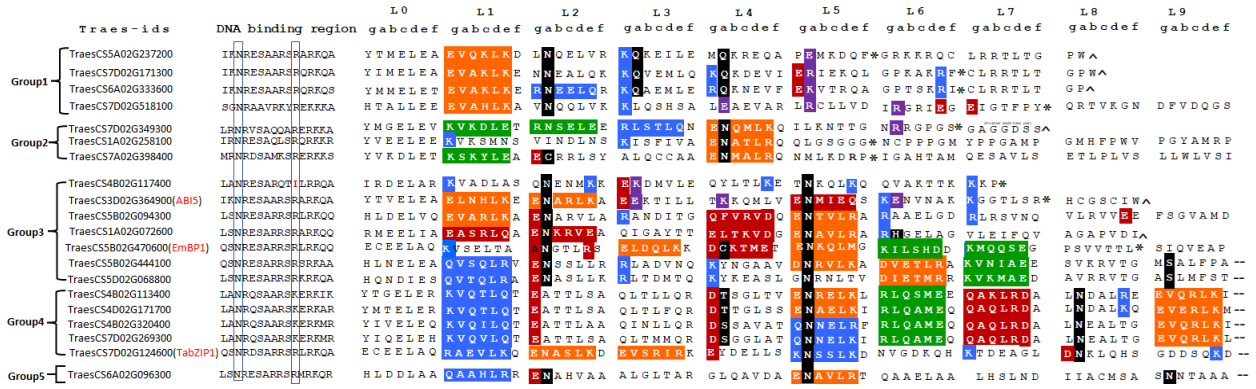
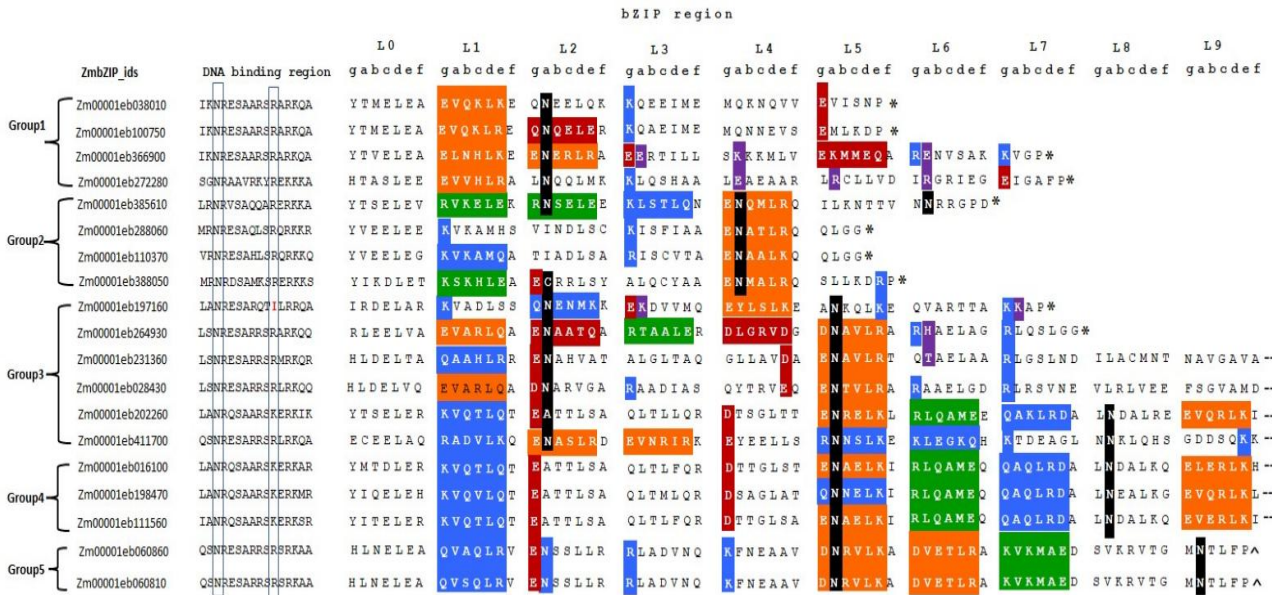


