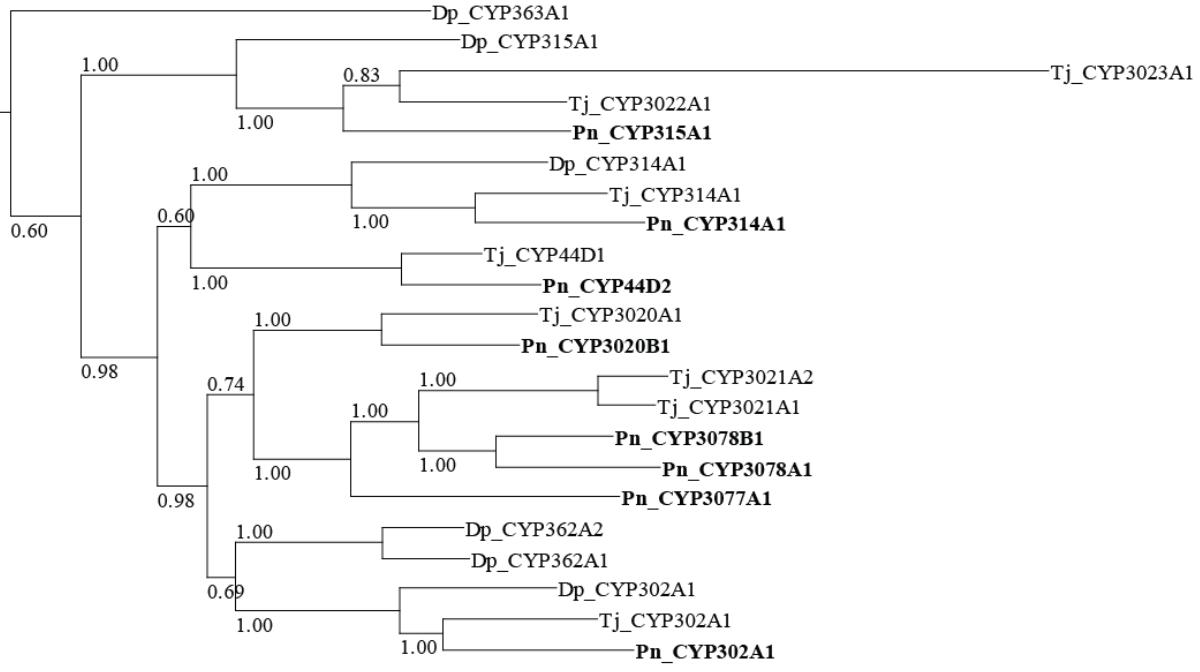
## Clan 2



## Clan 3

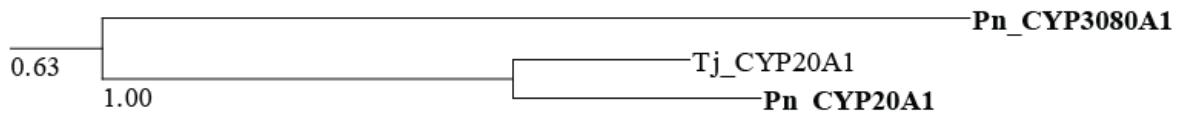


