Mixed Integer Linear Programming based machine learning approach identifies *regulators* of telomerase in yeast

Alexandra Poos[1,2,3], André Maicher[4,5], Anna Dieckmann[2,3], Marcus Oswald[1,2], Roland Eils[3,6], Martin Kupiec[5], Brian Luke[4,7] and Rainer König[1,2,3,*]

[1] Integrated Research and Treatment Center, Center for Sepsis Control and Care (CSCC), Jena University Hospital, D-07747 Jena, Erlanger Allee 101, Germany
[2] Network Modeling, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute (HKI) Jena, Beutenbergstrasse 11a, 07745 Jena, Germany
[3] Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
[4] Center for Molecular Biology at Heidelberg University (ZMBH), German Cancer Research Center (DKFZ)-ZMBH-Alliance, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany
[5] Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Ramat Aviv 69978, Israel.
[6] Department of Bioinformatics and Functional Genomics, Institute of Pharmacy and Molecular Biotechnology, and Bioquant, University of Heidelberg, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany
[7] Telomere Biology Group, Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany.

* To whom correspondence should be addressed. Tel: 004936415321189; Fax: 004936415320800; Email: rainer.koenig@uni-jena.de

Supplementary Material

Text S1. Model performances

To elucidate if rather small models or the larger models show a better performance, we assembled the performance results (Pearson correlation between prediction and measured expression value) of all cross-validation runs for each model size (from 1 until 10 regulators). The results are given in Figure S6, showing a quite homogenous performance with a tendency that the smaller models yielded the better results. Furthermore, Figure S7 shows a scatterplot of the predicted versus the experimental expression of *EST1* for the short *tlm* (red) and the control dataset (black) with a correlation of PCC r=0.51.

As typically, a standard (L2 based) linear regression coupled with analysis of Variance is employed for such a bottom up or top down approach to get models with different optimized sets of parameters, we calculated the adjusted R^2 for the complete short *tlm* dataset (Figure S8). We found increasing adjusted R^2 values until 11 regulators, after which a large decline occurred. This is in line with our approach of selecting models below 11 regulators. To relate the model results with the correlation analysis, we compared the beta values (median over all cross-validations) with the correlation between the regulator activities and *EST1* expression. We expected that a positive correlation between the regulator activities and *EST1* expression is positively correlated with the beta values for this regulator. Indeed, 10 out of the 12 significant regulators of *EST1* in the short *tlm* dataset showed this behavior (see Table S9 for all values). For the controls, this was much less (3 out of 12) reflecting the property of the selected TFs to better explain the dataset we foccused on (short *tlm* mutants). We further investigated how often each regulator was chosen over the cross-validation runs in the short *tlm* and the control datasets (Figure S3) as well as how often the regulator was chosen in models with one to 10 regulators (Table S10 and S11). For *EST1*, Sum1 and Hst1 were selected most often for the short *tlm* dataset and comparably unfrequently for the control dataset. This confirms that Sum1 and Hst1 were the most important regulators explaining *EST1* expression. These frequencies further showed that Sum1 and Hst1 were used mutually exclusive, in particular for the smaller models (1-5 regulators, see Table S12).

Text S2. Correlation analysis

If the activity of a regulator was very similar to the activity of another regulator in each of the investigated samples, the model may have difficulties distinguishing between them and may neglect one of these regulators causing false negatives. To identify such false negatives due to collinearity of the regulators' activity values, we (i) performed a correlation analysis between the activity values of each pair of the different regulators, (ii) calculated the correlations between the regulators' activitiy values and the gene expression of the putatively targeting *EST* genes, and (iii) simulated *in silico* knockouts to identify mutually exclusive regulators.

For (i) and (ii), the correlations were calculated for the short *tlm* mutants and the controls seperately. Then, the largest differences (between short *tlm* mutants and controls) were selected to obtain short *tlm* mutant specific regulators. A complete list of all positively correlating regulators with a difference in activity correlation larger or equal to 0.1 is shown in Table S8, some results are pointed out in the

following. For *EST1,* the correlation differences of regulator pairs were quite low, we found the highest activity correlation difference between Sum1 and Hst1 activities (correlation difference between short *tlm* mutants and the controls: PCC r=0.17). This is in agreement with the results from our modeling analysis (Table 1 in the main text). For *EST2* and *EST3* the differences were larger. The largest difference for *EST2* was r=0.45 for the pair Rtg3 and Rgt1, and for *EST3* it was r=0.37 for the pair Dig1 and Gln3. Our modeling approach found Rtg3 but not Rgt1 for *EST2* and Dig1 but not Gln3 hinting at Rgt1 and Gln3 being potential further candidates regulating *EST2* and *EST3*, respectively.
We then analyzed the correlation between activity of the regulators and the expression of the putative target genes (*EST*s). Again, we calculated the difference in correlation between short *tlm* mutants and controls. The results are given in Table S5, S6 and S7 for *EST1*, *EST2* and *EST3*, respectively. For *EST1* we found Hst1 and Sum1 with the highest correlation difference (Hst1-Est1: r=0.51, Sum1-Est1: r=0.62, Table S5) confirming our modelling results that these regulators are involved in *EST1* regulation. For *EST2*, we found Arg81, Nrg1, Tec1, Nrg2, Msn4 and Pdr3 with the highest correlation difference of which all were predicted by our model except of Nrg1. For *EST3*, Ume6 had the highest absolute correlation difference, in agreement to our modeling results. In summary, the results of the correlation analysis confirmed the modeling results, in particular for *EST1* and our most promising regulators of the modelling analysis, Hst1 and Sum1. We note that this analysis may serve the purpose of adding potential other candidates to the regulatos that we selected from our modeling analysis.
(iii) Further, we simulated a knockout of a specific regulator if we found this regulator's activity to correlate highly with another regulator. As a case study, we investigated *EST1* as the target gene and the pair of Sum1 and Hst1 as regulators. Indeed, when Sum1 was knocked out, Hst1 took over Sum1's function explaining the gene expression of *EST1* and was used distinctively more often by the models when compared to the non knockout models (P=5.49 E-30, Table S13). In turn, knocking out Hst1, Sum1 was used instead (P=2.55 E-40, Table S14). This is consistent with the literature, it was reported that Sum1 and Hst1 together with Rfm1 form a complex repressing genes through histone deacetylation (1-4). Hence, we suggest that Sum1 and Hst1 act synergistically also for the expression of the telomerase gene *EST1*. To further confirm this computationally, we followed up on this complex and built models with the combination Sum1-Hst1 (we multiplied the activities and used the square root of the product as activity of the complex) instead of the single regulators Sum1 and Hst1. In this case the combination Sum1-Hst1 was used instead of the single regulators (P=4.51 E-32, Table S15).

