

Supplementary Information for

**Convergent evolution of type I antifreeze
proteins from four different progenitors in
response to global cooling**

Laurie A. Graham and Peter L. Davies

Included herein are figures S1 to S6 and tables S1 to S3

```

cunner-a1 QQDLRIVLLGKTGSGKSTGNTILG-REAFRAKFSPPSSVTNTCKKVFFQYEDRIVSLIDTPGIFDTSITEVKLKEEIVKY 79
cunner-a2 QQDLRIVLLGKTGSGKSTGNTIVG-REAFRAKFSPPSSVTNTCKKVFFQYEDRIVSLIDTPGIFDTSITEVKLKEEIVKY 79
cunner-a3 QQDLRIVLLGKTGSGKSTGNTIVG-REAFRAKFSPPSSVTTKCKKVFFQYEDRIVSLIDTPGIFDTSITEVKLKEEIVKY 79
cunner-a4 QQDLRIVLLGKTGSGKSTGNTILG-REAFRAKFSPPSSVTNTCKKVFFQYEDRIVSLIDTPGIFDTSITEVKLKEEIVKY 79
cunner-a5 QEDLRIVLLGKTGSGKSTGNTILG-RAFAEAEFSPSSVTTKCRKVLQYKDRVRRLIDTPGIFDTSIKEVKLKEEIQKC 79
ballan-a1 QEDLRIVLLGKTGSGKSTGNTILG-NEVFFAKCSASSVTRTCREVVLYQKDRVRRLIDTPGIFDTSIAEVLKKEEITC 79
ballan-a2 QEDLRIVLLGKTGSGKSTGNTILG-NEVFFAEFSASSVTTKCKKVFFHYKDRGVSLIDTPGIFDTSITEVELK-EIEEC 78
ballan-a3 QEDLRIVLLGKTGSGKSTGNTILG-NEVFNTDVSPSSVTTKCKKVFFHYKDRMSVIDTPGIFDTSIAEVELKKEEIEKC 79
ballan-a4 QEDLRIVLLGKTGSGKSTGNTILG-REVFFAECSASSVTTKCRKVLFFHYKDRVRRLIDTPGIFDTSIAEVQLKEEIEKC 79
spotty-a TKDLRIVLLGKTGSGKSTGNTIAG-REVFFAGVSPSSSLTKCTKVMSSHHDGNSVSIIDTPGVFDTSTTNEELKDEIEKC 79
spotty-b SGGRLIVLVGKTGSGKSATGNTILG-AAAFKEDPSPVSVTKHCETQNGEVDGTLVQVIDTPGLFDTGITTEELKTRIEEC 79