# Clan MT



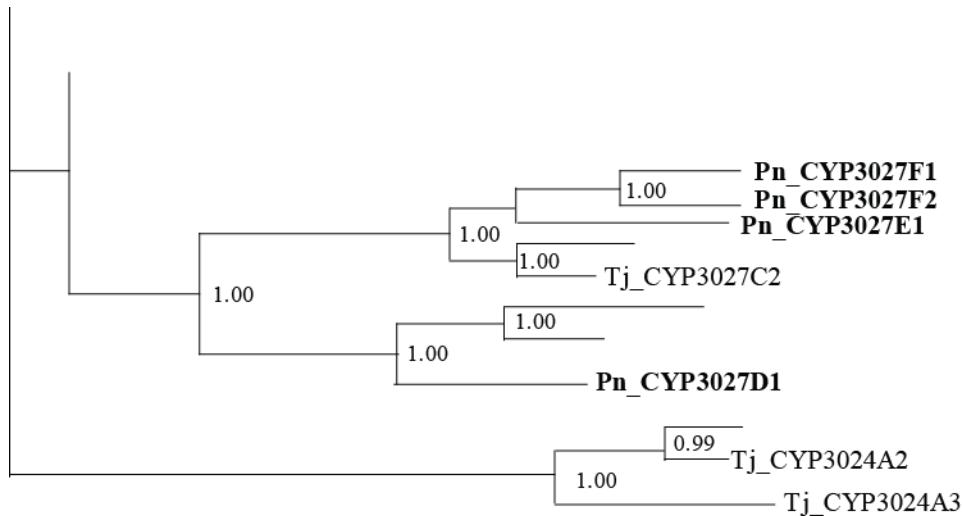
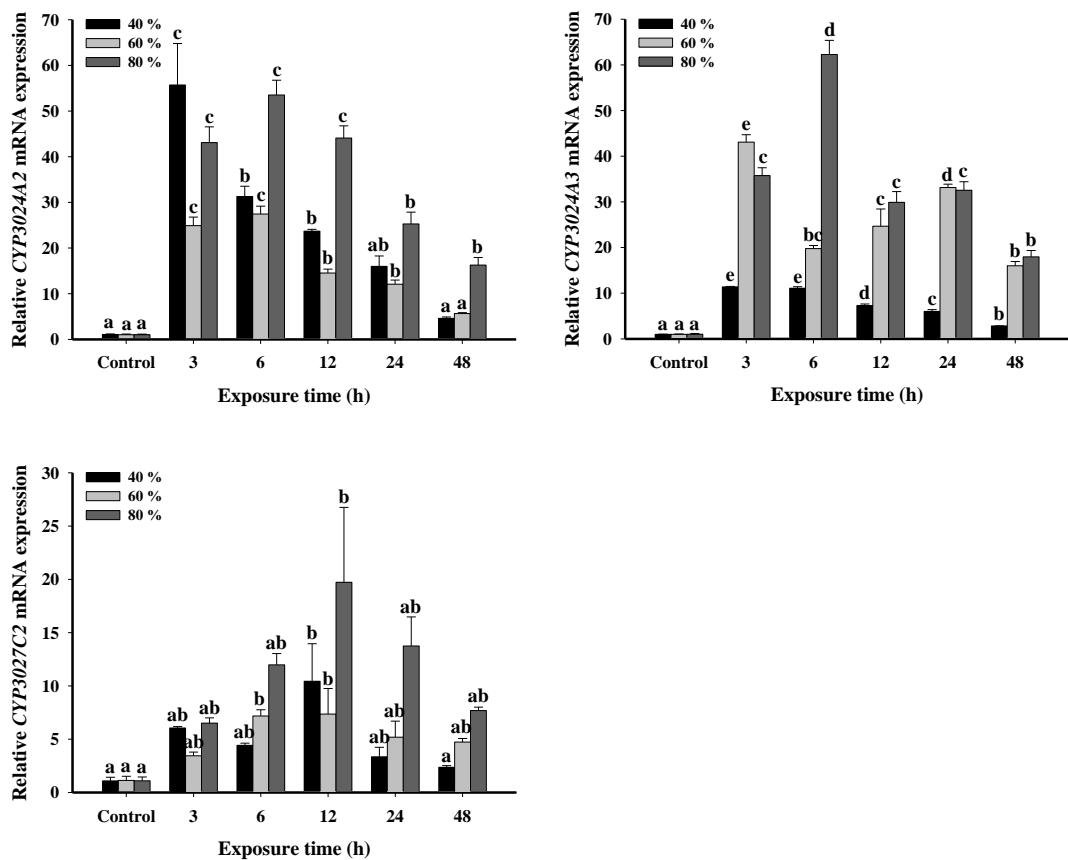
0.5