Text S3. Modeling a complex of Sum1 and Hst1

We further built models with only the regulators Sum1 and Hst1 as well as with a combination of both mimicking cooperative activity [as suggested by (5)]. We used the data of the short *tlm* mutants and estimated the performance using a ten-times sixfold cross-validation (as described in the main text, see section The machine learning approach). We investigated the new model starting with one parameter for the smallest possible model (only beta 0 and one of the three possibilities of Sum1, Hst and Sum1-Hst1). The results are shown in Table S16. For the smallest models, Hst1 was most often selected (43 out of 60). Restricted to two β-parameters (either two single regulators, or one single and

one Sum1-Hst1 combined) the optimizer chose the combination of Sum1-Hst1 together with Sum1 most often supporting the suggestion of a combined regulation of Sum1 and Hst1.

Table S16. Modeling results of the Sum1-Hst1 combined model

Regulator	One regulator model	Two regulator model	Three regulator model
Hst1	43*	28	60
Sum1	9	46	60
Hst1-Sum1	8	46	60

* Number of selections by the model out of 60 runs.

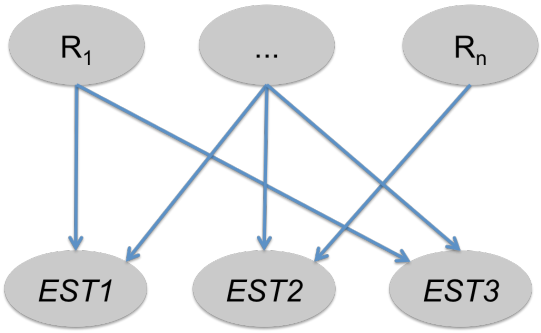

Figure S1. The regulatory model. The *EST* genes are regulated by regulators R_1 - R_n.

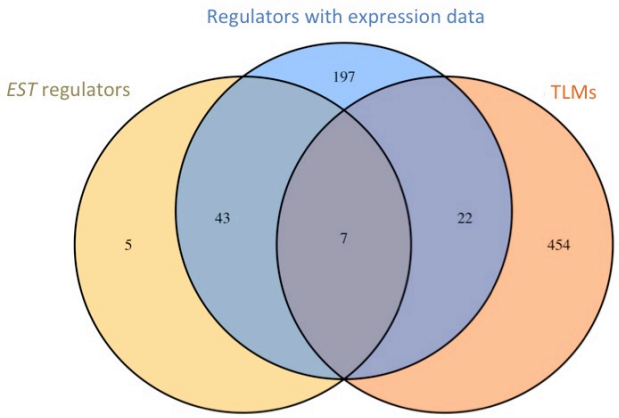

Figure S2. Overlaps between the TLM genelist (orange, data was taken from (6-10)), the yeast deletion strains of regulators of which we used the expression data (blue, data was taken from (11)), and regulators with know binding (taken from YEASTRACT) to at least one of the *EST* genes (*EST1, EST2, EST3*).

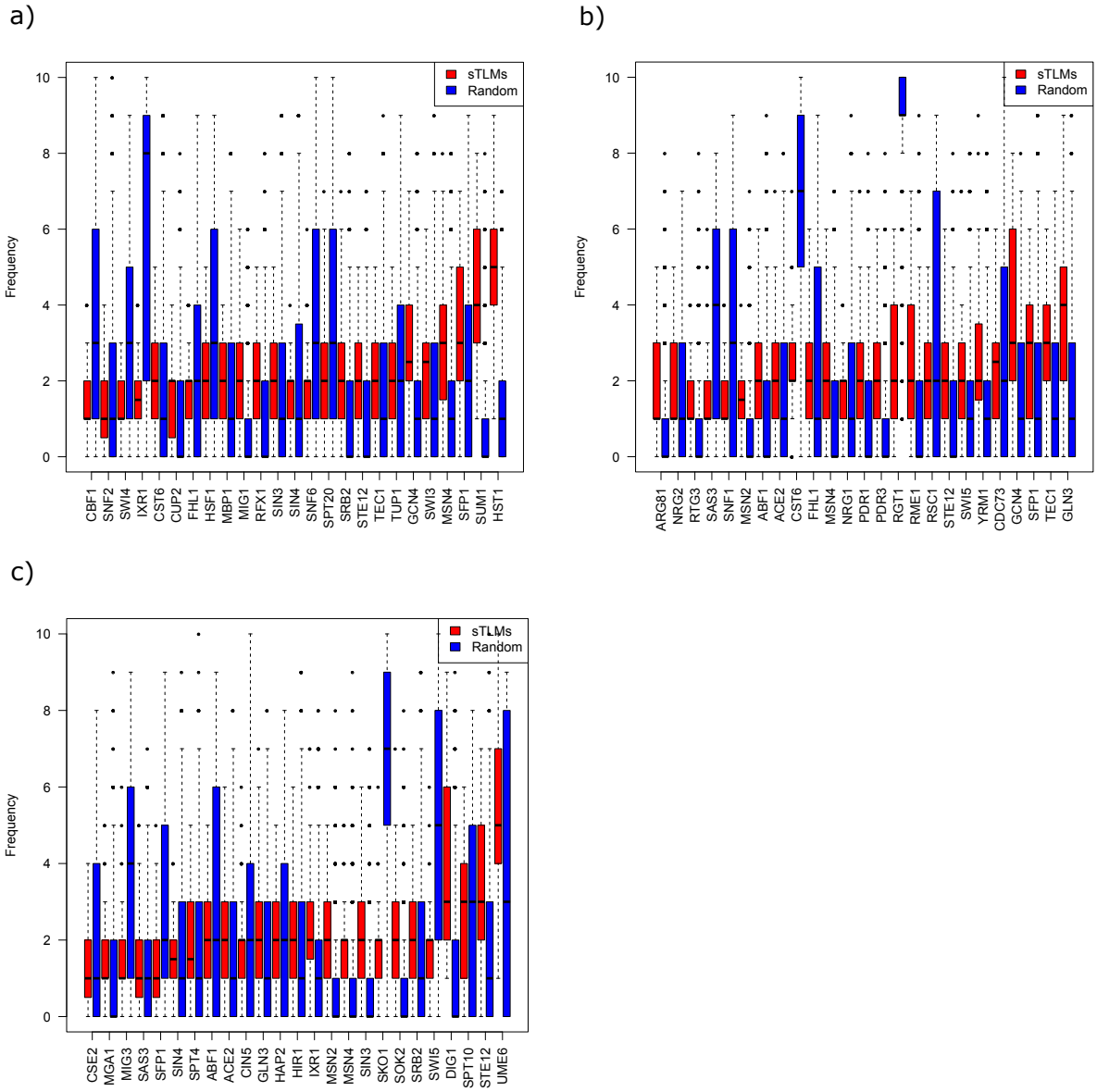

Figure S3. Regulator frequencies (of all 10 runs of each cross-validation) for the short *tlm* (red) and the control dataset (blue) over all cross-validation runs. a) *EST1*, b) *EST2*, c) *EST3.*

