

Supplementary Materials: Antimicrobial Peptide Arsenal Predicted from the Venom Gland Transcriptome of the Tropical Trap-Jaw Ant *Odontomachus chelifer*

Josilene J. Menk, Yan E. Matuhara, Henrique Sebestyen-França, Flávio Henrique-Silva, Milene Ferro, Renata Santos Rodrigues and Célio D. Santos-Júnior



Figure S1. Tridimensional structure prediction with marked cysteine residues of the putative poneratoxin ICK-like toxin homolog of *O. chelifer* (TRINITY_DN24839_c0_g1_i1.p1). Yellow: strand. Gray: turns. Purple: cysteine residues.

	1	10	20	30
1 Omo_Abaecin_A0A348G6C5	FRV	P P P R S E	P GGWKGPNEPQGPFN P KFG K
2 TRINITY_DN21408_c0_g1_i1.p1	IPT	P A P V A	P GAITAYPIQRPAY A PNPWVYPKPTTVQPI R	
3 TRINITY_DN16370_c0_g1_i1.p1	ISQ	P L P P T	P ...SGHVINKRVAPI P R...PVPTVIGLLM K	

Figure S2. Primary sequence alignment between two predicted AMP sequences from *O. chelifer* and the A0A348G6C5 abaecin from *O. monticola*. Similar residues are boxed. Identical residues are bolded. The red line indicates the PXP/PXXP motifs.

	1	10	20	30	40	50																																																			
Lni_Crustin4_A0A0J7L9H5	M	R	G	N	S	L	Y	L	L	L	F	V	I	L	T	K	A	Y	V	K	S	C	K	Y	W	C	K	M	G	H	H	R	Y	Y	C	C	P	N	G	...	K	N	D	S	S	S	E	H	S	W	H	S	F	L	F	P	W
TRINITY_DN17655_c0_g1_i1.p1	M	R	R	F	R	A	I	N	L	I	A	F	C	A	L	V	L	L	A	D	D	S	S	A	...	T	V	F	Y	N	L	K	N	Y	L	P	...	G	F	L	R	S	F	K	Q	L	K	T	L	R	Y	G	I	V	P	G	T
TRINITY_DN12961_c0_g1_i1.p1	...	M	A	R	L	I	N	V	V	M	I	I	G	M	V	L	V	S	V	F	A	A	E	...	E	L	Y	S	D	K	Y	D	D	I	D	V	M	G	I	L	T	N	D	R	L	R	D	Q	Y	F	K	C	F	M	D	...	
	60	70	80	90	100																																																				
Lni_Crustin4_A0A0J7L9H5	F	W	E	G	I	V	G	S	H	A	E	H	P	W	H	E	V	...	E	K	M	K	W	K	P	K	K	Q	C	P	P	L	R	P	Q	C	P	R	M	...	Y	E	W	Y	K	P	P	K	Y	C	D	S	D				
TRINITY_DN17655_c0_g1_i1.p1	L	W	G	C	P	G	N	I	A	S	N	...	Y	S	E	L	...	G	V	N	W	K	L	D	T	C	C	R	A	H	D	S	C	N	D	L	I	R	P	R	N	S	K	Y	G	L	Y	N	S	S	K	L	C	S	S		
TRINITY_DN12961_c0_g1_i1.p1	...	T	G	P	C	T	A	D	M	K	F	Y	K	D	I	I	G	E	G	V	V	T	K	C	K	K	C	T	E	K	Q	K	E	N	L	D	T	L	...	T	D	W	Y	T	K	N	R	L	D	D	W	N					
	110	120	130																																																						
Lni_Crustin4_A0A0J7L9H5	H	E	C	E	G	E	G	K	C	C	Y	D	V	C	L	E	H	K	T	C	K	T	A	E																																	
TRINITY_DN17655_c0_g1_i1.p1	L	L	C	Q	C	D	L	Q	F	Y	...																																								
TRINITY_DN12961_c0_g1_i1.p1	T	F	V	K	K	...	L	M	E	D	F	K	...																																												

Whole Body Body and Venom Gland

Figure S3. Global alignment between the reference crustin and the grouped candidate AMP sequences. The color specifies the source tissue. The gray line indicates the signal peptide of the reference AMP.

