

***Hm-MyD88* and *Hm-SARM*: Two key regulators of the neuroimmune system and neural repair in the medicinal leech.**

**Rodet F., Tasiemski A., Boidin-Wichlacz C., Van Camp C., Vuillaume C., Slomianny C., Salzet M.**

<i>HmMyD88</i>	-----MAEAMAQYHKLPEKHWKE	18
<i>XtMyD88</i>	-----MACGSSSLGSVDM	12
<i>BbMyD88</i>	-----MATNAPSQDTEGENY---	15
<i>CsMyD88</i>	-----MASFVPMKMDL	10
<i>HsMyD88</i>	-----MAGGPGAGSAAVSSST	17
<i>GgMyD88</i>	-----MATVFGAGSAPGPEPADL	19
<i>SpMyD88</i>	MNFNDLTVKLHLPDGERDCGAVVLSKRPAADSSNRDTQVCDKLSLSCAPAVSALEKATTTLEDGNSGCLLELMAAEIPEKEPKDENINKIE	90
<i>CfMyD88</i>	-----MAMADIENLIESM	13
<i>ScMyD88</i>	-----MKSRSQSHGLPDLSPNAHY	18
<i>DmMyD88</i>	-----MRPRFVCHQQHSAVSHYQPHSHFHHTHRHPNP PHHHHIYGATDVSYRRYRTAGMVVAEGVMDSGSGSGTGTGIGHF	78

**Death Domain**

<i>HmMyD88</i>	CDIPFKSVHHAARAWLSLHLNPERSTFGPENGNLQDWRGVAEELGFVPLQIENFORSHDELSEVMKNLRITKQEAVLGDVLLAIIIRCE	106
<i>XtMyD88</i>	NSIFLVALNYNVRHRLSLYLNPNVVA-----AGWTQLAEEEMGYDYLEIKNFERFPDCTMKLLEDWEKCCFRATVGGLEMLKME	93
<i>BbMyD88</i>	DSIFARYLGVQTRKKLALWLNPPRVGG-----TNWEDLADEMGHEYLEIENFKLKNSEPMYEVLLHVWSLQQ--NATIGHLIEMLKISIE	95
<i>CsMyD88</i>	DSVPLTALNVTVRKKLCLYLNPKNTVA-----ADWTAVAEAMDFTYLEIKNYETVRSFMAGVLDWEQARNKDSAVGKLLSIIILKLE	91
<i>HsMyD88</i>	SSLPLAALNMRVRRRLSLFLNVRTQVA-----ADWTAIAEEMDFEYLEIIROLETQADETGRLLDAWQGRP--GASVGRLLLELLTKLG	97
<i>GgMyD88</i>	HSVPMVALNYGVRRLCLYLNPRATA-----ADWTAIAEKLCHDYLEIRRLAALPDTAALLEEWQSRCPGATVGOLELLRLQLG	101
<i>SpMyD88</i>	RSMFATGIGWRRTNLSQYLRPPRPGS-----CDWRDMAEEMGHSYQLHIQNFALSDPVAKVLSAWCTKP--GSNVGKLLDITIEKIE	171
<i>CfMyD88</i>	KDVHVNLSLNGCRRIIAMIYLNPECSLIPDSDMFDNWCYMAELLNFSQPEIENMKRHKSEFTEMLHLWSTRNDPPEPKVGNLISFICKLE	101
<i>ScMyD88</i>	PLTFCSSIDPSAIDPEEIQIRIRRVQVT-----AGIDPEQLAKCLAGLSIAHSSPSRSENSTNSTASPOQT--PSGFSFSKQINHTSNK	99
<i>DmMyD88</i>	NETPLSALGIETRTQLSRMLNRKKVLRSEEGYQRDWRGISELAKQKG---FVDENANNPMDLVLIISWSRSPQTAKVGHLEHFLGIIID	163