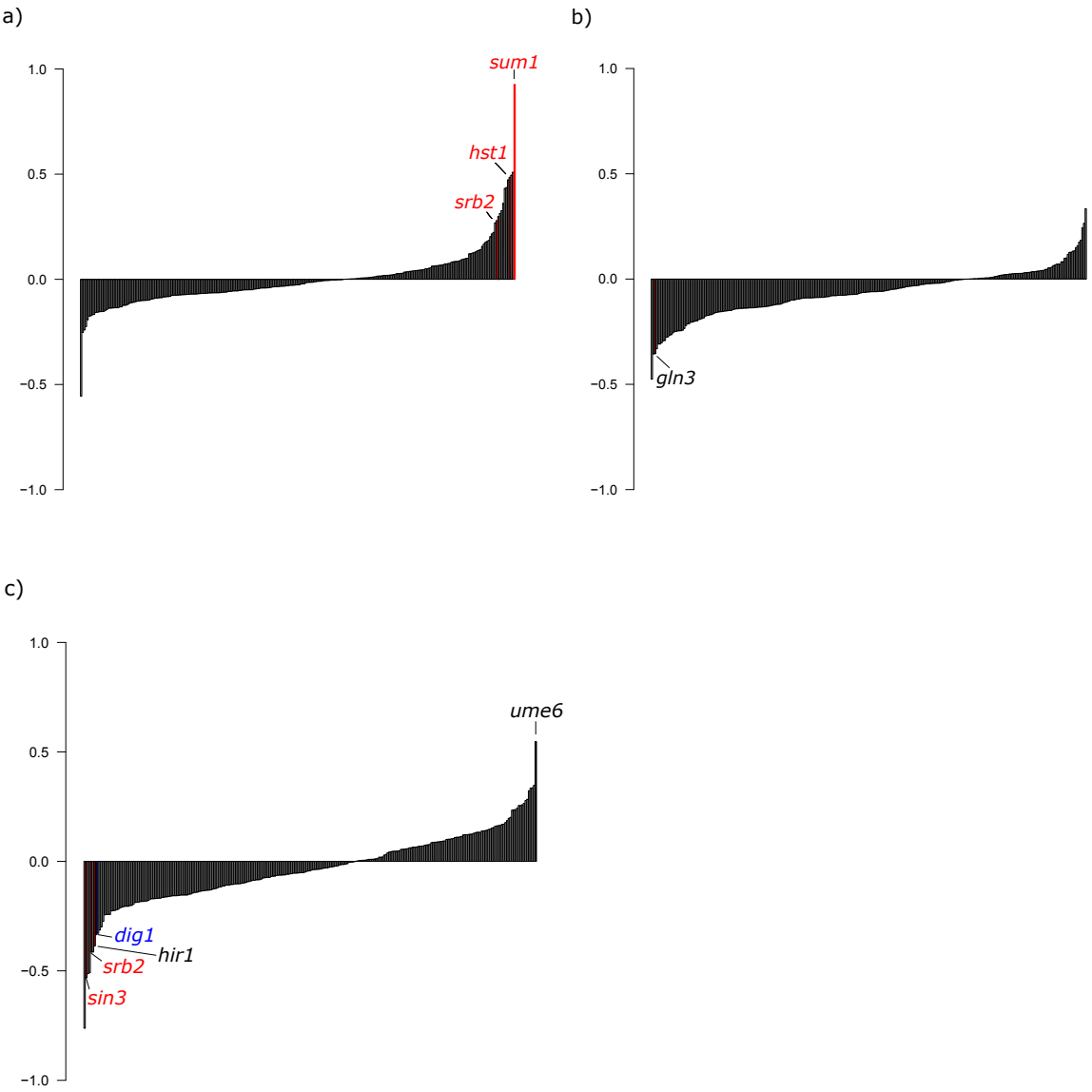

Figure S4. *EST1-3* expression (log-fold change) over all observed 269 regulator knockouts, ranked according to the expression levels of the corresponding *EST* gene; a) *EST1*, b) *EST2* and c) *EST3* (red: short *tlm*, blue: long *tlm* and black: control sample).

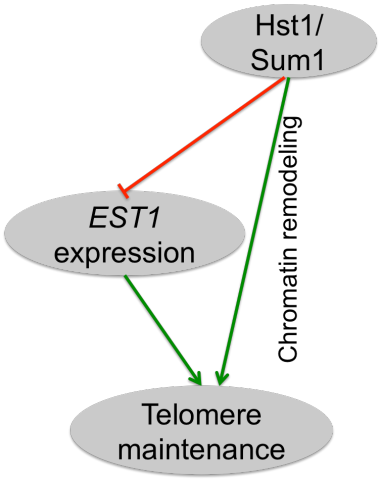

Figure S5. Incoherent feed forward loop, where Sum1/Hst1 and *EST1* positively regulate telomeres and Sum1/Hst1 negatively regulate *EST1*.

