BzpF is a CREB-like transcription factor that regulates spore maturation and stability in *Dictyostelium*

Eryong Huang[1], Shaheynoor Talukder[2], Timothy R. Hughes[2], Tomaz Curk[3], Blaz Zupan[3], Gad Shaulsky[1] and Mariko Katoh-Kurasawa[1,]*

Supplementary Materials Table of Contents

This supplement includes 2 supplementary data, 7 supplementary tables and 2 supplementary figures.

Supplementary datasets

Supplementary Tables

Supplementary Figure Legends

Fig. S1. Amino acid sequence alignment of *D. discodeum (Dd)* and *D. purpureum (Dp)* BzpFs. Predicted amino acid sequence of *Dp* BzpF is made from a contig assembled from RNA-seq reads (Parikh 2011). Two poly-Glutamine sequences (Q-rich), bZIP domain, CREB Signature sequence and Nuclear Localization Signal (NLS) are indicated. Asterisk shows the Serine residue in a PKA phosphorylation consensus sequence. Arrowhead indicates the end of the truncated BzpF peptide (1-137 amino acids) expressed as a consequence of deletion in *bzpF*$^{-}$. Note that the sequence of C-terminal half including bZIP domain and a short N-terminal sequence (1-56 amino acids) are conserved.

Fig. S2. *bzpF*$^{-}$ spores are compromised. We tested spores of AX4 and *bzpF*$^{-}$ from 2-day old fruiting bodies by phase-contrast microscopy and by staining with fluorescent dyes. Strain names are indicated on the left of each panel and the visualization method is on the top. (A) *Staining with FITC-lectin* to test the integrity of spores in AX4 and *bzpF*$^{-}$ with/without urea treatment. Bar = 20 μm. (B) *Calcofluor white staining* to test the integrity of the middle spore coat layer. Bar = 5 μm.

Table S1. *D. discoideum* strains

Strain name	**Strain Descriptor**	**Parental strain**	**Antibiotic resistance**
AX4	AX4 (Wild type)	AX3	
EH001	*bzpF*⁻	AX4	BlasticidinS
EH002	*bzpF*⁻/[act15]:bzpF(4-628):GFP	EH001	G418
EH003	*bzpF*⁻/[act15]:bzpF(4-628):GFP	EH001	G418

Table S2. qRT-PCR primers

C_1	bzpF *	F	GTCTTATTGAAGGTGACTCCA
		R	AAATTAAACCATTTGATTGTTTATTTTG
C_3	DDB_G0285833	F	CAAATAGTTTAGCTAGATTTGCA
		R	TTAAGCATCGTCAACAACTAAG
C_5	sigD	F	TTCAACTACTGGAAATACTGG
		R	CTAATACACAATGAGCTAAACC
C_6	DDB_G0267964	F	GTAAACAACATACAATAAAAACC
		R	TACCATTTTCTTTATCAAGTGG
C_8	DDB_G0267898	F	ATGTTACCCAAGGTAGTGC
		R	CCACGTGGATACCATAGC
C_12	DDB_G0270980	F	CAAGTTGAAAATGCAATAATGG
		R	ATACATGATGTTTTGCACTACC
C_13	DDB_G0271570	F	AGTTCAACAGGTAGTGCTGG
		R	TTCCAATGTTTCTTTATCATCACC
C_15	DDB_G0270474	F	AATTTCGATTCAAGTCACCA
		R	ATTGCATATAAAGTAGGTTTTGG
C_16	DDB_G0290419	F	TGGGTTCATTCAGTATTTTAGC
		R	AATATTTTCAGCACCAACAGC
C_17	DDB_G0293262	F	TTTCTAAAGGTGGTGATATTCC
		R	TTTGAGTGAAAACACCTTGCA
C_18	gxcl	F	AGAATTAAGAAAATCATTAGAACC
		R	ATTTGAGGTTGAGATTGTTCC

R_1	DDB_G0271972	F	GAGCTTTTTGTTCACTTACC
		R	ATGGAGATGGGGTACCAG
R_2	rnpA	F	TGGTGAGAGTAAAGGATTTGG
		R	TGGTGACATTAAGAAATCGATTG
R_3	sigK	F	TTCAATACACCATTCGTCCCT
		R	CAATTTAAGCAGTTGAGCGAAC
R_4	DDB_G0276285	F	TGAAACTTTGGCTTCTTGGTG
		R	TTACAAATGCAGCAACAATACC
R_5	DDB_G0269730	F	CAGTTGAAGAACAAGATAAACC
		R	CAGGAGAAGAAGTTTGTTTTGC

* Here we considered sequences that remained intact
in the knockout strains at the 5’ end of the gene.