<i>HmMyD88</i>	RYDVLEDFQIMTQLERCKPS--AFSPKFFAELQDENVVGSS-----NDENYWLTVGEADSG-----APEYFNAFVCSSEVD	174
<i>XtMyD88</i>	RNDILTLDFGLIEADCMKHLKKNAP---LPLQDDKVDSE-----SEQYRIKSDDPGSGS-----LPETFDAFICYCAQD	159
<i>BbMyD88</i>	RFDVLEETQSSLAKDISKYRERSSSP--MPVQVEVSTGNYPNVPTTSELHGITLQDDPYGV-----DKELFDAYVCKED	170
<i>CsMyD88</i>	RDDVVEDLRPLIDEDVRRYREKMAER---PLOVEEVDSC---VPRTPMSGITLEDDEPDG-----TPEKFDAFICYCHSD	160
<i>HsMyD88</i>	RDDVLELGLPSIEEDCKKYILKQQQEEAEKPLQVAAVDSS---VPRTAELAGITLDDPLGH-----MPERFDAFICYCSD	171
<i>GgMyD88</i>	RHDVLELGGVVEEDCKKYLRKQOEAEQFLOVEAVDSS---VFKTSELMGITTRDDPYGH-----GTEMFDAFICYCQKD	174
<i>SpMyD88</i>	RHDVHELPAFLEEDCKRWKRTQAEAR--DPIQVEEVTGN--FASTSDLLRGIITLNDSPS-G-----PEMFDAYVCFAMAD	244
<i>CfMyD88</i>	RFDVLESDRTMIERDVEKWRQNQLSLQKINDDPTFREPOSRRSPQSEETQTRADVGSKDEI-----YYDCFVLYNFDASDQ	178
<i>ScMyD88</i>	SKDLEKCLLESIQ-----KQQFDVFLSFAEAD	125
<i>DmMyD88</i>	RWDVCCDDIQENLAKDTQRFIMKQEQRTALVEACEPPPSDCFETNNNYSNNITVGGQSVQILSDEDQRCVQMGPPLRNACVLYAEAD	253

**TIR Domain**

	Box2      Box3	Box1
<i>HmMyD88</i>	LTIVKNNMTELEDKQFF--KLCIPDRDFLPGP-PKYEATTKLIQARCDKFIIVVLSPDFLKSDECAFMNLNVALALNFSARKKLIPIIYK	260
<i>XtMyD88</i>	ISFVQEMISRLSEQTDYK--LKLQVDRDVLPGT-CLWSTISELIEHRCKRMVVIISDDYLSSECDFOQTKFALSSEGARFKRLIIPVKYK	246
<i>BbMyD88</i>	RDFVIQMKRLESSEFGRRLLKLCIDDRDLLPGT-AYLTVTAELIENRCKRMVVVLSPEFLESPECDFQTKFALSSEGAKRRLIIPVMYK	259
<i>CsMyD88</i>	LEFVHEMIRQLENSNYK--LKLQVDRDVLPGS-CVWTTITSELIEQRCRMMVVVISEDYLDSEACDFQTKFALSSEGAEKRLIIPVYK	247
<i>HsMyD88</i>	IQFVQEMIRQLEQTYNR--LKLQVSDRDVLPGT-CVWSIASELIEKRCRMMVVVSDDYLSQKEDCFQTKFALSSEGAGHOKRLIIPKYK	258
<i>GgMyD88</i>	LOFVQEMIRLEQTEFK--LKLQVDRDVLPGT-CVWSISGELIEKRCRMMVVVSDDYLSSECDFOQTKFALSSEGARFKRLIIPVKCK	261
<i>SpMyD88</i>	LEFVQQLRSQLESEPHN--YKLCIDDRDLLPGG-SHALVTAETIKNRCNKMLVLSPEFLQSPSCDFQTKFAVSLSEFARKRRIIPILVK	331
<i>CfMyD88</i>	LEFVKQMSRILEGPEYN--FRLFLWRDLDVGT-AMHSVSAAIIEKKCRRCIVVLSKSEFNNSPAADFOQKFAHALSEGARKRVIPIVIE	265
<i>ScMyD88</i>	ERFAEEVQRRLIENAGVR--VFVPSAGLMSDR-AFHQETADIKQGCRTRVVILSPDYLSSSWCGYEANLAFTNSPDARNTTIPILYR	211
<i>DmMyD88</i>	IDHATEIMNNESEERYN--LRLFLRHRDMLMGVPEFHVQLSHFMATRCHLIVVLIIEEFLRSPPENTYLVNFTQKIQIENHTRKLIPIIYK	341