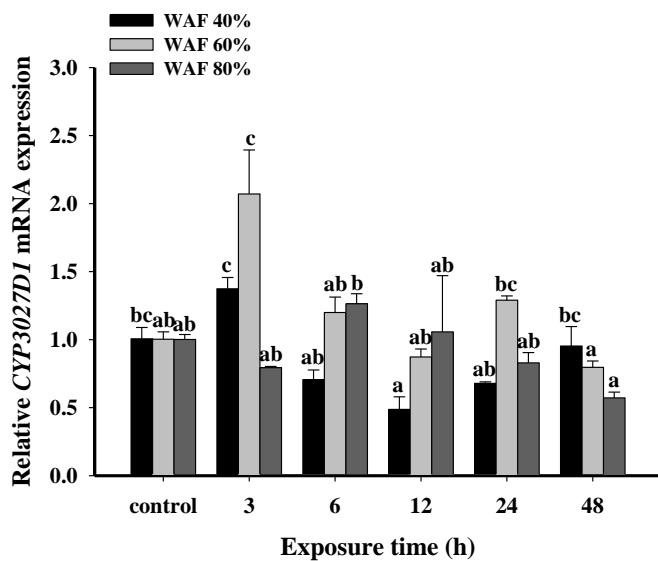
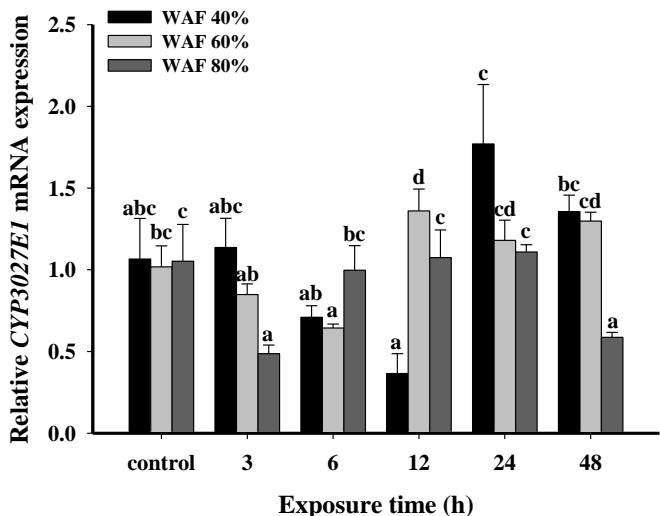
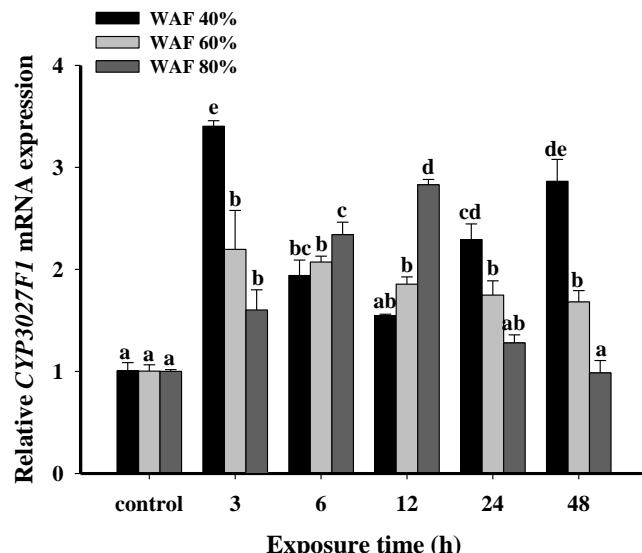
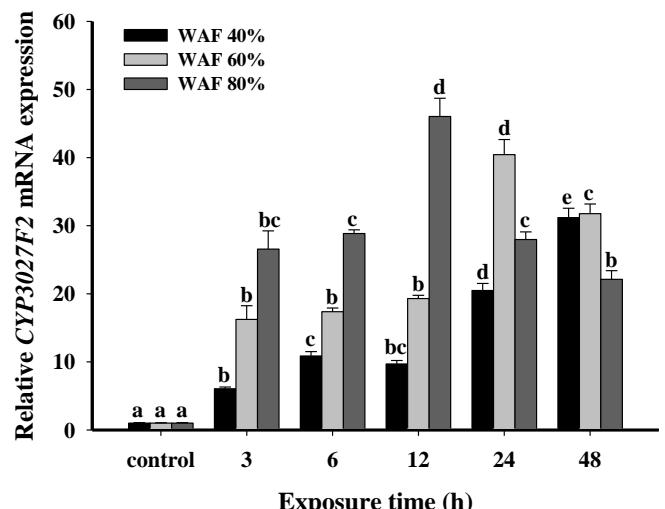
# Clan 20, 26



