

Supplementary Materials

In-Silico and Functional Characterisation of Nematode Galectin-3 and Galectin-9

Elizabeth Mullens¹, Gemma Zerna², Farah Ahmady-Nield^{1,3}, Travis Beddoe², David Piedrafita⁴ and Sarah Preston^{1,*}

¹ Institute of Innovation, Science and Sustainability, Federation University, Mount Helen, VIC 3353, Australia

² Department of Ecological, Plant and Animal Science, School of Agriculture, Biomedicine and Environment, Latrobe University, Bundoora, VIC 3083, Australia

³ Fiona Elsey Cancer Research Institute, Ballarat Central, VIC 3350, Australia

⁴ Institute of Innovation, Science and Sustainability, Federation University, Churchill, VIC 3842, Australia

* Correspondence: sj.preston@federation.edu.au

How To Cite: Mullens, E.; Zerna, G.; Ahmady-Nield, F.; et al. In-Silico and Functional Characterisation of Nematode Galectin-3 and Galectin-9. *Parasitological Science* **2025**, *1*(1), 4.



Copyright: © 2025 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Table S1. Protein sequences collected for analysis and naming key.

Organism	Search Method	Sequence Name-NCBI	Sequence name-Uniprot	Name (Research)	Protein Sequence
<i>Trichuris trichiura</i>	NCBI BLAST-Gal-9 and Gal-3	CDW59317.1 galectin [Trichuris trichiura]	A0A077ZG03	TTGO3	MLTTAEDLHKEWMKKLPYWAKLENPIQPGQSLIVKGLVCGDRFDINFTAGPNIP DHDNVVLHLSCRQNEKAIVLNSYEGGEWKKEVRFKTPFKVGEFDDVIRAHDDRR YEIFADHKHLGDFKHRLALTATTHVNIQGGCEVNAVLWEGNYYSVPYKTSITGF APGRRLFVSGIAVDSPKRFAVNFFSKDDIALHFNVRFDEKKVVRNTLKGGAWGK EDRVGGFPFEKKKNFDLLFVCEEDTIQIYINDDHFCYAHHIPPKQIDAIEIEGDIE LQLIHIE
		CDW52167.1 galactoside binding lectin family protein [Trichuris trichiura]	A0A077YVV9	TTVV9	MSIAVESLTRSLASDTSGCLTAMVLDENRIYAPKLP LISIIPNGMTAGKSILVKGL ITGKHRNNFAINLCCGLLTDGLHRDNIALHVNPRFDESEKLVLSNLIIDDTWGPEE RHKNPLRKGAQAFMIRITALPDSFKIAFNHDVTIEYLHRVPLSDIKTIQVEGCIVDA VEYFAAPPTALPFSLSVNYKEALPVLRFMTWVWLKSPYRIEMDFLAGDDIAFHSSV RFDEQTIVRNVQRRGKWEAEERLLDIFPFQPSCTYELEIHCLPEHFVVRVGGSHCY NFNHRLRPELIDRFNLKGNARLLKFCLE
		CDW55930.1 Gal-bind lectin domain containing protein [Trichuris trichiura]	A0A077ZBA2	TTBA2	MENVFVKKLDSCDNGLIIRIAGEVPEDARWFAINLQNGTVMHPRSDIVLHISPRFD GAFFRVVRNSLVNERWGNEETAGHFPPKLGEQFEVLILVEHEGYKIAVNGEHYA EYRHRLPFATVNAIVIDGDVTLYAVETQAVLKTVPSPAPVEDHYADSSTTATAHPG LPGSVPPAAPPYPVPSPFTPPYPPIPHGYVPVQARWACGLGNQAPVPPYPVPVSIPR TNENPQFEADSASPMPIPENSPIPRKRLSLFKAAGIGAAAGLGAYALGQLTYPGIA NFGEHILEARGAHQCDEEHTKWGSSDSDSSEED
		CDW53214.1 32 kDa beta galactoside binding lectin [Trichuris trichiura]	A0A077YZM7	TTZM7	MSVPYLSRLEGSCIQPGQSLHIRGVLLGEKFELNMMTGSKDNDMPFHMSFRRK DKTCVMNSMSNGQWGKEERKKISLKENEFPDLRIRAHDSKFEVFLNQKELCEYD YRQPLTSISFIRINGDVELNHVSWGKYYVPVPEAGIPDGFSPGKRLYVSGVPEKK AKRFVAVNILTSSGDIALHFNVRFDEKVRCAVVLFNTIVRNSQFGGTWQNEEREG KLTLEKDHAFDLMFVNEVHSFQVFINGIHWCAAHRTDPNSLKGLRIEGDVELQ GVHVK
		CDW53139.1 galectin 9B [Trichuris trichiura]	A0A077YZH6	TTZH6	MQARLAVISAENFTTGSVVLSVGEMAEDVKSILIINLVSKANKQLFHFNPRLKK KVYRNAKLNNWTWGKEEFDGDFPFSTASNFTIMIVCQTEGFRVFVDQAPKFFFAHR ETAEIFDITVYVEGNSSDIVINDVGYY
	Uniprot keyword search		A0A077Z3S5	TT3S5	MLSLRPSLASVLTFTPTSLSTTTVMIFSLWKTSRLWKFIPLFCLCYFLEAITFTYLISA VQSMERQFQIPSHLSGTLVASSDIGYIPTVILAHFGSRGNRPRWIGFGCLLISLACL LISLSGLTIPPADFGSGSNSLNASALESTIPKPLMLYSEIDTPSKLTSHMYLGCAFSR WLIENDLSPPSWAKVDEDWYQNSCQSSNRSTALRFINNSSFWDLLASQLEHNVG YGTFAKVLVEKVNAAQNYPYEDELRTENFASLLKQSKRPQALCSRLINHFRALIK SSQCATKETNRLAFSIVLMGMIAIGVGHSMPWSLGMPLIDDSVKKSKMPYFVGG VFFIRILGPMGLGFLGVCNRLYYNLKADIEVEPTDVTWIGAWWLGFLEIAVALFI PSLMLFFPHVELSLENKPEQSSATVAVQQKNSVAVKAVGDRKLTVCNSIAAEL KILFESLWLLIRLPVYTLCMIGRMIDVLAFFKGFFVFHSKYLETHYGLPQYQANAA MGLAGTVGFALGNIAGGVILRKLKLEGRKVVLVYLCGIFSVSATFVKLALKCSS TLNQLGKMGRENGFNFTTSCNGDCHCDNSMLLPVCSSSGTLYFSPCHAGCRSSQ RVDSLTNFTDCFCMDNGTIASRDFCRDDCDQPFVSYMLLWAATGLVAGTAMVP GLLIVLRYSLSHGGRLTIALLLDIGVWYFAKDLKIVADSDVKKAKDPPTNDSSSD PATFRKNLTNLGHLQEFSDVLLLTNAQTFSINLFVKNDIVWHLRAYYIQRVWV MDSLLKHRWQYQKEIAMKSSIPTRRPNLSLIIQGGIPLPWEASFAPLISVQVVEI VLAIQDPVTSVINLHVGGDIAWHFRAYYALKGWMNSNVGVDGWWQAERRMPM KLCVVPGRINVRIVIGPSKVQATYDNQTINDIRDLVGLPISHVTVTHDVILYKIRI YNMGRDMSLLPKISSLWKFLIAFCLCFFLETMTSTYLISCVSQIERQFQITSKLSGIL ISSGDAGYMTVVLLAYFGSKGNRPRWIGSGFLFTSVACFFIALPSLLFRPNQSHQ LAFSNDTFDNLKLPSELITYGEVDSVRKLLAHPHIGCSIRNWMLRDGVQIPKRWH