Table S3. DNA binding preference of BzpF (E-score > 0.4)

binding sequence	E-score
ATGACGTA	0.496
ATTACGTA	0.493
ACGTCATC	0.484
ACGTAATC	0.482
ATACGTAA	0.482
TACGTATA	0.474
ATACGTAT	0.473
CTTACGTA	0.470
TTACGTAA	0.466
CATACGTA	0.460
TGACGTAA	0.442
ACGTATAA	0.437
GCCACGTA	0.437
ATACGTCA	0.431
TGACGTCA	0.427
ACGTCATA	0.421
CGTCATCA	0.420
ATACGTGG	0.415
ACACGTAT	0.408
GTGACGTA	0.406

* For complete analysis,
see Supplementary datasets S1; excel file “datasets_S1_PBM.xls”.

Table S4. BzpF target candidates by computational prediction

#	DDB_G #	*Gene name*	score	P-value	annotation	logAX4/*bzpF*$^{-}$ (AX4 raw reads)
C_1	DDB_G0279529	*bzpF*	4.64	0.000	basic-leucine zipper (bZIP) transcription factor	**2.9 (918)**
C_2	DDB_G0283007		3.72	0.001	unknown	-0.3 (6808)
C_3	DDB_G0285833		2.77	0.004	unknown	**2.0 (202)**
C_4	DDB_G0292442		2.76	0.007	NAD+ kinase family protein, phosphorylates NAD+ to NADP+	0.9 (875)
C_5	DDB_G0267474	*sigD*	1.88	0.028	sig = SrfA-Induced Gene, spore coat protein	**2.1 (5089)**
C_6	DDB_G0267964		1.88	0.028	unknown	**1.6 (104)**
C_7	DDB_G0271326		1.88	0.028	contains a predicted cell-cell adhesion domain (Dicty_CAD)	0.8 (161)
C_8	DDB_G0267898		1.85	0.038	EGF(epidermal growth factor) -like domain-containing protein, two EGF repeats : expressed in pstAB cells	**1.3 (517)**
C_9	DDB_G0285813		1.85	0.038	coiled-coil domain containing protein, belongs to DUB812 family	-0.1 (251)
C_10	DDB_G0282127		1.84	0.064	unknown	0.5 (52)
C_11	DDB_G0291598		1.84	0.047	unknown	0.4 (2197)
C_12	DDB_G0270980		1.82	0.116	unknown	**2.9 (386)**
C_13	DDB_G0271570		1.82	0.116	unknown	**2.1 (116)**
C_14	DDB_G0290025		1.82	0.116	unknown	0.4 (70)
C_15	DDB_G0270474		1.35	0.003	unknown	**1.3 (112)**
C_16	DDB_G0290419		0.91	0.035	DUF298 family protein, DCUN1 domain-containing protein	**4.7 (562)**
C_17	DDB_G0293262		0.90	0.059	contains a predicted signal peptide, similar to DDB_G0268828	**1.4 (1361)**
C_18	DDB_G0290023	*gxcL*	0.89	0.079	RacGEF, pleckstrin homology (PH) domain-containing protein RhoGEF domain-containing protein	**2.3 (927)**

Table S5. Additional BzpF target candidates by transcriptome comparison

#	DDB_G#	*Gene*	annotation	logAX4/*bzpF*⁻ (AX4 raw reads)
R_1	DDB_G0271972		unknown	**6.8 (269)**
R_2	DDB_G0282293	*rnpA*	putative RNA binding protein (RNP) regulated by the MADS-box transcription factor SrfA	**3.9 (501)**
R_3	DDB_G0267476	*sigK*	EGF-like domain-containing protein regulated by the MADS-box transcription factor SrfA	**3.2 (8095)**
R_4	DDB_G0276285		unknown	**3.0 (2370)**
R_5	DDB_G0269730		unknown	**2.9 (1033)**
R_6	DDB_G0290703	*pks31*	putative polyketide synthase	2.8 (173)
R_7	DDB_G0274433		CBS (cystathionine-beta-synthase) domain-containing protein	2.5 (213)
R_8	DDB_G0271662	*pks10*	putative polyketide synthase	2.0 (180)
R_9	DDB_G0271196		contains a predicted cell-cell adhesion domain (Dicty_CAD)	1.9 (296)
R_10	DDB_G0288389		unknown	1.4 (117)
R_11	DDB_G0280581		unknown	1.4 (237)
R_12	DDB_G0268746		contains a UBX domain, found in ubiquitin-regulatory proteins	1.3 (365)
R_13	DDB_G0290075		protein phosphatase 2C	1.1 (473)

