# SUPPLEMENTARY PLOTS for paper Global mapping of DNA Conformational Flexibility on Saccharomyces cerevisiae 

by Giulia Menconi, Andrea Bedini, Roberto Barale and Isabella Sbrana

This document contains plots and graphs which is referred to in the text of main paper.

Table 1: Summary on flexibility data for each yeast chromosome: length, measurements over threshold $\theta$, regions and peaks. Finally a global view of genome covered by flexible zones is given.

| Chrm | Length (bp) | Peaks |
| :--- | :---: | :---: |
| I | 230208 | 2 |
| II | 813178 | 11 |
| III | 316617 | 7 |
| IV | 1531919 | 27 |
| V | 576869 | 12 |
| VI | 270148 | 5 |
| VII | 1090947 | 20 |
| VIII | 562643 | 11 |
| IX | 439885 | 4 |
| X | 745742 | 5 |
| XI | 666454 | 13 |
| XII | 1078175 | 19 |
| XIII | 924429 | 14 |
| XIV | 784333 | 5 |
| XV | 1091289 | 18 |
| XVI | 948062 | 10 |
|  | genome covered | $0.38 \%$ |



Figure 1: Snapshot of UCSC visualization of flexibility data for some peaks lying at $3^{\prime}$ UTR unidirectional intergenic regions: (a) peakX-5 between IME1 (YJR094C) and FIP1 (YJR093C); (b) peakIV-14 between DBF4 (YDR052C) and DET1 (YDR051C); (c) peakIV-5 between CDC53 (YDL132W) and LYS21 (YDL131W).


Figure 2: Snapshot of UCSC visualization of flexibility data for some peaks lying at convergent intergenic regions: (a) peakII-10 between ERV15 (YBR210W) and AME1 (YBR211C); (b) peakI-1 between SNC1 (YAL030W) and MYO4 (YAL029C); (c) peakIV-27 betweenDIG2 (YDR480W) and PHO8 (YDR481C).


Figure 3: Snapshot of UCSC visualization of flexibility data for some peaks lying at divergent intergenic regions: (a) peakXII-3 between PUF3 (YLL013C) and YEH1 (YLL012W); (b) peakX-3 between TDH2 (YJR009C) and MET3 (YJR010W).
peakIV-14 (unidirectional $\lll \lll$ ) : DBF4 (YDR052C) downstream - DET1 (YDR051C) upstream

| >Skud | TTCTTGCACTAATATTTTTTAGTTGCTCAAACAGAATGCTGTAGTAAA-TTTCTCACAGGATAAACTATAAATACTAGTTGCAATTAATGAGGTTGAGAAGCTGCCTGGTTTTTTTTTTTTCATG-TCTCGGCCTATGTAAGCTATCGGATTGAAACTGAA- |
| :---: | :---: |
| >Smik | TTCTTGGGCTTATGCTTTTCAGGTATTCGAATGGAATATTGTGGCATA-GCCTTTACGTAATAAACAACTAATGTAAGCAGCAACTAATGCGGTtGAGAAGCTT-----TTT--TTTTTTTCCTGCTTCACGGCCTGCGCAAGCTATCGGGTAGACATAGAA-GGTA- |
| >Spar | TTCTTATTCTAGTATTTCTGGGTTATTCGAATGGGATGCAGTATTATA-GCTCTCATATGTCAAATCATGACTACAAGCAGCAACCAATGCAGTTGAGAAGCTG-----CTTTATTTTTTTTATTCCTCGCGGCCTCTGCAGGATATCGGGTAGAAATGAAAAACTA- |
| >Sbay |  |
| >Scer | TTCTTTTACTAATACT-CTGAGTTATTCAAATGGAATGCTGTAGCGTAAGTTCTTTTATGTtGAACCATGAATACGAGCTGCAATtAATGCAGTTGAGAACTAG-----CTTTTTTTTTTTCATGCCTCTCGGTCTCTGCAGGATATCGGGTAAAAATGAAA-ACTG- |
| >Skud |  |
| >Smik |  |
| >Spar |  |
| >Sbay |  |
| >Scer |  |
| >Skud |  |
| >Smik |  |
| >Spar |  |
| >Sbay |  |

Figure 4: Alignment of $D B F 4-D E T 1$ unidirectional intergenic region for Saccharomyces sensu stricto species (peakIV-14). For each genomic sequence, motif AAWAAA is underlined and waved. Only for $S$. cerevisiae, TA repeats are underlined, poly(A) sites are overlined and $3^{\prime} \mathrm{UTR}$ is shown in bold.
peakIV-9 (CONV >>>> <<<<<<<) : RAD59 (YDL060W) downstream - TSR1 (YDL060W) upstream

|  | peakIV-9 (CONV $\ggg>\lll \lll)$ : RAD59 (YDL060W) downstream - TSR1 (YDL060W) upstream |
| :---: | :---: |
| >Skud |  |
| >Smik | AGTGATGAGAATATATACGTATATTGAT-----------------GTATATGTGATCTATTTGTATAATATAA-TGCTAATAAGAAAAGTATCAGTGGAGCTATTAGTAAAGATAAGTTGATATATGTACCACCCGCATCAATT---TAC- |
| >Spar | AGTGATAAAAATATACACGTATATATAC----------------------- |
| >Sbay | AGTGATAAAAATATGCACATATAGGTATGTATGTATGTA-----TGTGTGTGAGATCTATTTGTATAATATAATTACTAATAA-AACAGTATTGAGGGAGTTATGAGTAAAGCTAGGGCAATACATGTtTCGTACGTATCAAATAGATAC |
| >Scer |  |
| >Skud |  |
| >Smik |  |
| >Spar | -ATCACTTATT----GTGTAGAAATATTAATGCAT-TATGTGAAAATCATAAGTT-TTATCAAGCAAAAATGAATTTGCTGGCT--TtTTATTCTTTTCTTTTCTTTTCTT--------TTCTTTTTT |
| >Sbay |  |
| >Scer | -GCCACTTATT-----TTGTAGAGGTATTAATGCAT-TATGTGAAAATTATGACTTTTTATCAAGCAAA-ATAAATTTGCTAC--тTGT--GСССтTTTTCTTTCTTTT----------------TTTTT |

Figure 5: Alignment of RAD59-TSR1 convergent intergenic region for Saccharomyces sensu stricto species. peakIV-9 is on $3^{\prime}$ UTR of $R A D 59$. For each genomic sequence, motif AAWAAA is underlined and waved. Only for $S$. cerevisiae, TA repeats are underlined, poly(A) sites are overlined and $3^{\prime}$ UTR is shown in bold. The peak is not conserved in S. kudriavzevii.


Figure 6: Treemaps of the outcomes of REVIGO for Biological Process (Top) and Molecular Functions (bottom) GO terms, referring to 175 ORFs characterized by a peak in $3^{\prime} \mathrm{UTR}$.


Figure 7: Distribution of distance of peaks from nearest genomic regions with strong nucleosome depletion (the so-called boundaries). Bins are 1 kb and count for peaks within that distance from nearest boundary. Only 23 peaks are closer than 1000bp. Data are from: Field Y, Kaplan N, FondufeMittendorf Y, Moore IK, Sharon E, et al. (2008) Distinct Modes of Regulation by Chromatin Encoded through Nucleosome Positioning Signals PLoS Comput Biol 4(11): e1000216.