ballan-a2 SGGRLIVLVGKTGSGKSATGNTILG-AAAFKEDPSSISVTKNCATQSGEVDGTAVQVIDTPGLFDTDTITEKELKTRIEEC 79
ballan-b2 PGGRLIVLVGKTGSGKSATGNTILG-AAAFKEDPSPVSVTKNCATQSGEVDGTAVQVIDTPGLFDTDTITEKELKTRIEEC 79
cunner-b1 STGLRIVLVGKTGSGKSATGNTILG-RSVFFKEDPSPVSVTKHCETQSGEVDGTAVQVIDTPGLFDTGITTEELKTRIEEC 79
cunner-b2 PTGLRIVLVGKTGSGKSATGNTILGVLDFAFKEDPSPVSVTKHCCKNQSGEVDGTAVQVIDTPGLFDTDTITEELKTRIEEC 80
spotty-c1 QGDLRLILVGKTGSGKSAGNTILGYRYAYKTDISPESVTVGCHKEEVQDRGNLNFVIDTPGLFDTTRKTTEVKLKGDIIEC 80
spotty-c2 QGDLRLILVGKTGSGKSAGNTILGRYAFETDISPESVTVGCHKEEVQDRGNLLVIDTPGLFDTTRKTEDVKKGDIIEC 80
spotty-c3 QGDLRLILVGKTGSGKSAGNTILGRYAFETDISPESVTAGCHQEEVKDGEENLNFVIDSPGLFDTSKTTEDVKEDIEIDC 80
spotty-c4 QGHLRLILVGKTGSGKSAGNTILGRYAFETDISPESVTAGCHQEEVQDGEENLNFVIDSPGLFDTSKTTEDVKEDIEIDC 80

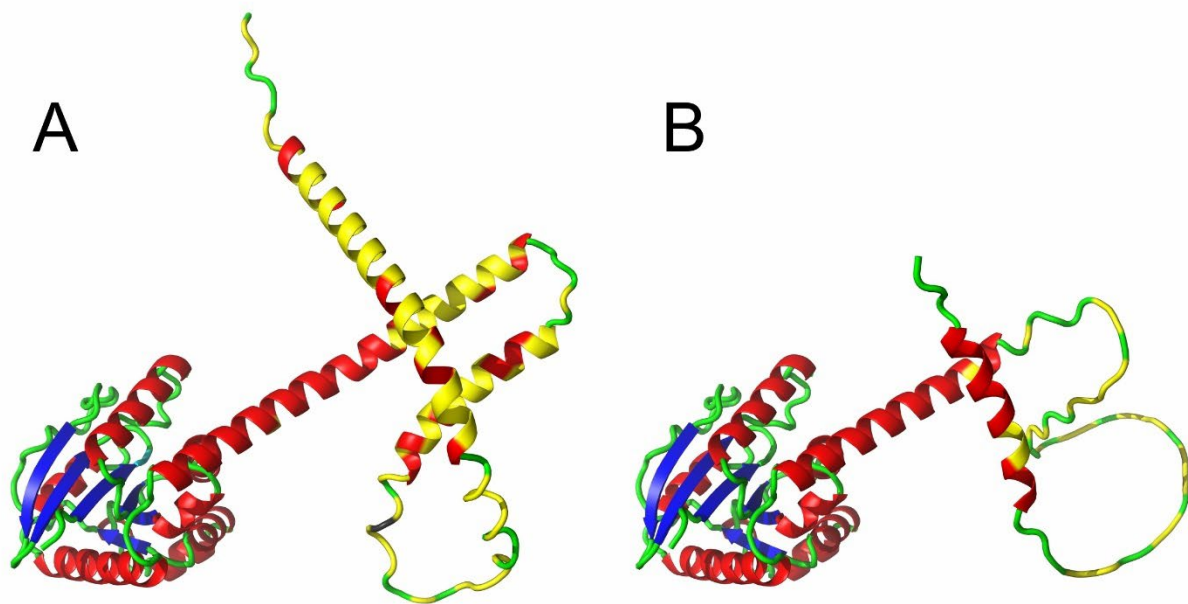
cunner-a1 ILLSVPGPHVFLLVIKLD-RFTEKEKNVAVKWIEDNFGAEASKYTLVLFTRGDELGQKSIEFTFLVESSDLRKFIVA 153
cunner-a2 ILLSVPGPHVFLLVIKLD-RFTEKEKNVAVKWIEDNFGAEASKYTLVLFTRGDELGQKSIEFTFLVESSDLRKFIVA 153
cunner-a3 ILLSVPGPHVFLVLIKLD-RFTEKEKNVAVKWIEDNFGAEASKYTLVLFTRGDELGQKSIEFTFLVESSDLRKFIVA 153
cunner-a4 ILLSVPGPHVFLLVIKLD-RFTEKEKNVAVKWIEDNFGAEASKYTLVLFTRGDELGQKSIEFTFLVESSDLRKFIVA 153
cunner-a5 ILLSVPGPHVFLLVIRLDVRFTEEEKNAFKWIDNFGAEASKYTLVLFTRGDELGQKSIEFTFLDTSSDLRDLIRT 154