		PFDQSVHLCVDDENMLTFDIPEPLRKMLLFYFSQTTTTEIFARHLVDAFNSIYAQP QNQVTWQHYYHKEKHLNLSRLPSYCSHLVNRVRLFADDOQHCEMRRTNEIAFVLV FIGMILIGIGHSVPWTLGVPLIDDCVKKNAPAYFSGVFFIKILGPMVGVFFLGGLCN HLYYNLSSDPAISYTDVNWIGAWWLGFLLIFFLLPAPAMMLFFLPRFDRVVKPSQ LGGKQPQVHPDPTNNVGFQHDSEKHHSPVLLELRRFVFALVLLFKSPVYMLSL VGRLEFELSYPKGFVFQTKYLESNFDLPVRVSTFYLGIFDVVVFAMGTLGSGFLIR KFRLQGRKAAAFVAACTGVCAVASVLQIFLSCNFAKKSAAAAHSLHTIDALTF NNSSSHCECDPHLLFPVCSSSGKAYFSPCQAGCQQFSPANLPKNFTDCFCVDSGS YVSQENCNDCRIPIIYFALISISGFASGISMLPSVFITLRSVPIHLKSVLGFSGFLL SLLSTLPSIMYGHVIDSTCLLWKKSCGKHGSCSLYDVDAIRTKHISLHASLRLLS FLFDMVIRWAKGLKFLEEPSFEEREPEEMTTLPAPAVQPKKKRCESLRVSART ARRTRKVSSETMTPIPLLLIIGCMCLCWHEKSIPIIITPGIQLFRKYYPWPQTRDIP GYDPYFTPYSNWGVLGKGFQPNPKHYEFMQPVGLPRFVSVILNGIRLGNNNGHMK MGYYLLQLKKEQFISIQLGFGFCNPQTNHSSKKSHTMSNLRYACDVLVFIANCHNQ FVFRYSLKTRNVFSPHERQSVAEENLVKRPLKITLPLGTTEWSTELEELAACKLAV ECIDEPYESQDGAIAIYVAKDKQGVKEVIAQLSEMAKKYNYDNNQCEQDDSAV CENYKQAIRLVGRRKIRLQNGNMQNSSELRGHSAHLRREEDKVRLEIFSKQAG QPYATGSQCTFCSATANCKDKLCCGEANPLTNSSCGEKPTNLPLVYHYNRN DTVLTSTKTRDQLMKKVYEDKGIFGYIAENNNTTDCHYLTPLTALKRPTVDYY VYVADEFKTVYESNGWDVVGELGYVVPSEGHCSAKLAYQFQRSTLNSALFYIT VDSMQAAEMLKKKAKQDAAHNGEQPRTLPISGRSNSHSMGGLLPFKIPFKHTG PSQNVVLELSLKAGASKIIDLNSDCTVVWRLEGNFRTKLLFIDLNLWEKAPLHAR TEGKLDNGVWKKQAQMGTDAPIDGKQMNITAVLDDDLVQFYNGHKVSPKYS LPKSEVKVITAGGNVESYEISIQSPSSKIDEDLVQKSITTDNIETDTEEMPRTIDE LWTFSSLINGQWTRNVNFRYRIPLTSGKEFLLNTTVQNGLAQMSLNGEKISPYK LPGKIVKNVTVTGDVKKCTVNVVHMKRIIKLHEYDPETVEEPATVPTTEQAWNE QEKYKGPASSVMLRSIQATVKLVKRPRYFYMRVKMENDKEWKLEAHLDDKE WTFSQLINQEWKQRLPFAYELPKRYNNVTLTVLAQDKMAQVLVNGKAISPKY QIPGLRVKNVTVAGDIASAVVRVAHVKRVARPTYVIQGGQETDITDDAKQQQEE TRKRVDLRRLQEKWTSKTLTNRVWEQQEDSFKSKLPLDDERTFNLLAKVKDR KMQLWLNKRKLPLTYTLPAERAKKIIIGGEVEAFIVNALHERIPTSELVSVDPK KKRRTDPFVRTTIQKGLLTEKEGADRPPEFSLLPRTEGTQILPKTATTNVALKKNA QNFRVDVKMGPSATLRDAPLRRPWTLSRIVDGNEEKITQFKPNLRMTDLKNPL TFTIENGVLQSVNGRNISSKFFPLPGNQMETIEVSGDVNYSVDTEYPEKKQVPSS EKGKDIEDKDTEKNKDDADASKQLSRPFKEGPLHHIRATVQLQPDASEFYMIHM DNNIAWKINVLNKTKEWTPSTLINNVWEQQEDTFKSTLRLNDKRTFNLLAKVKN RKMQLWLNKKLSPIYTLPGERAKKIVINGDVQAYNIDAFHKRTLPKPTGADSV DPQKKRKVDPLVRTTIKSETLTDKENPERPAEFSPWSSNEGSQVLPRATTNVAL KKNARNFRVDVKMGPRATLRLHAPLRRPWTLRSIIDGNEEKITQFKPNLRVTDLK NIPLTFTMDNGFAQLSVNGTNISKKFPLPGNRVEKVDVNGDVDSYSVDMEYPEA EQPVISEEETEGLRFPDENEIEDTNTKTSEKNEDDVKPLSYPPKRGPLRLIHATVQ MQPDPSPKFMIVRMDNNTAWRINTNMKTQRWTPETLVNNVWVQQKISFKSPLP LNDKRAFNLAKVTDREVELWVNKRKISPKYTLPGTRAEEKVVGIGDVFAFVNT LHRTLPTPIEAGSADPKKKRRKDPYVRRTIRSGPLEDDGDLGRYPEFSPWSNIGE TRMLPKTATANIALKKNPKGFRVDLNMGRGATLRLEAPLKRPKLSQIVDGNEE KSTEFKPNLQVIDLNNIPLTFTIDSGFVQLSVNGRNISSKFFPLPGNRLETIGVSGDV DSYSIDTEYPKADERVTSEKDTEHWRLPEKEIDIENTDTEITEDDEDVKEWKPS VIDNVWEQQDSFQSKPYIGSEKTFNFVANVKDGTQVQYLQNLTKRAVELPKKTT VNVALRKNAQNFRVDVKMGPRATLRDASLRRPWKLSQIIDNGEKEVTFQFKPN VRVTNLRKIRLTFTMDNGVVQLSVKGRNISKKFFPLPGNHVEAIEVTGDVGYYSV DTDLPVSDQTGPSETNVKSDAENGELLETRKVVRTTNDGADGENEGDAKGHPPK MSPKYESALRKIITKIKLPDTSFVLRMQQESGKAWQLDVLPTQIWRVSSLSG
A0A077YWX2	TTWX2	

