

Supplementary Material Table 1. AtbHLH proteins predicted to have DNA-binding capacity.

a. Putative E-box binders

Amino acid sequences of the bHLH domains of the proteins predicted to have E-box DNA binding capacity, based on the conservation of residues in positions E13 and R16 (highlighted in bold). Entry Number 139 lacks R16, but has a basic residue in this position and was kept in this category (See Functional Categories Section). Proteins are listed according to the Entry Number assigned and their AtbHLH Number and PID Number are also provided. Positions within the bHLH domain are numbered at the top of the figure, below the schematic that indicates the basic, helix and loop regions of the domain.

EN	AtbHLH Number	PID Number	BASIC		HELIX			LOOP		HELIX	
			10	E R 2	0	3	0	4	0	5	0
1	55	AAF79643	EPKNKRAKHK	ELERQRRQEN	TSLFKILRYL	LPSQYIKGKR	SSADHVLEAV	NYIKDLQKKI	K		
2	125	AAF75809	DRESKMKKHR	DIERQRRQEV	SSLFKRLRTL	LPFQYIQGKR	STSDHIVQAV	NYIKDLQIKI	K		
3	126	NM118673	NKKKKKLLHR	DIERQRRQEM	ATLFATLRTH	LPLKYIKGKR	AVSDHVNQAV	NFIKDTQTRK	T		
4	120	NP199992	TKKEKKLLHR	NIERQRRQEM	AILFASLRSQ	LPLKYIKGKR	AMSDHVNQAV	SFIKDTQTRI	K		
5	118	NM118672	...MEKLVHK	EIEKRRRQEM	ASLYASLRSL	LPLEFIQGKR	STSDQVKGAV	NYIDYLQRNI	K		
6	36	NM124557	...MEKMMHR	ETERQRRQEM	ASLYASLRSL	LPLHFIKGKR	STSDQVNEAV	NYIKYLQRKI	K		
7	100	AAC78547	PVVMKKLNHN	ASERERRKKI	NTMFSLLRSC	LPPNTQTKKL	SVSATVSQL	KYIPELQEQV	K		
8	38	CAB72167	PVVVKKLNHN	ASERDRRKKI	NTLFSLLRSC	LPASDQSKKL	SIPETVSKSL	KYIPELQEQV	K		
9	39	CAB72168	PVVVKKLNHN	ASERDRRKKI	NSLFSLLRSC	LPASGQSKKL	SIPATVSRSL	KYIPELQEQV	K		
10	101	NM120497	VVLEKKNLHN	ASERDRRKKI	NALYSSLRAL	LPLSDQKRL	SIPMTVARVV	KYIPEQKQEL	Q		
11	67	CAB71902	EIENQRMTHI	AVERNRRRQEM	NEHINSLRAL	LPPSYIQ.RG	DQASIVGGAI	NYVKVLEQII	Q		
12	57	CAB77716	EVENQRMTHI	AVERNRRRQEM	NEHLNSLRSL	MPPSFLQ.RG	DQASIVGGAI	DFIKELEQLL	Q		
