

Additional File 5. Repeat domains in MARTX-like genes of *X. innexi*

	XNC1_1381	XBJ1_1089	XIS1_650005	XIS1_640001	XIS1_640004	XIS1_650001	XIS1_630001	XIS1620012
A1	DTGGDLTVKGASGFTSINK	DTSGLTLTVKGGSGYTSINK	DTSGSLTVKGGSGYTAINK	DTSGNLSVKGAAGYAEINK		DTSGNLSVKGAAGYAEINK		
A2	TQSGPITFEGVSGGVKINHTG	TSDGTIDFAGAAGGLKINHT	THNGTIEFAGVAGGLNINHT	EKEGTISFAGAAGGAVIQHD		EKEGTISFAGAAGGAVIQHD		
A3	DKSGIFYSGAAGYNSITRK	GDKGRIDYSGASGYNGITRK	GDNGAVNYSVSGYNGITRK	GHSNGIHYSGVAAYNLYRR		GRSGNIHYSGAAAYNLYRR		
A4	GLQGDISFKGAGAYNELWHE	GWEGDFTFKGAGGYNELWHE						
A5	TNQGNSFAGAGGANKIDRTWFN	TNQGNFYFAGVGGGNKIDRTWFD						
A6	RYQGSHGDLTFKGAGAAIIISG	RYQGSHGDLTFNGAGAAIIISSR						
A7	VESGNVEFEAGAAANNIVRK	VESGDIVFAGAGAAANNLVRR						
A8	GKIGDVTLRGAGASNHVERIREADDVYS	GKTGNVTLTGAGASNRIERTRTDQDDVYS						
A9	ETKGNIRFEGVGGYNSLYSD	ETKGDIFEGAGGYNSLYSD	GVQGDIFRQAGGYNSLYSD	GHSGDITFKGAGGYNYLYSD		GHSGDITFKGAGGYNHLYSD		
A10	VAHGDIHFAGAGGYNQITRK	VAHGNIRFAGVGGYNHITRK	VEQGNIFHAGAGGYNRIIRK	VVNGNIHFVAGAGGYNQITRK		VVNGNIHFVAGAGGYNQITRK		
A11	ESSGDVHYNGAGGGNVIKSN	ESSGDVRYTGAGGGNVIQSN	ETSGDVRFEAGGGNVVKSND	DSNGDVRFEAGGGNVIKSN		DSSGDVRFEAGGGNVIKSN		
A12	VTRGNVYFNGAGIANIIHS	VTRGNVYFDGAGIANVIEHT	VTGNVYFSGAGIANVIEHT	VARGNVYFKGAGIANVILHN		VARGNVYFKGAGIANVILHN		
A13	SDFGNTEFNGAGGANVIKK	SEFGNTEFNGGAANVIVKK	SEFGNTEFNGAGGANVIKK	SEIGDTEFSGAGGANVVFCK		SEIGDTEFSGAGGANVVFCK		
A14	GKEGNLSFNGAGIANVLLHQ	GKEGNLTFNGVGIANVLLHQ	GKEGNLTFNGAGVANVLLHQ	GRKGHLNFRGAGIANVITHI		GRKGHLNFRGAGIANVITHI		
B1	SQRGDMNINAGGAANVLVR	GLRGDMVDVNAGGAANILVR	GKRGDMEVNAGGAANVLVR					
B2	VGDGRYLAHLLAVGNISIH	VGDGRYLAHLLAVGNISIH	VGDGRYLAHLLAVGNISIH					
B3	KGNGNSRISMGGGFNTHQTQ	KGNGNSRISMGGGFNTHQTQ	KGNGDSRISMGGGFNTHQTQ					
B4	IGNGDADFWTGVGGANVLTLQ	IGHGDADFWSGVGGFNVLTLQ	MHGDAVWWSGVGGANVLTS					
B5	TGNGDVSILLGGANVLTK	MKGHGVSSILLGGANVLTK	IGQGKVSALLGGANILTK					
B6	MGGGHLESAMFGGANIITHISND	VGEGLDESGLGGANIITHISND	MKGDLAAGLLGGANIITHISDS	SSTHQVRNPSAD		SSTHQVRNPSAD		
B7	NNQGTSDTKAFNTKAIALGGANVL	RETSNTKAIALGGANILTK	SDDTASDKTIRDKTIHYKTTAIALGGANVLTK	SRSGNTHVIALGGANVLTRQ		SQSGNTHVIALGGANVLTRQ		
B8	TKKKGKGDVQAVMGGGVNVL	KGKGDVLSVMGGGANVLTH	KGKGDLLSVMGGGANVLTH	QGNRVRFGVMGGGANVLTH		GNGRVRFGVMGGGANVLTH		
B9	THVGDGKTTGVMLGGANIL	VGDGKTTGILLGGANILTK	IGNGATSGVLLGGANILTK	IGNGVTTGVMLGGANVLTK		IGNGVTTGVMLGGANVLTKV		
B10	TKVNGDGTGIMFGLGNVL	VNGDGTGIMLGLGNILTH	VGDGDTTGIMLGLGNVLTH	VGDGETTGILFGLGNVLTH		GDGETTGILFGLGNVLTH		
