

## Prediction result for 9aaTAD (2006)

### 9aa TAD search result for E2A (TFE2):

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP}[ILVFWM]{KRHCGP}{CGP} {CGP}[ILVFWM]{CGP}{CGP} pattern and filter settings for TFE2.

#### P15923 TFE2\_HUMAN

18 - 26: DLLDFSMMF

594 - 602: QAVSVHNL



#### P98180 TFE2\_MESAU

19 - 27: DLLDFSMMF

589 - 597: QAVAVHSL



#### P15806 TFE2\_MOUSE

19 - 27: DLLDFSMMF

591 - 599: QAVAVHSL



#### P21677 TFE2\_RAT

19 - 27: DLLDFSMMF

589 - 597: QAVAVHSL



#### Q01978 TFE2\_XENLA

18 - 26: DLLDFSMMF

600 - 608: QAVSVHSL



### 9aa TAD search result for Rtg3:

Prosite Pattern searching was done with [DENQSTYGM] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {KRHCGP} [ILVFWM] [ILVFWMAY] {KRHCP} pattern and filter settings for Rtg3\_.

#### P38165 RTG3\_YEAST

27 - 35: ETLDFS LVT

189 - 197: SSINSØMMT



## Experimental data and loss-of-function mutants of E2A and Rtg3:

		Sequence fused to LexA				$\beta$ -galactosidase
<b>Putative 9aa TAD</b>				<b>D .F.. .F..</b>		<b>% of B51</b>
LexA-Gal4 (1-871)	B51	. .TGMFNTTTM	<b>D</b>	DVYN YLFD	Di .	100 $\pm$ 3%
LexA-Oaf1 (1035-47)	B12		<b>nn</b>	n LFDY DFLF	GNDFA .	256 $\pm$ 5%
LexA-E2A	B34		nn	n LLDF SMMF	PLPG .	11 $\pm$ 4%
LexA-E2A	B33		n	n nLDF SMMF	PLPG .	1 $\pm$ 1%
						<b>Activity*</b>
Gal4bd-E2A, wt,1-99	ref. *	MNQPQRMAPVGTDKELS	<b>D</b>	LLDF SMMF	PLP . .	+
Gal4bd-E2A, mut1	ref. *	MNQPQRMAPVGTDKpLS	<b>p</b>	LLDF SMMF	PLP . .	-
Gal4bd-E2A, mut2	ref. *	MNQPQRMAPVGTDKELS	<b>D</b>	rLDr SMMF	PLP . .	-
Gal4bd-E2A, mut4	ref. *	MNQPQRMAPVGTDKELS	<b>D</b>	LLpF SpMF	PLP . .	-
LexA-Rtg3, wt, 1-99	ref. *	. .AENQRLLDEL MNQTKVLQ	<b>E</b>	TLDF SLVT	PTP . .	+
LexA-Rtg3, F31A	ref. *	. .AENQRLLDEL MNQTKVLQ	<b>E</b>	TLDa SLVT	PTP . .	-

\*(Massari et al., 1999; Massari et al., 1996)

### Synthetic orthologs:

<b>Putative 9aa TAD</b>				<b>D .F.. .F..</b>	
TFE2_HUMAN (E2A)	MNQPQRMA	PVGTDKELS	<b>D</b>	LLDF SMMF	PLP . .
E12_BRARE	MAT	VGTDMELN	<b>D</b>	LLDF SIMF	PHP . .
E12_XENLA	MN QQQRMTA	VGTDKELS	<b>D</b>	LLDF SMMF	PLP . .
HTF4_HUMAN	MNPQQQRMAA	IGTDKELS	<b>D</b>	LLDF SAMF	SPP . .
ITF2_HUMAN	MH HQQRMAA	LGTDKELS	<b>D</b>	LLDF SAMF	SPP . .
PAN2_RAT	MMNQSQRMA	PVGS DKELS	<b>D</b>	LLDF SMMF	PLP . .
RTG3_YEAST	. .AENQRLLDEL MNQTKVLQ		<b>E</b>	TLDF SLVT	PTP . .
RTG3_SACPA	. .AENQRLLDEL MNQTKILQ		<b>E</b>	TLDF SLVA	PAP . .

## Supplementary Table 2

### 9aa TAD search result for p53:

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {CGP} {CGP} pattern and filter settings for primates and p53\_.

#### P53\_HUMAN (P04637)

17 - 25: **ETFSDLWKL** TADI (1-43aa) Loss-of function mutation\*: **ETFSDQSKL**  
 38 - 46: QAMDDLMLS  
 48 - 56: **DDIEQWLTE** TADII (43-73aa) Loss-of function mutation\*: **DDIEQSTE**



#### P53\_CERAE

17 - 25: **ETFSDLWKL**  
 38 - 46: QAVDDLMLS  
 48 - 56: **DDLAQWLTE**



#### P53\_MACFA

17 - 25: **ETFSDLWKL**  
 38 - 46: QAVDDLMLS  
 48 - 56: **DDLAQWLTE**



#### P53\_MACFU

17 - 25: **ETFSDLWKL**  
 38 - 46: QAVDDLMLS  
 48 - 56: **DDLAQWLTE**



#### P53\_MACMU

17 - 25: **ETFSDLWKL**  
 38 - 46: QAVDDLMLS  
 48 - 56: **DDLAQWLTE**



### Experimental data and loss-of function mutants of p53 and interaction with TAF9:

		Sequence fused to LexA	$\beta$ -galactosidase
Putative 9aa TAD		<b>D .F.. .F..</b>	% of B12
Oaf1 (1035-47)	B12	<b>nn n LFDY DFLF</b> GNDFA.	100 $\pm$ 3
LexA-xp53 (17-29)	B35	<b>ef TFED LWSL</b> LPDPL.	85 $\pm$ 2
			Activity*
p53, L22Q/W23S*		. . VEPPLSQ <b>E TFSD qSKL</b> LPEN . .	-
p53-TADI - TAF interaction *		- - <b>FSD LW</b> - -	

p53 - TADII (43-73)

.. D DIEQ WFTE ..

p53, W53Q/F54S\*

.. D DIEQ qSTE ..