→RDxLPGx←

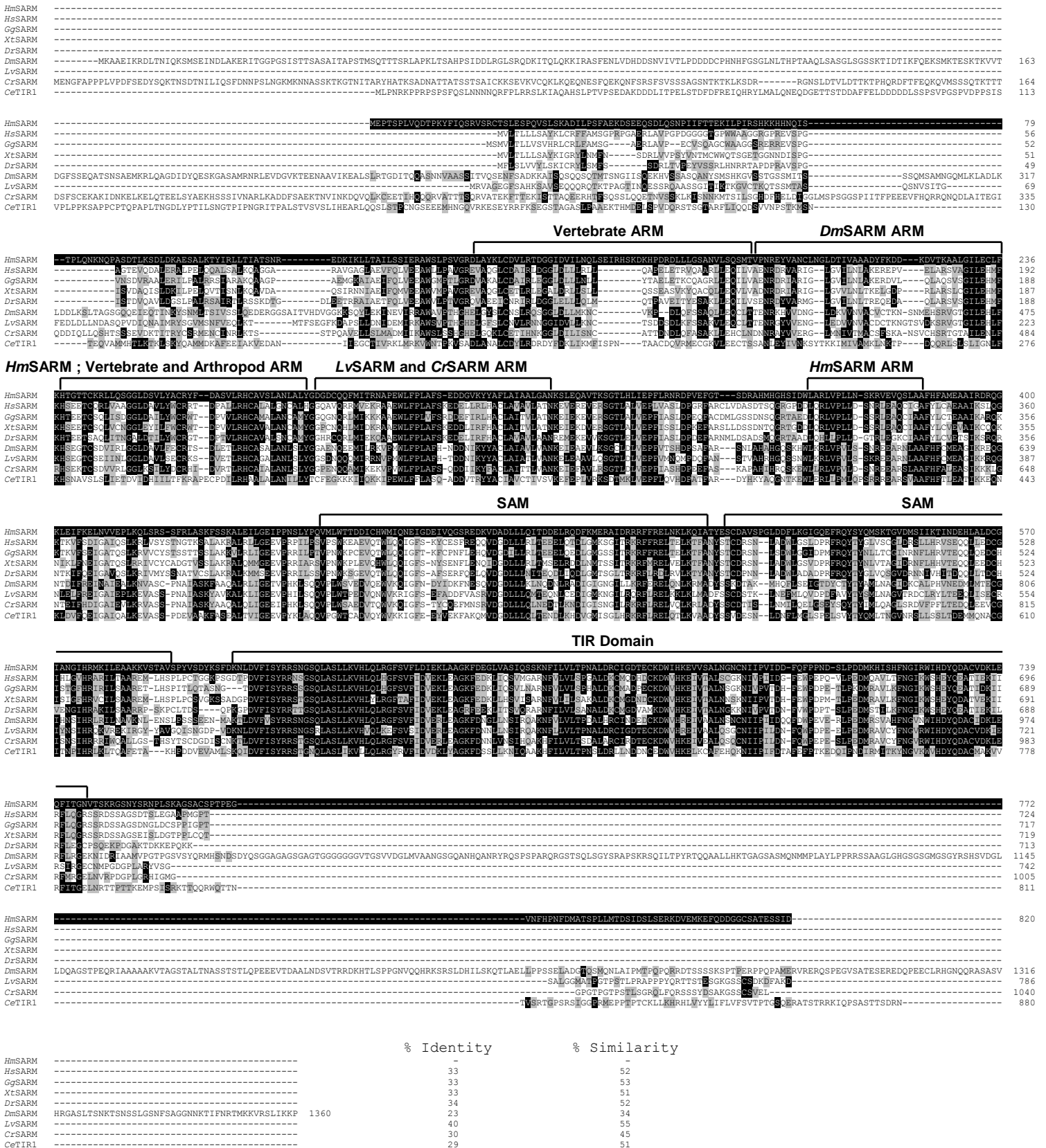
<i>HmMyD88</i>	RCE--IPPTLSFITLVDFTRG-----DVIGFWNRLSKSLITTGCLNTIISSSPSLPSLKDFOPTSQDTIVITLTKTKPRPGEFLEVEKGPMS	343
<i>XtMyD88</i>	PMKRPFPSILRFITLVDYTNP-----CTKGFWFERLAKALS-----	283
<i>BbMyD88</i>	QIE--VPQILRFVTVIDYVKE-----ELEERFWHRLSKALS-----	295
<i>CsMyD88</i>	PMKKPFPRILRYITVCDYTRQ-----STQAFWDRLAKSLSLP-----	285
<i>HsMyD88</i>	AMKKEFPSILRFITVCDYTNP-----CTKSFWFWRLAKALS-----	296
<i>GgMyD88</i>	TMKNEFPSILRFITLVDYTNP-----CTKKFWFWRLAKSLSLP-----	299
<i>SpMyD88</i>	PCD--LPLILRHITLVDYTRQ-----DLRPLWDRIRKAMSIR-----	367
<i>CfMyD88</i>	NIE--PPSILNFWAPAFYNY-----GIRDWQWRVAATIKSELRPDPEIWLPKVDANNIALDTSVITKRELWGITVATQIYPEEDFVGI	348
<i>ScMyD88</i>	PCE--VPAFLSHLHYLDYPRFKDDRKNCDQYFWFKRLYKSVTHHI-----	253
<i>DmMyD88</i>	TDMH--IPQTLGIYTHIKYAGD-----SKLNFWDKRLARS-----	425