13	70	AAC33499	EIESQRMTHI	AVERNRRRQEM	NVHLNSLRSL	IPSSYIQ.RG	DQASIVGGAI	DFVKILEQQL	Q		
14	97	BAB01355	EVEQRMTHI	AVERNRRRQEM	NEHLRVLRSL	MPSYVQ.RG	DQASIVGGAI	EFVRELEQLL	Q		
15	96	AAG51804	EIENQRMTHI	AVERNRRRQEM	NEYLAVLRSL	MPPYAAQ.RG	DQASIVGGAI	NYLKELEHHL	Q		
16	94	NM102098	EIENQRMTHI	AVERNRRRQEM	NEYLAVLRSL	MPSYAAQ.RG	DQASIVGGAI	NYVKELEHLL	Q		
17	71	NM124039	EAENQRMTHI	AVERNRRRQEM	NQHLVLRSL	MPQPFVH.KG	DQASIVGGAI	DFIKELEHKL	L		
18	99	BAB11554	DKENQRMNHI	AVERNRRRQEM	NHFLSILKSM	MPLSYSQ.PN	DQASIVGGAI	SYLKKLEQRL	Q		
19	98	BAB09783	QDQQQKMSHV	AVERNRRRQEM	NEHLTVLRSL	MPCFYVK.RG	DQASIVGGAI	EYISELQQLV	Q		
20	45	AAF30305MSHI	AVERNRRRQEM	NEHLKSLRSL	TPCFYIK.RG	DQASIVGGAI	EFIKELQQLV	Q		
21	95	AAG13058	EESPDHEIHI	WTERERRRQEM	RDMFSLKHAL	LPQ.LPP.KA	DKSTIVDEAV	SSIKSLEQTL	Q		
22	92	BAB11628	PEKERSRRHM	LKERTTREKQ	KQSYLALHSL	LPFA...TKN	DKNSIVEKAV	DEAIAKLQRL	Q		
23	10	AAD20667	GRGSRKSRTS	PTERERRVHF	NDRFFDLKNL	IPNP...TKI	DRASIVGEAI	DYIKELLRTI	E		
24	89	AAF80214	GRGSKRKRIF	PTERERRVHF	KDRFGDLKNL	IPNP...TKN	DRASIVGEAI	DYIKELLRTI	E		
25	91	AAD20666	GRGKRKNKPF	PTERERRRCHL	NEREVALKLL	IPSP...SKG	DRASILQDGI	DYINELRRRV	S		
26	19	AAC63587	RSPVLAKEHV	LAERKRREKL	SEKFIALSAL	LPGL...KKA	DKVTILDDAI	SRMKQLQEQV	R		
27	20	AAC63588	REPHLLKEHV	LAERKRREKL	NERLIALSAL	LPGL...KKT	DKATVLEDAI	KHLKQLQERV	K		
28	18	AAC63586	SSQNSAQDHI	LAERKRREKL	TQRFVALSAL	IPGL...KKM	DKASVLDGAI	KHIKYLQESV	K		
29	25	CAB38933	RNQSNAQDHI	LAERKRREKL	TQRFVALSAL	VPGL...KKM	DKASVLDGAI	KHIKYLQERV	Q		
30	2	AAB72192	HTRGKPGNHA	LSERKRREKL	NERFMTLRSL	IPSI...SKI	DKVSILDDTI	EYLQDLQKRV	G		
31	1	BAB08530	EARDETGSHA	VLEKKRREKL	NERFMTLRSL	IPSI...NKI	DKVSILDDTI	EYLQELERRV	Q		
32	42	CAC14865	RLPREDLSHV	VAERRRREKL	NEKFITLRSM	VPFV...TKM	DKVSILGDTI	AYVNHLRKRV	H		
33	14	AAB62853	KHHPAVLSHV	EAERQRREKL	NHRFYALRAI	VPKV...SRM	DKASLLSDAV	SYIESLKSKI	R		