-

\*(Uesugi et al., 1997; Choi et al., 2000; Venot et al., 1999; Lin et al., 1994)

## Synthetic orthologs:

Putative 9aa TADI	D .F.. .F..
P53_HUMAN	LSQ <b>E</b> TFSD LWKL LPEN
P53_RAT	LSQ <b>E</b> TFSC LWKL LPPD
P89002	LSQ <b>E</b> TFQR LWKL LPPE
P53_MOUSE	LSQ <b>E</b> TFSG LWKL LPPE
P53_FELCA	LSQ <b>E</b> TFSE LWNL LPEN
P53_CHICK	EPT <b>E</b> VFMD LWSM LPYS
P53_BRARE	EFA <b>E</b> LWEK NLII QPPG
P53_SALIR	LSQ <b>E</b> SFED LWKM NLNL
P53_XENLA	LSQ <b>E</b> TFED LWSL LPDP
P53_PLAFE	GSQ <b>D</b> SFSE LWAS VQTP
P53_ORYLA	ESQ <b>G</b> SFQE LWET VYPP
O88897	LSP <b>E</b> VFQH IWDF LEQP
O15351	DGG <b>T</b> TFEH LWSS LEPD
Q27937	NSQ <b>E</b> TFNL LWDS LEQV

9aa TAD	<u>p53 TADI</u> D .F.. .F..	<u>p53 TADII</u> D .F.. .F..
Loss off function mutants*	QS	QS
P53_MOUSE	MEEQSQDISLELPLSQ <b>E</b> TFSG LWKL LPPEDIL ----PSPHCDLDDL-L-P	<b>Q</b> DVEE <b>FF</b> EG --
P53_RAT	MEDSQSDMSIELPLSQ <b>E</b> TFSC LWKL LPPDDILPTTATGSPNSMEDLFL-P	<b>Q</b> DVAE <b>LL</b> EG --
P53_CRIGR	MEEPQSDLSIELPLSQ <b>E</b> TFSD LWKL LPPNNVLSTL--PSSDSIEELFLS-	<b>E</b> NVTG <b>WL</b> ED SG
P53_MESAU	MEEPQSDLSIELPLSQ <b>E</b> TFSD LWKL LPPNNVLSTL--PSSDSIEELFLS-	<b>E</b> NVAG <b>WL</b> ED --
p53_MERUN	MEEPQSDLSIEPPLSQ <b>E</b> TFSD LWKL LPPKNLKSAL----EPMEDLLL-P	<b>Q</b> DVTS <b>WL</b> GD --
P53_BOVIN	MEEQSAELNVEPPLSQ <b>E</b> TFSD LWNL LPENNLSSSEL--S-APVDDLL PY	<b>T</b> DVAT <b>WL</b> DE C-
P53_DELLE	MEEQSAELGVEPPLSQ <b>E</b> TFSD LWKL LPENNLSSSEL--S-PAVDDLLLSP	<b>E</b> DVAN <b>WL</b> DE R-
P53_MARMO	MEEAQSDLSIEPPLSQ <b>E</b> TFSD LWNL LPENNVLSPLV--S-PPMDDLLLSS	<b>E</b> DVEN <b>WF</b> DK -G
P53_CANFA	MEEQSELNIDPPLSQ <b>E</b> TFSE LWNL LPENNVLSSEL--C-PAVDELL-L-P	<b>E</b> SVVN <b>WL</b> DE DS
P53_CAVPO	MEEPHSDLSIEPPLSQ <b>E</b> TFSD LWKL LPENNVLSDSL--S-PPMDHLLLSP	<b>E</b> EVAS <b>WL</b> GE N-
P53_CERAE	MEEPQSDPSIEPPLSQ <b>E</b> TFSD LWKL LPENNVLSPL--PS-QAVDDLMLSP	<b>D</b> DLAQ <b>WL</b> TE DP
P53_MACMU	MEEPQSDPSIEPPLSQ <b>E</b> TFSD LWKL LPENNVLSPL--PS-QAVDDLMLSP	<b>D</b> DLAQ <b>WL</b> TE DP
P53_HUMAN	MEEPQSDPSVEPPLSQ <b>E</b> TFSD LWKL LPENNVLSPL--PS-QAMDDLMLSP	<b>D</b> DIEQ <b>WF</b> TE DP
P53_TUPGB	MEEPQSDPSVEPPLSQ <b>E</b> TFSD LWKL LPENNVLSPL--PS-QAMDDLMLSP	<b>D</b> DIEQ <b>WT</b> ED -P
P53_XENLA	M.EPSSETGMDPPLSQ <b>E</b> TFED LWSL LPD.....	

\*(Uesugi et al., 1997; Choi et al., 2000, Venot et al., 1999, Lin et al., 1994)

\*\*<http://www.sanger.ac.uk/cgi-bin/Pfam/pfamget.pl?acc=PB003515>, adapted alignment

## Supplementary Table 3

### Transcription factors interacting with TAF9\*:

	TF region interaction with TAF9:	
Putative 9aa TAD		D .F.. .F..
p53	SVEPPLSQ	E TFSD LWKL.
HSF1	SHGFSVDT	S ALLD LFSP.
NF-IL6	PASSGQHH	D FLSD LFSD.
NFAT1	PGGSPQ	D ELDF SILF DY.
ALL1	DDCGNILP	S DIMD FVLK.
ESX	TSSSSDEL	S WIEE LLEK.
NF-κB	EDFS	S IADM DFSA LLSQ.