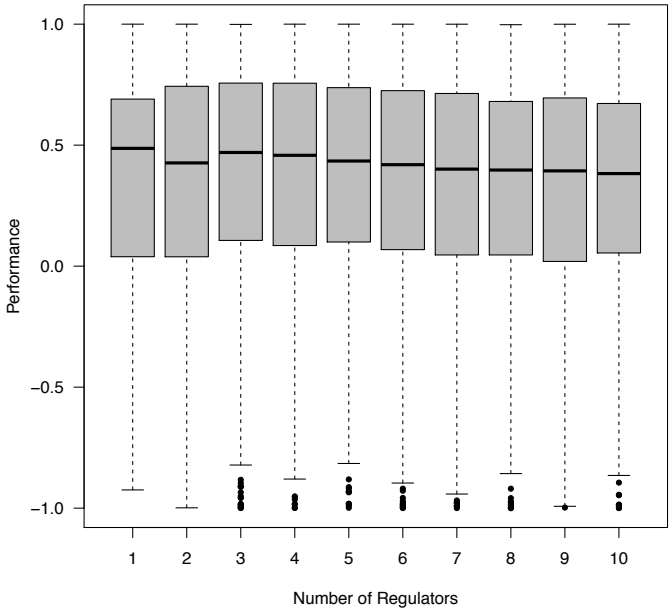

Figure S6. Pearson correlation of our modeling predictions with the measured expression values, shown for different sizes of the model, left: only one regulator, right, the maximal number of 10 regulators (shown for *EST1* regulators, short *tlm* mutants and control).

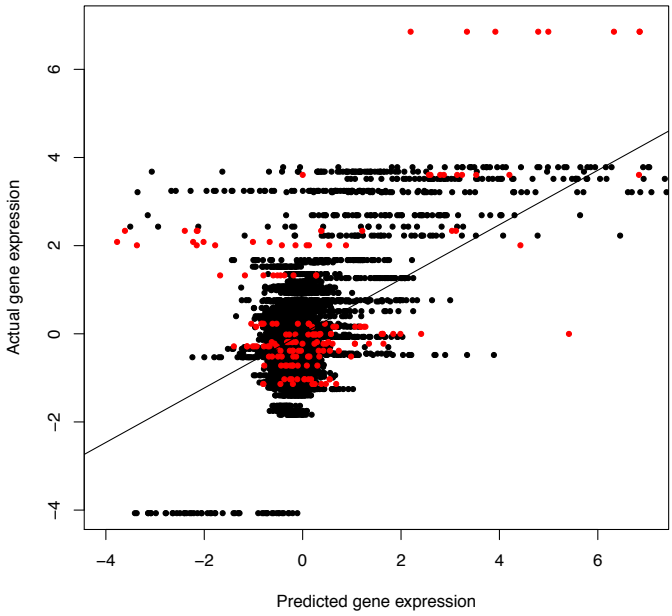

Figure S7. Actual *vs*. model predicted performance of *EST1* for the short *tlm* (red) and the control dataset over all cross-validation runs (cor (actual GE, predicted GE) = 0.51).

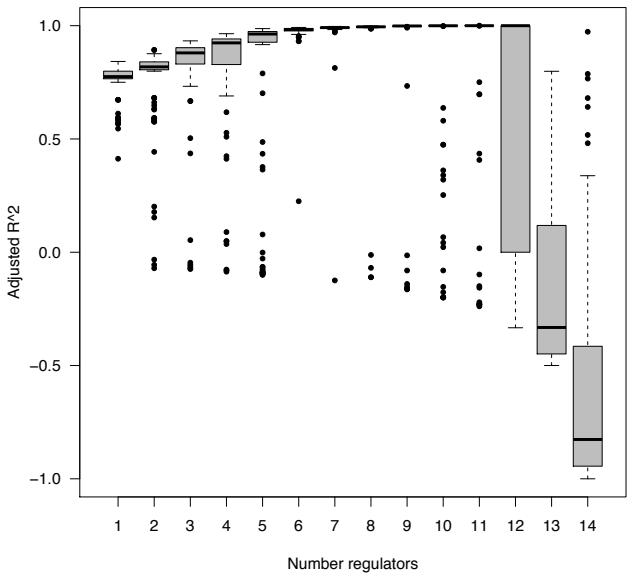

Figure S8. Adjusted R^2 calculation for all short *tlm* samples (similar to ANOVA)

Table S1. Corresponding genes of the investigated regulator (R) deletion strains of the dataset of Reimand and coworkers (11) and their telomere phenotype (6-10).

R	Group	R	Group	R	Group
CDC73	short	*ARG82*	control	*GCN4*	control
CST6	short	*ARO80*	control	*GCR1*	control
GAL11	short	*ARR1*	control	*GCR2*	control
HFI1	short	*ASH1*	control	*GLN3*	control
HST1	short	*ASK10*	control	*GTS1*	control
MOT2	short	*AZF1*	control	*GZF3*	control
MOT3	short	*BAS1*	control	*HAA1*	control
OPI1	short	*BDF2*	control	*HAC1*	control
PGD1	short	*CAC2*	control	*HAL9*	control
RPN4	short	*CAD1*	control	*HAP2*	control
RSC2	short	*CAF17*	control	*HAP3*	control
RTF1	short	*CAF4*	control	*HAP4*	control
SIF2	short	*CAT8*	control	*HAP5*	control
SIN3	short	*CBF1*	control	*HAT1*	control
SIR3	short	*CHA4*	control	*HAT2*	control
SRB2	short	*CIN5*	control	*HDA1*	control
SRB5	short	*CRZ1*	control	*HIR1*	control
SUM1	short	*CSE2*	control	*HIR2*	control
DIG1	long	*CUP2*	control	*HIR3*	control
HCM1	long	*CUP9*	control	*HMS1*	control
MET18	long	*DAL80*	control	*HMS2*	control
NUT1	long	*DAL81*	control	*HOG1*	control
RAP1	long	*DAL82*	control	*HPA2*	control
REB1	long	*DAT1*	control	*HSF1*	control
RIF1	long	*DOT5*	control	*HST3*	control
RIF2	long	*DOT6*	control	*HST4*	control
SRB8	long	*ECM22*	control	*IME1*	control
SSN2	long	*ESC2*	control	*INO2*	control
SSN3	long	*FKH1*	control	*INO4*	control
ABF1	control	*FKH2*	control	*ISW1*	control
ACA1	control	*FLO8*	control	*ISW2*	control
ACE2	control	*FZF1*	control	*IXR1*	control
ADA2	control	*GAL3*	control	*KAR4*	control
ADR1	control	*GAL4*	control	*KSS1*	control
AFT2	control	*GAL80*	control	*LEU3*	control
ARG80	control	*GAT1*	control	*MAC1*	control
ARG81	control	*GAT3*	control	*MAL13*	control