34	3	CAB78685	NGREEALNHV	EAERQRREKL	NQRFYALRAV	VPNI...SKM	DKASLLGDAI	TYITDMQKDI	O		
35	17	AAD20162	NGREEPLNHV	EAERQRREKL	NQRFYALRSV	VPNI...SKM	DKASLLGDAI	SYIKELQEKV	K		
36	5	BAB08920	NGREEPLNHV	EAERQRREKL	NQRFYALRSV	VPNV...SKM	DKASLLGDAI	SYINELKSKL	Q		
37	4	CAA17131	NGREEPLNHV	EAERQRREKL	NQRFYALRAV	VPNV...SKM	DKASLLGDAI	SYISELKSCL	Q		
38	6	BAA25078	NGREEPLNHV	EAERQRREKL	NQRFYALRAV	VPNV...SKM	DKASLLGDAI	AYINELKSKV	V		
39	13	AAF97322	NGRAEALNHV	EAERQRREKL	NQRFYALRSV	VPNI...SKM	DKASLLGDAV	SYINELHAKL	K		
40	28	BAA97217	HGRDKPLNHV	EAERMREKL	NHRFYALRAV	VPNV...SKM	DKTSLLEDAV	CYINELKSKA	E		
41	35	BAA97365	AASSPASKNI	VSERNRRQKL	NQRLFALRSV	VPNI...TKM	DKASIVKDAI	SYIEGLQYEE	K		
42	27	CAB43668	ATSPASSKNV	VSERNRRQKL	NQRLFALRSV	VPNI...SKL	DKASIVKDSI	DYMQLIDQE	K		
43	29	AAC98450	KTKTDRSRTL	VSERRRRGRM	KDKLYALRSL	VPNI...TKM	DKASIVGDAV	LYVQELQSQA	K		
44	33	AAF78492	KKKGMPAKNL	MAERRRRKKL	NDRLYMLRSV	VPKI...SKM	DRASILGDAI	DYKELLQRI	N		
45	116	NM113586	KKKGMPAKNL	MAERRRRKKL	NDRLYMLRSV	VPKI...SKM	DRASILGDAI	DYKELLQRI	N		
46	61	CAB89386	KLEGQPSKNL	MAERRRRRKL	NDRLSLLRSL	VPKI...TKM	DRASILGDAI	DYMKELLDKI	N		
47	93	NM125962	KLEGQPSKNL	MAERRRRRKL	NDRLSLLRSL	VPKI...SKM	DRASILGDAI	DYMKELLDKI	N		
48	21	AAC64222	SGKGSQAKNL	MAERRRRRKL	NDRLYALRSL	VPRI...TKL	DRASILGDAI	NYVKELQNEA	K		
49	22	CAB79132	EDENFKSPNL	EAERRRREKL	HCRLMALRSH	VPIV...TNM	TKASIVEDAI	TYIGELQNNV	K		
50	90	AAD39586	PTENFKSKNL	HSERKRREKL	NQAMYGLRAV	VPKI...TKL	NKIGIFSDAV	DYINELLVEK	Q		
51	41	BAA97026	GPSATQLQHM	ISEKRREKL	NSFQALRSL	LPPG...TKK	DKASVLSIAR	EQLSSLQGEI	S		
52	134	CAB80472	SKEVAACKHS	DAERRRRLRI	NESFATLRTI	LPNL...VKQ	DKASVLTGETV	RYFNELKMMV	Q		
53	30	AAF07352	KALAASKSHS	EAERRRRERI	NNHLAKLRSL	LPNT...TKT	DKASLLAEVI	QHVKELKRET	S		
54	32	BAA95758	KALAASKSHS	EAERRRRERI	NTHLAKLRSL	LPNT...TKT	DKASLLAEVI	QHMKELKRQT	S		