ballan-a1 ILLSVPGPHVFLLVIRLDVRFTEEEKNVVKKWIDNFGAEASKYTLVLFTRGDELGQKSIEFTFLDSSDLKELIRT 154
ballan-a2 ILLSVPGPHVFLLVIRLDVRFTEEEKNVVKKWIDNFGAEASKYTLVLFTRGDELGQKSIEFTFLDSSDLKELIRT 153
ballan-a3 ILLSVPGPHVFLLVIRLDVRFTEEEKNVVKKWIDNFGAEASKYTLVLFTRGDELGQKSIEFTFLDSSDLRDLIRM 154
ballan-a4 ILLSVPGPHVFLLVIRLDVRFTEEKNAVKKWIDNFGAEASKYTLVLFTRGDELGQKSIEFTFLDSSDLKELIRT 154
spotty-a IMMSLPGPHVFLLVISLDARFTEEEKKAVTWIKDNFGAEASKYTLVLFTRGDELGQKSIEFTFLKESPELQEVIRT 154
spotty-b VKMSVPGPHAFLLVIRLVGRFTEERNNAVKKWIDNFGEDASMYTIMLFTCKD---QAKADNALKECKELRRLSIT 151
ballan-b1 VKMSVPGPHAFLLVIRLVGRFTEERNNAVKKWIDNFGEDASMYTIMLFTCKD---QAKADNALKECKELRRLSIT 151
ballan-b2 VKMSVPGPHAFLLVIRLVGRFTEERNNAVKKWIDNFGEDASMYTIMLFTCKD---QAKAEDALKECKELRRLQDS 151
cunner-b1 VKMSVPGPHAFLLVIRLVGRFTEERNNAVKKWIDNFGEDASMYTIMLFTCKD---QAKADNALKECKELRRLSIT 151
cunner-b2 VKMSVPGPHAFLLVIRLVGRFTAERNNAVKKWIDNFGEDASMYTIMLFTCKD---QANVDDALKDCKELRRLTDS 152
spotty-c1 VQSVPGPHAFLLVISLKAFTTEEEKASVKKWIDNFGPKSSLYTMILFTHADLLGKTVEDFVRESKHLQKLITE 155
spotty-c2 VQSVPGPHAFLLVISLKAFTTEEEKASVKKWIDNFGPKSSMYTMILFTHADLL----- 134
spotty-c3 IVQSVPGPHAFLLVISLKAFTTEEEKASVKKWIDNFGPKSSMYTMILFTHADLLGKTVEDFVRESKHLQKLITE 155
spotty-c4 VQSVPGPHAFLLVISLKAFTTEEEKASVKKWIDNFGPKSSMYTMILFTHADLLGKTVEDFVRESKHLQKLITQ 155