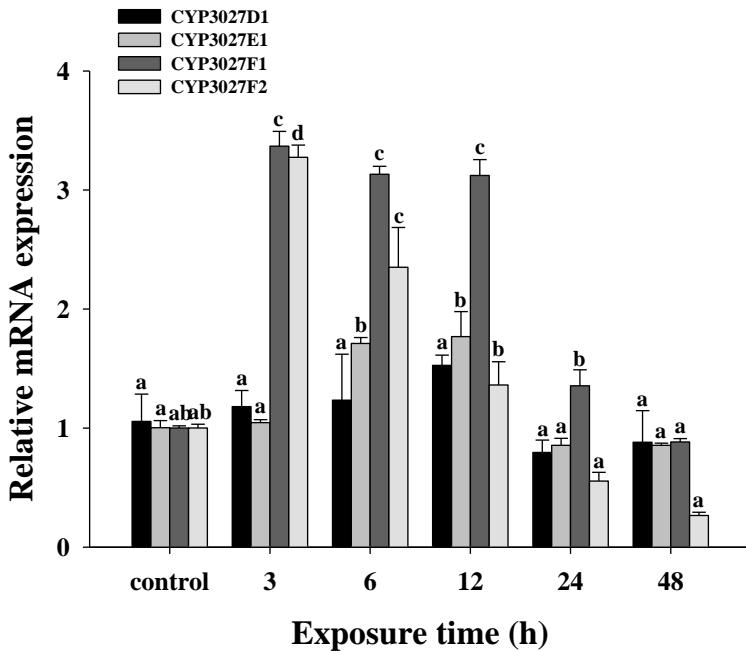
0.5

A)