### Co-localization of region for TAF9 interaction and 9aa TAD in TFs:

TF	Accession Number	Length of TF	Region involved* in TAF9 interaction	9aa TADs prediction by pattern	9aa TADs considered by RC	Predicted 9aa TAD in TAF9 interaction region
p53	P04637	393 aa	5 aa	3	2	E TFSD LWKL
VP16	P04486	479 aa	4 aa	7	1	Q MFTD ALGI
HSF1	Q00613	529 aa	17 aa	12	4	S ALLD LFSP
NF-IL6	P17676	345 aa	17 aa	2	1	D FLSD LFSD
NFAT1	Q13469	925 aa	17 aa	2	1	D ELDF SILF
NF-κB	Q04206	551 aa	17 aa	7	2	S IADM DFSA
ALL1	Q03164	3969 aa	17 aa	2	2	S DIMD FVLK
ESX	P78545	371 aa	17 aa	6	2	S WIEE LLEK

\*(Uesugi et al., 1997; Choi et al., 2000; Venot et al., 1999; Lin et al., 1994)

### Transcription factor HSF1, loss-of function mutants:

Putative 9aa TAD		D .F.. .F..	Activity*
HSF1 - TAF9 interaction	SHGFSVDT	S ALLD LFSP	
Ga14BD-HSF1 (371-430)	..HGFSVDT	S ALLD LFSP	SVTVPD..
D416K	..HGFSVDT	S ALLK LFSP	SVTVPD..
F418A	..HGFSVDT	S ALLD LASP	SVTVPD..

\*(Brown et al., 1998)

### Transcription factor HSF1, synthetic orthologs:

Putative 9aa TAD

D .F.. .F..

HSF1_HUMAN	GFSVDT	S ALLD LFSP	SVTVPD
HSF1_CHICK	GFSVDT	T ALLD LFSP	SMTVTD
HSF_XENLA	SFSVDT	S ALMD LFSP	SLGIPD
AF159134_1	SINFDS	S PLFD IFSS	AASDVD
HSF2_HUMAN	QFSIDP	D LLVD LFTS	SVQMNP
HSF2_CHICK	QFSIDP	D LLFD LFTS	SVQMNP
HSF3_CHICK	QYSFGS	E AFSD VFNP	ELPALD
HSF_DROME	GVAIDQ	N MLMG LFND	SDLMDN

## Transcription factor NF-IL6, synthetic orthologs:

Putative 9aa TAD		D .F.. .F..		
NF-IL6 - TAF9 interaction	PASSGQHH	D	FLSD	LFSD
CEBB_HUMAN	PAPASSGQHH	D	FLSD	LFSD
CEBB_CHICK	PACSSGQ--	D	FLSD	LFAD
CEBB_BOVIN	PVPASSGQHH	D	FLSD	LFSD
CEBA_HUMAN	AYIDPAAFND	E	FLAD	LFQH
CEBA_BOVIN	AYIDPAAFND	E	FLAD	LFQH
CEBA_MOUSE	AYIDPAAFND	E	FLAD	LFQH
CEBD_HUMAN	AVPTLELCHD	E	LFAD	LFNS
CEBD_MOUSE	AVPTLELCHD	E	LFAD	LFNS
CEBE_HUMAN	LSAYIESGEE	Q	LLSD	LFAM
CEBE_RAT	LSAYIESGEE	Q	LLSD	LFAM

## 9aa TAD search result for NF-IL6:

Prosites Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {CGP} {CGP} pattern and filter settings for particular TF, restrictive for vertebrata/human.

**P49715**                    **CEBPA\_HUMAN**  
 75 - 83:                    **DEFLADLFQ**  
 76 - 84:                    EFLADLFQH  
 329 - 337:                EQLSRELDT



**P17676**                    **CEBPB\_HUMAN**  
 116 - 124:                **DFLSDLFSD**  
 318 - 326:                EQLSRELST



**P49716**                    **CEBPD\_HUMAN**  
 79 - 87:                    **DELFADLFN**  
 80 - 88:                    ELFADLFNS  
 217 - 225:                QEMQQLVE



**Q15744**                    **CEBPE\_HUMAN**  
 51 - 59:                    **EQLLSDLFA**  
 52 - 60:                    QLLSDLFAV  
 203 - 211:                DSLEYRERR  
 251 - 259:                EQLTQELDT





### Transcription factor NFAT1, synthetic orthologs:

Putative 9aa TAD		D	.F..	.F..
NFAT - TAF9 interaction	PGGSPQ	D	ELDF	SILF DY
NFC2_HUMAN	PGHEPGGSPQ	D	ELDF	SILF DYEYLN
NFC2_MOUSE	PGHEPGGSPQ	D	ELDF	SILF DYDYLN

### 9aa TAD search result for NFAT1:

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {CGP} {CGP} pattern and filter settings for particular TF, restrictive for vertebrata/human.

**Q13469**            **NFAC2\_HUMAN**  
 26 - 34:            **DELDFSILF**  
 906 - 914:         **DDVNEIIRK**



**Q60591**            **NFAC2\_MOUSE**  
 26 - 34:            **DELDFSILF**  
 1030 - 1038:      **GMVLLTLMH**  
 1031 - 1039:      **MVLLTLMHH**



### Transcription factor ALL1, synthetic orthologs:

Putative 9aa TAD		D	.F..	.F..
ALL1 - TAF9 interaction	DDCGNILP	S	DIMD	FVLK
HRX_HUMAN	SDDCGNILP	S	DIMD	FVLK NTPSMQA
gb AAC41377.1	SDDCGNILP	S	DIME	FVLN TPSMSM
MLL2_HUMAN	RARPPEDLP	S	EIVD	FVLK N LGGPG

### Transcription factor ESX (ELF3), synthetic orthologs:

Putative 9aa TAD		D	.F..	.F..
ESX - TAF9 interaction	TSSSSDEL	S	WIIE	LLEK
ESX_HUMAN (E74F3)	DLTSSSSDEL	S	WIIE	LLEK DGMAFQ
ESX_MOUSE	DLTSNNSDEL	S	WIIE	LLEK DGMSFQ

## Transcription factor NF-κB (TF65), synthetic orthologs:

Putative 9aa TAD		D	.F.D	.F..	.F..
hTF65 - TAF9 interaction	ED FS	S	IADM	DFSA	LLSQ
TF65_MOUSE	GDED FS	S	IADM	DFSA	LLSQ ISS.
TF65_CHICK	TEDS LP	S	LGDL	DFSA	FLSQ FPSS.
TF65_XENLA	REE IHLT	S	LFEL	DFSS	LLSN MK.