```

Supplementary Figure 1: Alignment of the GTPase domain of the GIMAP isoforms from cunner, ballan wrasse, and spotty, coloured using ‘Color Align Properties’ from the Sequence

Manipulation Suite (1). This alignment was used to generate the phylogenetic tree shown in Fig.

3.



Supplementary Figure 2: AlphaFold2 prediction of the structures of A) Ballan wrasse GIMAP-a1 and B) Cunner GIMAP-a5 rendered using cartoon mode in PyMOL (2). The GTPase domain (globular domain at left) was predicted with high confidence while the prediction of the C-terminal extension was with low confidence. Helices are red, sheets cyan, and loops green, with Ala residues in the extension in yellow.

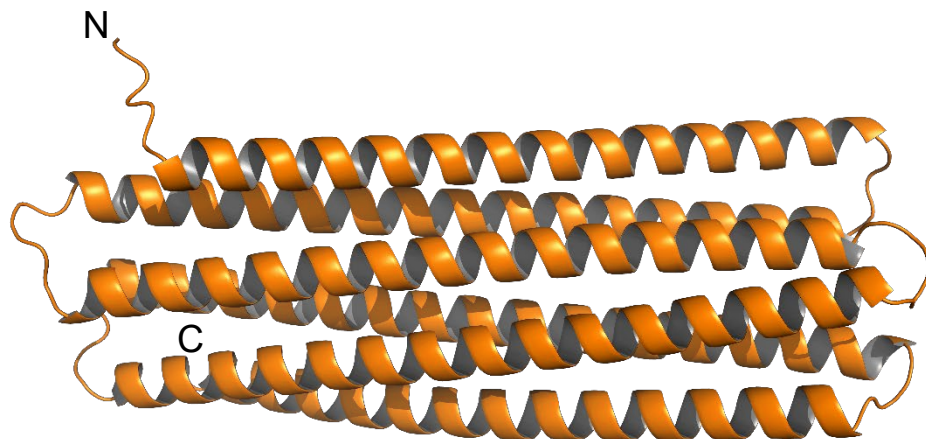
A

MDAANTPATEAAAAAAAAANAAATAKATAATAAAACGAITTAEEAAEAAKTASEAAAN 60
 AYTSATATVAAAIGTLAAPSAAVAAANAATAAAAAAATVAAPKAVDAAAATPTAATAAS 120
 TAAASKPPAASTAATPPLPPPPFEGGPKQEIAAGRSCRHTMTDTTDWPTDTQRPTQLTDG 180
 HIISGV* 186

B

MEYDDTPAAKAAAAAVAAAAAASAAATAAAAAKAAVDTAAAAAASAASTGAPNAAA 60
 IAAAKTAATAAAAAAKTAATATAAAKATAAAAAAAKATAAAAYAAKATAAATAPDAAA 120
 IAAAKTAATAAAAAAATAAAATAAAKATAAAAAAATATATAAAAAAATATAPDAAAIA 180
 AAKTAATAAAAAAATAAAATAAAKATAAAAAAATATAAAAAAASASATAATPAQRAAA 240
 NVTAAATAAASAVAAAAAATTNIAANISEASAATVAKTVVAAKAAATAVTPNIGAITAAT 300
 AASATADAAAKAARATADAAATKAAAAAVTAKAAAAALAL* 341

C



Supplementary Figure 3: Sequences of putative AFP pseudogenes and a structural model. A) Translation of Tanaka CM070098 bases 3,865,061 to 3,865,621. The region of homology matching 1 to 120 of 137 a.a. of AFP6 is underlined. B) Translation of dusky JBEEID010000351 bases 69,858 to 70,883. The region of homology from a.a. 4 to the end of the other dusky AFPs is underlined. C) AlphaFold3 model of dusky sequence from B (3).

A

AAVI**AAAAAVN****AAAA**VNATANAAAAANAAVRAAAVNAAAVIAAANA**KAAATTAANAA** 60
ANAAAVRAAAANADSNAAATTAAANAAAVRAAASNTAAKAAAKAAATTAAANAAADVRAAA 120
ANAAASNAAAKAAATTTATANAAAIIRAAAAANAAASNAAAKAAATTAAANATAVRAAASNAAAK 180
AAATTAAANAAATRDAATTAAANAAAIIRDAATTAAANAAAVRAAATTAAANAAATRDAA 240
TAAANAAAIIRAAATTAAANAAANAAAARAAAANAA 276

B

AISRAXPSSAARSSRAARPSSAARPSXAARPSSAARPSXAARPSSAARPSSAARPSSAA 60
RSSRAARVAATARAATARAATARAARAXAAASASARADAAASASARADAAASASARADAA 120
ASASARADAAASASARADAAASASARADAAASASARADAXARADAXASASARADAXASAX 180
ADAAAXASAGAAAAA*SRMXAAAA*ASRMXAAAA*ASRMVAAAA*ASRVXAAAA 234

Supplementary Figure 4: Coding potential of low-complexity sequences. A) Translation spanning the end of exon 3 (**bolded**) into intron 3 of PARP12, JBEEID010000351 bases 79,339 to 82,433. B) Translation spanning bases 439 to 1141 of the consensus transposon sequence DF000003478.1 [Copia-8-I_DR] (4). The regions corresponding to the nucleotide alignments in Supplementary Fig. 5 are underlined.