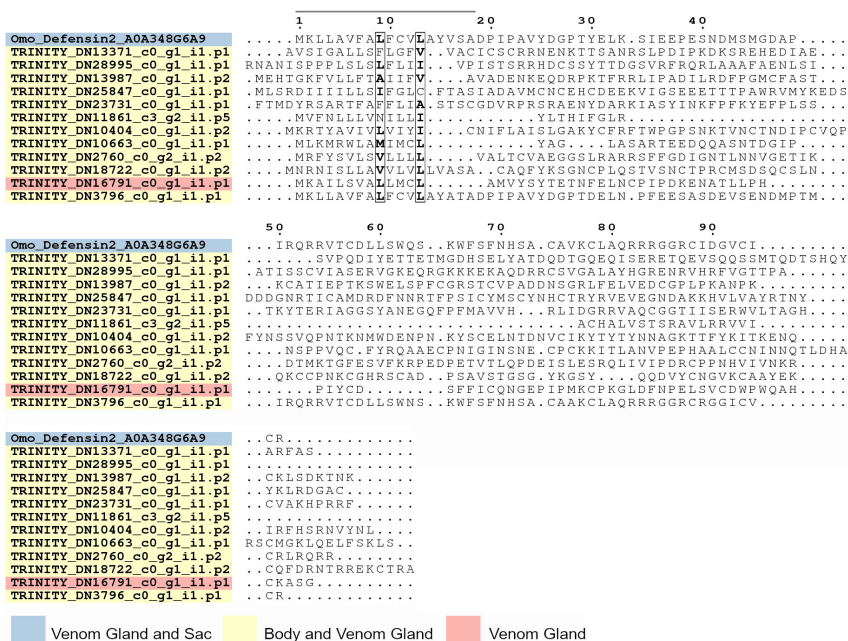


Figure S4. Global alignment between the reference defensin and the grouped candidate AMP sequences. The color specifies the source tissue. The gray line indicates the signal peptide of the reference AMP.