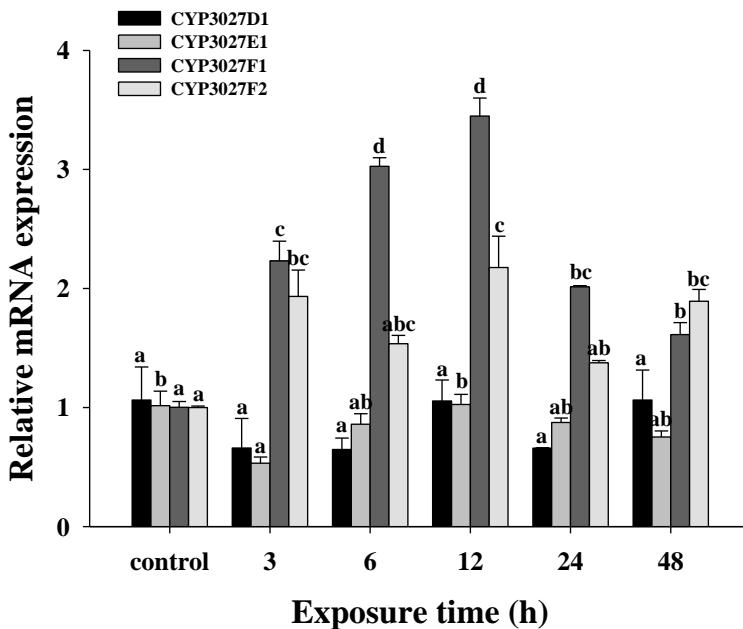
B) *Tigriopus japonicus*

**A)****B)****C)****D)**

**A) 2-ETHYL-Phenanthrene  
0.1mg/L (NOEC)**



**B) 1,8-dimethyl-9H-fluorene  
0.2mg/L (NOEC)**



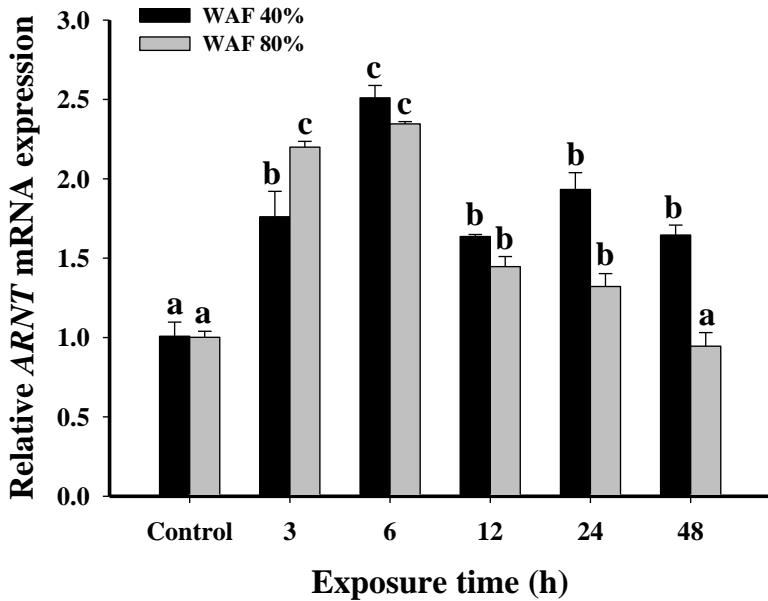
## A) Aryl hydrocarbon receptor (AhR)

1 ATGAATTCTTCGCCAATCCTCTGAGCACAGTGATGCCACCAAAAGGCAGAACAGATCT  
 M N S S P N P L S T V Y A T K R R R R S  
 61 AATAAAAGCCTGCGGCCACCCAAACAAAGGACAGCTCGAGCCAGGCAAAGCAACCC  
 N K S L R P T P T K D S S S Q A K S N P  
 121 TCCAAACGTCACAGGGAACGACTCAATGGCGAACACTGACCACTTGGCTCACTTCTGCC  
 S K R H R E R L N G E L D H L A S L L P  
 181 TTCGAGCAGAACATTCTGTCCAAGCTGGACAGGGCTGTCCATCCTCAGATTGGCTGTGCC  
 F E Q N I L S K L D R L S I L R L A V A  
 241 TTCCCTCAGGACCAAGAACCTACTTCAAGTGGCCATCGAACCGCAGCGGGGAGCAGCTGGTG  
 F L R T K T Y F Q V A M Q R E R D E L V  
 301 GACGGGCTTACCGCAGAGATTCTCACCTACGACAACCATCTCTGGACGGAGACATC  
 D G P Y R R D F L T Y D N H L L D G D I  
 361 TTCCCTCAGGCTCTAACGGGTTTTGATGATCCTGACTGCAACGGGAGCTTTCTTT  
 F L Q A L N G F L M I L D C N G E L F F  
 421 GCCACGACACAAATAGAAAAGCTATCTAGGGTTCATCGATGTAATTCAACAGTCT  
 A T H T I E T Y L G F H Q S D V I H Q S  
 481 GTGTATGAGCTAGTCCATTAGAGGATCGCGAGGACTGCAACGTCAGCTCGCTGGAC  
 V Y E L V H S E D R E E L Q R Q L G W D  
 541 TCGTCTCGGCCAACAGGATTGGCATCGGCTACAGGGAGCTGCTCGCTCGGACCAT  
 S F L G Q E D S G I G L Q E L L A S D H  
 601 CATACCTGCTCGACAGATCTTACCGTTGATCATCGAGTCGCTCGACAACACCTCG  
 H H L L D R S F T V R F R C L L D N T S  
 661 GGATTCTCAGGCTGGATATTGGGGAGATCAAGGTGCTCACGGCAAACAAAGAAG  
 G F L R L D I R G K I K V L H G Q N K K  
 721 AGTGAGGAGGCCTGGGGCTGTTCTGCTGTGACGCCCTCGGCCCCGCGCTCGCTG  
 S E E A P L A L F C L C T P F G P P S L  
 781 CTCGAGATCCCGCACAGGAGGTGATGTTCAAGTCGAAGCACAAGCTGACTTGGCTC  