B11	THVNGQTLGVMMVSAGNIF	VNGQTLGVMAAGNIFTK	VGDGQTLGVMAAGNIFTK	VNGGSTLGAMGAVGNVFTK		VNGGSTLGAMGAVGNVFTK		
B12	TKVGNDDTTIAMIGAGNIF	VGEGTTIAMVAGNIFTH	VGDGETLAAMLAFGNLFTH	VGDGETVAALLAVGNIFTH		VGDGETVAALLAVGNIFTH		
B13	THVGEENAWALMGGAGNIF	VQQGNAWALMGGGNIPTK	VGDGNTYALMAGGNIFTK	IGQGDAYALMAGGGNVFTK		IGQGDAYALMAGGGNVFTK		
B14	TKVNGDALALMLAFGNVF	VGDGDALALMLAAGNVFTH	VGDGNALALMLAFGNVFTH	VGDGNALALMVALGNVFTH		VGDGNALALMVALGNVFTH		
B15	THVGDGMSVALMIAKGNIA	VGDGMSVALMVAKGNIATK	VGDGLSVALMIAKGNIAATK	VGDGLSVALMVAKGNLVTK		VGDGLSVALMVAKGNLVTK		
B16	TKVNGDVLVSAMIGKGNIF	VGHGEMLSAMIGGNILTK	AGNGDALAAMVGKGNIFTQ	VNGDGLAAMVGMGNLMTH		VNGDGLAAMVGMGNLMTH		
B17	TQIGHGSTFAAMIGGANVL	IGNGDTFAAMLGKANILTK	IGHGSTFAAMIGGANILTK					
B18	TKVGDDLTAAALMIGGANIY	VGNGLTAALMVSEANIYTH	VGNDLTAALMIGKANIYTH					
B19	THVGKTSIGLFGGSANIM	VGDGTSIGLFGGSVNMVTK	VGEGETSIGLFGGTANIMTK					
B20	TKVGDDTLAAMFGKANIM	VNGGTTLAAMFGKANIMTH	VGDGTTLAAMFGKANIMTH					
B21	THVGDDTLGVLALGKANIV	VGDGLTGVLALGKANIVTK	VNGGMTGVLALGKANIVTK					
B22	TKVGDDFMGVVAASEANIVTH	VGNDFMGVVAASEANIVTH	VNGFMGVVAASEANIVTH					
B23	VGDGTTAALLSGKNILTK	VGGGTTAALLSGKNILTK	IGEGTTAALLSKGNILTK					
B24	VGDGTTVGLLKSEIGNIMTH	VGDGTTVGLLVSKIGNIMTH	VGDGTAVGLLISKIGNIMTH					
B25	LGDGTTVGFAGKGANIITK	LGDGTTIGFAGKEANIITK	IGDGATVGFAGKEANIITK				VGLLVSELGNSLTH	
B26	VGDGLGINAAWGKANIMTH	SGDGLGINAAWGKANIMTH	VGDGTGVNAVWGKANILTH				VGDGLAVNAAWGKANILTH	
B27	VGDGDRYNFAKGEANIITK	IGEGDRYNFAKGEANIITK	YGDGDRYNFAKGEANIITK				VGHGDHYNFAKGGQANGVSK	
B28	IGDGGQEVTVVQQGANIVTH	VGDGGQEVTVVQQGANIVTH	VGDGGQEVTVVQQGANIITH				IGDGGQMVSMQGT	

							NIITH	
B29	VGNDDYTGAWGKANVITK	VGNDDYTGAWGKANVITK	IGNDDYTGTWGGKANVITK				VQGDDYTGVWVK ANVVTK	
B30	VGDGRNVVLAKGEANIVTQ	VGDGRNVVLAKGDANIVTQ	VGNGRNVVLAKGDANIVTQ				VGRGRQVTLAKGDA NLVTQ	
B31	IGNGDSFNLWSQGNIVTK	IGNGDSFNLWSQGNVVTK	VEGDSFNLWSQGNVVTK				VEGDSLNLWSQS NVVTK	
B32	VGDGIQVTAAGGEGNITTT	VGDGMQVTAAGKGNITTT	VGDGIQATVAKGKNITTT				VGDGMQVTAAGKGS NITTT	
B33	AGDGLNVTAVHGLNINTK	VGNGLSVTTVHGLNVNTK	VGNGLSVTAVHGLNVNTK				VGHGLSVVA	
B34	VGDGVSVNVAWGKLVNTR	VGDGVSVNVAWGKLVNTR	VGDGVSVNVAWGKYNVNTK					
B35	VGDGLNVSMKGKGNANIR	VGDGLNVSMKGQANANIR	VGDGLNVSMKGTGNANIR					
B36	VGDGLNINASYARNNVAIQ	VGDGLNINASYARNNVAIQ	VGDGLNINASYARNNVAIQ					
B37	VNGDFYSLFAVESNTESNK LD	VNGDFYSLAVAESNTESNKL S	VNGDFYSLAVAASNTESNKL G					
B38	ALFANIKQTLGSGSQGINYL VN	ALFGNVKQTLGSGSQGISH LVN	ALFDNVKQTLGSGSQAISYLVN					