	1	10	20	30	40	
Omo Hymenoptaecin A0A348G6C2MKLLVSL	LAALSCAIVYASADLT	GDQVLPQSPG	VPVLPHTSRFP	READP	
TRINITY_DN10137_c1.g1.i1.p5FFFFF	FFFFFLCDSP	VPNLNCRW	
TRINITY_DN1540_c0.g2.i1.p1AAVF	ALLCVTDVTAI	RSDDGCGN	
TRINITY_DN1886_c2.g8.i1.p8ERERERE	GENVCMCV	CVCVCVCARA	CVCTR	
TRINITY_DN4362_c0.g1.i1.p2MKLIT	LMIAVVLST	TVTGTGD	
TRINITY_DN9325_c1.g1.i1.p2NRLTNV	VFLSV	LGTELMKMTV	CPYKCKRCSL	
TRINITY_DN11193_c0.g4.i1.p2RYFHCS	LRFPPL	CLASDFIVCFDL	SRPKKEKKLVISKRYKINTK	
TRINITY_DN28762_c0.g1.i1.p2MKAIVV	VLAI	CFAAALGEFT	DEEKEKLKTYKESCISE	TGVDPNVVKNA	
TRINITY_DN6307_c0.g1.i1.p1ARQTA	LVLF	FGDFIVARAR	DKTSDDLRTA	
TRINITY_DN12005_c4.g1.i1.p15MLCV	CMCV	CAIQVDER	EK	
TRINITY_DN1280_c0.g1.i1.p7MLCDMF	QIDAVAFV	ALLINQYSSCRPI	SHLHHIRCAL	
TRINITY_DN26077_c0.g1.i1.p1IHPTV	FSVLE	WLGYNCSAIN	PCYIALF	
TRINITY_DN168_c0.g1.i1.p2NSTFILY	DRIT	GMIVALFI	VIAVTSSSR	
TRINITY_DN22786_c0.g1.i1.p1LFLSL	LSL	FFSICLSSTTL	LLFRSARTQSAENNKD	VEGQRLLRVQFAG	
TRINITY_DN23188_c0.g1.i1.p1FARIF	VF	FFSLLTSLVK	MATASITGV	
TRINITY_DN1980_c0.g1.i1.p1IVLV	LVLL	LECUVVIGLAC	GFRDAEKG	
TRINITY_DN8792_c0.g1.i1.p3DIRDWR	LAL	LIFLSTREFT	DGMTSS	
TRINITY_DN5860_c0.g1.i1.p2PPIDDEY	MGGICAA	FLFQLSFATISTT	IVSGMAEERCNF	
TRINITY_DN26054_c0.g1.i1.p1HVS	TGCVVLL	LFYKCVLPST	LVSFSGSTLVRGRVDS	
TRINITY_DN4617_c0.g1.i1.p1SPSASAST	MHRPRGR	QSFLE	CLFPYTSCEF	
TRINITY_DN15646_c0.g1.i1.p3MRCVRC	WV	VIKCVFACLT	GRGTGGV	
TRINITY_DN1241_c0.g1.i1.p2MTAFLH	GLMIV	LI	IINAFTLIGS	
TRINITY_DN26693_c0.g1.i1.p3MEENEVR	GILV	FLVLAACIGPAL	SITSACR	
TRINITY_DN23671_c0.g1.i1.p2GTRAL	ALAL	ALSICIOGCI	VRRGRCCAM	
TRINITY_DN1222_c0.g1.i1.p3PLNARN	VAS	FLSLLLEALR	AD	
TRINITY_DN13204_c0.g1.i1.p1CVVLV	LVV	VVVVIVKVT	DTQEAASRASW	
TRINITY_DN10866_c0.g4.i1.p9TNSSVP	VNGLV	ISVLCYFAAPFS	PVSVRPERF	
TRINITY_DN21005_c2.g1.i1.p4SLSLSL	LSLSL	SLFLLTFSSLFF	SDPRVQIVRAIIAISSLDERHSGE	
TRINITY_DN13289_c0.g1.i1.p2MRATW	LV	LILAEILPPLSL	LGEAVCHRN	
TRINITY_DN1481_c0.g1.i1.p6MYLMLG	MALIAM	CFNLMQEEVI	
TRINITY_DN18603_c0.g1.i1.p1MLAAS	PILLV	LYLSFPRSAGFF	REGSLFLRRLSRTASFF	LNLRNHTF	
TRINITY_DN24006_c0.g1.i1.p1MKLLV	LV	CLVFVAISV	ATGNRYFAK	
TRINITY_DN1844_c0.g3.i3.p1MKFYV	CL	LLVIMILSLV	PEDLCQE	
TRINITY_DN1674_c8.g1.i1.p17MWCRL	TL	LLLVCCCHLAA	ASKRIRVDGK	
TRINITY_DN27940_c0.g1.i1.p1MAPARLL	LL	LVSLMLVFAN	
	50	60	70	80	90	100
Omo Hymenoptaecin A0A348G6C2	QNSISFQGSQPLSGPNRQPTW	DLNVNVRNIANNDRS	RTDIFGGIGKAPGQSAQPHIG		
TRINITY_DN10137_c1.g1.i1.p5	KERRRKEQSAASSRLACFVSTV	VLVCPVETEKYR	
TRINITY_DN1540_c0.g2.i1.p1	GMFRCDNGKCIQSILVCDYRGDCDDNSDEM	QSCPPDCDFGQI	
TRINITY_DN1886_c2.g8.i1.p8	RTTCVMT	HVRSICV	CVVLACVNV	
TRINITY_DN4362_c0.g1.i1.p2	DEYIHLFGKSCSDAPP	CPDGRPCVMAPPRCN	VGTGKDLVPTCGRR	..
TRINITY_DN9325_c1.g1.i1.p2	DEFPHEQLLQOHR	LYHHGARERARS	NARWNETILSF	FTYTCTPC	HDVTTYTKPGNLK
TRINITY_DN11193_c0.g4.i1.p2	HEKKVNRVNSSQ	ETTRGKISV	VYILLHL	SLAADGKKNC	SVHQTGLGKARNNRND
TRINITY_DN28762_c0.g1.i1.p2	KEGMIDESDEKL	ACFSTCLF	KKFGVMKENG	IDITD	TARSKISSNV
TRINITY_DN6307_c0.g1.i1.p1	DFLRLVERRPKID	TPRLKIAKLARPCAPRW		