AFP3	ATGGCACCTGCTACCCCGCC--CAG--AAAGCCGCTGCCACC---GCCGCTGACGCCACCGCCAAAGCCG	64
PARP12	AAAGCCGCTGCCACCACCGCCACCGCCAACGCCGCCGCCATTAGAGCCGCTGCCGCCAACGCT---GCCT	67
AFP3	CCAACATCGCCGCAAAGTCTCTG---AAAGCGCCGCTGCCACCGCCAGAGCCGCTGTAGGCACCGGTGC	131
PARP12	CCAACGCCCGCCGCAAAGCCGCTGCCACAACGCCGCCGCCAACGCCA---CCGCCGTTAGAGCCGCTGC	134
AFP3	-----CGCTGCCGCCGCTGCCTCTGCCACCGCCGACGCCGCCGCCAAAGCCGCCAGACCACCGCCGAC	195
PARP12	CTCCAACGCCCGCCGCCAAAGCCGCTGCCACCAC---CGCTGCCGCCAACGCCGCC---GCCAC--CAGA-	195
AFP3	GCCGCTGACAAAGCCGCCCGCTGCCAACGCCGTCTTTAA	237
PARP12	GACGCTGCCA---CCACCGCCGCCGCCAACGCCGCCGCCATT	234

170/252 = 67%

AFP3	ATGGCACCTGCTACCCCCGCCAGAAAGCC---GCTGCCACCGCCGCTGACGCCACCGCCAAAGCCGCCA	67
<i>DrCopia</i>	TCCAGAGCTGCCAGAGTCGCTGCCACCGCCAGAGCTGCCACCGCCAGAGCTGCCACCGCCAGAGCTGCCA	693
AFP3	-----ACATCGCCGCCAAAGTCTCTGAAAGCGCCGCTGCCACCGCCAGAGCCGCTGTAGGCACCGGTGC	131
<i>DrCopia</i>	GAGCCGNCGCCGCCAGCGCCTCAGCCAGAGCCGACGCCGAGCCAGCGCCTCAGCCAGAGCCGACGC	763
AFP3	CGTGCCGCCGCTGCCTCTGCCACCGCCGACGCCGCCGCCAAAGCC---GCCAGAGCCACCGCCGA----	194
<i>DrCopia</i>	CGCA---GCCAGCGCCTCAGCCAGAGCCGACGCCGAGCCAGCGCCTCAGCCAGAGCCGACGCCGAGCC	830
AFP3	--CGCCGCTGACAAAGCCGCCGCCGCTGCCAACGCCGTCCTTTAA	237
<i>DrCopia</i>	AGCGCCTCAGCCAGAGCCGACGCCGAGCCAGCGCCTCAGCCAGA	875
		158/255 = 62%

[illegible]

Supplementary Figure 5: Alignment of Dusky snailfish AFP3 with low-complexity sequences. A)

AFP3 aligned with intron 1 of PARP12, JBEEID010000351 bases 34,522 to 34,758 (top) and 79,726 to 79,959 (bottom) or with B) the consensus sequence for the *Danio rerio* transposon DF000003478.1 [Copia-8-I DR] (4) or with C) (GCC)n.

A

PARP12	TCCGTGATACTGAAGTTCATTCTCGAAAACCAAGGAGCCGTCGACACCGATTTTTTTGACGTGTAACCTCG	70
Near AFP4	TCCGTGATACTGAAGTTCATTCTTGAAAACCAAGGATCCGTCGACACAGATTTTTTTGACGTGTAACCTCG	263
PARP12	GCTCGGGCGACGTTATTGATGACGTCATCCGCAACCGGGATAGGTTTGCCCTGTGTGAGCCCTTCGGGCA	140
Near AFP4	GCTCAGGCGACGTTATTGATGACGTCATCCGCAACCGGAAGAGGTTTGCCCTGTGTGAGCCCTTCGGGCA	193
PARP12	GTCGAAGGTGGTGGCCCGGACCCGGATGAAGCTGTGCCGGGTCGCCGAGTGCCCGGGCTCCTGTGACATG	210
Near AFP4	GTCGAAGGTGGTGGCCCGGACCCGGATGAAGCTGTGCCGGGTCCTCGAGTGCCCGGGCTCCTGTGACATG	123
PARP12	CTGCACCTGTGCAGACAGTTCCTCTTCACCGGAACTGCCCGTTCAACCAGAGGAGgtaaacagactcga	280
Near AFP4	CTGCACCTGTGCAGACAGTTCCTCTTCACCGGAACTGCCCGTTCAACCAGAGGAGgtaaacagactcga	53
PARP12	gcgcgaggcgtgtggagaggcctgcgacctggcttcaactctgtccctgttg	332
Near AFP4	gcgcgaggcgtgtggagaggcctgcgatctggcttcaactctgtccctgttg	1