			GTWKKHPSFKNYVPLAAGQEFDLTGNIINGLIQFDKKKRPYTLADAELTLRPNGR HIYIRIPMEKNTWEMITADAKTNEWTCSMINGERLREVSYQSNHPLQYGKKFPL LIFVKAGVGQVGPDMEKAYTDKYALPGRSVKGIMITGDVDVHYLSAVNYPRDS MDSVVLPSWKPTIHTEDIVTEIKDHWEESNDYKNALAKRIQYITEAMPDPSAGM IFMTMKVNPNAKKFHIVAEMDEKKIGWILDAQLPELHWVLRVIDGKFKTVGNY TSKTPMRQDQKILLLTQIHNGHLKFIIDTVAVPLKCKLPATIVRNVIALGDVESFEP TIVYWKNSSKSPQRYDDGSFGQTKIRPEELIINAGSPSSAIGRAKLTEILEVSKAPPK KGQQEKL RDMQSIHATVRLEPNAESFAVHIKLDDETEWKA VVNVLTLRWTISSF VNGRWIESLAFKSRQPLIQGATFKLIHIDIYTAQILINGQVISPKCPLLGESAKQVTI SGDVDVRAIRTVYWNILTGPEPLQAFGQQQKSPRSVIEKVYGEIMREKQEYFYS KPRGILNSRRPEELKWLINIPMKLLPWADKFGVSIKMDLHVWLQVLVANVKTNEW ILSSYINGTEKNKVSFKNPVPLTTGKLFRLTISIGSGESEVMLNKKVSPVYAVPG RSLGRVVVTGDVIAQEIRSTQLRREPLITTG DYLYGYVLKWETLKKIYDTPVDD KKRKRSLKVEGEYVDITESKILLFTRVLLKKTRKAFTVDFIMKVSVTNPLGTTSM EASIQGIKNTIDKVSQVLLNVAEYRLPYSVSLRDINGVQYLYLRLKFKPGAKAISL KLFSGHKVIYNDQVQTTNKLRSKAHVDRVIDGDVILYKMOVAGTEKSIMDHVPK KERKIFLYYNQLSKMKKQGEAMTLEIILYVLKREHGKKINLEPIALPYYPFKRN VVANFAMKLAFKQNAAGTFSIDFMIDNEMVLRVQGTAKNSLNAENKHALRLK ENRVDTMLIKETENKPFVDVLVILYRSRIYITYGEALRVYEVFGERRLVNGIKISGD VNLYETSPGVIEKVKRKRRT
A0A077Z0U1	TT0U1	MEAA LARSFNATYLEFYRGPEDIAFCITGNYGAKLWMIKSLINKTWIGPAKYED MKVTAEPGRMFATKVVVQNGRVQLSIDNQSFNFVLP LPARPITMAV VNGDFNLY KAEARTLSRCQNAEYCLVRTVRNDIGDYVSQIVGKSNMRSLAFVTEYKVNGKW KLQRTVNMTVPIQKGRMFNVTA VLKDGKVKA TVDNQIYDVLVPVGYKPKITKVS RGLFALYSMNVTTLTYANSRDTKDSLERRRPNGVISEIAEVNEKRYLALPYRAHF AKFGKSLLFSMQVSLLPNAKVTLIDVYSNFDIVWHIRGDYANSRWLNSWIDKK WEKDRIEQMKIAKVDGRLFNVTIKIEDGKVITIDNQTIVDNVAQRDYPVTSVNV VGDFFLHGAEAVVLKEDPIMTEREKARIESFIRERNPFLENGVPSGLNALPFWLP FKLPSATQFLFMQTSCLKPKPYSTIIDFYVGDAIVWHFRGYDSKLWVLNWKRSKL WQGQRYFQMSVDHTAGKTFNVTVMNGGYVEARFGNQTIYDFRDLDPDKPLTGA LIHGDFLLQKAIVLPMESNYVPRSDADENAGRQKRLVEMEWKVKPLSLPITT KFEPPQYKQILYLQVSLKPNRHSLIQFYANENLVWYIRGVYENKTWVLSQIEN DLPVKRYRKMEIAGVQERIFNVTVALDGMNVKATIDNQTIYDIVDLTVPVTSVTV DGDFTLYEINTGKVERDPLATKRGVPGTDNTQPA AHPKEARENKNRLKQTEKQ WMLPPVPETKEGSDERMVLLKLQELPGAKEFEVDLEAADGKLWTIRIVLEKETW TVGSLVKGVWRRNSSGILKFPLKVAGGFPVKLRKQNGTILAFVDKIGNNLKLQLP EKVLQRVSVKGDVIAREARLLQRLNDEVNRSIDPHCEDGRILEAGESLMLVAIA DDSAPSNKTKAISAIKLPKPEAKRFHMKVSMNDGNAWRLVADLDKDEWNYSTL IDETWVNRGSFKSSIRLADKGVFSLRSVMGNDSEFYLLDDRKVPWPYSMPGSFA KHVELSGDVEKRLVSVENLPSTAQSNEIVWHDDDEIEINERREELMSKELQKQAE DSSRISKSNVMSGKNMRQA FGMNRVSSVKEWLLPPTPKELRTNDTGREIFVKFE PLPETKRFDLDLEINETVAW TMRVDLEKNFTTCGVLVEGKWHKKFLTALGKPLK VGSASVVKLKLENGTVQVFLSGTTKSPKVTLPGETVHHVALIGNVVSHGARSKL FSQSEQQLKKSLESESPQEEVDGSDDLMEWEAAEPLLPILVTNTTESSNATKTVTA VGKVKDKGKLIRVDLKMDSGSASRFQLDLTSNDWYFNTFADNVWEKRGPFKSPI SLAGGDAFKLEARMKDGV AELYLNERKLSPSYRLPGRIVQKATMSGDIQGRSITI EDLEWMSQTEETIWHFAGVDSKVQSEQEKEQNLKSKEQESSPKMTGHQHYVPA HQQKKFGLLVGIVARFILKPD AKKFQLYLETDDNKAWRLYAPLGSKDWMFGTF ANGTWQKQKHFE GPKLKAGQEFLFNLFVNDGIATVYVDTHEMPVKYELPGLTA KRFIVDDVTNNSYSKMRLESMITSSKQILRNDEQSTTTEEASGGSLSRPNTQVCTL STLINGVWQKEMSFKSNVSLATQETLKLSAKIDDRFVQIFGKNKSISPKYEVLGKT VEEVLAEGDVETYSIKVAHYR WEEKQPVGTDHPLENSLKSQGDGREFSDSEFAL	

					APASYKFLRAIRITIGLNYIAKKFRVDVIMDQQTIWRLDANMQSREWMLSSRING AWKKQVKFNCVPLAGGNVFLHTASVEDGFANLLLNRTVLASKYSLPGRFAKN VTVSGDIVSYSDLEHAKGVLVPSDENVTQTTYSGKNKNKLPNRSGDSGRPEWT FSLMVNGLWKKPLTFKHDLRLEKTKTLNALVTIERDFVQVTLDNRTVSPKYPM KLHSLCVFVKGDVSKYSVLALLQKWKKGSPDNASRQGMKKLTPATSEKVSLS KRKNDIMAFKWNSTTVSMLYAIVTLKPVQQLIIFYAEMDKGAAWTFQIDTVTQT AIFRLRLNGAWQTAMKFENALPPNYARNFTFIAVVENGITIEFTVFGKKIKPEVFLP ARVFNNLWIDGADIKFTWHSIYWKKTLMTSRNRDGSFSQYDVGPDDFIRVAMKF PLPSVDGRKNNIRLLQSIVANMKLKEKRNKLFMKLKFDNGAIWRLDIHLSTGKW TAKSTAGDVRSVELKTETHFLATSTFNLTANVEYGVAAQIYLNKGQVATRFPLPGE IVEATTFGGEIEKCEISTVSLNEITLNGRWYESIPLEYGTVPYRFLNNLITGIYGIPI KANLHFLYDPPLNLSNRESTVDIFLTIKLAPAARNFQIDMEIDDSSKMRLQFNVVT NKWVLTSFIGGVWSRNISLQNYMDLSQQESLTLAAAIRNDVVMFLLAGQRISV YQLPGKILRGIAADGDLSCQVRTVQWKGFEFYKTVKMLNDVPPIARRLLQSME YAKPSEVF
Species	Source	NCBI acc#	Logan et al./ UniProt gene name	Name (Research)	Protein Sequence
<i>Necator americanus</i>	(Logan et al. [48])	XP_013308007.1	NAME_12433/ W2TUL5_NECAM	NA077	MPEKKGKRFINILKKNGDIALHFNPRFDEKAVVRNLSISNEWGNEEREGKMPFE KAVGFDLEIKNEEYAFQIMVNGERFTSFahrLEPHELNGLQIGGDVEITGIQMH MTSFLKWYEHNKPPYRSVLQEKIEPGQTLIVKGSTIDESQRFTINLHKSADFSG NDVPLHISVRFDEGKIVMNTFSNGEWGKEERKSCPIKKGDSFDIRAHDDRQFVI IDQKEFKDYEHLRPLTSTITLSIDGDLYLNVHVWGGKYYYVSDHRKSLCFTSEIRS
		XP_013307933.1	NAME_12704/ W2TXA4_NECAM	NA933	MVRLARSKSTSVSKSAVLSVKLPQHLQNSMTEKSYVSSFYLSIGHGLQVYQVFR AALTATSANSRGKAPYSGHFRGRSAHLRSPQPALVVLVSSRMPTLWHFAKYINI PYRSKLTEKIEPGQTLIIRGRTSEEAKRFNINLHKDSPDFSGNDVPLHVSVRFDEGK LVFNSYTKGAWGKEERQKNPVKKGDGFDIRAHDNKFTVSINRKEVKSFEHRIP LQHVTLSIDGDVVLNVHVQWGGKYYPPVYESGIAENGLIPGKTLVIYGTPEKKA KKFNINLLKKNKDIALHFNPRFDEKAKGFICGKIKPGCVIRNSLVNGEWGNEERE GKNPFEGVGFDELVKNEEYAFQIFVNGERFASYAHRIDPHEIGGLQIHGDVELT GIQIVGN
		XP_013299521.1	NAME_08305/ W2T6P9_NECAM	NA521	MIGAGVGMGSPVLYNSATPAGLRFAGFENGKRLRLVLIPLPEGNKFAVNFRTPY DIAFHFNPRFDEDCVVSNSSTRNGGWLNEERTNPNPQVGKVVYTFEVSNGYQIQV LLNGVPFTTFAERLPGHDINSVEEGGVHVHSAHVF
			NAME_11043	NALOG2	MAHPTDAPIPVPYTTKLGQPLLPQQLDIHGKINSASRVEINLLHGAAQIDPGQA VLHVNLRFDEKKIVMNTYMDGAWGKEERESMPFKQGEKFDLKMRLVDGMEI WCDDKKVHLYKHRMPFDQIEYLQIKGDCCTLSGVHWGGKFYQVPWETGFPDGH RASQRIMLHAIPKGDWRNLDLLGRGGDILFHFNPRFKDKQIFVNNERIGTFQHRT ENPTQDYIGMRIDGIEVTNVEFV
		XP_013304533.1	NAME_01439	NA533	MFSGQPDQQLVINSYINNRWGAERFPNPFEIGQPFIQIRILVLENYFKIAVNGKHIC DYPHRVPVEEVKTIYVGGNIRVDFIEFQPPPEAGSSYTG
	Keyword search- uniprot and NCBI BLAST of gal-3 and 9	XP_013304534.1	W2TMU3_NECAM	NA534	MRYPAENGLHDLAKALRQPVSKTASKKKKMNQELPFVSAIVGGGLFPGRAIVVSG MVLQSFASDMKRFYIDLCCGLLIQGDHMDNKAHLFNPRFDTGGGW
		XP_013308382.1	W2TVN3_NECAM	NA382	MRTHFRSLSPNATMKWLLVLLVAYAFADEDSKESSEKETKKNYKYIGEONQR LPFKTRVSEAFVVGQTIHAVGNLTDKPTRVDFNFHKGAAKDADLPLHFSIRFDEG IFSGKFVYNTFKDGNWSDNEQRISNPFKAGQEFDLRVRILDGKFQVFANRVEVG FEQRQPLDGDIDHVSIRGDLTKLRLFHYGGRIFFPNPYMAVAELKPGKRLDISALPTG KRVNVNLYRENKEYALQVSIRFNEGAIVRNAMNNVWGKEEREGDMPISKGEV FDLTIINEEFSQIFFNGKRFATFAHRGSPTDIKTLEIDGDVEIHTVTINDAPGI
		XP_013292547.1	W2SL40_NECAM	NA547	MVLAHPTDAPIPVPYTTKLGQPLLPQQLDIHGKINSASRVEINLLHGAAQIDPG QAVLHVNLRFDEKKIVMNTYMDGAWGKEERESMPFKQGEKFDLKMRLVDGMEI EIWCDDKKVHLYKHRMPFDQIEYLQIKGDCCTLSGVHWGGKFYQVPWETGFPDG