EN	AtbHLH Number	PID Number	BASIC		HELIX			LOOP	HELIX				
			1	0	E	R	2	0	3	0	4	0	5
55	107	CAC00740	KALASLRNHK	EAERKRRARI	NSHLNKLRLK	LSCN...	SKT	DKSTLLAKVV	QRVKELKQQT	L			
56	106	AAD11998	RALAALRNHK	EAERRRRERI	NSHLNKLNRV	LSCN...	SKT	DKATLLAKVV	QRVRELKQQT	L			
57	51	AAD25935	KAESLSRSHR	LAEKRRRDRI	NSHLTALRLK	VPNS...	DKL	DKAALLATVI	EQVKELKQKA	A			
62	103	CAA20199	IDYIKFLQEQ	ITEKVRKEKL	GDRITALQQL	VSPF...	GKT	DTASVLHDAI	DYIKFLQEQI	D			
69	130	AAB88652	KRGCATHPRS	IAEVRRTTRI	SERMRKLQEL	VPNM...	DKQT	NTSDMLDLAV	DYIKDLQRQY	K			
70	122	AAG50543	KRGCATHPRS	IAEVRRTKI	SERMRKLQDL	VPNM...	DTQT	NTADMLDLAV	QYIKDLQEQV	K			
71	80	AAG12608	KRGCATHPRS	IAEVRRTRI	SDRIRRLQEL	VPNM...	DKQT	NTADMLEEAV	EYVKALQSQI	Q			
72	81	CAB78042	KRGCATHPRS	IAEVRRTRI	SDRIRKLQEL	VPNM...	DKQT	NTADMLEEAV	EYVKVLQRQI	Q			
73	129	AAC64303	KRGFATHPRS	IAERERTRI	SGKLLKLQEL	VPNM...	DKQT	SYADMLDLAV	EHIKGLQHQV	E			
74	128	AAF29386	KRGCATHPRS	IAERERTRI	SGKLLKLQDL	VPNM...	DKVT	SYSMDLDLAV	QHIKGLQHQI	Q			
75	136	AAG28811	KRGQATDSHS	IAEVRREKI	NERLKCLQDL	VPGC...	YKAM	GMAVMLDVLQ	NQIEFLSMKL	S			
76	50	AAF24852	RRGQATDSHS	IAEVRRGKI	NERLKCLQDI	VPGC...	YKTM	GMATMLDEII	NYVQSLQNV	E			
77	44	AAF25996	RRGQATDSHS	IAEVRRGKI	NERLRCLQDM	VPGC...	YKAM	GMATMLDEII	NYVQSLQNV	E			
78	75	NP564229	KRGQATDSHS	IAEVRREKI	NERLKCLQDL	VPGC...	YKAM	GMAVMLDVI	DYVRSLQNI	E			
79	64	AAD15506	RRGQATDRHS	LAERARREKI	SKMKYLQDI	VPGC...	NKVT	GKAGMLDEII	NYVQCLQRQV	E			
80	58	CAB80320	RRGEATDRHS	LAERARREKI	SKMKCLQDI	VPGC...	NKVT	GKAGMLDEII	NYVQSLQQV	E			
81	79	BAA97208	RRGQATDRHS	LAERARREKI	SEKMTALQDI	IPGC...	NKII	GKALVLDEII	NYIQSLQRQV	E			
82	49	AAF07355	RRGQATNSHS	IAEVRREKI	SERMKFLQDL	VPGC...	NKVT	GKAVMLDEII	NYVQSLQRQV	E			
83	76	AAG29214	RRGQATNSHS	IAEVRREKI	SERMKFLQDL	VPGC...	DKVT	GKAVMLDEII	NYVQSLQCCI	E			
84	63	CAA18832	RRGQATDSHS	IAEVRREKI	SERMKFLQDI	VPGC...	DKIT	GKALMLDEII	NYVQSLQRV	E			
85	62	AAF02164	RRGQATDSHS	IAEVRREKI	SERMKLLQDL	VPGC...	NKVT	GKALMLDEII	NYVQSLQRV	E			
86	78	BAB10689	RRGQATDSHS	IAEVRREKI	GERMKLLQDL	VPGC...	NKVT	GKALMLDEII	NYVQSLQRV	E			
87	77	BAB01846	RRGQATDSHS	IAERARREKI	SERMTLLQDL	VPGC...	NRIT	GKAVMLDEII	NYVQSLQRV	E			
88	31	BAA87957	RRGQATDSHS	LAERARREKI	SERMKILQDL	VPGC...	NKVI	GKALVLDEII	NYIQSLQRQV	E			
89	137	AF428350	RRGQATDSHS	IAEVRREKI	SERMRTLQNL	VPGC...	DKVT	GKALMLDEII	NYVQTLQRV	E			
90	74	AAC34336	RRGQATNSHS	IAEVRREKI	SERMRLQEL	VPGC...	NKVV	IGAVMLDEII	NYVQSLQQV	E			
91	60	CAB67608	RRGQATDSHS	IAERARREKI	NARMKLLQEL	VPGC...	DKIQ	GTALVLDEII	NHVQSLQRV	E			