R	Group	R	Group	R	Group
MAL33	control	*RFX1*	control	*SNF6*	control
MBF1	control	*RGM1*	control	*SOK2*	control
MBP1	control	*RGT1*	control	*SPS18*	control
MCM1	control	*RIC1*	control	*SPT10*	control
MDS3	control	*RIM101*	control	*SPT2*	control
MET28	control	*RIS1*	control	*SPT20*	control
MET31	control	*RLF2*	control	*SPT23*	control
MET32	control	*RLM1*	control	*SPT3*	control
MGA1	control	*RLR1*	control	*SPT4*	control
MGA2	control	*RME1*	control	*STB1*	control
MIG1	control	*ROX1*	control	*STB2*	control
MIG2	control	*RPD3*	control	*STB3*	control
MSI1	control	*RPH1*	control	*STB4*	control
MSN1	control	*RPI1*	control	*STB5*	control
MSN2	control	*RSC1*	control	*STB6*	control
MSN4	control	*RTG1*	control	*STP1*	control
MSS11	control	*RTG3*	control	*STP2*	control
MTH1	control	*RTT107*	control	*STP4*	control
NDT80	control	*SAS3*	control	*SUT1*	control
NGG1	control	*SAS4*	control	*SUT2*	control
NOT3	control	*SAS5*	control	*SWI3*	control
NRG1	control	*SDS3*	control	*SWI4*	control
OAF1	control	*SEF1*	control	*SWI5*	control
PDR1	control	*SET2*	control	*SWI6*	control
PDR3	control	*SFL1*	control	*TAF14*	control
PDR8	control	*SFP1*	control	*TBS1*	control
PHD1	control	*SIN4*	control	*TEC1*	control
PHO2	control	*SIP3*	control	*THI2*	control
PHO23	control	*SIP4*	control	*TIS11*	control
PHO4	control	*SIR1*	control	*TOS8*	control
PIB2	control	*SIR2*	control	*TUP1*	control
PIP2	control	*SKN7*	control	*TYE7*	control
POP2	control	*SKO1*	control	*UGA3*	control
PPR1	control	*SMK1*	control	*UME1*	control
PUT3	control	*SMP1*	control	*UME6*	control
RCS1	control	*SNF1*	control	*UPC2*	control
RDR1	control	*SNF11*	control	*WAR1*	control
RDS1	control	*SNF2*	control	*WTM1*	control
RDS2	control	*SNF5*	control	*WTM2*	control

R	Group	R	Group	R	Group
XBP1	control	*YER051W*	control	*YKR064W*	control
YAP1	control	*YER130C*	control	*YLR278C*	control
YAP3	control	*YER184C*	control	*YML081W*	control
YAP5	control	*YFL044C*	control	*YMR075W*	control
YAP6	control	*YFL052W*	control	*YNR063W*	Control
YAP7	control	*YGL131C*	control	*YOX1*	Control
YBL054W	control	*YGR067C*	control	*YPL230W*	Control
YBR033W	control	*YGR089W*	control	*YPR022C*	Control
YBR239C	control	*YHP1*	control	*YPR196W*	Control
YDR026C	control	*YIL130W*	control	*YRR1*	control
YDR049W	control	*YJL103C*	control	*ZAP1*	control
YDR266C	control	*YJL206C*	control	*ZDS1*	control
YDR520C	control	*YKL005C*	control	*ZMS1*	control
YER028C	control	*YKL222C*	control		

Table S2. Putative regulators of the *EST* genes (taken from YEASTRACT).

Regulator	***EST*s**
Msn4	*EST1, EST2, EST3*
Sfp1	*EST1, EST2, EST3*
Ste12	*EST1, EST2, EST3*
Abf1	*EST2, EST3*
Ace2	*EST2, EST3*
Cst6	*EST1, EST2*
Fhl1	*EST1, EST2*
Gcn4	*EST1, EST2*
Gln3	*EST2, EST3*
Ixr1	*EST1, EST3*
Msn2	*EST2, EST3*
Sas3	*EST2, EST3*
Sin3	*EST1, EST3*
Sin4	*EST1, EST3*
Srb2	*EST1, EST3*
Swi5	*EST2, EST3*
Tec1	*EST1, EST2*
Arg81	*EST2*
Cbf1	*EST1*
Cdc73	*EST2*
Cin5	*EST3*
Cse2	*EST3*
Cup2	*EST1*
Dig1	*EST3*
Hap2	*EST3*
Hir1	*EST3*
Hsf1	*EST1*
Hst1	*EST1*
Mbp1	*EST1*
Mga1	*EST3*
Mig1	*EST1*
Mig3	*EST3*
Nrg1	*EST2*
Nrg2	*EST2*
Pdr1	*EST2*
Pdr3	*EST2*
Rfx1	*EST1*
Rgt1	*EST2*
Rme1	*EST2*
Rsc1	*EST2*
Rtg3	*EST2*
Sko1	*EST3*
Snf1	*EST2*
Snf2	*EST1*
Snf6	*EST1*
Sok2	*EST3*
Spt10	*EST3*
Spt20	*EST1*
Spt4	*EST3*
Sum1	*EST1*
Swi3	*EST1*
Swi4	*EST1*
Tup1	*EST1*
Ume6	*EST3*
Yrm1	*EST2*

Table S3. Significant regulators of *EST* genes comparing *tlms* (short + long) *vs*. control samples

	Regulator	Z-score*	Significance (P)**	Number of targets
EST1	Sum1***	6.85	3.76 E-30	579
	Hst1***	3.61	2.51 E-28	219
	Rfx1	-0.45	5.71 E-22	660
	Mig1	0.14	1.56 E-17	423
	Srb2***	2.08	5.23 E-10	785
	Sfp1	3.24	6.56 E-9	4199
	Ste12	-****	8.89 E-6	3673
	Cup2	-0.30	2.09 E-2	548
EST2	Nrg2	-	6.11 E-23	331
	Pdr1	0.31	1.28 E-17	1318
	Gcn4	-0.22	1.77 E-16	2712
	Rtg3	0.16	5.33 E-9	646
	Cdc73***	-3.79	2.23 E-8	757
	Yrm1	-	6.11 E-6	2509
	Tec1	-0.22	1.05 E-5	3669
	Msn2	0.05	2.83 E-5	3260
	Arg81	0.42	9.24 E-5	335
	Ste12	-	1.12 E-2	3673
	Sfp1	-0.06	1.43 E-2	4199
	Gln3	-2.67	1.57 E-2	981
	Ace2	-1.70	1.63 E-2	4683
	Rme1	-0.31	3.22 E-2	399
	Abf1	-0.29	4.38 E-2	2715
	Pdr3	-0.44	4.38 E-2	929
	Nrg1	0.07	4.48 E-2	686
	Msn4	-1.17	4.83 E-2	2483
EST3	Dig1***	-1.87	1.15 E-36	334
	Sok2	0.28	4.39 E-23	2160
	Cin5	1.81	1.03 E-21	2062
	Sin3***	-3.11	1.41 E-20	1759
	Ste12	-	9.37 E-14	3673
	Sin4	2.31	1.58 E-10	2144
	Msn2	-0.46	2.72 E-6	3260
	Spt10	-2.37	1.46 E-5	1691
	Msn4	0.62	3.19 E-4	2483
	Srb2***	-2.38	9.73 E-4	785
	Gln3	-1.50	4.05 E-2	978