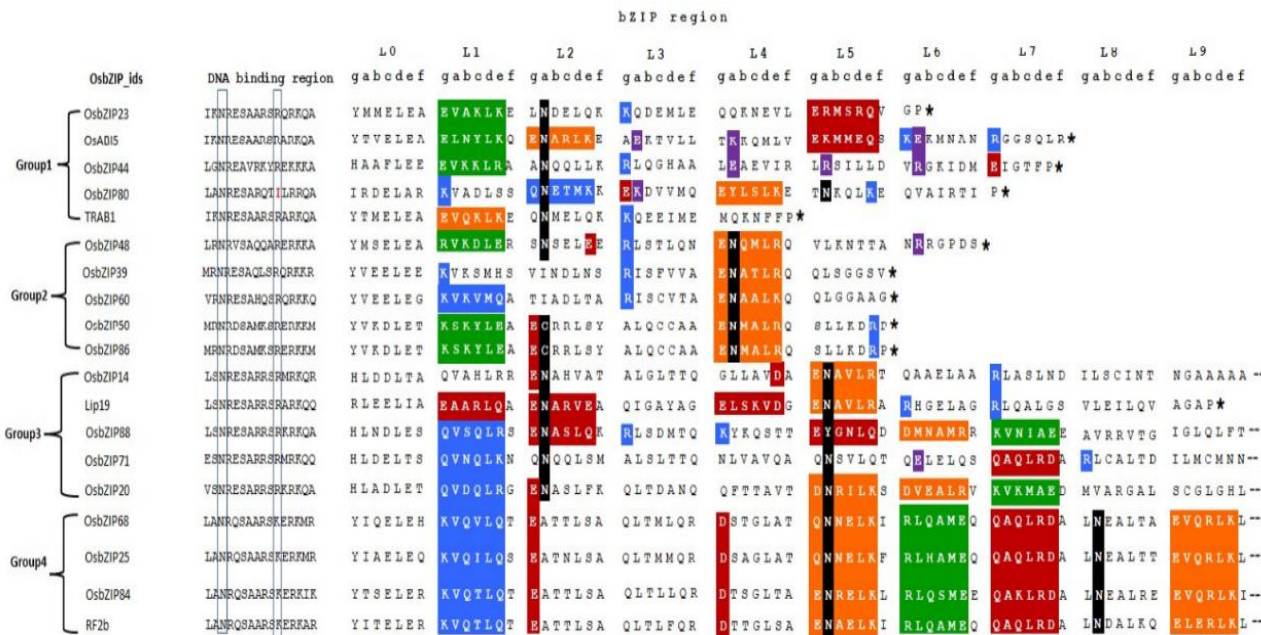
(A)



(B)



(C)



(D)