DUF1	315 AA (1998-2312)	314 AA (1972-2285)	312 AA (1706-2017)		337 AA (542-878)			337 AA (548-884)
RID	639 AA (2313-2951)	637 AA (2286-2922)	646 AA (2018-2663)		631 AA (879-1509)			637 AA (885-1521)
ABH		201 AA (2923-3123)			199 AA (1510-1708)			202 AA (1521-1722)
PMT	460 AA (2952-3412)							
MCF	445 AA (3413-3858)	486 AA (3124-3609)			467 AA (1709-2175)			468 AA (1723-2190)
CPD	191 AA (3859-4050)	192 AA (3606-3797)	202 AA (2664-2865)		187 AA (2176-2362)			190 AA (2191-2380)
B39	HTAVDAKAQGAIAENVETF	HIDTDAKAKGAIAENAETF	RTGAETETKGAIAENTATF		RRNVNTAATGAIAENN EVF			SLRSSTNSNKSRRKHPV HGVNTDKNAIADNTEVF
B40	VGDGEFTAVNWGTTNVGIK	AGDGEFTVINWGTNNVGIK	VGDGEFTAVNWGTTNVGIK		IGDGEFTSINWGTNNV SIK			AGDGEFTSINWGTNNV GVK
B41	MGTGGFKSLAFGDNNVMVH	VGTGGFKSLAFGDNNVMVH	VGTGGFKSLAFGDNNVMVH		LGTGGFKSLVFGDNNV MVH			VGTGGFKSLAFGDNNV MVH
C1	SYDSPDWLTTQNNQ	SFESPDWLTAQDQ	GLSGSADWLTAQEQQWLE		SFSSPGWLAVQEEQW TIA			SYASPDWLSEQEAQ
C2	GGQGADIQVTTGNWNFMF	GGQGADIQVTTGNWNFMF	GGQGADIQVTTGNWNFMF		GGQGADIQVTTGSWN FMF			GGQGADIQVTTGNWNL MF
C3	GHLKGNGDINVSLGNYNFNW G	GHLQGDGDMHVS LGNYNFN W	GHLQGDGDIHVS LGNYNFNW		GHLQGDSDIHVS LGNY NFW			GHLQGDGDIHVS LGNY NFW
C4	GGDKDLGAYLGNNNFWG	GGDKDLGAYLGNNNFW	GGDKDLGAYLGNNNFW		GGDKDLGAYLGNN NFW			GGDKDLGAYLGNNN FW
C5	GRGDDVFYSVGTSNIFTG	GGRGDDVYYSVGTSNIFT	GGRGDDVFYSVGTSNIFA		GGRGDDVFYSLGT SNI FT			GGRGDDVYALGT SNIFT
C6	GAGNDTGVMGRENMMFG	GGEENDMGVLMGRENMMF	GGEENDIGVLMGRENMMF		GGEENDTGVL MGRENMMF			GGEENDTGVL MGRENI MF
C7	GAGNDVAVLAGRINYAYM	GGAGNDVAVLAGRINYAY	GGAGDDIAGRINYAY		GGAGDDVAVLAGRIN YAY			GGAGNDVAVLAGRIN YAYM
C8	GEGDDQVFAFEGEGMIDA	MGDGDQVVFEGEGVIE	MGDGDQVFAFEGEGMIE		MGKNDQVFAFEGEG VIE			GDGDQVFAFEGEGV IEA
C9	GKGRDYIVASGNFNRIEA	GGKGRDYIVASGNFN SIH	GGKGSYIVASGNFN RVS		AGEGRDYIVASGN FNRIE			GKGRDYIVASGN FNRIE
C10	GEDQDYVVTIGNNNQVDL	AGEGQDYVVAIGNNNQVS	AGEDQDYVVTIGNNNQVD		SGKDQDYVVAIGNNN QVG			GGEDQDYVVAIGNNNQ IDL
C11	GEGNDFARVFGNYNRIDG	LGEGDDFATVFGN YNNID	LSGDDFATVFGN YNRID		LAEGNDFATVFGND NRIE			GDDHDFATVFGN YNRID
C12	NTGNNSIKLMGYHAMING	GNAGNNAIKLMGYH AVIN	GNTGRNAIKLMGYH AVIN		GNEGNDSIKLMGYH ALING			GGAGNDVAVKLMGYH ALIN
C13	GTGNDHLIADVVSQFSQFDG	GGTGDDHLIANVASKFSQ LN	GGTGNDHLIADVVSQFSR FN		GEGNDHLIADVVSQFS QLN			GGAGDDHLIADVVSQFS QFD
C14	GDGDDLLVLGGYQNSFK	GGDGDLLVLGGYQNNFK	GGEENDLLVLGGYQNRFK		GGEENDLLVLGGYQ NRF			GGEENDLLVLGGYQNR FE
C15	GGTGVNSYVSGDIVNVVE D	GGTGVNSFVSGDIVNVVE D	GGSGVNSFVSGDIVNV VE		GGKGVNSFVSGDIV NIVED			GGAGVNSYVSGNVID NVVED