* Effect of the knockout of the regulator on the expression of the *EST* genes (positive z-score = upregulation of the corresponding *EST* gene; negative z-score = downregulation of the corresponding

EST gene); ** Multiple testing corrected (Benjamini-Hochberg); *** red: short *tlm* mutant, blue: long *tlm* mutant; **** For some genes, no expression data was available

Table S4 is in a separate file of the Supplementary Material and includes the correlation between all regulator activities.

Table S5. Correlations between *EST1* expression and the correpsonding regulator activities in the short *tlm* dataset and the control dataset (non-*TLM* dataset)

Regulator	Correlation sTLMs	Correlation controls	Correlation difference
Sum1	0.54	-0.08	0.62
Hst1	0.87	0.35	0.52
Rfx1	-0.11	0.16	-0.27
Snf2	-0.10	-0.37	0.27
Snf6	-0.11	-0.34	0.23
Ixr1	0.26	0.05	0.21
Tup1	-0.13	-0.34	0.21
Spt20	-0.20	-0.41	0.21
Msn4	0.20	0	0.20
Mig1	0.03	0.22	-0.19
Srb2	0.02	-0.17	0.19
Swi3	-0.16	-0.35	0.19
Tec1	0.34	0.17	0.17
Cst6	-0.16	-0.33	0.17
Ste12	0.18	0.02	0.16
Sin4	-0.22	-0.37	0.15
Cup2	-0.01	0.09	-0.10
Hsf1	-0.24	-0.33	0.09
Sfp1	-0.15	-0.07	-0.08
Fhl1	-0.38	-0.32	-0.06
Mbp1	0.07	0.12	-0.05
Cbf1	-0.18	-0.14	-0.04
Swi4	-0.04	-0.06	0.02
Gcn4	-0.09	-0.11	0.02
Sin3	-0.27	-0.29	0.02

Table S6. Correlations between *EST2* expression and the corresponding regulator activities in the short *tlm* dataset and the control dataset (non-*TLM* dataset)

Regulator	Correlation sTLMs	Correlation controls	Correlation difference
Arg81	0.41	-0.19	0.60
Nrg1	0.63	0.16	0.47
Tec1	0.60	0.14	0.46
Nrg2	0.73	0.28	0.45
Msn4	0.49	0.04	0.45
Pdr3	0.52	0.09	0.43
Rgt1	-0.40	-0.01	-0.39
Ste12	0.43	0.06	0.37

Gln3	-0.50	-0.18	-0.32
Sas3	-0.61	-0.31	-0.3
Msn2	0.23	-0.03	0.26
Rme1	0.36	0.11	0.25
Cst6	-0.51	-0.27	-0.24
Pdr1	0.17	-0.05	0.22
Swi5	-0.37	-0.16	-0.21
Abf1	-0.27	-0.13	-0.14
Rtg3	-0.48	-0.36	-0.12
Fhl1	-0.25	-0.16	-0.09
Snf1	0.13	0.04	0.09
Ace2	0.00	-0.09	0.09
Yrm1	-0.17	-0.09	-0.08
Rsc1	0.14	0.07	0.07
Cdc73	0.08	0.14	-0.06
Gcn4	-0.01	-0.07	0.06
Sfp1	-0.17	-0.16	-0.01

Table S7. Correlations between *EST3* expression and the corresponding regulator activities in the short *tlm* dataset and the control dataset (non-*TLM* dataset)

Regulator	Correlation sTLMs	Correlation controls	Correlation difference
Ume6	-0.25	0.19	-0.44
Sin3	-0.02	0.41	-0.43
Hap2	0.02	0.43	-0.41
Sfp1	0.02	0.42	-0.40
Spt4	0.03	0.41	-0.38
Sas3	-0.19	0.19	-0.38
Sin4	-0.10	0.21	-0.31
Ixr1	-0.15	0.16	-0.31
Srb2	-0.19	0.11	-0.30
Ace2	0.08	0.37	-0.29
Spt10	0.20	0.45	-0.25
Msn2	0.10	0.31	-0.21
Abf1	0.15	0.35	-0.20
Gln3	0.25	0.38	-0.13
Sok2	-0.02	0.11	-0.13
Ste12	0.14	0.26	-0.12
Msn4	0.21	0.30	-0.09
Mig3	0.14	0.22	-0.08
Mga1	0.48	0.41	0.07
Dig1	0.62	0.68	-0.06
Cin5	0.40	0.46	-0.06
Cse2	0.14	0.20	-0.06
Hir1	0.50	0.55	-0.05
Sko1	0.40	0.35	0.05
Swi5	0.26	0.24	0.02

Table S8: Positive correlation differences (diff ≥ 0.1) of regulator activities between the short *tlm* and the control dataset; the correlation differences between significant hits of the corresponding *EST* gene are marked in bold.