TRINITY_DN12005_c4.g1.i1.p15	EGGKGIRDKDK	KILKIKTQ	LLTIAKKRHSA	
TRINITY_DN1280_c0.g1.i1.p7	NCTQSYTLARKP	SNFCKQKRNQAKL	MDAFATTC	VARDAQCPRYL
TRINITY_DN26077_c0.g1.i1.p1	SKDFRFAPKRI	ICKCFCKRRANT	LRGSDGQLAMRNSP	SPS
TRINITY_DN168_c0.g1.i1.p2	DYFRARAVDKL	CKRTSAATMTSLNRGT	VIATPRSL	TREKOPRA
TRINITY_DN22786_c0.g1.i1.p1	RSSRKWSLSS	ASKETSASYDIL	DLRLSSSRKSR	RRPKVAGGT	RFPAPEKSSRIE
TRINITY_DN23188_c0.g1.i1.p1	SFYARAALHAAT	WCYTVIHRARY	IEVSTFE	VDIFASCTKRKN
TRINITY_DN1980_c0.g1.i1.p1	QDFLKSTIKYY	ATTADKAET	VTVAWDG	IMSQFHCCGVES
TRINITY_DN8792_c0.g1.i1.p3	KLQKRIAASVM	RCGKKKSMVR	SK
TRINITY_DN5860_c0.g1.i1.p2	KAYCIFSFLNT	IVYCLPAGV	VWGDHGFLNR	MGVVDIAGSGPVHLV
TRINITY_DN26054_c0.g1.i1.p1	RTDLALS	SRASQFSR	KISRRANCCG	KDRNVPRHFPT	MRCEFLAIFARI
TRINITY_DN4617_c0.g1.i1.p1	LRPAAVIT	TGGAETG	VILAPTL	RATNSLIGN	NGGDTARK	RGRGEPVNLIR
TRINITY_DN15646_c0.g1.i1.p3	FVYRLQNDEDS	QDSSASEALDHRAQ	IYRGTSRPT	ITRLSALRPTW
TRINITY_DN1241_c0.g1.i1.p2	TANHPIFPRHY	PTTSKIAST	PLYTTPRI	IDTFAK	SEVKCPTGQKDK
TRINITY_DN26693_c0.g1.i1.p3	NVVCISRITMRA	PRDIPDRP	VDVFGHAISM	QIVLSFADDT
TRINITY_DN23671_c0.g1.i1.p2	HSVCGVTTRDT	SRVLKACQLEVH	PELVHTSA	FDGFFERPRAVK
TRINITY_DN1222_c0.g1.i1.p3	QTSRPGAAVVT	KRRNTL	LGSGKKIK
TRINITY_DN13204_c0.g1.i1.p1	KKRLTR	TNSPPG	INGPKKRPL	LIRTIKQADR	KFLRPDNRAA	AAKRLHASTFV
TRINITY_DN10866_c0.g4.i1.p9	EIFQVTRKKKS	AI	
TRINITY_DN12005_c2.g1.i1.p4	RAFPSSHVPRE	LLYPNGPCARASVFGK	GELVRGES	MIYTSRRDR
TRINITY_DN13289_c0.g1.i1.p2	DTDLSAGTNTG	QORSGETSSRRR	RELAF	PKGSFAFV
TRINITY_DN1481_c0.g1.i1.p6	STTRSFLVPD	LLRCP	SVAAPPEL	GDPRRSQSR
TRINITY_DN18603_c0.g1.i1.p1	QGSCCPALPV	QFCGHSCF	VDSHCAGIG	KCCPTQ	CGGSI
TRINITY_DN24006_c0.g1.i1.p1	EDGRSRLGE	LRNYP	IVVMLSK	AKPVCEGGQRLAMEK
TRINITY_DN1844_c0.g3.i3.p1	SYTVSAYQED	CRKYVMCSE	GRCKLETCE	TYFFDP	ITSTCTH
TRINITY_DN1674_c8.g1.i1.p17	EGMPLSFFKRTRP	QSKCP	PGVYRHTRTGCREIL
	110	120	130	140	150	160
Omo Hymenoptaecin A0A348G6C2	IQHERNLGRNGFIRG	SGQLQPGYGG	RGLTPSFGVTGGL	RFRREAEADAK	DEDDTELIEE	
TRINITY_DN10137_c1.g1.i1.p5	
TRINITY_DN1540_c0.g2.i1.p1	
TRINITY_DN1886_c2.g8.i1.p8	
TRINITY_DN4362_c0.g1.i1.p2	
TRINITY_DN9325_c1.g1.i1.p2	
TRINITY_DN11193_c0.g4.i1.p2	
TRINITY_DN28762_c0.g1.i1.p2	
TRINITY_DN6307_c0.g1.i1.p1	
TRINITY_DN12005_c4.g1.i1.p15	
TRINITY_DN1280_c0.g1.i1.p7	
TRINITY_DN26077_c0.g1.i1.p1	
TRINITY_DN168_c0.g1.i1.p2	
TRINITY_DN22786_c0.g1.i1.p1	
TRINITY_DN23188_c0.g1.i1.p1	
TRINITY_DN1980_c0.g1.i1.p1	
TRINITY_DN8792_c0.g1.i1.p3	
TRINITY_DN5860_c0.g1.i1.p2	
TRINITY_DN26054_c0.g1.i1.p1	
TRINITY_DN4617_c0.g1.i1.p1	
TRINITY_DN15646_c0.g1.i1.p3	
TRINITY_DN1241_c0.g1.i1.p2	
TRINITY_DN26693_c0.g1.i1.p3	
TRINITY_DN23671_c0.g1.i1.p2	
TRINITY_DN1222_c0.g1.i1.p3	
TRINITY_DN13204_c0.g1.i1.p1	
TRINITY_DN10866_c0.g4.i1.p9	
TRINITY_DN12005_c2.g1.i1.p4	
TRINITY_DN13289_c0.g1.i1.p2	
TRINITY_DN1481_c0.g1.i1.p6	
TRINITY_DN18603_c0.g1.i1.p1	
TRINITY_DN24006_c0.g1.i1.p1	
TRINITY_DN1844_c0.g3.i3.p1	
TRINITY_DN1674_c8.g1.i1.p17	
TRINITY_DN27940_c0.g1.i1.p1	

