

Data Set 2

RSA-TOOLS Results

LORE search: Using 750 bp of promoter and ACTCAACAA sequence as a seed

Training set: *IZH4*, *IZH2*, *HSP26*, *YGL039w*, *ERG3*, *PIR3*, *YO338w*, *HSP30*, *COS10*,
NCE103, *AHP1*, *YGR161c*, *PDR3*, *OLE1*, *MGA2*, *HSP104*

Matrix Length: 9

Complement strand: include as a single sequence

At least one match in each sequence

MATRIX 3

Number of sequences = 16

Unadjusted information = 33.5596

Sample size adjusted information = 33.5596

$\ln(P \text{ value}) = -229.618$ $P \text{ value} = 1.89667\text{E-}100$

$\ln(\text{expected frequency}) = -120.061$ $\text{expected frequency} = 7.21564\text{E-}53$

A | 11 0 2 0 16 15 0 15 13

C | 0 13 0 16 0 0 14 0 0

G | 1 2 2 0 0 0 2 1 2

T | 4 1 12 0 0 1 0 0 1

1|1 : 1/1 ACTCAACAA

2|2 : 2/595 TGTCAACAG

3|3 : -3/636 GCTCAACAA

4|4 : -4/544 ACTCAAGAA

5|5 : 5/707 ATTCAACAA

6|6 : -6/254 TCGCAACAA

7|7 : 7/135 AGACAACAA

8|8 : 8/69 TCTCAAGAA

9|9 : 9/644 ACACAACAA

10|10 : 10/615 ACTCAACAG

11|11 : -11/283 ACTCATCAA

12|12 : -12/560 TCTCAACAA

13|13 : 13/496 ACTCAACGA

14|14 : -14/300 ACGCAACAA

15|15 : 15/408 ACTCAACAA

16|16 : 16/565 ACTCAACAT

Map	Type	ID	Strand	Start	End	Sequence	Score	ln(P)
TIS11	site	patser	R	-304	-296	tctcTTCTTGACAgcac	5.72	-7.74
NCE103	site	patser	R	-182	-174	tctcTCGATGAGTaatt	5.74	-7.76
FIT2	site	patser	D	-698	-690	actgACTCAACGTtgat	5.88	-7.84
HSP26	site	patser	D	-16	-8	acaaACTAAACAAatta	5.95	-7.9
UBS1	site	patser	D	-709	-701	gtcgAATCAACAAcgaa	6.15	-8.08
YOL101c	site	patser	D	-220	-212	ggcaGCTCAACATtatt	6.18	-8.13
YOL101c	site	patser	D	-157	-149	atggTGTC AACAGctaa	6.19	-8.14
ERG3	site	patser	R	-55	-47	cagtTTGTTGCATttgt	6.3	-8.25
AHP1	site	patser	D	-50	-42	gaaaTTTCAACAAacca	6.65	-8.61
PIR3	site	patser	D	-617	-609	cttaAGACAACAacggg	6.87	-8.87
YOL101c	site	patser	R	-197	-189	aagcCTGTTGAGCtcaa	7.11	-9.05
OLE1	site	patser	R	-508	-500	aaagCTGTTGAGCgagc	7.11	-9.05
YOR338w	site	patser	D	-683	-675	agacTCTCAAGAAatccg	7.51	-9.28
COS10	site	patser	D	-111	-103	gcgcTGTC AACAAattd	7.58	-9.35
YGL039w	site	patser	D	-45	-37	tgagATTCAACAAAaat	7.62	-9.43
NCE103	site	patser	R	-469	-461	ttggTTGATGAGTgacc	7.88	-9.66
MGA2	site	patser	D	-187	-179	aaaaACTCAACATtttc	8.02	-9.82
ERG3	site	patser	R	-498	-490	ccttTTGTTGCGAttgt	8.06	-9.93
YGR161c	site	patser	D	-256	-248	ctgtACTCAACGAtcat	8.2	-10.02
HSP26	site	patser	R	-208	-200	cacgTTCTTGAGTggta	8.48	-10.11
YOL002c	site	patser	R	-116	-108	tccgTTGTTGAGCtctt	8.5	-10.26
HSP30	site	patser	D	-108	-100	tacaACACAACAaaaa	8.66	-10.59
COS10	site	patser	D	-137	-129	taaaACTCAACAGctgc	8.95	-10.85
PDR3	site	patser	R	-452	-444	gagaTTGTTGCGTaagc	9.03	-11.07
FTR1	site	patser	D	-521	-513	acgtTCTCAACAAtttt	9.37	-11.36
AHP1	site	patser	R	-192	-184	cataTTGTTGAGAggtt	9.37	-11.36
OLE1	site	patser	D	-344	-336	gaacACTCAACAAacct	10.34	-12.05
YMR045C	site	patser	R	-550	-542	tcgtTTGTTGCCTtccc	7.23	-9.18
YMR046C	site	patser	R	-550	-542	tcgtTTGTTGCCTtccc	7.23	-9.18
YJR026W	site	patser	D	-392	-384	tttaTCGCAACAcaat	8.06	-9.93
YJR027W	site	patser	D	-392	-384	tttaTCGCAACAcaat	8.06	-9.93