		bZIP region									
		L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
		gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef
HvbZip_ids	DNA binding region										
Group1	Hv7HG0713820	LANVSAQAARERKA	YMGEELE	AVKDLLE	RNSELEE	RLSTLQN	ILKNTTG	NKRGPF*			
	Hv6HG0619650	IRNRESAARSQRKQS	YMMEELE	EVAKLKE	RNQEQER	RQKNEVF	EKVSQRQ	GP*			
	Hv3HG0300770	IRNRESAARSARFQA	YTVELEA	ELNHLKE	ENARLKA	EKTILL	TKKQMLV	EKMMEQS	KENVNAK	KG*	
	Hv5HG0486380	IRNRESAARSARFQA	YTMELEA	EVQALAD	LNEELVK	KQTEILK	MQRREAP*				
	Hv3HG0295240	LGNREAVRYQRKKA	HAAPLEE	EVKKLRA	ANQQLLR	RLQGHAA	LEAEVAR	LTSLLLD	VRGKIDE	AEIGALP*	
Group2	Hv4HG0378580	SGNRAAVRYEERKA	HTASLEE	EVVHLRA	LNQQLMK	KLQSHAA	LEAEAR	LRCLLVD	IRGRIEG	EGAFPP*	
	Hv1HG0065020	IRNRESAQLSRQRKR	YVEELE	KVKSMNS	VINDLNS	KISFIVA	ENATLRQ	QLGNNGG*			
Group3	Hv2HG0119190	VNRRESAHLSPQRKQ	YVEELE	KVKAMQA	TIADLST	RISCVTA	ENAAKQ	QLAGAGG*			
	Hv1HG016310	LSNRESAARSARFQA	RMEELIA	EASRLQA	ENARVEA	QIGAYTT	ELTKVDG	ENAVLRA	RHGELAG	RLQALGG*	
	Hv7HG0686580	LANQSAARSKERFMR	YIQEHL	KVQVLQT	EATTLA	QLTMLQR	DSGGLAT*				
	Hv5HG0443520	LSNRESAARSRLRQK	HLDELVQ	EVARLKA	ENARVLA	RANDITG	QFVRVDQ	ENTVLRA	RAAELGD	RLRSVNQ	VLRVVE
	Hv5HG0433250	LSNRESAARSRLRQK	HQSDIES	QVTQLRA	ENASLIN	RLTDMTQ	KYKEASL	GNNRLTV	DIETMR	KVNVL	
Group4	Hv6HG0570630	VNRRESAARSRLRQK	HLVELET	QVDQLRG	DNASIFK	QLTDANQ	QFTTAVT	ENRILKS	DVEALRV	KVLAED	MYARGAL
	Hv5HG0449210	QSNRESAARSRLRQK	ECEELAQ	RAEVLKQ	ENASLKD	EVSRIRK	ETDELIS	KNNSLKD	NIGDKQH	KTDEAGL	HNELOTS
	Hv4HG0349900	LANQSAARSKERKIK	YTGELE	KVQTLQT	EATTLA	QLTLLQR	DFSGLTV	ENRELKL	RLQSMEE	QAKLRDA	LNDALRE
	Hv4HG0351410	LANQSAARSKERFAR	YMTLE	KVQVLQT	EATTLA	QLTLFQR	DTTGLSS	ENAELEI	RLQAMEQ	QAQLRDA	LNDALRG
	Hv2HG0112060	LANQSAARSKERFAR	YITELER	KVHTLQT	EATTLA	QLTLFQR	DTTGLSS	ENAELEI	RLQAMEQ	QAQLRDA	LNDALRG
Group5	Hv7HG0723770	QSNRESAARSRLRKA	HLNELEA	QVSQLRV	ENSSLR	RLADVQ	KFNEAAV	DNVLA	DVETLRA	KVMAAE	SVERVTG
	Hv6HG0554560	IRNRESAARSRLRQR	HLDDLAA	QAAHLR	ENAHVAA	ALGLTAR	GLLAVDA	ENAVLRT	QAELAA	RLQSLND	IIACMSA

(E)