\*2nd (overlapping) 9aa TAD motif in brown

## Transcription factor NF-κB (TF65), experimental data:

Putative 9aa TAD		D	.F..	.F..	.F..	Activity*	
LexA-TF65 (537-47)	B49	n	IADM	DFSA	LLS.	+	
Gal4-TF65 (521-51)	ref. *	LSGDEDFS	S	IADM	DFSA	LLSQISS.	+

\*(Schmitz et al., 1995)

## 9a AD search result for STAT1/3 (requested by P. Kovarik 2006):

ScanProsite Pattern searching (<http://www.expasy.ch/tools/scanprosite/>) was done with [DE]

{NQKRHCGP}{ILVFW}{NQKRHCGP}{NQKRHCGP}{NQKRHCGP}{ILVFWM}{NQKRHCGP}{KRHCGP} optimized pattern and filter settings for vertebrata and STAT3\_.

[DE] {NQKRHCGP}{ILVFW} {NQKRHCGP}{NQKRHCGP}{NQKRHCGP}{ILVFWM}{NQKRHCGP}{KRHCGP} [MDENQSTYG]{KRHCGP}{ILVFWM} {KRHCGP} {CGP} {CGP}{ILVFWM}{CGP}{CGP}

P40763 STAT3\_HUMAN  
753 761: E SLTF DMEL

## Synthetic orthologs:

STAT3:		D	F	F		D	F	F
<a href="#">STAT3_HUMAN/687_763</a>	DL PMSP	R	TLDS	LMQFGNNGE	GAEPSAGGQF	E	SLTF	DMEL TS
<a href="#">Q8CFJ6_MOUSE/169_245</a>	DL PMSP	R	TLDS	LMQFGNNGE	GAEPSAGGQF	E	SLTF	DMDL TS
<a href="#">STAT3_MOUSE/687_763</a>	DL PMSP	R	TLDS	LMQFGNNGE	GAEPSAGGQF	E	SLTF	DMDL TS
<a href="#">STAT3_RAT/687_763</a>	DL PMSP	R	TLDS	LMQFGNNGE	GAEPSAGGQF	E	SLTF	DMDL TS
<a href="#">Q9PVX8_XENLA/687_762</a>	DL PMSP	R	TLDS	LMQFP	GE GADSSAGNQG	E	TLTF	DMEL TS
<a href="#">Q7ZXX3_XENLA/687_759</a>	DL PMSP	G	TFDS	VMQFP	GE GSESGNGNQF	E	TLTF	DVDL PS
<a href="#">Q90Y16_TETFL/688_764</a>	DLFPMSP	R	TLDS	LMHN	E A EANPGH L	D	SLTL	EMDV ASPM
<a href="#">O13133_ONCMY/688_767</a>	DLFPMSP	R	TLNS	LMHN	K A EANPGP L	D	SLTL	DMEL SSDVASPM
<a href="#">O93599_BRARE/688_782</a>	DLFPMSP	R	TLDS	LMHN	E AAEANPGP L	E	SLTL	DMEL SS
STAT1:		D	F	F		D	F	F
<a href="#">Q8JGN0_XENLA/712_746</a>	FLLPMSP	D	EFDE	VARL	VGP	A	EFDT	VMCS Y
<a href="#">Q8C3V4_MOUSE/713_749</a>	NLLPMSP	E	EFDE	MSRI	VG	P	EFDS	MMST V
<a href="#">Q8C8M3_MOUSE/719_755</a>	NLLPMSP	E	EFDE	MSRI	VG	P	EFDS	MMST V
<a href="#">STAT1_MOUSE/713_749</a>	NLLPMSP	E	EFDE	MSRI	VG	P	EFDS	MMST V

<a href="#">Q9D323_MOUSE/713_749</a>	NLLPMSP	<b>E</b>	<b>EFDE</b>	MSRI	VG	<b>P</b>	<b>EFDS</b>	MMST	V
<a href="#">Q8C497_MOUSE/713_749</a>	NLLPMSP	<b>E</b>	<b>EFDE</b>	MSRI	VG	<b>P</b>	<b>EFDS</b>	VMST	V
<a href="#">Q9QXK0_RAT/713_749</a>	NLLPMSP	<b>E</b>	<b>EFDE</b>	MSKI	VG	<b>S</b>	<b>EFDS</b>	MMSA	V
<a href="#">STAT1_HUMAN/713_750</a>	NLLPMSP	<b>E</b>	<b>EFDE</b>	VSRI	VGS	<b>V</b>	<b>EFDS</b>	MMNT	V

## Supplementary Table 4

### 9aa TAD search result for KLF4/2:

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {CGP} {CGP} pattern and filter settings for KLF4/2.