	Regulator 1	Regulator 2	Correlation Difference	Correlation sTLMs	Correlation Controls
EST1	**Hst1**	**Sum1**	**0.17**	**0.77**	**0.60**
	Sfp1	**Srb2**	**0.16**	**0.85**	**0.69**
	Mig1	Tec1	0.14	0.74	0.60
	Mbp1	**Mig1**	**0.13**	**0.86**	**0.73**
	Srb2	**Swi3**	**0.12**	**0.91**	**0.79**
	Srb2	Snf6	0.12	0.95	0.83
	Srb2	Cst6	0.10	0.94	0.84
EST2	Rtg3	Rgt1	0.45	0.58	0.13
	Cdc73	Rtg3	0.34	0.30	-0.04
	Arg81	**Rme1**	**0.26**	**0.54**	**0.28**
	Nrg2	**Tec1**	**0.27**	**0.77**	**0.50**
	Rtg3	**Gln3**	**0.26**	**0.70**	**0.44**
	Nrg2	**Pdr3**	**0.21**	**0.78**	**0.57**
	Msn4	**Nrg2**	**0.20**	**0.68**	**0.48**
	Cdc73	Rsc1	0.18	0.96	0.78
	Arg81	**Ste12**	**0.18**	**0.72**	**0.54**
	Cdc73	Sfp1	0.17	0.70	0.53
	Abf1	Rgt1	0.16	0.82	0.66
	Cdc73	Sas3	0.17	0.51	0.34
	Nrg2	**Ste12**	**0.16**	**0.59**	**0.43**
	Arg81	Cdc73	0.15	0.06	-0.09
	Arg81	**Msn4**	**0.15**	**0.67**	**0.52**
	Arg81	**Nrg2**	**0.15**	**0.36**	**0.21**
	Arg81	**Pdr3**	**0.14**	**0.54**	**0.40**
	Rtg3	Sas3	0.12	0.57	0.45
	Arg81	**Tec1**	**0.11**	**0.59**	**0.48**
EST3	Dig1	Gln3	0.36	0.73	0.37
	Srb2	**Hir1**	**0.34**	**0.58**	**0.24**
	Dig1	**Srb2**	**0.32**	**0.44**	**0.12**
	Dig1	Sin4	0.24	0.59	0.35
	Dig1	Swi5	0.24	0.47	0.23
	Srb2	Hap2	0.24	0.82	0.58
	Srb2	Spt4	0.23	0.83	0.60
	Srb2	Sas3	0.19	0.86	0.67
	Dig1	Abf1	0.15	0.45	0.30
	Srb2	Sfp1	0.16	0.85	0.69

	Srb2	Spt10	0.15	0.81	0.66
	Dig1	Sas3	0.11	0.55	0.44
	Sin3	Sas3	0.12	0.85	0.73
	Dig1	**Sin3**	**0.11**	**0.66**	**0.55**
	Sin3	**Hir1**	**0.10**	**0.67**	**0.57**

Table S9. Median beta values of short *tlm* and control models, and Pearson correlation between regulator activities and *EST1* expression

Regulator	Correlation sTLMs	Median betas sTLM models	Correlation controls	Median betas control models
Sum1*	0.54	4.24	-0.08	1.63
Hst1*	0.87	2.08	0.35	-1.44
Msn4	0.20	38.05	0	-0.73
Mig1	0.03	5.96	0.22	-0.99
Gcn4	-0.09	-34.39	-0.11	5.64
Ste12	0.18	22.67	0.02	-0.86
Rfx1	-0.11	-5.75	0.16	0.46
Srb2*	0.02	0.99	-0.17	-1.05
Sfp1	-0.15	-47.78	-0.07	-4.66
Cup2	-0.01	4.08	0.09	-0.90
Swi3	-0.16	24.82	-0.35	1.79
Mbp1	0.07	1.41	0.12	-1.98

*red: short *tlm* mutant, blue: long *tlm* mutant

Table S10. Frequency of the regulators in the models for the short *tlm* models, from 1 (left) to 10 regulators (right)

Regulator	1	2	3	4	5	6	7	8	9	10
Cbf1	0	0	0	7	7	9	13	11	15	18
Cst6	1	0	0	7	12	14	14	17	18	30
Cup2	0	0	0	2	10	15	11	16	16	19
Fhl1	0	1	7	6	8	8	10	20	16	30
Gcn4	0	8	25	21	16	16	23	17	28	21
Hsf1	0	0	3	9	8	10	15	17	19	25
Hst1	58	11	29	38	27	28	28	31	24	32
Ixr1	0	0	7	5	1	7	15	17	23	25
Mbp1	0	1	5	5	9	13	18	22	25	19
Mig1	0	0	4	4	14	17	14	19	25	24

Msn4	0	2	26	25	22	19	21	18	24	27
Rfx1	0	12	1	3	3	10	14	19	22	29
Sfp1	0	3	3	18	35	31	27	33	30	26
Sin3	0	0	8	3	7	12	12	17	24	24
Sin4	0	0	1	4	2	8	12	19	27	24
Snf2	0	0	3	2	6	9	15	14	15	17
Snf6	0	0	0	6	5	12	17	19	27	21
Spt20	0	13	3	7	13	10	16	17	23	26
Srb2	0	23	13	14	6	13	18	14	17	18
Ste12	0	0	1	6	16	19	17	23	23	25
Sum1	1	45	24	18	29	30	32	29	35	20
Swi3	0	1	2	6	19	20	21	22	16	31
Swi4	0	0	0	6	3	3	7	16	20	22
Tec1	0	0	4	8	15	11	12	17	20	27
Tup1	0	0	11	10	7	16	18	16	18	20

Table S11. Frequency of the regulators in the models for the controls, from 1 (left) to 10 regulators (right)

Regulator	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Cbf1	28	19	217	216	236	251	266	284	304	305
Cst6	0	52	58	73	111	129	172	197	243	263
Cup2	21	0	10	15	29	63	102	126	153	194
Fhl1	1	18	18	47	104	145	177	219	261	296
Gcn4	0	2	26	36	72	88	137	148	180	210
Hsf1	26	41	94	174	215	253	278	302	326	328
Hst1	2	9	4	20	36	63	98	129	157	227
Ixr1	1	339	345	358	375	388	414	413	416	417
Mbp1	0	2	35	58	62	107	143	169	206	237
Mig1	3	0	2	14	29	48	67	88	115	128
Msn4	0	3	36	47	80	81	91	129	151	186
Rfx1	0	1	29	25	46	64	84	102	132	171
Sfp1	0	62	130	121	156	190	196	229	251	274
Sin3	1	62	34	73	107	145	148	177	202	235
Sin4	13	21	23	89	126	151	156	178	215	240
Snf2	117	53	82	98	96	114	142	164	191	228

Snf6	87	256	269	255	231	220	214	224	228	244
Spt20	75	116	101	151	201	237	275	295	302	313
Srb2	1	27	5	28	46	63	91	124	146	177
Ste12	0	2	9	32	48	63	98	128	157	169
Sum1	1	3	5	23	34	52	74	100	135	172
Swi3	190	38	51	60	86	103	125	142	159	181
Swi4	5	3	96	166	194	233	274	299	314	328
Tec1	3	1	36	88	116	160	187	206	223	247
Tup1	25	70	85	133	164	189	191	228	233	230

Table S12. Frequency of the regulators Sum1 and Hst1 individually or combined (Hst & Sum1) in the short *tlm* models, from 1 (left) to 10 regulators (right)