Table S6. Differentially-expressed genes

transcription factor	**bzpF ***	G0291348	**gtaE**	**gtaG**									
	2.9	1.2	**1.1**	**1.0**									
antioxidant	**trxD**	**gcsA**											
	2.5	**1.7**											
cytokinin synthesis	**iptA**												
	2.6												
Pplyketide synthase	**pks31***	G0287501	**pks24**	**pks10***	**pks5**	**pks23**							
	2.8	2.3	**2.2**	**2.0**	**1.6**	**1.5**							
SrfA-Induced Gene	**sigL-1**	**rnpA***	**sigK***	**sigM**	**sigD***	**sigI**							
	4.8	**3.9**	**3.2**	**2.2**	**2.1**	**1.2**							
coat protein	**sigD***	**cotE**	G0282407										
	2.1	**1.7**	1.5										
germination	**gerA**	G0287549	**gerB**	**gerC**									
	5.9	5.8	**5.1**	**4.9**									
Lipoprotein / oxysterol binding / cysteine proteinase	**expl2**	**osbK**	**cprF**										
	2.8	**1.7**	**1.0**										
cellulose-binding domain	**ecmL**	G0281233	G0278437	G0278399	**staA**	G0268756	G0278439	G0278423	**staB**	G0282589	G0278519	G0289363	G0278537
	26.8	9.6	9.0	8.7	**7.3**	6.5	6.5	5.8	**5.8**	3.9	3.6	3.5	1.7
glycoside hydrolase domain	G0284239	G0271314	G0284295	G0268446									
	5.5	3.1	1.3	1.0									

1) Green: preferential pst expression, Red: preferential psp expression, * : dCRE motif in the promoter region, underlines: spore and sporulation defect mutant, numbers: log AX4/*bzpF*$^{-}$

2) ***gcsA*$^{-}$:** **abolished sporulation | aberrant spore coat morphology | abolished cellulose biosynthesis | abolished stalk formation | abolished stalk cell differentiation**

***cotE*$^{-}$:** **aberrant spore coat morphology**

***iptA*$^{-}$:** **decreased spore viability**

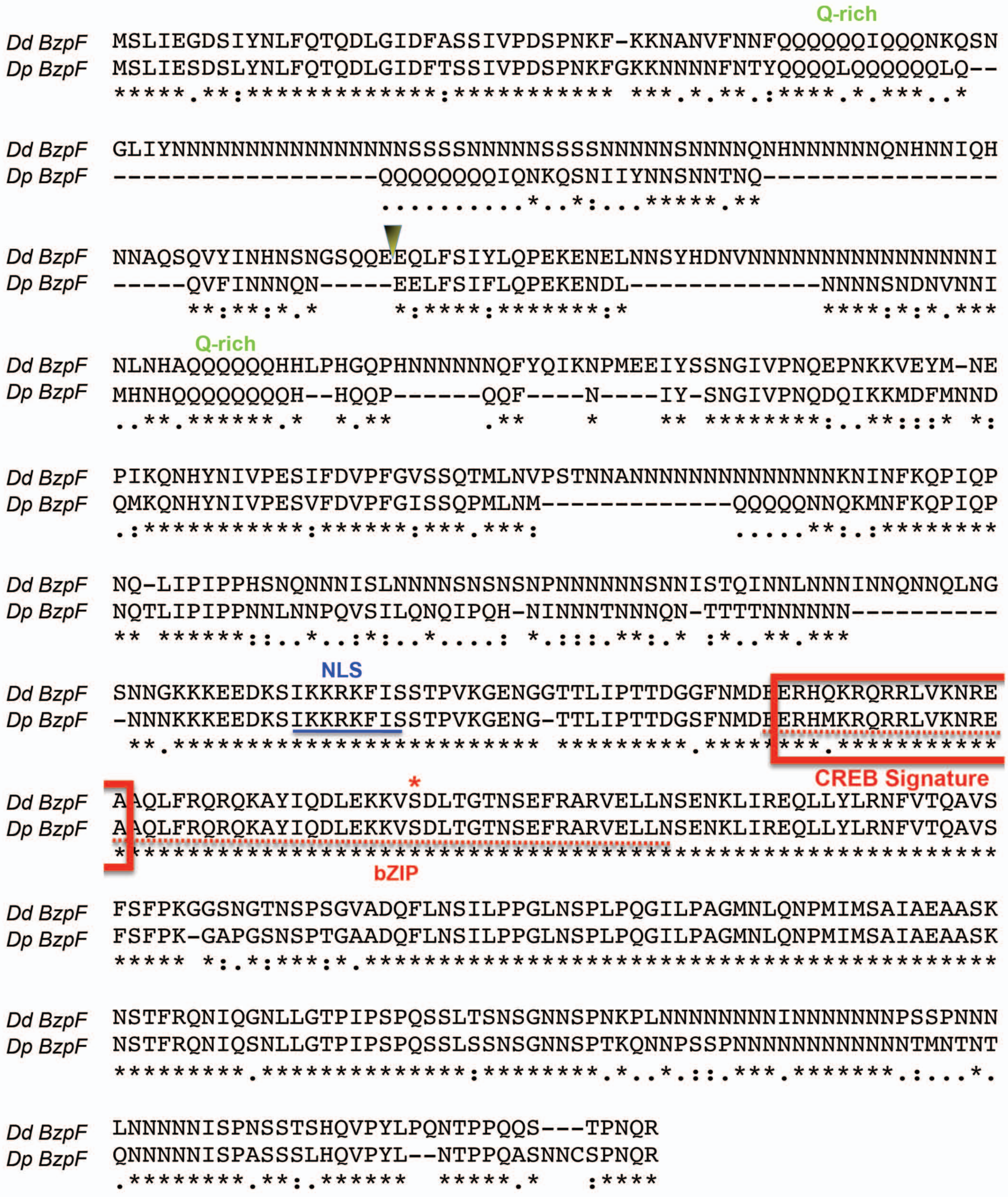
Table S7. Differentially-expressed genes (prespore, prestalk marker genes)