 L E I P H K E V M F K S K H K L D F G F  
 841 GTCTCCATGGACCAGCGGGCAAGGAGATGCTTCAACTACGACGAGGAAGAGATGTCAGAC  
 V S M D Q R G K E M L Q Y D E E M S D  
 901 ATTGGAGGCTACGACCTCGTCCACACAGACGACCTCCCTACGTCGCCAGGCCACCAA  
 I G G Y D L V H H D D L A Y V A S A H Q  
 961 GAACTGCTGAAACGGGTGCGAGCGGGATGATGCCCTACAGACTCCAAGCAAAGATCAA  
 E L L K T G A S G M I A Y R L Q A K D Q  
 1021 AAGTGGCAATGGCTCAGACCTCATCGAGATTAGTGTACAAAAACTCCAAGCCGGACTTT  
 K W Q W L Q T S S R L V Y K N S K P D F  
 1081 ATCATCTCCACGCAACGGCCACTGATGGAGGAGGGCCGAGATTGCTGGCAACAGA  
 I I S T H R P L M E E E G R D L L G K R  
 1141 ACGATGGACTTAAAGTGCACCTACTGGACGCTGGCTTAACAGTCCTACTCAACGAC  
 T M D F K V T Y L D A G L T S A Y F N D  
 1201 GGCACCCGCCAACCTGAACGGCTGGCCACGGCCAGCGTCAACGCCAACCTGGT  
 G D P P N S N G S A T S G S V N A N S G  
 1261 AGCGCGGCTCGAGCAAAGTGCACAAAGCGTACAAAGCGCATCTCGAGATTCTCAGC  
 S G G S S K V P K R Y K A H L R D F L S  
 1321 TCGTCCGGAACGAAACGCTCCAAAACCACACGCTTCCGTTCCGCTCGACATTAC  
 S C R T K R S K N H N A S V S A S D I Y  
 1381 GTGGACCCGGCGGCCGCGGCCCTCGCCCTACGCCACAGCTACTCCGCCCCG  
 V D P A A A A A S A L P Y A T V Y S A P  
 1441 CCCTCGGCCGCGCCGCATACGCCACCGACAGCCCTCTGTACATGCGACAAACATCACC  
 P S A A P A Y A T D S P L Y M Q Q N I T  
 1501 TCGTCTCCATATCGACCATACCCGTGGACATAAGGTATTCTCTCGAGATTG  
 S S P Y Q T I Y P V D N R Y F S S E Y L  
 1561 GCCAGCTACAGATCTCTACCGCCACCTACTACCCAGAACATCGCAGCTAGTGCCTCATGCT  
 A S Y R S L T A T Y Y P E Y A A S A H A  
 1621 GCTGCTGGCTACATGGCAATGGCTACTTTGACGTGGCTCGCCGAGGTGCTCACCTCC  
 A A G Y I G N G Y F D V A S P R S L T S  
 1681 ATCCCCTATGACCGGCCCTCGCAATTCAACCGCTACTTTGAGCGGACAAAGTCGACTGC  
 I P Y D A A S Q F N R Y F E A D K V D C  
 1741 AAGTACAACGTGCCAACCTCGAATGCCAAAAGCGGAATTC  
 K Y N V P N S E S P K A E F