**O43474**                    **KLF4\_HUMAN**  
 88 - 96:                    EEFN~~D~~LLDL  
 92 - 100:                   **DL**LDL**DF**IL  
 109 - 117:                   ESVAATVSS



**Q9Y5W3**                    **KLF2\_HUMAN**  
 39 - 47:                    D~~D~~ENSVLDF  
 42 - 50:                   **NS**VLD**F**ILS  
 43 - 51:                    SVLDF~~F~~LSM



### Experimental data, loss-of function mutants of KLF4:

Putative 9aa TAD	D . F . . . F . .	Activity*
Gal4bd-hKLF4 (1-109)	. . RRETEEFN <b>D</b> LLDL <b>DF</b> IL SN.	+
EEE93, 95, 96VVV	. . RR <b>v</b> T <b>v</b> v <b>F</b> N <b>D</b> LLDL <b>DF</b> IL SN.	-
DDD99, 102, 104VVV	. . RRETEEFN <b>v</b> LL <b>v</b> L <b>v</b> FIL SN.	-

\*(Geiman et al., 2000; Zhang et al., 2000)

### Synthetic orthologs /paralogs:

Putative 9aa TAD	D . F . . . F . .
KLF4_HUMAN	EEFN <b>D</b> LLDL <b>DF</b> IL SNSL
DANRE	DDLS <b>K</b> FLDL <b>EF</b> IL SNTV
Q90XF0	EDLD <b>K</b> YLDL <b>EF</b> IL ANTA
XENLA	DDL <b>G</b> <b>K</b> FVDL <b>DF</b> IL AHTS
Q90XE6	DDSE <b>G</b> CWDM <b>EF</b> LL SDWA
Q90XE7	EDLS <b>N</b> YLDL <b>EF</b> IL ANTT
KLF2_HUMAN	DDL <b>N</b> <b>S</b> VLDF <b>IL</b> SM GLDG
KLF2_MOUSE	EDLN <b>N</b> VLDF <b>IL</b> SM GLDG

## Supplementary Table 5

### Experimental data, loss-of function mutants of E1A:

Putative 9aa TAD			D .F.. .F..	Activity*
Gal4bd-E1A (1-29)	MRTEMTPLVL	SYQEA	D DILE HLVD NFFNE .	+
Gal4bd-E1A (1-29) P7A	MRTEMTaLVL	SYQEA	D DILE HLVD NFFNE .	-
Gal4bd-E1A (1-29) Y12A	MRTEMTPLVL	SaQEA	D DILE HLVD NFFNE .	+
Gal4bd-E1A (1-29) D16A	MRTEMTPLVL	SYQEA	a DILE HLVD NFFNE .	-
Gal4bd-E1A (1-29) D17A	MRTEMTPLVL	SYQEA	D aILE HLVD NFFNE .	+
LexA-Gal4 (860-71), B1		nnnTM	D DVYN YLFD Di .	+

\*(Lipinski et al., 1998)

### Synthetic orthologs:

Putative 9aa TAD				D .F.. .F..	
E1A_ADECT	MKYTI	VP APR	NLH	D YVLE LLEE	WQPD
E1A_ADECR	MKLTLEP	APR	CLH	E YVSQ LLED	WQPE
E1A_ADEC2	MKYTI	VP APR	NLH	D YVLE LLEE	WHPD
E1A_ADE41	MRM	LPDFF	TGNW	D DMFQ GLEE	AEHPF
E1A_ADE40	MRM	LPDFF	TGNW	D DMFQ GLEE	TEYVF
E1A_ADE12	MRTEM	TP LVL	SYQEA	D DILE HLVD	NFFNE
E1A_ADE05	MRHII	CHGG	VIT	E EMAA SLLD	QLIEE
E1A_ADE02	MRHII	CHGG	VIT	E EMAA SLLD	QLIEE
E1A_ADEM1	MSRLRLSLSSRVW	LAA	Q	EATR NVSE	DPV
E1A_ADE07	MRHLRFLPQ	EIISSETGI	E	I LEF VVNT	LMGD
E1A_ADE04	MRHLRDLDPDEEIIIA	SGS	E	I LEL VVNA	TMGD
E1A_ADES7	MRHLAE	MIS	E	L LDL GLDT	IDGW

## Supplementary Table 6

### 9aa TAD search result for ETV1/PEA3/ERM (ETS family):

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {CGP} {CGP} pattern and filter settings for ETV1 and PEA3.

#### P50549 ETV1\_HUMAN (ER81)

44 - 52: **EELFQDLSQ**  
 51 - 59: **SQLQETWLA**  
 336 - 344: **QLWQFLV~~AE~~**  
 339 - 347: **QFLV~~AE~~LDD**  
 436 - 444: **TDMERHINE**



#### Q9PUQ1 PEA3\_BRARE

49 - 57: **EDLFQDLSQ**  
 56 - 64: **SQLQEFWLT**  
 351 - 359: **QLWQFLV~~AE~~**  
 354 - 362: **QFLV~~AE~~LDD**



#### P41161 ETV5\_HUMAN (Ets-related protein ERM)

44 - 52: **EELFQDLSQ**  
 51 - 59: **SQLQEAWLA**  
 369 - 377: **QLWQFLVTL**  
 372 - 380: **QFLVTL~~AE~~LDD**



### Experimental data, loss-of function mutants of ERM:

Gal4bd-ERM-TAD (1-72)*	.. <b>E</b> <b>ELFQ</b> <b>DLSQ</b> ..
L50P, loss-of mutant*	.. <b>E</b> <b>ELFQ</b> <b>DPSQ</b> ..
F47P, loss-of mutant*	.. <b>E</b> <b>ELLQ</b> <b>DLSQ</b> ..

\* ERM-TAD (1-72) activates transcription in both mammalian and yeast (Defossez et al., 1997)

### Synthetic orthologs:

PEA3_BRARE	ELPPQES <b>E</b> <b>DLFQ</b> <b>DLSQ</b> LQ
------------	---------------------------------------------

Q96AW9_HUMAN	SLPPLDS E DLFQ DLSH FQ
ETV4_MOUSE	SLPPSDS E DLFQ DLSH FQ
Q9PUQ2_BRARE	TELAQDT E ELFQ DLSQ LQ
ERM_HUMAN	TDLAHS E ELFQ DLSQ LQ
ETV1_HUMAN	RDLAHS E ELFQ DLSQ LQ