<i>HmMyD88</i>	ILGSSAGEDSGKEMSESAAVQPTKEEEFIDVSTEVVSVSHISYKFFSFTQCCKSSSK-----	402
<i>XtMyD88</i>	-----	
<i>BbMyD88</i>	-----	
<i>CsMyD88</i>	-----	
<i>HsMyD88</i>	-----	
<i>GgMyD88</i>	-----	
<i>SpMyD88</i>	-----	
<i>CfMyD88</i>	PDVDAQRKKHKNKNNKGRGE-----	367
<i>ScMyD88</i>	-----	
<i>DmMyD88</i>	KVPEAETTIVSVSGDTGSPLEPKPKKKDRFLRRITHSFGKTARSDGASGKTLRHAHSVSTINVTERERTLSASSNISTTSESKSFIK	515

	% Identity	% Similarity
<i>HmMyD88</i>	-	-
<i>XtMyD88</i>	36	53
<i>BbMyD88</i>	34	52
<i>CsMyD88</i>	33	53
<i>HsMyD88</i>	32	51
<i>GgMyD88</i>	32	50
<i>SpMyD88</i>	26	42
<i>CfMyD88</i>	23	42
<i>ScMyD88</i>	21	41
<i>DmMyD88</i>	18	37

**Supplementary Material 1: Comparison of the amino acid sequence of *Hirudo medicinalis* MyD88 (*HmMyD88*) with vertebrate and invertebrate MyD88 orthologs.**

Alignment was generated by using the program CLUSTAL W. Identical residues with *HmMyD88* are shaded in black, homologous residues in grey. Gaps (-) were introduced to optimize sequence alignment. →RDxLPGx← indicates the key motif for MyD88 dimerization and its interaction with TLRs. Accession numbers in EMBL/Genbank: *XtMyD88*: *Xenopus tropicalis* (NP\_001016837); *BbMyD88*: *Branchiostoma belcheri* (ABQ32299); *CsMyD88*: *Cynoglossus semilaevis* (ACU31062.1); *HsMyD88*: *Homo sapiens* (Q99836.1); *GgMyD88*: *Gallus gallus* (A5HNF6.1); *SpMyD88*: *Strongylocentrotus purpuratus* (XP\_780590); *CfMyD88*: *Chlamys farreri* (ABB76627); *SdMyD88*: *Suberites domuncula* (Q4W1E7); *DmMyD88*: *Drosophila melanogaster* (NP\_610479).



**Supplementary Material 2: Comparison of the amino acid sequence of *Hirudo medicinalis* SARM (*Hm-SARM*) with vertebrate and invertebrate SARM orthologs.**

Alignment was generated by using the program CLUSTAL W. Identical residues with *HmSARM* are shaded in black, homologous residues in grey. Gaps (-) were introduced to optimize sequence alignment. Accession numbers in EMBL/Genbank: *HsSARM*: *Homo sapiens* (Q6SZW1); *GgSARM*: *Gallus gallus* (A5HNF6.1); *XtSARM*: *Xenopus tropicalis* (XM\_002937143.2); *DrSARM*: *Danio rerio* (B3DK97); *DmSARM*: *Drosophila melanogaster* (Q6IDD9); *LvSARM*: *Litopenaeus vannamei* (G8GV23); *CrSARM*: *Carcinoscorpius rotundicauda* (A9X3T4); *CeTIR-1*: *Caenorhabditis elegans* (Q86DA5).