323/332 = 97%
expect = 8e-165

B

PARP12	gatacgtgcgcgtgctttccgtctccacggaactggctgctgtcgggaggcggaaggagacgcggaagga	70
Near AFP6	gaaacgtgcgcgtgctcttcgtctccacggaactggctgctgtcgggaggcggaaggagacgcggaagga	70
PARP12	gacgcggaaggagcATGGAGTCCGTGATACTGAAGT	106
Near AFP6	g-cccgg-----ccATGGGGTCTGTGATACTGAAGT	100

93/106 = 88%
expect = 2e-34

C

PARP12	tgttatgcaagcctcacgcggacaccttcaggctaattgaacacgtgacgtaggaacgcgctccgagggc	70
Near ΨAFP	tgttatgcaagcctcacgcggacaccttcaggctaattgaacatgtgacgtagtgaagcgcctccggggc	70
PARP12	acattggcaacctgtgatattgggctgtgtgtaataaagtgaattgaagtgacgcctcagctttagagt	140
Near ΨAFP	acattggtaacctgtgatattgggctttgt-ttaataaagtgaatttaagtgacgcctcagctttagagt	139
PARP12	tcta	144
Near ΨAFP	tata	143

132/144 = 92%
expect = 4e-57

Supplementary Figure 6: Sequence alignments between PARP12 and AFP intergenic regions. A)

A) A portion of dusky PARP12 exon 1 coding (uppercase) and intron 1 (lowercase) aligned to a segment ~3.5 kb downstream of dusky AFP4. JBEEID010000351 bases 76,820 to 77,151 (top) and 45391 to 45060 (bottom). B) A portion of dusky PARP12 exon 1 coding (uppercase) and flanking 5' sequence aligned to a segment ~2.6 kb downstream of Tanaka AFP6.

JBEEID010000351 bases 76,730 to 76,835 (top) and CM070098 bases 3,858,144 to 3,858,243 (bottom). C) A portion of dusky PARP12 intron 1 aligned to a segment ~1 kb upstream of dusky ΨAFP. JBEEID010000351 bases 77,303 to 77446 (top) and 68,751 to 68,893 (bottom).

Supplementary Table 1: Verification of GIMAP annotation, reannotated sequences are below.

Sequence	GenBank Accession	Status
Ballan GIMAP-a1	XP_020509819.1	Second intron unsupported
Ballan GIMAP-a2	XP_029137074.1	Second intron unsupported
Ballan GIMAP-a3	XP_020509833.1	Verified
Ballan GIMAP-a4	XP_020509831.1	Verified
Ballan GIMAP-b1	XP_020509835.1	Verified
Ballan GIMAP-b2	XP_029137070.1	Verified
Spotty GIMAP-a	XP_034543013.1	Supported
Spotty GIMAP-b	XP_034544901.1	Earlier start codon not consistent
Spotty GIMAP-c1	LOC117816646	Only isolated exons predicted, listed as pseudo
Spotty GIMAP-c2	LOC117815504	Only isolated exons predicted, listed as pseudo
Spotty GIMAP-c3	LOC117815504	Only isolated exons predicted, listed as pseudo
Spotty GIMAP-c4	XP_034544162.1	First exon not consistent

>Ballan GIMAP-a1
MALRSTKVCSPDIQEDLRIVLLGKTGSGKSSTGNTILGNEVFKAACASSVTRTCREVVLYKDRRVRLIDTPGIFDTSIAEVKLKEEIIETCILLS
VPGPHVFLLVIRLDVRFTEEEKNVVWKIKDNFGAEASKYTLVLFTRGDELGQKSIETFLEDSSDLKELIRTCNARYIVFDNKSMMNRTQVNDLFEKI
DEIVQLNGGHYTSSIIYAEARQRIKSAEWWNTCVQYVPLLSAGAAAAAAGAAAVAIGTGASAAAIGAGAAAAAAGAAAVAIGAGATAADGAGAAA
AAVGVGAAAAAGIGAAAAAIGAGAAAAVGAAAAAAGAAAAAAGAAAAAIGAAAGPRAAG