Regulator	1	2	3	4	5	6	7	8	9	10
Hst1	58	11	29	35	27	22	22	23	17	25
Sum1	1	45	24	15	29	24	26	21	28	13
Hst1 & Sum1	0	0	0	3	0	6	6	8	7	7

Table S13. Significant regulators of *EST1* after a simulated *SUM1* knockout

	Regulator	Z-score*	Significance (P)**	Number of targets
EST1	Hst1***	3.61	5.49 E-30	219
	Msn4	-0.63	9.49 E-20	2483
	Gcn4	-0.13	1.39 E-17	2712
	Rfx1	-0.45	3.44 E-14	660
	Mig1	0.14	4.91 E-10	423
	Srb2***	2.08	9.65 E-10	785
	Sfp1	3.24	7.58 E-7	4186
	Mbp1	0.76	7.71 E-7	662
	Sin3***	2.01	4.43 E-5	1759
	Ste12	-	2.50 E-3	3640
	Tec1	1.52	1.93 E-2	3653
	Cup2	-0.30	2.10 E-2	548

* Effect of the knockout of the regulator on the expression of the *EST* genes (positive z-score = upregulation of the corresponding *EST* gene; negative z-score = downregulation of the corresponding *EST* gene); ** Multiple testing corrected (Benjamini-Hochberg); *** red: short *tlm* mutant, blue: long *tlm* mutant

Table S14. Significant regulators of *EST1* after a simulated *HST1* knockout

	Regulator	Z-score*	Significance (P)**	Number of targets
EST1	Sum1***	6.85	2.59 E-40	579
	Rfx1	-0.45	1.22 E-15	660
	Gcn4	-0.13	3.40 E-12	2712
	Srb2***	2.08	1.02 E-9	785
	Msn4	-0.63	2.71 E-8	2483
	Sin3***	2.01	2.13 E-7	1759
	Mig1	0.14	5.35 E-7	423
	Mbp1	0.76	5.74 E-6	665
	Sfp1	3.24	7.06 E-6	4199
	Ste12	-	6.43 E-5	3673

* Effect of the knockout of the regulator on the expression of the *EST* genes (positive z-score = upregulation of the corresponding *EST* gene; negative z-score = downregulation of the corresponding *EST* gene); ** Multiple testing corrected (Benjamini-Hochberg); *** red: short *tlm* mutant, blue: long *tlm* mutant

Table S15. Significant regulators of *EST1* using the regulator complex Sum1-Hst1 instead of the single regulators Sum1 and Hst1

	Regulator	Z-score*	Significance (P)**	Number of targets
EST1	Sum1-Hst1***		4.51 E-32	
	Msn4	-0.63	1.72 E-15	2483
	Mig1	0.14	4.66 E-8	423
	Rfx1	-0.45	1.17 E-7	660
	Gcn4	-0.13	2.68 E-5	2712
	Srb2***	2.08	1.33 E-3	785
	Mbp1	0.76	4.97 E-3	665
	Sin3***	2.01	6-33 E-3	1759
	Fhl1	-	2.53 E-2	1164
	Cst6	2.34	2.61 E-2	2601
	Tec1	1.52	2.65 E-2	3653
	Cup2	-0.30	2.98 E-2	548

* Effect of the knockout of the regulator on the expression of the *EST* genes (positive z-score = upregulation of the corresponding *EST* gene; negative z-score = downregulation of the corresponding *EST* gene); ** Multiple testing corrected (Benjamini-Hochberg); *** red: short *tlm* mutant, blue: long *tlm* mutant

References

1. Li, M., Valsakumar, V., Poorey, K., Bekiranov, S. and Smith, J.S. (2013) Genome-wide analysis of functional sirtuin chromatin targets in yeast. *Genome biology*, **14**, R48.
2. Bedalov, A., Hirao, M., Posakony, J., Nelson, M. and Simon, J.A. (2003) NAD+-dependent deacetylase Hst1p controls biosynthesis and cellular NAD+ levels in Saccharomyces cerevisiae. *Molecular and cellular biology*, **23**, 7044-7054.
3. McCord, R., Pierce, M., Xie, J., Wonkatal, S., Mickel, C. and Vershon, A.K. (2003) Rfm1, a novel tethering factor required to recruit the Hst1 histone deacetylase for repression of middle sporulation genes. *Molecular and cellular biology*, **23**, 2009-2016.
4. Zill, O.A. and Rine, J. (2008) Interspecies variation reveals a conserved repressor of alpha-specific genes in Saccharomyces yeasts. *Genes Dev*, **22**, 1704-1716.
5. Lai, X., Schmitz, U., Gupta, S.K., Bhattacharya, A., Kunz, M., Wolkenhauer, O. and Vera, J. (2012) Computational analysis of target hub gene repression regulated by multiple and cooperative miRNAs. *Nucleic acids research*, **40**, 8818-8834.
6. Askree, S.H., Yehuda, T., Smolikov, S., Gurevich, R., Hawk, J., Coker, C., Krauskopf, A., Kupiec, M. and McEachern, M.J. (2004) A genome-wide screen for Saccharomyces cerevisiae deletion mutants that affect telomere length. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 8658-8663.
7. Ben-Shitrit, T., Yosef, N., Shemesh, K., Sharan, R., Ruppin, E. and Kupiec, M. (2012) Systematic identification of gene annotation errors in the widely used yeast mutation collections. *Nature methods*, **9**, 373-378.
8. Gatbonton, T., Imbesi, M., Nelson, M., Akey, J.M., Ruderfer, D.M., Kruglyak, L., Simon, J.A. and Bedalov, A. (2006) Telomere length as a quantitative trait: genome-wide survey and genetic mapping of telomere length-control genes in yeast. *PLoS Genet*, **2**, e35.
9. Shachar, R., Ungar, L., Kupiec, M., Ruppin, E. and Sharan, R. (2008) A systems-level approach to mapping the telomere length maintenance gene circuitry. *Molecular systems biology*, **4**, 172.
10. Ungar, L., Yosef, N., Sela, Y., Sharan, R., Ruppin, E. and Kupiec, M. (2009) A genome-wide screen for essential yeast genes that affect telomere length maintenance. *Nucleic acids research*, **37**, 3840-3849.
11. Reimand, J., Vaquerizas, J.M., Todd, A.E., Vilo, J. and Luscombe, N.M. (2010) Comprehensive reanalysis of transcription factor knockout expression data in Saccharomyces cerevisiae reveals many new targets. *Nucleic acids research*, **38**, 4768-4777.