<i>H. medicinalis myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	
1	G gtatgtg Glu	tttcag AA Glu	1	
2	G gttaac Val	aattaag TT Val	1	
3	G gtgcgta Ala	tttttag CT Ala	1	
4	AG gtaggtc Arg	ttatcag G Arg	2	
5	A gtaagtc Ser	tatttag GT Ser	1	

**Deuterostoma  
Vertebrates**

**Protostoma  
Ecdyzozoa**

<i>H. sapiens myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	<i>H. medicinalis</i> I/E junctions
1	G ataaaa Glu	cccaca AG Glu	1	
2	G gtaaggg Gly	caccac GG Gly	1*	i3
3	AG gttggct Arg	cccctag G Arg	2*	i4
4	G gtaagct Gly	cttcag GT Gly	1*	i5

<i>D. melanogaster myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	<i>H. medicinalis</i> I/E junctions
1	G gtaggca Ala	tcaaag CC Ala	1	
2	AGG gtgagtg Arg	cttacag CTA Leu	3	
3	G gtgagta Glu	tccccag AG Glu	1*	i4
4	AC gtaagt Thr	tatgcag T Thr	1	

<i>D. rerio myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	<i>H. medicinalis</i> I/E junctions
1	G gtacgtc Asp	tgacag AT Asp	1*	i1
2	G gtcagtg Gly	tcaacag GA Gly	1*	i3
3	AG gcatgtt Arg	ttcttag G Arg	2*	i4
4	G gtacaaa Gly	ttctcag GA Gly	1*	i5

<i>L. vannamei myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	<i>H. medicinalis</i> I/E junctions
1	AG gtgaata Arg	tttacag A Arg	2	
2	T gtaatat Tyr	tttacag AT Tyr	1*	i1
3	G gtaagga Asp	attgcag AT Asp	1*	i3
4	AAG gtaagt Lys	cttcag CTG Leu	3	
5	AG gtaacgt Ser	cttcag C Ser	2	
6	GAT gtaagt Asp	tcagcag TCT Ser	3	

<i>G. gallus myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	<i>H. medicinalis</i> I/E junctions
1	G gtgagcc Glu	tttcag AG Glu	1	
2	G gtgagtg Gly	cttacag GG Gly	1*	i3
3	AG gtgagtg Arg	tctgaag G Arg	2*	i4
4	G gtaagat Gly	cttcag GG Gly	1*	i5

<i>X. tropicalis myd88</i>				
Intron	5'-Donor	3'-Donor	Intron phase	<i>H. medicinalis</i> I/E junctions
1	G gtaagta Glu	gttacag AG Glu	1*	i1
2	G gtacgcc Gly	cttacag GG Gly	1*	i3
3	CG gtaaggg Arg	gttatag G Arg	2*	i4
4	G gtaagat Gly	cttacag GG Gly	1*	i5