>Ballan GIMAP-a2
MALRSTKVCSPDIQEDLRIVLLGKTGSGKSSTGNTILGNEVFKAEFSSASVTCTCKVFFHYKDRGVSLIDTPGIFDTSITEVELKEIEECILLSV
PGPHVFLLVIRLDVRFTEENKNNVWKIKDNFGAEASKYTLVLFTRDQLGQKSIETFLEESSDLKELIRTCNSRYIVFDNKSMMNRTQVNDLFEKID
EIVQLNGGHYTSSIIYEEARQRIKSDKRWNFLIHQMSPLSTSMKP

>Spotty GIMAP-b
MAEAIPLREKKQKTLPRSDGCGLNFDALDLSGGLRIVLVGKTGSGKSATGNTILGRAAFKEDPSPVSVTKHCETQNGEVDGTLVQVIDTPGLFDTG
ITEELKTRIEECVKMSVPGPHAFLLVIRLGVRFTEERNNAVKWIQDNFGDDASMYTILMFTCKDQAKADNALKECKELRRLSITFGRRYHAFNNND
ADDRVQVNELITMIKDMIQDNGGKHYTNEMYERAQRKLREEEERRKQEEEEKKEERKVWDAEREKQEKERAKKKNVRIKNIRVASAAAVLVVAGV
VIAVGASNTVALALGAPVLVLGVLGCLAAICVWKGIRCKAKKGVVPV

>Spotty GIMAP-c1
MLRIKAWIPGCTLLVLVTLGCGTAHCQSQLKDSHPHQDLRLILVGKTGSGKSASGNTILGYRYAYKTDISPESVTVGCHKEEVQDRGRNLFVIDTPG
LFDTRKTTEGLKGDIEECVAQSVPGPHAFLLVISLKFRTTEEEKASVKWIQDNFGPKSSLYTMILFTHADLLEGKTVEDFVRESKHLQKLITECGGP
IPLAHQ

>Spotty GIMAP-c2
MLRIKAWIPGCTLLVLVTLGCGTAHCQSQLKDSHSHQDLRLILVGKTGSGKSASGNTILGRRYAFETDISPESVTVGCHKEKVQDRGRNLLVIDTPG
LFDTRKTTEDEVKGDIEKCVVQSVPGPHAFLLVISLKFRTTEEEKASVKWIQDNFGPESSMYTMILFTHADLL

>Spotty GIMAP-c3
MLRIKAWIPGCTLLVLVTLGCGTAHCQSQLKDSHSHQDLRLILVGKTGSGKSASGNTILGRRYAFETDISPESVTAGCHQEEVKDGERNLFVFDSPG
LFDTSKTTEDVKEDIEDCIVQSVPGPHAFLLVISLKFRTTEEEKASVKWIQDNFGPKSSMYTMILFTHADLLGEKTVEDFVRESKHLQKLITECGGR
YHSLINDNKQSRDQVRQLIDKIDTMVESNGGSHYTSEMYQRAQEQVEEERRREEEQEKEEAKRAKCKAVMMALAGITGGSFLFPSYLLTLVAGGLGP
EALDCIKDMFK

>Spotty GIMAP-c4
MLRIKAWIPGCTLLVLVTLGCGTAHCQSQLKDSHSHQDLRLILVGKTGSGKSASGNTILGRRYAFETDISPESVTAGCHQEEVKDGERNLFVIDSPG
LFDTSKTTEDVKEDIEDCIVQSVPGPHAFLLVISLKFRTTEEEKASVKWIQDNFGPKSSMYTMILFTHADLLGEKTVEDFVRESKHLQKLITECGGR
YHSLINDNKQSRDQVRQLIDKIDTMVEFNGGSHYTSEMYQRAQEIEEERRWEEEQEKEEAKRAKCKAVMVALAGIAGVSYLFPYLLTLVAGTFGP
EAYDCIKYMFK

Supplementary Table 2: Presence or absence of type I AFP sequences in sequence databases from fishes in family Liparidae.