## Supplementary Table 7

### Experimental data, loss-of function mutants of plants HSF - HsfA:

Putative 9aa TAD	D .F.. .F..	Activity*
HsfA2ΔC323	..MEPVAD D IWEE LLSE DL..	+
HsfA2ΔC323, ME289VD	..vdPVAD D IWEE LLSE DL..	+
HsfA2ΔC323, A293D	..MEPVGD D IWEE LLSE DL..	+
HsfA2ΔC323, W297A	..MEPVAD D IaEE LLSE DL..	-
HsfA2ΔC323, W297L	..MEPVAD D I1EE LLSE DL..	+
HsfA2ΔC323, W297F	..MEPVAD D IfEE LLSE DL..	+
HsfA2ΔC323, IW296WI	..MEPVAD D wiEE LLSE DL..	+
HsfA2ΔC323, L300P	..MEPVAD D IWEE pLSE DL..	-
HsfA2ΔC323, L300A	..MEPVAD D IWEE aLSE DL..	-
HsfA2ΔC323, L301A	..MEPVAD D IWEE LaSE DL..	-

\*(Doring et al., 2000)

### Conservation of putative 9aa TAD in synthetic plants HSF orthologs:

Putative 9aa TAD	D .F.. .F..
Q9LUH8_ARATH	KENNEIYGE G FWED LLNE GQN
Q9M1V5_ARATH	DGVERELDD G FWEE LLMN NEN
Q9SV12_ARATH	LVVERELDD G FWEE LLSD ESL
Q6VBB2_ORYSA	IGPQGELND N FWEE LLNE GLV
Q6F388_ORYSA	KPSNGELNE D FWED LLHE GGL
Q8H7Y6_ORYSA	ASNETELTD D FWEE LLNE GAR
Q84T61_ORYSA	TQKLPALID S FWEQ FLVA SPL
Q84MN7_ORYSA	GGGGGGDTE S FWMQ LLSL GLE
Q7XEJ7_ORYSA	TNGQAEGLD D FWAE LLVE DFT
Q564C8_SOYBN	DIDDLLNPN H FWDD ILRT PVS
Q4L0F7_MEDSA	ADIDLLSN S IWDD LLQT PIP
O82042_PEA	DPSANESNL S DWEE LLNQ KLV
HSF30_LYCPE	GTDMEPVAD D IWEE LLSE DLI



## Supplementary Table 8

### Artificial transactivation domains - AH, loss-of function mutants:

Putative 9aa TAD	D .F.. .F..	Activity*
AH	ELQEL Q ELQA LLQQ Q	+
AH II (1)	EVSEV L ELEA LLQQ Q	-
AH I (b7)	ELQEL Q WLQA LLQQ Q	+
AH II (b3)	ESYEL Y ELIA LLQQ Q	+
AH II (b12)	ELLEET S EWYA LLQQ Q	+
Oaf1 (1023-47)	YFGGL D LFDY DFLF GNDFA.	+

\*(Lu et al., 2000)

### Artificial transactivation domain - P201, loss-of function mutants:

Putative 9aa TAD	D .F.. .F..	Activity*
P201	..KALLT G LEVQ DYLL PTCIP	+
P201, L2A	..KALLT G aFVQ DYLL PTCIP	+
P201, F3A	..KALLT G LaVQ DYLL PTCIP	-
P201, V4A	..KALLT G LEaQ DYLL PTCIP	+
P201, Q5A	..KALLT G LEV a DYLL PTCIP	+
P201, D6A	..KALLT G LEVQ aYLL PTCIP	+
P201, Y7A	..KALLT G LEVQ DaLL PTCIP	-
P201, L8A	..KALLT G LEVQ DY aL PTCIP	+
P201, L9A	..KALLT G LEVQ DYLa PTCIP	+
P201, P10A	..KALLT G LEVQ DYLL aTCIP	+
P201, T11A	..KALLT G LEVQ DYLL PaCIP	+
P201, C12A	..KALLT G LEVQ DYLL PTaIP	+
P201, I13A	..KALLT G LEVQ DYLL PTCaP	+
P201, P14A	..KALLT G LEVQ DYLL PTCIa	+

\*(Giniger and Ptashne, 1987; Ruden, 1992)

## Supplementary Table 9

### 9aa TAD search result for EBNA2:

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {CGP} {CGP} pattern and filter settings for EBN2\_.

**P12978**                      **EBN2\_EBV**  
456 - 464:                      **ESWDYIFET**



### Experimental data, loss-of function mutants of EBNA2:

Putative 9aa TAD	D .F.. .F..	Activity*
Gal4-EBNA2 (449-62)	DLD E SWDY IFET TE .	+
Gal4-EBNA2 (449-83)	EDL- E SWDY IFET TE . .	+
Gal4-EBNA2 (449-83) W454A	EDL- E SaDY IFET TE . .	-
Gal4-EBNA2 (449-83) W454L	EDL- E S1DY IFET TE . .	-
Gal4-EBNA2 (449-83) W454F	EDL- E SFDY IFET TE . .	+

\*(Cohen, 1992; Cohen et al., 1991)

### Synthetic orthologs:

Putative 9aa TAD	D .F.. .F..
EBNA2, RHESUS	IEPPDLEDID E NWDD IFNV SDSEH
EBNA2, PAPIO	LQSP D N QWED IFNF SESSS
EBNA2, HUMAN, typ1	IDPA DLD E SWDY IFET TESPS
EBNA2, HUMAN, typ2	LEPA ELD E SWEG IFET TESHS

## Supplementary Table 10

### 9aa TAD search result for VP16:

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {CGP} [ILVFWM] {C} {C} pattern and filter settings for ATIN\_HHV11 and ATIN\_HHV1F.