			HLRASQRIMLHAIPKGDWRNLDLLGRGGDILFHFNPFRKDKQIFVNNERIGTFQH RTENPTQDYIGMRIDGEIEVTNVEFV
XP_013292055.1	W2SJW5_NECAM	NA055	MPFKQGEKFDLKMVRVLVDGMEIWCDDKKVHLYKHRMPFDQIEYLQIKGDCTLS GVHWGGKFYQVPWETGFPDGHLLRASQRIMLHAIPKGDWRNLDLLGRGGDILFH FNPRFKDKQIVRNSYKGGIWNKEEREGPFPFEKERGFDLTIENEPYSMQVILFGF
XP_013305618.1	W2TQJ3_NECAM	NA618	MMDHIRFNINLHKDSPDFSGNDVPLHVSVRFDEGKLVFNSYTKGAWGKEERQK NPVKKGDGDIRIRAHDNKFTVSINRKEVKSFEHRIPLQHVTLSIDGDVVLNHV QWGGKYYPVPYESGIAENGLIPGKTLVIYGTPEKKAKK
XP_013293541.1	W2SNR3_NECAM	NA541	CDFQVPVYESGIAAGFQVDKTLIFGTVEKKAKRFNVNLLRRNGDIALHFNPRFD EKAVIRNALAANEWGNEEREGKMPFEKGVGFDLAIKNESYAFQVSVMTFSSRFS SRFQIFVNGERFTSFAHRSDP
XP_013302504.1	W2TH03_NECAM	NA504	MYTVNCPSPVPAIPICEPLHPGCFIDVYGNVENGNFVVELLSGPHIVLHVNRFRQL DEQEVVMNSSSNGVWGEEIDVNGQHLANYRHRFPMESVQAIGLKGDMCIERIEF DGFLFHSGWDYDTDYGHSGYKAYGTDYQAPVKFSDATLMGILFPSQISWKGQF RIGTKVA
XP_013308204.1	W2TY78_NECAM	NA204	MHQVLHPPVPSAIPHEQLRDGCKIDVHGRVEHGHKNFAIELLSGPHIVLHVNR FHHEHVVMNSCSHGAWGPEVRHHNPLHHSEHFHLHHVHSHFDIDVNGHKL GTFQHRFPVESVQALGLKGDVHVEKILFSGFHFQKNWNATVDFGHGGFIGYGT SYAPPAFSGGNQYNAYH
XP_013299437.1	W2T5A6_NECAM	NA437	MHIIEFPVIPCTTPIYNGFNHDAVVDIHGDLLQGPPQNGFVIEFPSREGIAFHMSVRM GLYGGENTIVLNNMRNGSFQIEERHYNVFLGAPFHVQIKSHSSHYSIHVNGNHL AHFHHREDPCHVTALTIRGDVRIQKIHFEHFIGMNGGGFQVGSSSAMVVAAPPPP PAVVIAPTPAPAVVIAPRPAPLVEVVLPPRPPIIEVVAPRPAPVIVVAPRRPIAPVV VINKGGHHHHNRHH
XP_013304533.1	W2TME8_NECAM	NA535	MFSGQPDQQLVINSYNNRWGAERFPNPFEIGQPFQIRILVLENYFKIAVNGKHIC DYPHRVPVEEVKTIYVGGNIRVDFIEFQPPVSVCVFVTESYVDVFYVSIVGSLGH P
XP_013304626.1	W2TL92_NECAM	NA626	LFKINLKNGEDIIHFNPRLKDEILVFNSFYKGEWQEEERASVIFPFERKKIYTFEF VASGNNTVYVYVNGQLLYEFRERQSGTNVASIYVGGDIDIHSVHVF
XP_013304627.1	W2TJZ4_NECAM	NA627	MKRRFLIRFRRRHFLSETGKKKLSAVIFTVVRMYRAYEYLEYVERPAPGPTA PPVSKFSNEVFNPNTNPVELPISGFGNGHRLRVVLKTLDKRNEQFSVNLKNGNDILL HFNPRLKDNVIVFNSFSNGEWQHEERPSIVFPFERKKIYTIEFVATTDSSALVNLKD GDDILLHFNPRLKDEVIVFNVFTDGEWQHEERPSIVFPFQKDQIYTIELVATTDNS ALIYVNGRFLYEFQRRESGQRASSVEVDGDVFIHSVHVT
XP_013294971.1	W2SSL8_NECAM	NA971	MNNNVWGKEEREGDMPISKGEVFDLTIINEEFSQIFNGKRFATFAHRGSPTDIK TLEIDGDVE
XP_013305619.1	W2TQT8_NECAM	NA619	MRPAFDTVICAAIPQRNVPMIFVNGERFASYAHRIDPHEIGGLQIHGDVELTGIQIV GN
XP_013304671.1	W2TLD5_NECAM	NA671	MAYRSLCHFSWLLPHSGYPLRRHLIRIFPRIFLISLAASIFCPFSEALDSRCSPRG VPLTFPVSSFAIGDRIRVLLPSFVTNHKFYIDLKSGSEVILRFEARTLMDEGDTTF SAFVVFNTRHKGAWESGEQSADRCFYIYQVYTIDFAPSGHYSIYVRVNGRTIHE FRERHLGFKVSSFEIAGDIELSLQCLRSTLFGIITIDILINTARIFSSRMTANSTNDN DNERSPVVEFATEVFRPNAPFEIPVTEFTHVRRIRAVLRCLATLFYRFYYINFKDG KDIVMQMSVKIMHKTIAFNSFYNGHWQIEETLSYPCPIEMAKTYTIDFVPTRFHSI FVYINGRFTYEFRRERQPGFKFSSVEITGDVEIHSVHVS
XP_013294790.1	W2SSR7_NECAM	NA790	MAQTFSIAPIFASLSSDSCHGASNRYLSESEWESARLLSVRHFSQFRRSDMAAPQR FPLPNMPFRREVHSGCAIGTSISLRAQAFQAKERSFTIQLSTAKDIAVLVTVPAK NGRMAVCARVGGKFSSEVDKPFLLPYEMKFTLHLRLAQFATEIYFNDNHMLDFV HRVNP AEIKEVHIEGPLIVEEVFTPPQGSSLDPVPSYEQATNMGSDPIKEFRHLEI GPPSSSLTPVPAPFPGSSTSPVLIPTPGPPQIGLSSSVVNQPIAGIPLSASTSTNPFVG SSSGTFPSPSPAGQLNSKSAIYEKNPQYVPTAFEQAQRYGTSSSTTTTPYYVTSPTP

				TPHTPAPATSSGSSAMPVLQPTSFQAQPLQPTPYSSSQALQIPISYGPTTNTPYP QVMPQPHTSNQPFPTAPGMNPYGYQQYPTVVPGPQGMMPMQPHYNPYQQQLPLQ YTVPSAPYVPGYAPQHYPGYPVVYVYSGCHVSFQMNFNSISPFGNFLLDWGYLIHL NLKLFISFPPEKTEAVAYSILPLRIGELIVGEQFLWFSFLSGKLYFVAQKSRRTVQ ENNAKQSSSDGVGSKDPHLNLQSSFTTSRQRHSSILDASGMNPAVEETNIDQTTI QLLIRIELIELLMQNLSKKAQPEPGAVDVSSARESYPEFVKQIQEQQLPRLAGIARSP EVTSALKNPRVLMALRNLIQLSMSTIQREAPMLAEAFRSRLAAQSGATLPKRDTAT SASATSTCISDPVLRREGDANVSRDITLCLQYEGELDKLAKMGFTDKKRNLTAI RASN
	XP_013304532.1	W2TJR0_NECAM	NA532	MSEVIPFSWTLPAELNGFCSPQSVRFTLTPFMSAKRFNCNLNLAGNEYLFHFRVDF RNASEKYSKVDVDGVHCVKFKYRPGDDPTLIDRVTVEGDCVLQRFVHRV

Table S2. NCBI BLASTp output for gal-3 query.