PSP genes

DDB_G ID	gene name	(logAX4/F-)
DDB_G0276469	syn16B	3.40
DDB_G0284677	expl2	2.80
DDB_G0280761	DDB_G0280761	2.29
DDB_G0274801	DDB_G0274801	2.21
DDB_G0291396	DDB_G0291396	2.10
DDB_G0288817	osbK	1.73
DDB_G0277903	cotE	1.72
DDB_G0272218	DDB_G0272218	1.47
DDB_G0282407	DDB_G0282407	1.46
DDB_G0293806	DDB_G0293806	1.22
DDB_G0274121	abcA4	1.17
DDB_G0278963	DDB_G0278963	1.11
DDB_G0284909	DDB_G0284909	1.09
DDB_G0278507	DDB_G0278507	1.09
DDB_G0276375	DDB_G0276375	1.05
DDB_G0268076	DDB_G0268076	1.05
DDB_G0272156	DDB_G0272156	1.01
DDB_G0279973	DDB_G0279973	-1.17
DDB_G0293048	DDB_G0293048	-1.50

PST genes

DDB_G ID	gene name	(logAX4/F-)
DDB_G0292636	DDB_G0292636	3.19
DDB_G0267476	sigK	3.16
DDB_G0270980	DDB_G0270980	2.85
DDB_G0270942	DDB_G0270942	2.51
DDB_G0271662	pks10	2.01
DDB_G0273213	aslA-1	1.85
DDB_G0283885	DDB_G0283885	1.74
DDB_G0278823	DDB_G0278823	1.67
DDB_G0293864	DDB_G0293864	1.49
DDB_G0273357	DDB_G0273357	1.37
DDB_G0275687	abcG18	1.34
DDB_G0275265	etnkB	1.32
DDB_G0269530	DDB_G0269530	1.32
DDB_G0283811	mcfR	1.29
DDB_G0267898	DDB_G0267898	1.26
DDB_G0272464	DDB_G0272464	1.23
DDB_G0289605	DDB_G0289605	1.16
DDB_G0269132	ecmB	1.03
DDB_G0270756	gtaG	1.03
DDB_G0284675	DDB_G0284675	-1.05
DDB_G0278951	DDB_G0278951	-1.18
DDB_G0281011	DDB_G0281011	-1.38



```
                                                                Q-rich
Dd BzpF MSLIEGDSIYNLFQTQDLGIDFASSIVPDSPNKF-KKNANVFNNFQQQQQQIQQQNKQSN
Dp BzpF MSLIESDSLYNLFQTQDLGIDFTSSIVPDSPNKFGKKNNNNFNTYQQQQLQQQQQQLQ--
        *****.**:*************:*********** ***.*.**.:****.*.***..*  

Dd BzpF GLIYNNNNNNNNNNNNNNNNSSSSNNNNNSSSSNNNNNSNNNNQNHNNNNNNQNHNNIQH
Dp BzpF ------------------QQQQQQQQIQNKQSNIIYNNSNNTNQ----------------
                          ..........*..*:...*****.**                

Dd BzpF NNAQSQVYINHNSNGSQQEEQLFSIYLQPEKENELNNSYHDNVNNNNNNNNNNNNNNNNI
Dp BzpF -----QVFINNNQN-----EELFSIFLQPEKENDL-------------NNNNSNDNVNNI
             **:**:*.*     *:****:*******:*             ****:*:*.***

          Q-rich
Dd BzpF NLNHAQQQQQQHHLPHGQPHNNNNNNQFYQIKNPMEEIYSSNGIVPNQEPNKKVEYM-NE
Dp BzpF MHNHQQQQQQQQH--HQQP------QQF----N----IY-SNGIVPNQDQIKKMDFMNND
        ..**.******.*  *.**      .**    *    ** ********:..**:::* *:

Dd BzpF PIKQNHYNIVPESIFDVPFGVSSQTMLNVPSTNNANNNNNNNNNNNNNNKNINFKQPIQP
Dp BzpF QMKQNHYNIVPESVFDVPFGISSQPMLNM-------------QQQQQNNQKMNFKQPIQP
        .:***********:******:***.***:             .....**:.:********

Dd BzpF NQ-LIPIPPHSNQNNNISLNNNNSNSNSNPNNNNNNSNNISTQINNLNNNINNQNNQLNG
Dp BzpF NQTLIPIPPNNLNNPQVSILQNQIPQH-NINNNTNNNQN-TTTTNNNNNN----------
        ** ******::..*..:*:..*....: *.:::.**:.* :*..**.***          

                        NLS
Dd BzpF SNNGKKKEEDKSIKKRKFISSTPVKGENGGTTLIPTTDGGFNMDEERHQKRQRRLVKNRE
Dp BzpF -NNNKKKEEDKSIKKRKFISSTPVKGENG-TTLIPTTDGSFNMDEERHMKRQRRLVKNRE
         **.************************* *********.********.***********
                                                     CREB Signature
                            *
Dd BzpF AAQLFRQRQKAYIQDLEKKVSDLTGTNSEFRARVELLNSENKLIREQLLYLRNFVTQAVS
Dp BzpF AAQLFRQRQKAYIQDLEKKVSDLTGTNSEFRARVELLNSENKLIREQLLYLRNFVTQAVS
        ************************************************************
                          bZIP

Dd BzpF FSFPKGGSNGTNSPSGVADQFLNSILPPGLNSPLPQGILPAGMNLQNPMIMSAIAEAASK
Dp BzpF FSFPK-GAPGSNSPTGAADQFLNSILPPGLNSPLPQGILPAGMNLQNPMIMSAIAEAASK
        ***** *:.*:***:*.*******************************************

Dd BzpF NSTFRQNIQGNLLGTPIPSPQSSLTSNSGNNSPNKPLNNNNNNNNINNNNNNNPSSPNNN
Dp BzpF NSTFRQNIQSNLLGTPIPSPQSSLSSNSGNNSPTKQNNPSSPNNNNNNNNNNNNTMNTNT
        *********.**************:********.*..*.::.***.*******.:...*.

Dd BzpF LNNNNNISPNSSTSHQVPYLPQNTPPQQS---TPNQR
Dp BzpF QNNNNNISPASSSLHQVPYL--NTPPQASNNCSPNQR
        .********.**:.******  *****.*   :****
```

Fig. S1. Amino acid sequence alignment of *D. discodeum (Dd)* and *D. purpureum (Dp)* Bzp Predicted amino acid sequence of *Dp* BzpF is made from a contig assembled from RNA-seq reads (Pa 2011). Two poly-Glutamine sequences (Q-rich), bZIP domain, CREB Signature sequence and Nuc Localization Signal (NLS) are indicated. Asterisk shows the Serine residue in a PKA phosphorylat consensus sequence. Arrowhead indicates the end of the truncated BzpF peptide (1-137 amino ac expressed as a consequence of deletion in *bzpF*⁻. Note that the sequence of C-terminal half including b domain and a short N-terminal sequence (1-56 amino acids) are conserved.