## B) Aryl hydrocarbon receptor nuclear translocator (ARNT)

1 ATGGCCATGCCAACCATGATGCAGCACCAAGCAGCACCAACCCCCGAGGTGGCG  
 M A M P T H M M Q H H Q Q H H H P E V A  
 61 CAGCTGCACCAGCCGCCGGTCACCAACCAAGCAGCCGTCGTGCACCCACGTGCAGGCCAAG  
 Q L H Q P A G H H Q P V V H H V Q A K  
 121 CGGCCCTGGCCGAGCGCAGCAGCAGCCTAAATCAGATCTGGACGACCTCGACGCC  
 R P R P Q P Q Q Q P Q N Q I L D D L D G  
 181 GACCCCATGCAGCAGAAAGTACGGCGCATGGACGAGCCGCATCTGACCCGGCTGGGCC  
 D P M Q Q K Y G R M D A A A S D P A A A  
 241 GCGGCCGCGGCCCTCGCCGCTCCCGCTCCAGCTGGAGAACCCGGGGTCGGTGGGCCG  
 A A A A L A A S A S Q L E N P G S V G P  
 301 GAGAAGGAGAAGTACGCGGGAGAACACTGCGAGATCGAGCGGGCGGGAAACAAG  
 E K E K Y A R E N H C E I E R R R R N K  
 361 ATGTCGGCCTACATCACGGAGCTGCGACATGGGCCACGTGCAACGCGCTGGCGCG  
 M S A Y I T E L S D M V P T C N A L A R  
 421 AAGCCGACAAAGTGACCATCTGCGCATGGCGTGGCGCACATCAAGCAGCTGCGCGG  
 K P D K L T I L R M A V A H I K Q L R G  
 481 CTGCCGCCACGCCGCCAACGACTCGTCGATAAGCCGTCGTTCTGACGGACCG  
 L P A T P P P N D S S Y K P S F L T D A  
 541 GAGCTGAAGCACCTGATCTGGAGGCCGAGCGGGTCCCTGTTCTGTTCTGCGAC  
 E L K H L I L E A A D G F L F V V S C D  
 601 AGCGGCCGCGTGATCTACGTCGACTCGGTGACGCCGGTGCTCAACCACAGCAGGCC  
 S G R V I Y V S D S V T P V L N H T Q A  
 661 GACTGGTCGGCACCTCCATCTACGACCACGTCCACCCCGAGGACACGGACAAGGTGAAG  
 D W F G T S I Y D H V H P E D T D K V K  
 721 GAGCAGCTGTCGACCTCCGGACTCTCCGGATCTTCAGCAAGCAACAACGGCCGATCCTC  
 E Q L S T S P D S S G S S S N N G R I L  
 781 GATCTCAAGACGGGACGGTAAGAAGGAGGGCAGCCAGACTCGATGCCGCTCGCATG  
 D L K T G T V K K E G S Q S S M R L C M  
 841 GGCTGCCGCCGGGCTCATCTGCCATGCCGCTCGCGCAGGTCACCCGGCTCCCTA  
 G S R R G F I C R M R L G A V N P A S L  
 901 GGCTACATGACCGTGTGCGCACGCCAACGCTGCTGCCGCGCGAGCGTACCGGGCAG  
 G Y M T R V R T R N V L R G A S V T G Q  
 961 AGCGACGGGACGGCAGCTTACCGCTGTCCTACTGCACGGGCTACATCAAGAACGGCG  
 S D G H G D F T V W H C T G Y I K N W P  
 1021 CCCAGGGCCTGCCGCCAACACCCACTGCCAACAGGAGAACGGCCACCACCGAC  
 P Q G L A A H H L H E D N G H H P D  
 1081 GCGTCGTCCTGCTGCCCTGTCGGCATCGGCCCTCCAGGTACCTCCATGCCAACAGC  
 A S S C C L V A I G R L Q V T S M P N S  
 1141 CACGACCTTCCGGAGCCGGGATCGGAGTGCAGGCCAGGAGTTCCCTCGGACACTCC  
 H D L S G A G G S E C A Q E F V V S R H S  
 1201 ATGGCGCAAGTTCAGTTCTGCGACCGCGCTCATCCCCCTCATGGGCTACAGCCG  
 M D G K F T F C D Q R V I P L M G Y S P  
 1261 CCCGACCTCCCTGGCAAGTCTGCTGACTTCATCCACGCCAGGGACACATG  
 P D L L G K S C F D F I H A E D Q G H M  
 1321 AAGGAGAGCTCGACCGAGGTGCTCAAGATGAAGGGCCAGGTATGAACCTCATGACCGA  
 K E S F D Q V V K M K G Q V M N F M Y R  
 1381 TTCAGGGCAAAGTCAACGACTGGATCTGGCTCCGACCCGATTGCCCTCTCAAC  
 F R A K S N D W I W L R T T A F A F L N  
 1441 CCCTACACCGACGACATGGAGTACGTCGTCGACCAACTCCACCAACGCCAAGGGCAC  
 P Y T D D M E Y V V C T N S T N A K G N  
 1501 AGCACGAAACAGGCCCTCACGCCAACACCTCAACAGCCACGGACGAGTACGGCCCTCG  
 S S N S A S T P T T S T A T D E Y R P P  
 1561 AACACGCGGCCGCCCTCGGCTCGACTACAGCATCGG  
 N N A A A A S G L D Y S I

A)



B)