**Supplementary Material 3: Intron-exon boundaries of the leech, vertebrates and ecdyzozoa *myd88* genes.**

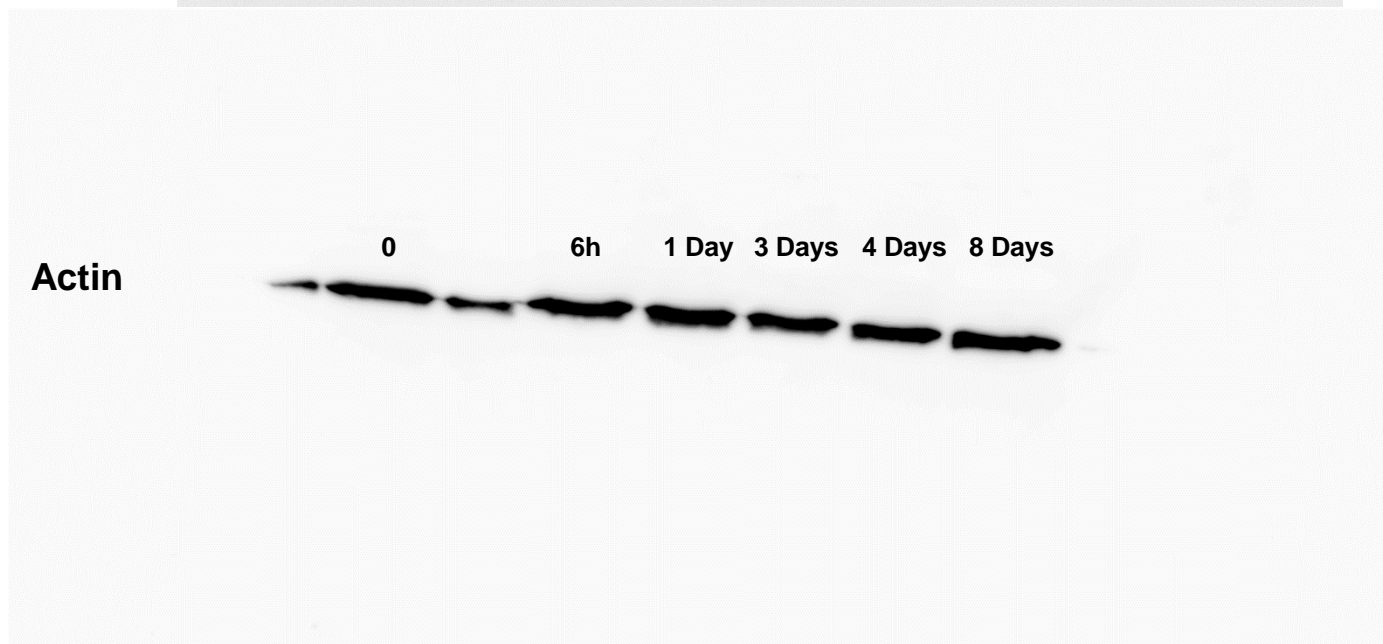
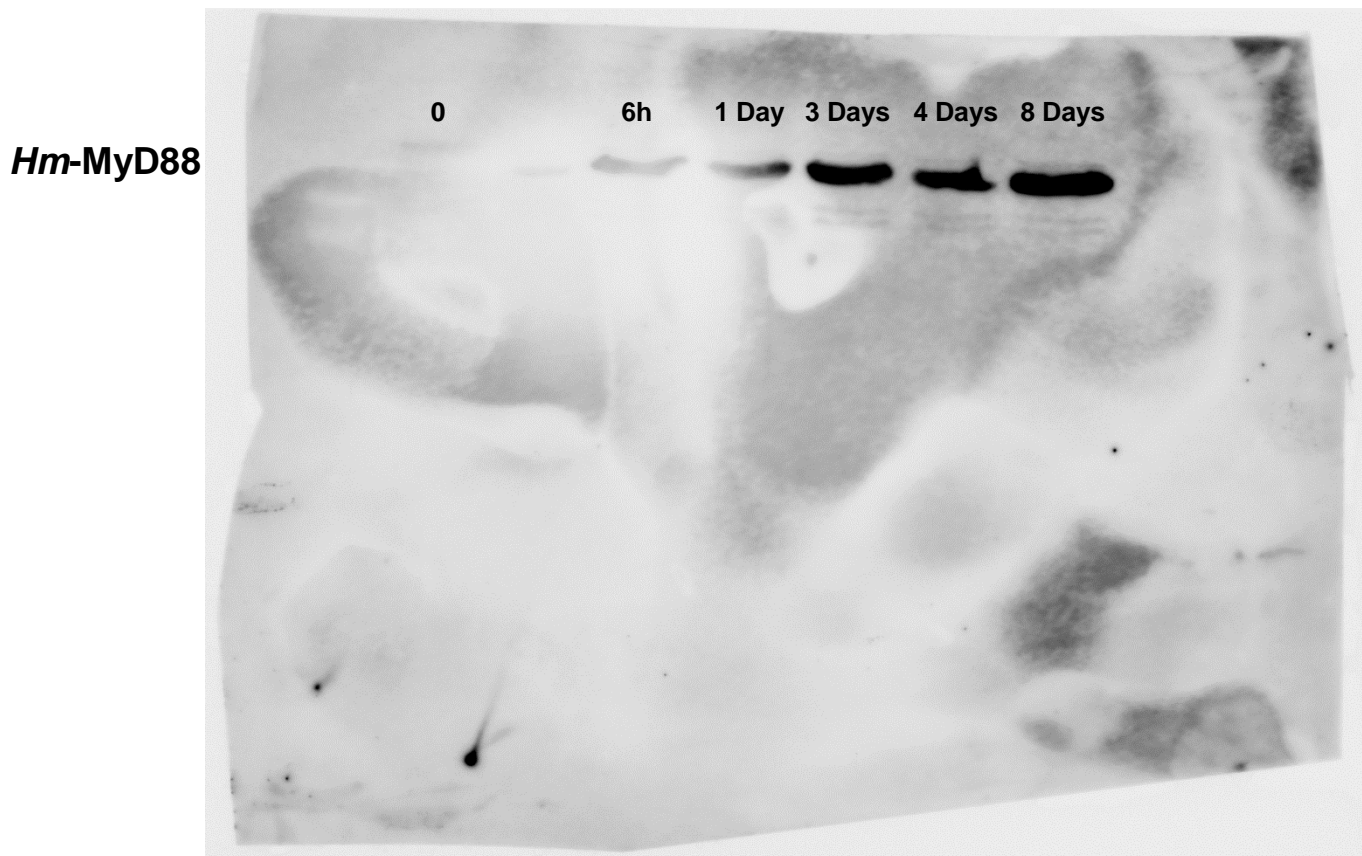
The sequence of each of the intron-exon boundaries is shown, as well as the codons for the amino acid residues. Uppercase and lowercase letters represent nucleotides in the exons and introns, respectively. Asterisks (\*) indicate that the position of an intron-exon junction is the same in leech and a vertebrate or an ecdyzozoa gene. The column entitled “*H. medicinalis* I/E junctions” point out which leech intron-exon junction is conserved in vertebrate or ecdyzozoa genes.