Database Accession	Type	Species	Tissue	Collection Date	Location	Result
GCA_040955725	Genome	<i>Liparis gibbus</i>		2023	Nova Scotia	+++
GCA_036178185	Genome	<i>L. tanakae</i>		2018	China Yellow Sea	+++
SRLO00000000	Genome	<i>L. tanakae</i>		2017	Sea of Japan	+++
SRR22396815	SRA DNA	<i>L. gibbus</i>		2011	Bering Sea	+++
SRR12518887	SRA DNA	<i>L. liparis</i>			North Sea	+++
SRR22396704	SRA DNA	<i>L. tunicatus</i>		2010	Bering Sea	+++
SRR19970642	SRA RNA	<i>Pseudoliparis swirei</i>	Skin		Mariana Trench	No obvious AFP
SRR19963796	SRA DNA	<i>P. swirei</i>			Mariana Trench	No obvious AFP
RZNG00000000	Genome	<i>Pseudoliparis</i> sp.			Yap Trench	No obvious AFP
SRR7206520	SRA RNA	<i>P. amblystomopsis</i>	Liver		Mariana Trench	No obvious AFP
SRR7223797	SRA DNA	<i>P. amblystomopsis</i>			Deep N. Pacific	No obvious AFP
SRR12575254	SRA RNA	<i>P. amblystomopsis</i>	Skin		Yap Trench	No obvious AFP
SRR21844227	SRA DNA	<i>Paraliparis pectoralis</i>		1996	Oregon Coast	No obvious AFP
SRR22396715	SRA DNA	<i>Pseudosquilla</i> sp.		2015	N. Carolina Coast	No obvious AFP
SRR22396812	SRA DNA	<i>Careproctus phasma</i>		2011	Bering Sea	No obvious AFP
SRR22396818	SRA DNA	<i>Careproctus scottae</i>		2011	Bering Sea	No obvious AFP
SRR22396822	SRA DNA	<i>Crystallichthys cyclospilus</i>		2010	Aleutian Islands	No obvious AFP

Supplementary Table 3: Characteristics of transposable element sequences flanking snailfish AFP sequences shown in Fig. 5B.

Transposon	Match identifier in Dfam (4)	Expect Value ¹
TE1	DF003572163.2 [TcMar-ISRm11-6_AstCal]	7.4e-62
TE2-5'/3'	DF003571736.2 [hAT-Ac-7_AstCal]	3e-19/2.4e-59
TE3	DF000002590.2 [DNA8-25_DR]	1.9e-12
TE4 (AFP)	DF000003478.1 [Copia-8-I_DR]	1.1e-12
TE4 (PARP12)	DF000003478.1 [Copia-8-I_DR]	7.3e-15
TE5	DF003571937.2 [Unknown-64_AstCal]	2.5e-06
TE6	DF000002858.1 [hAT-N46_DR]	3.2e-05
TE7	DF000003937.1 [hAT-N131_DR]	2.6e-15
TE8	DF000003589.1 [Gypsy150-I_DR]	6e-20
TE9	DF000002933.1 [TDR8C]	7.7e-39
TE10	DF000003081.1 [Academ-N1_DRe]	8.9e-14
TE11	DF000003108.1 [hAT-N31_DR]	9.7e-13

¹ The lowest expect value is given when multiple occurrences of varying length were found.

References

1. P. Stothard, The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* **28**, 1102, 1104 (2000).
2. L. Schrödinger, & DeLano, W. (2020) The PyMOL Molecular Graphics System, Version 2.5.2 Schrödinger, LLC. (<https://pymol.org/>).
3. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
4. J. Storer, R. Hubley, J. Rosen, T. J. Wheeler, A. F. Smit, The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mob DNA* **12**, 2 (2021).