Species	Accession	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len
<i>N. americanus</i>	XP_013292547.1	galactoside-binding lectin [Necator americanus]	70.9	106	54%	6.00×10^{-15}	30.5	242
	XP_013308007.1	galactoside-binding lectin [Necator americanus]	66.6	66.6	38%	1.00×10^{-14}	34.74	109
	XP_013293541.1	galactoside-binding lectin [Necator americanus]	66.2	66.2	39%	4.00×10^{-14}	37.37	130
	XP_013304533.1	galactoside-binding lectin [Necator americanus]	60.5	60.5	30%	3.00×10^{-12}	39.24	114
	XP_013304627.1	galactoside-binding lectin [Necator americanus]	59.3	59.3	49%	8.00×10^{-11}	33.59	263
	XP_013299521.1	galactoside-binding lectin [Necator americanus]	54.7	54.7	55%	7.00×10^{-10}	27.59	145
	XP_013292055.1	galactoside-binding lectin [Necator americanus]	52.8	86.6	52%	6.00×10^{-9}	35.35	161
	XP_013308382.1	galactoside-binding lectin [Necator americanus]	48.5	93.2	52%	5.00×10^{-7}	28.7	328
	XP_013294971.1	galactoside-binding lectin [Necator americanus]	40.8	40.8	25%	1.00×10^{-5}	29.69	65
<i>T. trichiura</i>	CDW55930.1	Gal-bind lectin domain containing protein [Trichuris trichiura]	74.3	74.3	45%	5.00×10^{-16}	34.45	311
	CDW59317.1	galectin [Trichuris trichiura]	73.2	73.2	50%	1.00×10^{-15}	34.11	280
	CDW52167.1	galactoside binding lectin family protein [Trichuris trichiura]	64.3	126	41%	2.00×10^{-12}	36.9	308
	CDW53214.1	32 kDa beta galactoside binding lectin [Trichuris trichiura]	63.5	63.5	50%	2.00×10^{-12}	31.11	278
	CDW53139.1	galectin 9B [Trichuris trichiura]	53.9	53.9	28%	9.00×10^{-10}	37.14	136
	CDW55425.1	clathrin heavy chain [Trichuris trichiura]	34.3	34.3	25%	0.026	36.51	1685

Table S3. NCBI BLASTp output for gal-9 query.

Species	Accession	Description	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len
<i>N. americanus</i>	XP_013308007.1	galactoside-binding lectin [Necator americanus]	93.2	170	59%	8.00×10^{-24}	43.81	109
	XP_013293541.1	galactoside-binding lectin [Necator americanus]	87.8	156	58%	2.00×10^{-21}	44.44	130
	XP_013307933.1	galactoside-binding lectin [Necator americanus]	83.6	166	76%	1.00×10^{-19}	32.87	167
	XP_013292547.1	galactoside-binding lectin [Necator americanus]	76.6	173	85%	2.00×10^{-16}	31.08	242
	XP_013299521.1	galactoside-binding lectin [Necator americanus]	73.9	132	62%	3.00×10^{-16}	39.22	145
	XP_013305618.1	galactoside-binding lectin [Necator americanus]	67.8	130	58%	4.00×10^{-14}	34.58	146
	XP_013304533.1	galactoside-binding lectin [Necator americanus]	65.1	118	46%	2.00×10^{-13}	42.86	114
	XP_013308382.1	galactoside-binding lectin [Necator americanus]	68.9	68.9	38%	2.00×10^{-13}	31.88	328
	XP_013304534.1	hypothetical protein NECAME_02044 [Necator americanus]	55.1	99.4	32%	4.00×10^{-10}	44.26	99
	XP_013292055.1	galactoside-binding lectin [Necator americanus]	53.5	140	67%	6.00×10^{-9}	33.01	161
	XP_013304626.1	galactoside-binding lectin [Necator americanus]	51.2	99	53%	1.00×10^{-8}	31.37	103
	XP_013304627.1	galactoside-binding lectin [Necator americanus]	48.9	97.8	33%	5.00×10^{-7}	31.91	263
<i>T. trichiura</i>	CDW59317.1	galectin [Trichuris trichiura]	86.7	172	78%	8.00×10^{-20}	36.3	280
	CDW55930.1	Gal-bind lectin domain containing protein [Trichuris trichiura]	83.6	162	69%	1.00×10^{-18}	36.96	311
	CDW52167.1	galactoside binding lectin family protein [Trichuris trichiura]	80.5	219	78%	2.00×10^{-17}	29.61	308
	CDW53214.1	32 kDa beta galactoside binding lectin [Trichuris trichiura]	76.6	201	73%	3.00×10^{-16}	32.64	278
	CDW53139.1	galectin 9B [Trichuris trichiura]	64.3	109	49%	4.00×10^{-13}	35.11	136

Table S4. Full results of Pairwise alignment of protein BLAST results for NALOG1 query.

Sequence Name	p-Value
EYC23915.1_hypothetical_protein_Y032_0014g2212_Ancylostoma_ceylanicum	0.0246478873
WKX96556.1_hypothetical_protein_Q1695_012748_Nippostrongylus_brasiliensis	0.0405405405
EYC23914.1_hypothetical_protein_Y032_0014g2212_Ancylostoma_ceylanicum	0.0430107527
EYC23917.1_hypothetical_protein_Y032_0014g2212_Ancylostoma_ceylanicum	0.0454545455
KJH52076.1_galactoside-binding_lectin_Dictyocaulus_viviparus	0.0545454545
AAF63404.1_galectin_Haemonchus_contortus	0.0549828179
AEK98125.1_galectin_3_Angiostrongylus_cantonensis	0.0574324324
WKX96555.1_hypothetical_protein_Q1695_012748_Nippostrongylus_brasiliensis	0.0584192440
CAJ0607339.1_unnamed_protein_product_Cylicocyclus_nassatus	0.0658307210
EYC23913.1_hypothetical_protein_Y032_0014g2212_Ancylostoma_ceylanicum	0.0668896321
VDO66889.1_unnamed_protein_product_Heligosomoides_polygyrus	0.0709459459
KAE9421383.1_hypothetical_protein_Angca_007301_Angiostrongylus_cantonensis	0.0716845878
CDJ84887.1_Galectin_domain_containing_protein_Haemonchus_contortus	0.0718750000
VDM61903.1_unnamed_protein_product_Angiostrongylus_costaricensis	0.0813559322
EYC23916.1_hypothetical_protein_Y032_0014g2212_Ancylostoma_ceylanicum	0.0827586207
WKX96557.1_hypothetical_protein_Q1695_012748_Nippostrongylus_brasiliensis	0.0927152318
WKX96558.1_hypothetical_protein_Q1695_012748_Nippostrongylus_brasiliensis	0.0931034483
CAI4227323.1_unnamed_protein_product_Auanema_sp._JU1783	0.1326164875
CAI4227322.1_unnamed_protein_product_Auanema_sp._JU1783	0.1408934708
CAI4227324.1_unnamed_protein_product_Auanema_sp._JU1783	0.1408934708
PIO64995.1_galactoside-binding_lectin_partial_Teladorsagia_circumcincta	0.1418181818
EGT56354.1_CBN-LEC-3_protein_Caenorhabditis_brenneri	0.1549295775
ULU04255.1_hypothetical_protein_L3Y34_017208_Caenorhabditis_briggsae	0.1642857143
NP_001022533.1_Galectin_Caenorhabditis_elegans	0.1655405405
CAJ0919981.1_unnamed_protein_product_partial_Mesorhabditis_belari	0.1756756757
CAI5441466.1_unnamed_protein_product_Caenorhabditis_angaria	0.1758620690
ULU04256.1_hypothetical_protein_L3Y34_017208_Caenorhabditis_briggsae	0.1780821918
NP_496159.1_32_kDa_beta-galactoside-binding_lectin_lec-3_Caenorhabditis_elegans	0.1780821918
UMM16258.1_hypothetical_protein_L5515_013350_Caenorhabditis_briggsae	0.1780821918
VDM41798.1_unnamed_protein_product_Toxocara_canis	0.1911764706
CAJ0577880.1_unnamed_protein_product_partial_Mesorhabditis_spiculigera	0.2027027027
CAB3404474.1_unnamed_protein_product_Caenorhabditis_bovis	0.2175324675
NP_001263668.1_Galectin_Caenorhabditis_elegans	0.2606060606
CAI2333305.1_unnamed_protein_product_Caenorhabditis_sp._36_PRJEB53466	0.2628398792
EFO86405.1_CRE-LEC-3_protein_Caenorhabditis_remanei	0.2696969697
AAD11972.1_galectin_Haemonchus_contortus	0.2877697842
PIC44116.1_hypothetical_protein_B9Z55_004596_Caenorhabditis_nigoni	0.2890173410
AAC47546.1_galectin_Teladorsagia_circumcincta	0.2913669065
ULU04257.1_hypothetical_protein_L3Y34_017208_Caenorhabditis_briggsae	0.2919075145