92	7	AAD25805	RRGQATDPHS	IAERLRRERI	AERIRSLQEL	VPTV...	NK.T	DRAAMIDEIV	DYVKFLRLQV	K			
93	59	CAB80752	RRGQATDPHS	IAERLRRERI	AERIRSLQEL	VPTV...	NK.T	DRAAMIDEIV	DYVKFLRLQV	K			
94	69	CAA18195	RRGQATDPHS	IAERLRRERI	AERMKSLQEL	VPNG...	NK.T	DKASMLDEII	DYVKFLQLQV	K			
95	66	AAD03387	RRGQATDPHS	IAERLRRERI	AERMKALQEL	VPNG...	NK.T	DKASMLDEII	DYVKFLQLQV	K			
96	82	BAA97525	RRGQATDPHS	IAERLRRERI	AERMKSLQEL	VPNT...	NK.T	DKASMLDEII	EYVRFLLQV	K			
98	73	BAB10945	KRNIDAQFHN	LSEKRRRSKI	NEKMKALQKL	IPNS...	NK.T	DKASMLDEAI	EYLKQLQLQV	Q			
99	24	CAB80359	KRCRAAEVHN	LSEKRRRSRI	NEKMKALQSL	IPNS...	NK.T	DKASMLDEAI	EYLKQLQLQV	Q			
100	8	AAC33213	KRSRAAEVHN	LSERRRRDRI	NEKMRALQEL	IPNC...	NK.V	DKASMLDEAI	EYLKSLQLQV	Q			
101	15	AAD24380	KRSRAAEVHN	LSERKRRDRI	NERMKALQEL	IPRC...	NK.S	DKASMLDEAI	EYMKSLQLQI	Q			
102	9	AAD22130	RRSRAAEVHN	LSERRRRDRI	NERMKALQEL	IPHC...	SK.T	DKASILDEAI	DYLKSLQLQL	Q			
103	65	CAB86934	RRSRAAEVHN	LSERRRRDRI	NERMKALQEL	IPHC...	SR.T	DKASILDEAI	DYLKSLQMLQ	Q			
104	119	CAA22971	KRSRAADMHN	LSERRRREKI	NERMKTLQEL	LPRC...	RK.V	DKVSMLEDVI	EYVKSLQLQI	Q			
105	138	CAA2297B	KRSRAEMHN	LAERRRREKI	MERNKTLQQL	LPRC...	NK.V	DKVSMLEDVI	EYVKSLQMI	Q			
106	56	CAA22972	KRSRTAEMHN	LAERRRREKI	NEKMKTLQQL	LPRC...	NK.V	EKVSTLDDAI	EYVKSLQSQI	Q			
107	23	CAB81467	KRSRAAIMHK	LSERRRRQKI	NEMMKALQEL	LPRC...	TK.T	DRSSMLDDVI	EYVKSLQSQI	Q			
108	16	CAB80763	KRSRAAAIHN	QSERKRRDKI	NQRMKTLQKL	VPNS...	SK.T	DKASMLDEVI	EYLKQLQAQV	Q			
109	72	BAB08482	RRGRAAAIHN	ESERRRRDRI	NQRMRTLQKL	LPTA...	SK.A	DKVSILDDVI	EHLKQLQAQV	Q			
110	124	AAC34226	KRKRSTEVHK	LYERKRRDEF	NKKMRALQDL	LPNC...	YK.D	DKASLLDEAI	KYMRTLQLQV	Q			
125	102	AAF07356	KASAIRSKHS	VTEQRRSKI	NERFQILREL	IPNS...	EQKR	DTASFLLVEI	DYVQYLQEKV	Q			
126	46	CAB93714	KLNTPRSKHS	ATEQRRSKI	NDRFQMLRQL	IPNS...	DQKR	DKASFLLVEI	EYIQFLQEKA	H			
127	141	BAB08642	NRNSCRSKHS	ETEQRRSKI	NERFQSLMDI	IPQNQNDQKR		DKASFLLVEI	EYIHFLQEKV	H			
133	105	BAB09934	CESSATSSKA	CREKQRDRDL	NDFKFMELGAI	LEPG...	NPPKT	DKAAILVDAV	RMVTQLRGEA	Q			
134	115	AAG50538	ESCTGSNSKA	CREKQRDRDL	NDFKFTLSSV	LEPG...	RTPKT	DKVAIINDAI	RMVNQARDEA	Q			
135	34	BAA95734	SCSKPG.TKA	CREKLRREKL	NDFKMDLSSV	LEPG...	RTPKT	DKSAILDDAI	RVVNQLRGEA	H			
136	104	CAB78483	SCSRGGGTKA	CRELRLREKL	NERFMDLSSV	LEPG...	RTPKT	DKPAILLDDAI	RILNQLRDEA	L			
137	11	AAL55718	KKEAVCSQKA	EREKLRDKL	KEQFLELQNA	LDPN...	R.PKS	DKASVLTDTI	QMLKDVMMNV	D			
138	121	NM112876	DVSARKSQKA	GREKLRREKL	NEHFVELGNV	LDPE...	R.PKN	DKATILTDTV	QLLKELTSEV	N			
139	47	NM114632	GKVPKRINKA	VRERLKRHL	NELFIELADT	LELN...	Q.QNS	GKASILCEAT	RFLKDVFGQI	E			