#### ATIN\_HHV1F (P04486)

71 - 79: DTWNE~~D~~LFS  
 75 - 83: EDLFSAL~~P~~T  
 140 - 148: EALSRRFFHA  
 166 - 174: SALYRYLRA  
 323 - 331: GYFMV~~L~~RA  
 429 - 437: DDFDL~~D~~M~~L~~LG  
 466 - 474: QM~~F~~T~~D~~ALGI

#### ATIN\_HHV11 (P06492)

2 - 10: D~~L~~~~L~~VDELFA  
 6 - 14: DELFA~~D~~MNA  
 82 - 90: DTWNE~~D~~LFS  
 86 - 94: EDLFSAL~~P~~T  
 151 - 159: EALSRRFFHA  
 177 - 185: SALYRYLRA  
 334 - 342: GYFMV~~L~~RA  
 440 - 448: DDFDL~~D~~M~~L~~LG  
 477 - 485: QM~~F~~T~~D~~ALGI

### Experimental data, loss-of function mutants of VP16 and interaction with TAF9:

Putative 9aa TAD:	D .F. . .F. .				
VP16-AD - TAF9 interaction	D	F	L	D	Activity*
VP16	MADFEFE	Q M <del>F</del> TD ALGI	DEYGG.		+
VP16	MADFEFE.				-
VP16	MADFE <del>S</del> E	Q <del>t</del> sTD ALGI	DEYGG.		-

\*(Sullivan et al., 1998; Uesugi et al., 1997)

### Synthetic orthologs:

Putative 9aa TAD	D .F. . .F. .				
ATIN_HHV1F	ALDMADFEFE	Q M <del>F</del> TD ALGI	DEY GG.		
ATIN_HHV11	ALDMADFEFE	Q M <del>F</del> TD ALGI	DEY GG.		
ATIN_HHV2H	ALDVDDFEFE	Q M <del>F</del> TD AMGI	DDF GG.		
Q9ELS1	SLDMADFEFE	Q M <del>F</del> TD ALGI	DEF GG.		
Q805Y1	SLDMADFEFE	Q M <del>F</del> TD ALGI	DDI GG.		

\*<http://www.sanger.ac.uk/Software/Pfam/data/jtml/full/PF02232.shtml>

## Supplementary Table 11

### Experimental data, loss-of function mutants of Leu3:

Putative 9aa TAD		D	.F..	.FD.	F..	.F..*	Activity*
LexA-Leu3 (860-886)	B43	.GWDNWES	D	MVWR	DVDI	LMN EFAFNPKV.	++
LexA-Leu3 (860-875)	B44	.GWDNWES	D	MVWR	DVDI.		-
LexA-Leu3 (860-878)	B45	.GWDNWES	D	MVWR	DVDI	LMN.	+/-
LexA-Leu3 (864-878)	B46	.WES	D	MVWR	DVDI	LMN.	-
LexA-Leu3 (868-878)	B48	nn	n	MVWR	DVDI	LMN.	-
Leu3	ref.*	.GWDNWES	D	MVWR	DVDI.		+
Leu3	ref.*	.GWDNWES	D	MV.			-

\*(Wang et al., 1997)

### Synthetic orthologs:

Putative 9aa TAD		D	.F..	.FD.	F..	.F..*
<i>Saccharomyces cerevisiae</i>	WDNWES	D	MVWR	DVDI	LMN	EFAF NPKV.
<i>Candida glabrata</i>	WDNWES	D	VVWK	DVDI	LMN	EFAF NPTL.
<i>Kluyveromyces lactis</i>	WDNWES	D	LVWK	DVDI	LMN	EFAF NPTL.
<i>Ashbya gossypii</i>	WDEWES	D	LIFK	DVGI	LMN	EFAF NPEVL.
<i>Candida albicans</i>	IDNFDH	D	MLWK	DVDS	VMN	DFGF HTS.
<i>Yarrowia lipolytica</i>	MDNWDL	D	SIFK	DVDS	IMN	DFGF RADEAGVAVNE.
<i>Saccharomyces castellii</i>	FDNWES	E	MVWK	DVDI	LMN	EFAF NPTV.
<i>Saccharomyces castellii</i>	LETLES	E	NIWK	DVDM	LMN	EFAF NPTI

\*2nd (overlapping) 9aa TAD motif in brown

## Supplementary Table 12

### Experimental data, loss-of function mutants of Gcn4:

Putative 9aa TADI	D .F.. .F..	Activity*
Gcn4 (1-100,170-281)	..LD D AVVE SFFS S.	++
Gcn4 (1-94,170-281)	..LD D AVV.	-
Gcn4 (1-100,170-281) F97A	..LD D AVVE SaFS S.	+
Gcn4 (1-100,170-281) F98A	..LD D AVVE SFaS S.	+
Gcn4 (1-100,170-281) F97AA	..LD D AVVE SaaS S.	-
Putative 9aa TADII	D .F.. .F..	
Gcn4 (1-17,95-117,170--281)	.ESFFSSTDSPMFYEYENLEDNS.	-
Gcn4(1-17,95-134, 170-281)	.ESFFSSTDSPMFYEYENLEDNS K EWTS LFDN DIP..	+
Gcn4 (1-17,95-134, 170-281) W120A	.ESFFSSTDSPMFYEYENLEDNS K EaTS LFDN DIP..	-
Gcn4(1-17, 95-134, 170-281) WLF120A	.ESFFSSTDSPMFYEYENLEDNS K EaTS aaDN DIP..	-

\*(Drysdale et al., 1995; Hope et al., 1988; Hope and Struhl, 1986), TADI: 1-100 aa, TADII: 95-134 aa, DNA binding domain: 170-281 aa

### 9aa TAD search result for Gcn4:

Prosite Pattern searching was done with [MDENQSTYG] {KRHCGP} [ILVFWM] {KRHCGP} {CGP} {KRHCGP} [ILVFWM] [ILVFWMAY] {KRHCP} pattern and filter settings for Gcn4.