ARM19984.1_galectin-9_Toxascaris_leonina	0.2949640288
AEK98124.1_galectin_2_Angiostrongylus_cantonensis	0.2949640288
AAC47547.1_galectin_Teladorsagia_circumcincta	0.2985611511
ULU04258.1_hypothetical_protein_L3Y34_017208_Caenorhabditis_briggsae	0.2988826816
ULU04259.1_hypothetical_protein_L3Y34_017208_Caenorhabditis_briggsae	0.2988826816
KHN77810.1_32_kDa_beta-galactoside-binding_lectin_lect-3_Toxocara_canis	0.2989323843
5GLT_A_Tl-gal_with_LNT_Toxascaris_leonina	0.3098591549
QHR84805.1_galectin_Dirofilaria_repens	0.3202846975
AAF37720.1_galectin_Dirofilaria_immitis	0.3238434164
AAF37721.1_galectin_Brugia_malayi	0.3250000000
CAB3404475.1_unnamed_protein_product_Caenorhabditis_bovis	0.3315789474
CAB3404476.1_unnamed_protein_product_Caenorhabditis_bovis	0.3487179487

No sequence from NCBI BLAST, returned a 100% similarity for NALOG1, but a large number of sequences returned >90% similarity. *p*-value in this case represents percentage difference in amino acid residues.

Table S5. Results of Pairwise alignment of protein BLAST results for NA521 query.

Sequence Name	<i>p</i> -Value
XP_013299521.1_galactoside-binding_lectin_Necator_americanus	0.0000000000
EYC36147.1_hypothetical_protein_Y032_0929g3083_Ancylostoma_ceypanicum	0.3169014085
EYC36146.1_hypothetical_protein_Y032_0929g3083_Ancylostoma_ceypanicum	0.3241379310
CAJ0598549.1_unnamed_protein_product_Cylicocyclus_nassatus	0.3623188406
WKX98653.1_hypothetical_protein_Q1695_013946_Nippostrongylus_brasiliensis	0.4166666667
WKX98652.1_hypothetical_protein_Q1695_013946_Nippostrongylus_brasiliensis	0.4166666667
VDL74252.1_unnamed_protein_product_Nippostrongylus_brasiliensis	0.4236111111
CAD6185709.1_unnamed_protein_product_Caenorhabditis_auriculariae	0.4444444444
EGT47829.1_hypothetical_protein_CAEBREN_05779_Caenorhabditis_brenneri	0.4513888889
CDJ91991.1_Galectin_domain_containing_protein_Haemonchus_contortus	0.4620689655
BAA09794.1_galectin_Caenorhabditis_elegans	0.4689655172
NP_497215.1_Galectin_Caenorhabditis_elegans	0.4689655172
3VV1_A_Crystal_Structure_of_Caenorhabditis_elegans_galectin_LEC-6_Caenorhabditis_elegans	0.4689655172
CAJ0580711.1_unnamed_protein_product_partial_Mesorhabditis_spiculigera	0.4791666667
PAV63311.1_hypothetical_protein_WR25_02416_Diploscapter_pachys	0.4861111111
CAI2349067.1_unnamed_protein_product_Caenorhabditis_sp._36_PRJEB53466	0.4861111111
CAI5444705.1_unnamed_protein_product_Caenorhabditis_angaria	0.4930555556
XP_002642976.1_Protein_CBR-LEC-6_Caenorhabditis_briggsae	0.5103448276
CAB71314.1_galectin_Haemonchus_contortus	0.5104895105

Sequence XP_013299521.1 is the source sequence of NA521 and therefore has 100% similarity, all other sequences are <70% similarity.

Table S6. Results of Pairwise alignment of protein BLAST results for TTZM7 query.

Sequence Name	p-Value
CDW53214.1_32_kDa_beta-galactoside-binding_lectin_Trichuris_trichiura	0.000000000
KHJ45589.1_galactoside-binding_lectin_Trichuris_suis	0.0074626866
KFD54272.1_hypothetical_protein_M513_04814_Trichuris_suis	0.0147601476
OUC47071.1_putative_galactoside-binding_lectin_Trichinella_nativa	0.2659176030
KRX86931.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_pseudospiralis	0.2686567164
KRX15368.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_nelsoni	0.2686567164
KRZ90043.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_sp._T8	0.2686567164
KRY89534.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_pseudospiralis	0.2686567164
KRX35770.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_murrelli	0.2686567164
KRZ09574.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_pseudospiralis	0.2686567164
KRZ90044.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_sp._T8	0.2686567164
KRY52250.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_britovi	0.2686567164
XP_003373758.1_32_kDa_beta-galactoside-binding_lectin_(Galectin-1)_partial_Trichinella_spiralis	0.2686567164
VDO94143.1_unnamed_protein_product_Soboliphyme_baturini	0.2723880597
NP_001022531.1_Galectin_Caenorhabditis_elegans	0.4606741573
VDM48758.1_unnamed_protein_product-Toxocara_canis	0.4756554307
EFO86405.1_CRE-LEC-3_protein_Caenorhabditis_remanei	0.4765342960
KAJ1358019.1_galactokinase_Parelaphostrongylus_tenuis	0.4831460674
KAE9553103.1_hypothetical_protein_FO519_003692_Halicephalobus_sp._NKZ332	0.4868913858
AAF63405.1_galectin_Haemonchus_contortus	0.4943820225
O44126.1_RecName: Full=32_kDa_beta-galactoside-binding_lectin_AltName: Full=Galectin-1_Haemonchus_contortus	0.4943820225

Sequence CDW53214.1 is the source sequence of TTZM7 and therefore has 100% similarity, the next two lowest *p*-value sequences are from *Trichuris suis* (pig whipworm).

Table S7. Results of Pairwise alignment of protein BLAST results for TTG03 query.

Sequence Name	p-Value
CDW59317.1_galectin_Trichuris_trichiura	0.000000000
KFD53595.1_hypothetical_protein_M513_05511_Trichuris_suis	0.0369003690
KRY91701.1_32_kDa_beta-galactoside-binding_lectin_lec-3_Trichinella_pseudospiralis	0.3955223881
KRY59236.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_britovi	0.4014598540
KRX78986.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_sp._T6	0.4051094891
KRZ57281.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_nativa	0.4064748201
KRX20929.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_nelsoni	0.4087591241
KRZ68313.1_32_kDa_beta-galactoside-binding_lectin_lec-3_partial_Trichinella_papuae	0.4100719424
KRZ13803.1_32_kDa_beta-galactoside-binding_lectin_lec-3_Trichinella_zimbabweensis	0.4100719424
XP_003381656.1_beta-galactoside-binding_lectin_Trichinella_spiralis	0.4100719424
KRZ57282.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_nativa	0.4100719424
KRY32824.1_32_kDa_beta-galactoside-binding_lectin_lec-3_partial_Trichinella_spiralis	0.4115523466
KRX96905.1_32_kDa_beta-galactoside-binding_lectin_partial_Trichinella_pseudospiralis	0.4136690647
KRZ44440.1_32_kDa_beta-galactoside-binding_lectin_lec-3_partial_Trichinella_pseudospiralis	0.4136690647
KRX78985.1_32_kDa_beta-galactoside-binding_lectin_Trichinella_sp._T6	0.4136690647
XP_003139211.1_beta-galactoside-binding_lectin_Loa_loa	0.5311355311
KHJ45589.1_galactoside-binding_lectin_Trichuris_suis	0.5395683453

Sequence CDW59317.1 is the source sequence of TTG03 and therefore has 100% similarity, the next lowest *p*-value sequences are from *T. suis* (pig whipworm). All other sequences have <60% similarity.

Supplementary material 8. Graphical breakdowns of in-silico filtering steps, for galectin-3 and galectin-9 homologs.