Total: 109 proteins

b. Non E-box binders: proteins with a basic region, but not predicted to have E-box binding capacity.

The proteins listed in this table have a considerable basic region (first 17 positions) and might bind DNA, but do not contain both of the amino acids considered to be required to recognize and E-Box (E13 and R16) (see Functional Categories Section). Proteins are listed according to the Entry Number assigned, and their AtbHLH Number and PID Number are also provided to identify them.

EN	AtbHLH Number	PID Number	BASIC		HELIX		LOOP		HELIX	
			10 E R 2	0 3 0	0 4 0	0 5 0	0 6 0			
61	113	BAA99700	GKR N QEKSSV	GI A KV R KERL	GERIAALQQL	VSPY...GKT	DAASVLHEAM	GYIKFLQDQI	Q	
67	135	AAF15922	RRSRSRQSSG	TSE R ISEDQI	NDLI I KLQQL	LPRS...DKV	SAARVLQDTC	NYIRNLHREV	D	
68	26	AAK15282	REVPSVTRKG	SKRR R RDEKM	SNKMRKLQQL	VPNC...HK.T	DKVSVLDKTI	EYMKNLQLQL	Q	
111	132	BAC10690	KRKRNAEAYN	SPE R NQRNDI	NKKMRTLQNL	LPNS...HK.D	DNESMLDEAI	NYMTNLQLQV	Q	
117	37	CAB62312	NVRISKDPQS	VAARH R RERI	SERIRILQRL	VPGG...TK.M	DTASMLDEAI	HYVKFLKKQV	Q	
123	53	AAC12822	SKKPTLSSQS	LAAR G RRRI	AEKTHELGKL	IPGG...NK.L	NTAEMFQAAA	KYVKFLQSQV	G	
124	52	AAD25754	TKKRELSAQ	TAARK R RRRI	TEKTQELGKL	IPGS...QK.H	NTAEMFNAAA	KYVKFLQAQI	E	
131	145	BAB10287	RISFLKR S KL	SSNKIGEEKI	FETVSL L RSV	VPGE...ELV	DPILVIDRAI	DY L KSLKMEA	K	
132	108	NM102341	KSSDKSDHDT	LLKK R RRERI	RRQLET L KEI	TPNC...PQS	DINAILDCVI	EY T NNLRLAH	Y	
142	147	BAA94998	RSKKQRATVL	RLKAKGLPAV	QRKVKVLSRL	VPGC...RK.S	LPV.VLEETT	DYIAAMEMQI	R	
143	148	AAF63634	RSRKRRVSVL	RLNKKSIPDV	NRKVRVLGRL	VPGC...GKQS	VPV.ILEEAT	DYIQALEMQV	R	

Total: 11 proteins

Total number of proteins predicted to have some type of DNA binding capacity: 120 proteins.