**P03069**                      **GCN4\_YEAST**  
 63 - 71:                      SNLDFDFAL  
 91 - 99:                      DAVVESFFS



### Synthetic orthologs:

Putative 9aa TADI	D .F.. .F..	
<i>Saccharomyces cerevisiae</i>	ELD D AVVE SFFS SSTD	
<i>Saccharomyces mikatae</i>	ELD A AVVE SFFS SSTD	
<i>Saccharomyces kluyveri</i>	QLN S AVVD TFFS SSTD	
<i>Saccharomyces castellii</i>	ELD S NVVD AFFS SSTD	
<i>Candida glabrata</i>	DID A AVVD AFFS SSTD	
<i>Kluyveromyces lactis</i>	ELN A SIVE TFFD AASSTD	
<i>Ashbya gossypii</i>	DLD S AVEN FFGS SSD	
<i>Debaryomyces hansenii</i>	NIH S SVLD SVFS ST	
<i>Candida dubliniensis</i>	QIHS S VLES VFST N	
<i>Candida albicans</i>	QIHS S VLES GFST N	
Putative 9aa TADII		
<i>Saccharomyces cerevisiae</i>	DNS K EWTS LFDN DIP	
<i>Ashbya gossypii</i>	ADP K TWSS LFDD DIP	
<i>Candida glabrata</i>	DNP E NWTS LFDD DVE	
<i>Kluyveromyces lactis</i>	GGV E TWTS LFDN DIP	




## Supplementary Table 13

### 9aa TAD search result for yeast Pho4:

Prosite Pattern searching was done with [MDENQSTYG]  
{KRHCGP}[ILVFWM]{KRHCGP}{CGP} {CGP}[ILVFWM]{CGP}{CGP} pattern and  
Pho4 sequence input.

P07270                      **Pho4**  
75 - 83:                      **DDLDRAFEL\***



\*) The transactivation domain was localized to 75–85 aa in previous study (Jayaraman et al., 1994; Ogawa and Oshima, 1990)

### Experimental data, loss-of function mutants of Pho4:

Putative 9aa TAD	D .F.. .F..	Activity*
Pho4 (75-83)	D DLDR AFEL V	100%
Pho4 (75-83) F78A	D DLAR AFEL V	75%
Pho4 (75-83) F79A	D DLDA AFEL V	180%
Pho4 (75-83) F79P	D DLDP AFEL V	7%
Pho4 (75-83) F81A	D DLDR AAEL V	17%
Pho4 (75-83) F81W	D DLDR AWEL V	90%
Pho4 (75-83) F81Y	D DLDR AYEL V	71%
Pho4 (75-83) F81P	D DLDR APPEL V	14%

\*(McAndrew et al., 1998)

Remarks: similarity Pho4 to Tea1 9aa TAD (Gal4 family)

9aa TAD of Pho4:                      D DLDR AFEL

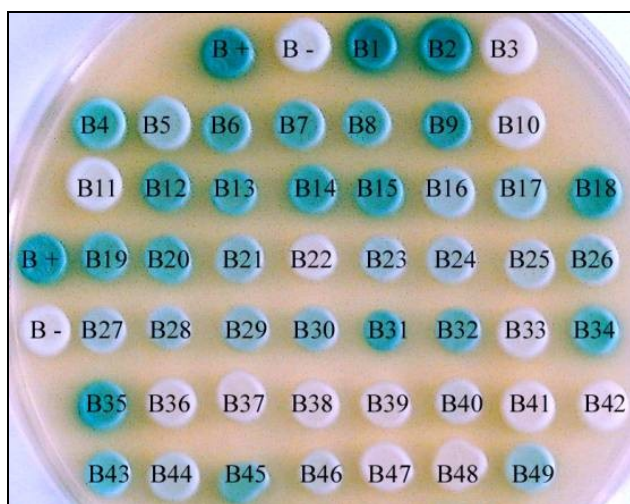
9aa TAD of Tea1:                      D DLIR ELFG

## Supplementary Table 14

### Promiscuity of the Oaf1 9aaTAD

Construct	ID	D	.F..	.F..	Activity
LexA-Oaf1(902-1047)	B66	. .NTPFPGYFGGL D LFDY DFLF GNDFA			++
LexA-Oaf1(902-1036)	B101	. .NTPFPGYFGGL D LF			-
LexA-Oaf1(902-1047)Y38Q	B21	. .NTPFPGYFGGL D LFD <sup>q</sup> DFLF GNDFA			++
LexA-Oaf1(902-1047)DYD37GGG	B22	. .NTPFPGYFGGL D LF <sup>gg</sup> gFLF GNDFA			-
LexA-Oaf1(902-1047)N44K	B23	. .NTPFPGYFGGL D LFDY DFLF G <sup>k</sup> DFA			++
LexA-Oaf1(902-1047)N44Q	B24	. .NTPFPGYFGGL D LFDY DFLF G <sup>q</sup> DFA			++
LexA-Oaf1(902-1047)D45E	B25	. .NTPFPGYFGGL D LFDY DFLF GN <sup>e</sup> F A			++
LexA-Oaf1(902-1047)D37A	B26	. .NTPFPGYFGGL D LF <sup>a</sup> Y DFLF GNDFA			++
LexA-Oaf1(902-1047)D37E	B27	. .NTPFPGYFGGL D LF <sup>e</sup> Y DFLF GNDFA			++
LexA-Oaf1(902-1047)D39A	B28	. .NTPFPGYFGGL D LFDY aFLF GNDFA			++
LexA-Oaf1(902-1047)D39E	B29	. .NTPFPGYFGGL D LFDY eFLF GNDFA			++
LexA-Oaf1(902-1047)F40A	B30	. .NTPFPGYFGGL D LFDY DaLF GNDFA			++
LexA-Oaf1(902-1047)F42A	B31	. .NTPFPGYFGGL D LFDY DFLa GNDFA			+++
LexA-Oaf1(902-1047)FF40, 2AA	B32	. .NTPFPGYFGGL D LFDY DaLa GNDFA			+++

## Supplementary Figure



**Transactivation of 9aa TAD in yeast:** The reporter activity in yeast strains B1-49 was co-monitored on the YPD X-gal plate beside to the liquid  $\beta$ -galactosidase measurement. B- represents strain with empty vector pBTM116



(negative control) and B+ represents strain with pBTM116-VP16 (robust viral transactivator, positive control).  $\beta$ -galactosidase activity correlates with intensity of blue staining.