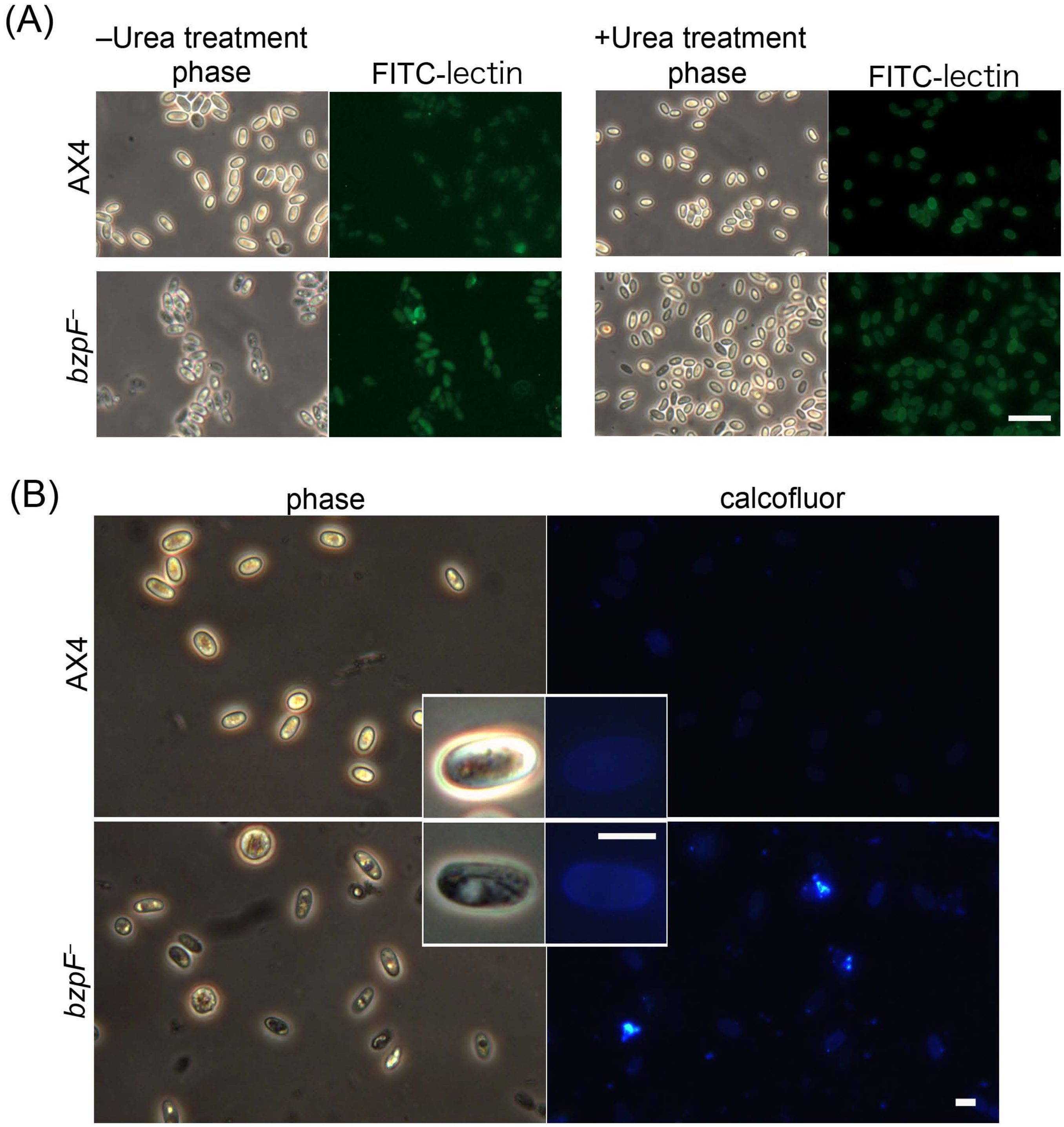


Fig. S2. *bzpF*⁻ spores are compromised. We tested spores of AX4 and *bzpF*⁻ from 2-day old fruiting bodies by phase-contrast microscopy and by staining with fluorescent dyes. Strain names are indicated on the left of each panel and the visualization method is on the top. (A) *Staining with FITC-lectin* to test the integrity of spores in AX4 and *bzpF*⁻ with/without urea treatment. Bar = 20 μm. (B) *Calcofluor white staining* to test the integrity of the middle spore coat layer. Bar = 5 μm.