Venom Gland and Sac
Body and Venom Gland
Venom Gland

Figure S5. Global alignment between the reference hymenoptaecin and the grouped candidate AMP sequences. The color specifies the source tissue. The gray line indicates the signal peptide of the reference AMP.

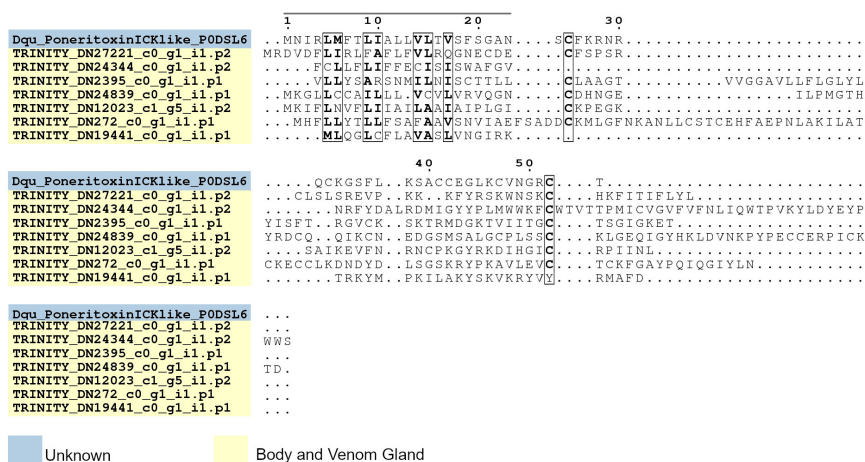


Figure S6. Global alignment between the reference ICK-type AMP and the grouped candidate AMP sequences. The color specifies the source tissue. The gray line indicates the signal peptide of the reference AMP.

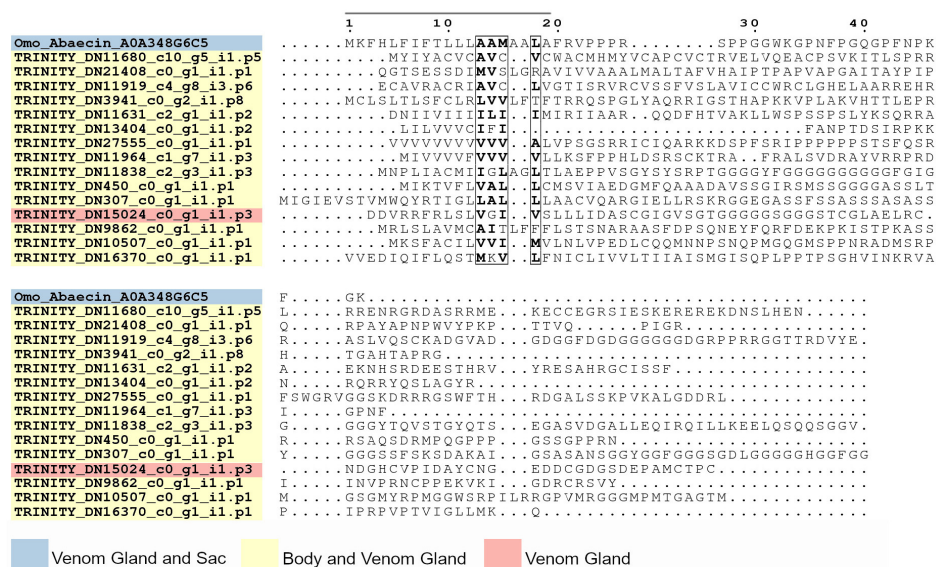


Figure S7. Global alignment between the reference abaecin and the grouped candidate AMP sequences. The color specifies the source tissue. The gray line indicates the signal peptide of the reference AMP.