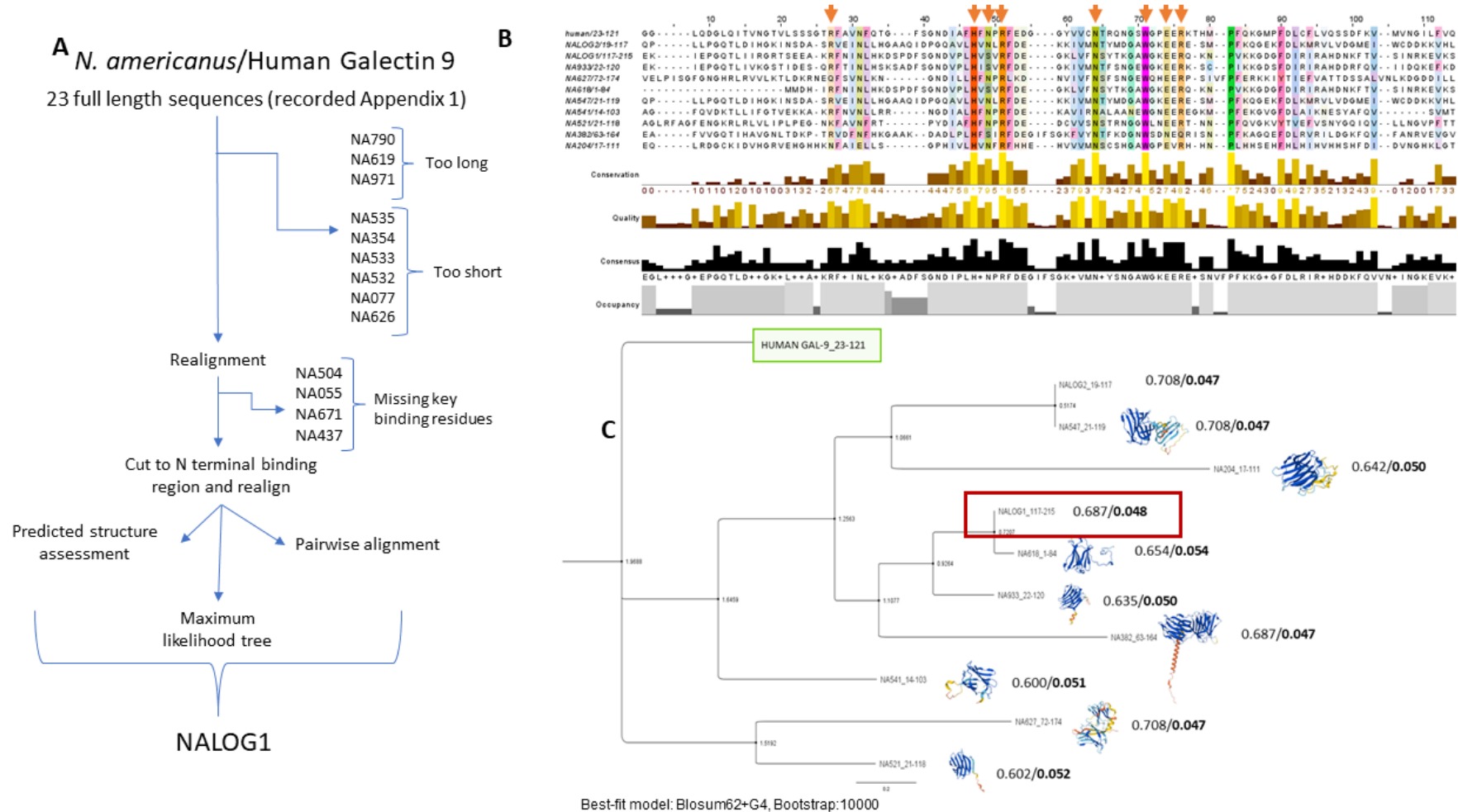


Figure S1. Overview of process for *N. americanus* galectin and galectin 9 comparison. Twenty-three full length sequences were recovered as listed in Tables S1–S3 and then imported into MEGA11. Sequences were removed based on length and absence of key binding residues (A). Alignments were imported into Jalview and cut to binding regions (B) before analysis in IQtree, comparison of the predicted structure (AlphaFold), and computation of *p*-distance in MEGA11 (*p*-distance/Standard error). Sequence NALOG1 was determined a good candidate for expression as a galectin-9 homologue (highlighted by red box in tree diagram) (C).

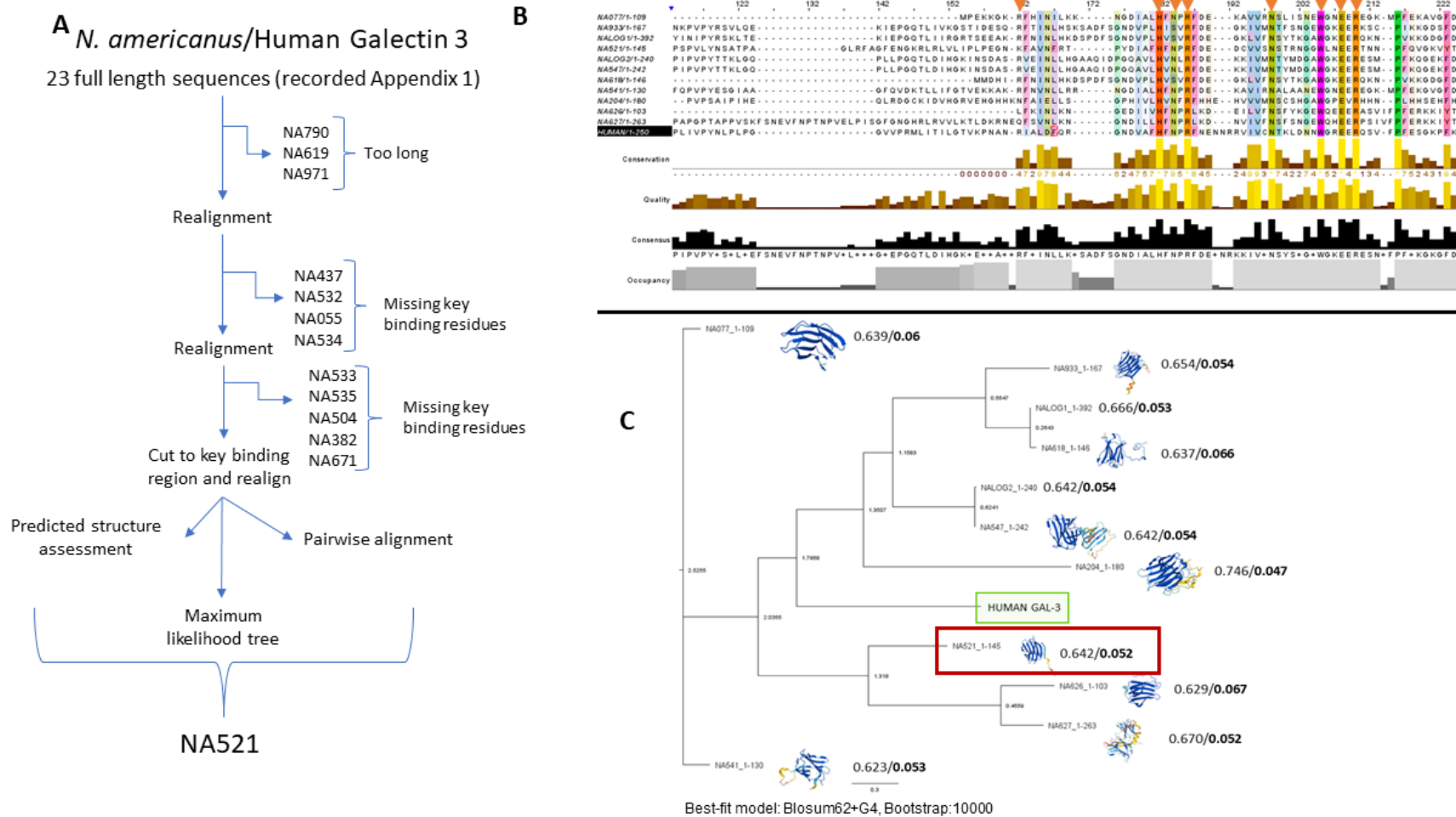


Figure S2. Overview of process for *N. americanus* galectin and galectin-3 comparison. Twenty-three full length sequences were recovered as listed in Tables S1–S3 and then imported into MEGA11. Sequences were removed based on length and absence of key binding residues (A). Alignments were imported into Jalview and cut to binding regions (B) before analysis in IQtree, comparison of the predicted structure (AlphaFold), and computation of p -distance in MEGA11 (p -distance/Standard error). Sequence NA521 was determined a good candidate for expression as a galectin-3 homologue (highlighted by red box in tree diagram) (C).