<i>H. Medicinalis sarm</i>			
Intron	5'-Donor	3'-Donor	Intron phase
1	CAG gtatcac	gtttcag GGT	-
2	GAG gttcgaa Glu	aattcag ATT Ile	3*
3	CAG gtgatat Gln	attttag AAT Asn	3
4	AG gtctgtg Arg	gttttag A Arg	2*
5	A gttggca Lys	cttttag AG Lys	1*
6	AG gtgaatg Ser	atcaaag T Ser	2*
7	AAG gtaaaat Lys	tttccag GAG Glu	3*
8	AG gttcatg Arg	ggtgcag A Arg	2*

<i>H. Sapiens sarm</i>			
Intron	5'-Donor	3'-Donor	Intron phase
1	CG gtgagcg Arg	tgggcag A Arg	2
2	AAG gtgggtg Lys	tccctag GTG Val	3 *
3	CGG gtagagt Arg	ccccag GAG Glu	3
4	AG gtacgga Arg	ggtgcag G Arg	2*
5	G gtcagcc Glu	actcag AA Glu	1*
6	AG gtgagga Ser	cttgcag T Ser	2*
7	AAG gtaggtg Lys	tctccag GAG Glu	3*
8	AA gtgagcc Lys	cccacag G Lys	2*

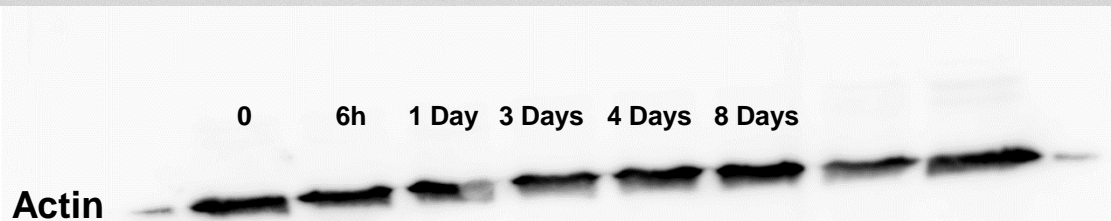
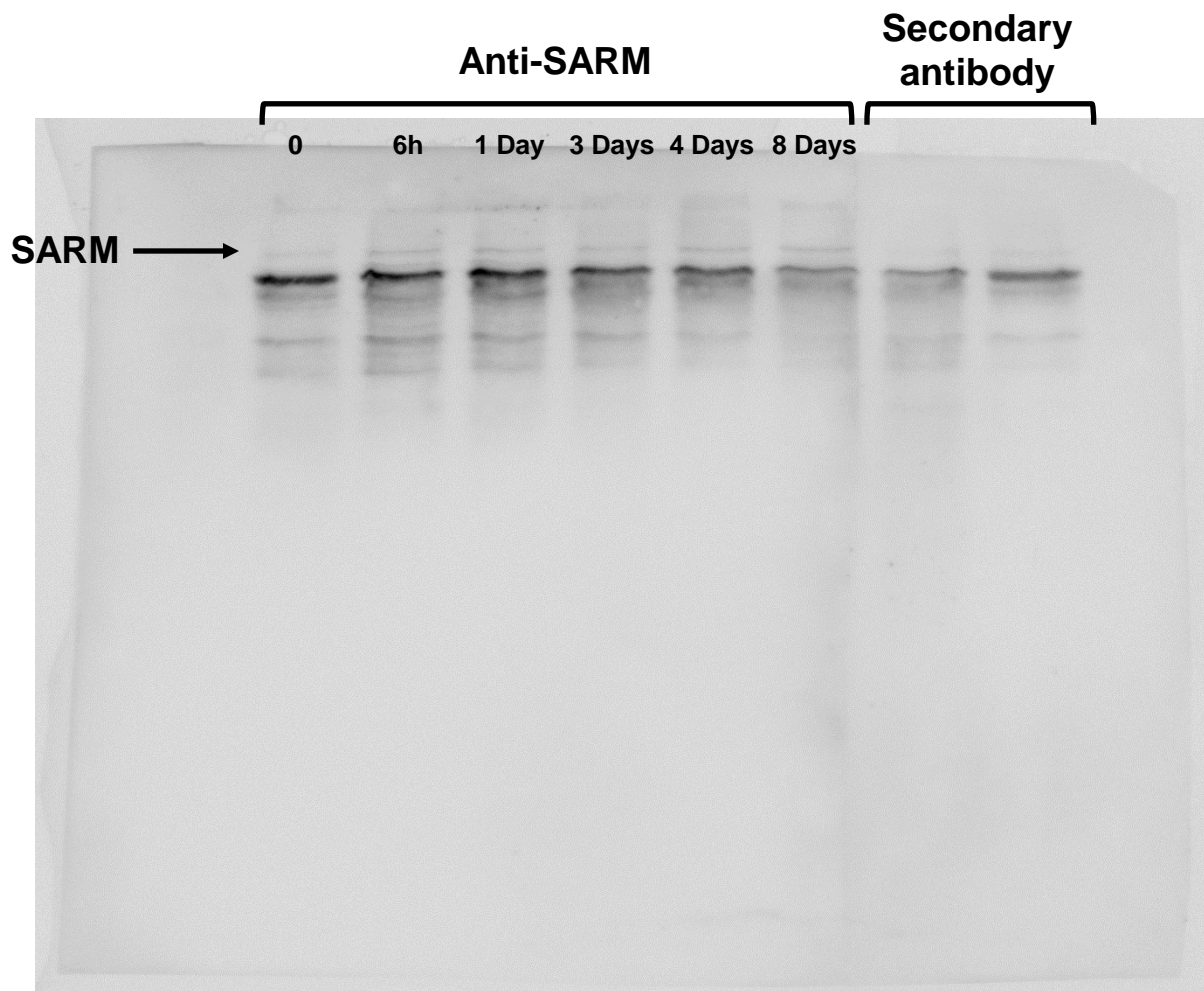
**Supplementary Material 4: Intron-exon boundaries of the leech and human *sarm* genes.**

The sequence of each of the intron-exon boundaries is shown, as well as the codons for the amino acid residues. Uppercase and lowercase letters represent nucleotides in the exons and introns, respectively. Asterisks (\*) indicate that the position of an intron-exon junction is the same in leech and human genes and has the same intron phasing.



**Supplementary Material 5: Unprocessed original scans for the blots in figure 10 (A) Kinetics of expression in the course of neural regeneration.**

*Hm-MyD88* protein expression was assessed by western blot in isolated nerve cords cultured *ex vivo* for up to 8 days under sterile conditions. Samples were collected at 0 h (T0), 6 h, 1 day, 3 days and 8 days. Membrane was incubated with anti-*Hm-MyD88*. To assess that an equal amount of proteins was loaded on gels, membrane was stripped and re probed with anti-Actin.



**Supplementary Material 6: Unprocessed original scans for the blots in figure 10 (B) Kinetics of expression in the course of neural regeneration.**

*Hm*-SARM protein expression was assessed by western blot in isolated nerve cords cultured *ex vivo* for up to 8 days under sterile conditions. Samples were collected at 0 h (T0), 6 h, 1 day, 3 days and 8 days. Membrane was incubated with anti-SARM. As a control, a membrane was incubated with the secondary antibody alone. To assess that an equal amount of proteins was loaded on gels, membrane was stripped and reprobed with anti-Actin.