		bZIP region									
		L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
		gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef	gabdef
AtbZIP_ids	DNA binding region										
Group1	ABF2	IRNRESAARSARFQA	YTVELEA	EVAKLKE	ENDELQR	RQARIME	MQRNQET	EMRNLLQ	GG*		
	AB15	IRNRESAARSARFQA	YTVELEA	ELNQLKE	ENAEQLH	ALAELE	RKRQVQK	TPIEFAL	LRWLQYF	GQKRTTK	WNCWMV
	AtbZIP31	LSNRESAARSQKFLQ	YMINLEL	KINFLN	KNASIFE	KIKLLEN	DKTMRMN	EKKEIMI	RIESLEQ	HAELRDA	LNEHLHV*
	AtbZIP24	CGNREAVRYEERKA	RTAYLED	EVMRLOS	LNEQFLR	KLQSQEM	VETELIR	LRALLVE	MQGRIEV	ELCSFSF	QKQCNGS
	ABF1	IRNRESAARSARFQA	YTVELEA	EIESLKL	VNQDLQR	KQAEIMK	THNSELK	AEFSKQP	LLARRQC	LRRTLGT	P
Group2	AtbZIP60	VNRDAVRSERKKE	YVQDLER	KSKYLE	ECERLGR	MLECFVA	ENQSLRY	CLQKNG	NNTTMS	KQESAVL	LLESLLL
	AtbZIP17	MRNRESAQLSRQRKH	YVEELE	KVRNMHS	TITDLNG	KISYFMA	ENATLRQ	QLGG*			
	AtbZIP28	IRNRESAQLSRQRKQ	QTEELER	KVKSMNA	TIAELNG	KIAYVMA	ENVALRQ	QMAVASG	AP*		
	HY5	LANVSAQAARERKA	YLSLEEN	RVKDLLE	RNSELEE	RLSTLQN	ENQMLRH	VRFDLT	VDLTCNV	NAFAICF	LIYHSTG^
	AtbZIP62	LANRESARQIRRRQA	MCEELSK	KAADLT	ENENLR	EKDWALK	EFQSLE	INKHLKE	QVLKSVK	P*	
Group3	AtbZIP63	LSNRESAARSARFQA	HLSELET	QVSQLRV	ENSKLMK	GLTDVTQ	TFNDASV	ENRVLKA	NIELTRA	KVKMAE	TVKRLTG
	AtbZIP11	LSNRESAARSRLRQK	LLDDLTA	QVNHLLK	ENTEIVT	SVSITTQ	HYLTVEA	ENSVLRA	QLDELNH	PPMGMY	IEFIDS
	bZIP53	ISNRESAARSRLRQK	QLGDLIN	EVTLLKN	DNARITE	QVDEASK	KYIEMES	KNNVLRA	QASELTD	RLRLSLS	VLEMVE
	bZIP16	QSNRESAARSRLRQK	ECDELAQ	RAEVLNE	ENTNLRA	EINKLKS	QCEELTT	ENTSLKD	QLSLFP*		
	AtbZIP1	LSNRESAARSRLRQK	LMEDTIH	EISSLER	RIKENSE	RCRAVQ	RLDSVE	ENAGLRS	EKIWLSS	YVSDLEN	MIATTS
Group4	AtbZIP29	LANQSAARSKERFMR	YIVELEH	KVQTLQT	EATTLA	QLTLLQR	DMGMLTN	QNNELKF	RLQAMEQ	QARLRDA	LNEALNG
	AtbZIP52	LANQSAARSKERFAR	YIQELER	KVQSLQT	EATTLA	QLTLYQR	DFNGLAN	ENELKL	RLQAMEQ	QAQLRNA	LNEALRG
	VIP1	LANQSAARSKERKIR	YTGELE	KVQTLQN	EATTLA	QVTMLQR	GTSELNT	ENKHLKM	RLQALEQ	QAELRDA	LNEALRD
	AtbZIP10	LANQSAARSKERFMR	QTSDEL	QVNDLKS	EHSLLK	QLSNMNH	KYDEAAV	GNILKA	DIETLRA	KVKMAE	TVKRVGT
											MNP^

- Attractive acidic-basic pairs
- Attractive basic-acidic pairs
- Repulsive basic pairs
- Repulsive acidic pairs

Supplementary figure 2: Amino acid sequence of orthologue bZIPs in reference plant species:

Protein sequences are aligned into sub-families demarcated with groups in first column clustered based on the similarity in sequence. Second column contains protein ids. Third column has n-terminal region of bZIP domain and c-terminal leucine zipper region is divided into heptads (a,b,c,d,e,f and g). Primarily, the placement of N at 'a' position was considered for classification and it was coloured black at the respective 'a' position. Secondly, for the g-e' interactions, we used four colours. The attractive acidic-basic pairs ($E \leftrightarrow R'$, $E \leftrightarrow K'$, $D \leftrightarrow R'$, $D \leftrightarrow K'$) were coloured orange, and the attractive basic-acidic pairs ($R \leftrightarrow E$, $K \leftrightarrow E'$) were coloured green. We coloured the repulsive basic pairs ($K \leftrightarrow K'$, $R \leftrightarrow K'$) with blue and repulsive acidic pairs ($E \leftrightarrow E'$, $D \leftrightarrow D'$, $E \leftrightarrow D'$) with red colour. We have coloured the polar residues black at 'a' and 'd' position and charged residues at 'a' and 'd' positions were coloured purple. The candidate genes have been divided into five groups. We have used references plant species *T aestivum* (A), *Z mays* (B), *O sativa* (C), *H vulgare* (D) and *A thaliana* (E) for the preparation of this figure.