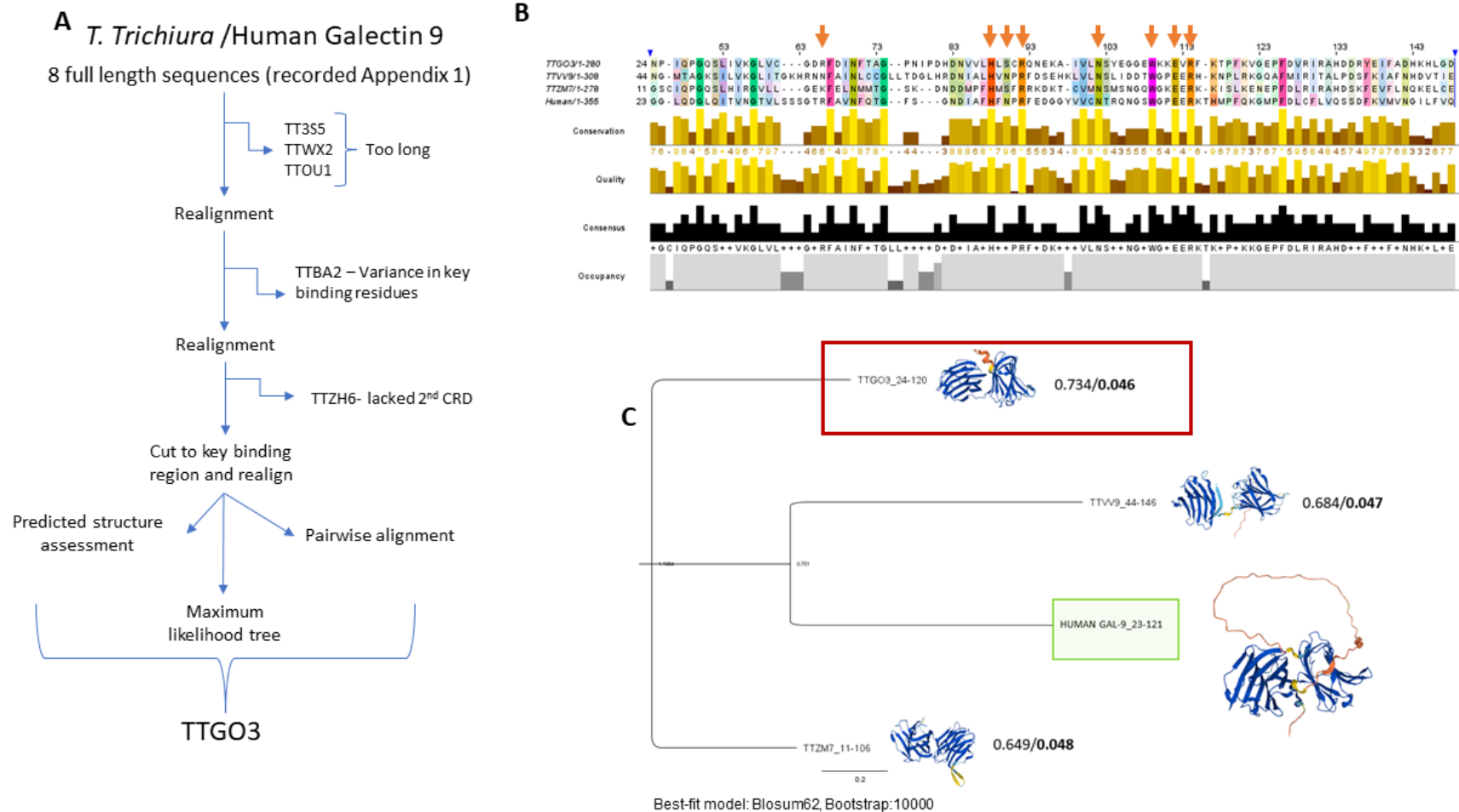


Figure S3. Overview of process for *T. trichiura* galectin and galectin-9 comparison. Eight full length sequences were recovered as listed in Tables S1–S3 and then imported into MEGA11. Sequences were removed based on length and absence of key binding residues (**A**). Alignments were imported into Jalview and cut to binding regions; (**B**) before analysis in IQtree, comparison of the predicted structure (AlphaFold), and computation of p -distance in MEGA11 (p -distance/Standard error). Sequence TTGO3 was determined a good candidate for expression as a galectin-9 homologue (highlighted by red box in tree diagram) (**C**).

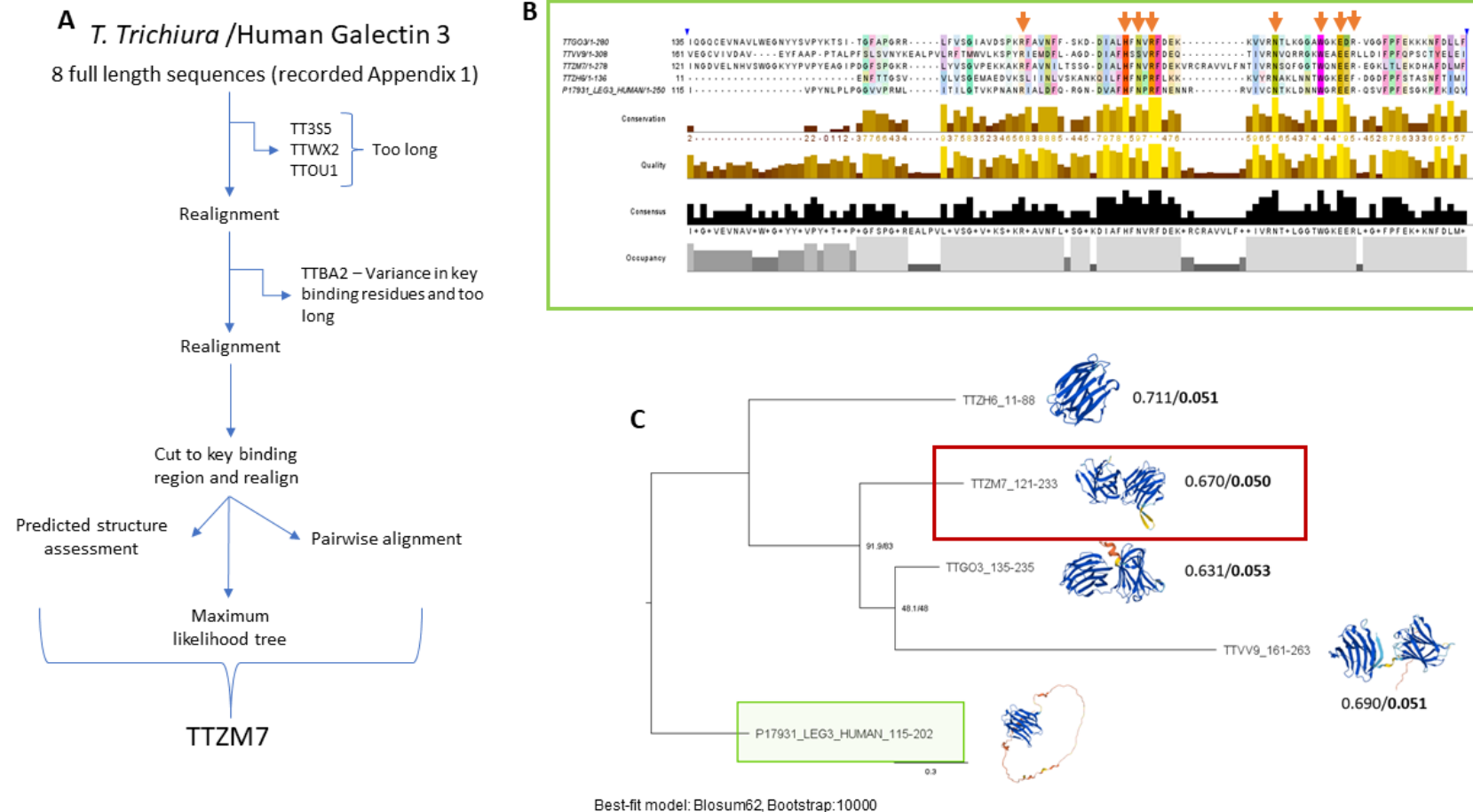


Figure S4. Summary of processing of *T. trichiura* sequences for galectin 3 homologue. Eight full length sequences were recovered as listed in Tables S1-S3 and then imported into MEGA11. Sequences were removed based on length and absence of key binding residues (A). Alignments were imported into Jalview and cut to binding regions (B) before analysis in IQtree, comparison of the predicted structure (AlphaFold), and computation of p -distance in MEGA11 (p -distance/Standard error). Sequence TTZM7 was determined a good candidate for expression as a galectin-3 homologue (highlighted by red box in tree diagram) (C).