Aft1p-binding site search: Using 750 b.p. of promoter and TGCACCCA sequence as a seed

Training set: *FIT3*, *FIT2*, *TAF1*, *TIS11*, *ENB1*, *ARN1*, *FTR1*, *FRE1*, *SIT1*, *FET3*, *HMX1*

Matrix Length: 8

Complement strand: include as a single sequence

At least one match in each sequence

MATRIX 1

Number of sequences = 11

Unadjusted information = 29.6925

Sample size adjusted information = 29.6924

$\ln(P \text{ value}) = -132.3$ $P \text{ value} = 3.49169E-58$

$\ln(\text{expected frequency}) = -59.2477$ $\text{expected frequency} = 1.85808E-26$

A		0	1	0	11	0	0	0	6
C		0	0	11	0	11	11	11	0
G		0	10	0	0	0	0	0	0
T		11	0	0	0	0	0	0	5

1 1	:	1/1	TGCACCCT
2 2	:	-2/566	TGCACCCT
3 3	:	3/485	TGCACCCT
4 4	:	-4/303	TGCACCCA
5 5	:	5/502	TGCACCCA
6 6	:	6/558	TGCACCCA
7 7	:	7/491	TGCACCCA
8 8	:	-8/385	TGCACCCT
9 9	:	9/528	TACACCCA
10 10	:	-10/562	TGCACCCT
11 11	:	11/500	TGCACCCA

Map	Type	ID	Strand	Start	End	Sequence	Score	$\ln(P)$
FIT2	site	patser	D	-250	-243	caaaTGCACCCCcct	7.41	-9.81
ENB1	site	patser	D	-138	-131	caatTGCACCCCtctg	7.41	-9.81
FTR1	site	patser	D	-176	-169	tgaaTGCACCCGattd	7.41	-9.81
YLR205c	site	patser	R	-583	-576	ggtgCGGGTGCAagtt	7.41	-9.81
YLR205c	site	patser	R	-539	-532	agtcAGGGTGTAtgtg	7.82	-10.05
FIT2	site	patser	R	-674	-667	gtagTGGGTGTActgt	7.99	-10.41
FRE1	site	patser	D	-224	-217	ctaaTACACCCAattd	7.99	-10.41
FIT3	site	patser	D	-286	-279	ttttTGCACCCTgttd	10.28	-10.97
FIT3	site	patser	R	-186	-179	ggaaAGGGTGCAaaaa	10.28	-10.97
FIT2	site	patser	R	-333	-326	gagcAGGGTGCAattd	10.28	-10.97
FIT2	site	patser	D	-267	-260	acaaTGCACCCTaagt	10.28	-10.97
FTR1	site	patser	R	-367	-360	ggccAGGGTGCAgaaa	10.28	-10.97
SIT1	site	patser	R	-190	-183	tttgAGGGTGCAatct	10.28	-10.97
FIT2	site	patser	R	-133	-126	taacTGGGTGCAgtat	10.46	-11.66

TAF1	site	patser	R	-449 -442aaaaTGGGTGCAgttc	10.46 -11.66
TIS11	site	patser	D	-250 -243aaaaTGCACCCAcatt	10.46 -11.66
TIS11	site	patser	R	-175 -168cagtTGGGTGCAAttgt	10.46 -11.66
ENB1	site	patser	D	-194 -187ttatTGCACCCAaatg	10.46 -11.66
ARN1	site	patser	D	-261 -254aaaaTGCACCCAtaaa	10.46 -11.66
FET3	site	patser	D	-252 -245aaagTGCACCCAtttg	10.46 -11.66
YOL106w	site	patser	D	-415 -408aactTGCACCCTtgaa	10.28 -10.97