Long-range single-molecule mapping of chromatin accessibility in eukaryotes

Zohar Shipony[1,*], Georgi K. Marinov[1,*], Matthew P. Swaffer[5], Nicholas A. Sinnott-Armstrong[1], Jan M. Skotheim[5], Anshul Kundaje[1,2], and William J. Greenleaf[1,3,4,#]

[1] *Department of Genetics, Stanford University, Stanford, CA 94305, USA*
[2] *Department of Computer Science, Stanford University, Stanford, CA 94305, USA*
[3] *Department of Applied Physics, Stanford University, Stanford, CA 94305, USA*
[4] *Chan Zuckerberg Biohub, San Francisco, California, USA*
[5] *Department of Biology, Stanford University, Stanford, CA 94305, USA*
[*] *These authors contributed equally to this work*
[#] *Corresponding author*

Abstract

Active regulatory elements in eukaryotes are typically characterized by an open, nucleosome-depleted chromatin structure; mapping areas of open chromatin has accordingly emerged as a widely used tool in the arsenal of modern functional genomics. However, existing approaches for profiling chromatin accessibility are limited by their reliance on DNA fragmentation and short read sequencing, which leaves them unable to provide information about the state of chromatin on larger scales or reveal coordination between the chromatin state of individual distal regulatory elements. To address these limitations, we have developed a method for profiling accessibility of individual chromatin fibers at multi-kilobase length scale (SMAC-seq, or Single-Molecule long-read Accessible Chromatin mapping sequencing assay), enabling the simultaneous, high-resolution, single-molecule assessment of the chromatin state of distal genomic elements. Our strategy is based on combining the preferential methylation of open chromatin regions by DNA methyltransferases (CpG and GpC 5-methylcytosine (5mC) and N^6-methyladenosine (m^{6}A) enzymes) and the ability of long-read single-molecule nanopore sequencing to directly read out the methylation state of individual DNA bases. Applying SMAC-seq to the budding yeast *Saccharomyces cerevisiae*, we demonstrate that aggregate SMAC-seq signals match bulk-level accessibility measurements, observe single-molecule protection footprints of nucleosomes and transcription factors, and quantify the correlation between the chromatin states of distal genomic elements.

The packaging of DNA by nucleosomes into chromatin is a major organizing principle of genome organization in eukaryotes. The majority of the genome is tightly packaged by nucleosomal particles that wrap around DNA (usually ∼147bp), thus making it inaccessible to binding by most regulatory proteins. Conversely, regions of open chromatin tend to be strongly associated with regulatory elements (REs), such as enhancers, promoters, and insulators, and nucleosomes often exhibit characteristic occupancy patterns in their vicinity. These biological properties have proven highly useful for identifying candidate such elements (cREs), and in turn for understanding the functional organization of genomes. Regions of open chromatin have greatly increased sensitivity to cleavage by nucleases such as DNAse I, as already noted nearly four decades ago for promoter and enhancer elements around individual genes[1–3]. Subsequent advances in microarray[4,5] and DNA sequencing technologies[6,7] have enabled DNAse hypersensitivity-based mapping of cREs genome-wide. Similarly, digestion of DNA is inhibited by nucleosome occupancy, and MNase digestion of chromatin (MNase-seq) has accordingly become a widely used tool to map nucleosome positioning throughout genomes[9]. More recently, the Tn5 transposase has also been used as a facile probe of chromatin accessibility[8]. However, while short read-based assays provide immensely useful information about the identity of cREs and positioned nucleosomes, they give little insight into the long-range physical organization of individual chromatin fibers. Because cleavage-based approaches remove the linkage between distal segments of DNA molecules, classical functional genomic assays return estimates of the relative chromatin states only at short localized regions of the genome. To what extent these states are coordinated across distance and what the distribution of

chromatin states looks like at the kilo– to multikilobase scale is largely unknown. Well established tools exist for mapping direct interactions between distal genomic elements in the context of three-dimensional genome architecture[10–12] but they only capture pairwise interactions within large cell populations, not the chromatin state of the interacting regions, and they are often of limited resolution. Similar limitations apply to emerging methods for studying local chromatin structure[13]. Super-resolution microscopy using highly multiplexed fluorescent probes is a powerful tool for revealing the folding of the chromatin fiber at the single cell level[14], but this approach is limited to a small number of loci and does not provide high-resolution base-pair level information about regulatory states. Cryo-electron microscopy-based approaches[15] allow direct observations of the nucleosomal state along individual chromatin fibers, but at present the underlying DNA sequence cannot be linked to these measurements. More recently, long-read sequencing was used to map MNase cleavage events at multinucleosomal lengths[22]; however, such approaches only provide information about two points in genomic space leaving the state of chromatin in the intervening sequence unknown.

These technological limitations prevent the investigation of the manner in which chromatin states of adjacent elements are correlated and how such coordination might play a role in gene regulation (for example, by creating self-reinforcing or mutually exclusive epigenetic states). To address these technological limitations, we have developed SMAC-seq, a versatile, single molecule method that directly assays long-range nucleosome positioning and accessibility states within the chromatin fiber. We use this method to study chromatin architecture and co-accessibility states in the yeast *Saccharomyces cerevisiae*, both under normal growth and conditions of cellular stress. We use SMAC-seq to assess the degree of coordination between the positions of nearby nucleosome particles, enumerate mutually exclusive regulatory states along individual loci, and observe coordinated changes in nucleosome positioning and chromatin accessibility upon transcriptional activation. SMAC-seq also allows for footprinting of transcription factor (TF) occupancy, and provides strand-specific information about the exposure of DNA bound to nucleosomes and other proteins. We expect future applications of, improvements on, and extensions of the SMAC-seq approach to enable novel insights into the dynamics of chromatin states in the context of a wide variety of experimental systems and biological questions.

Results

SMAC-seq maps chromatin accessibility and nucleosome positioning at the multi-kb scale

SMAC-seq is built on the conceptual foundations of the NOMe-seq assay[16,17]. In its original form, NOMe-seq relies on the preferential modification of bases within accessible DNA with M.CviPI (a GpC-specific 5mC methyltransferase), followed by bisulfite conversion and Illumina-based sequencing readout. A more recent variation, dSMF[18] (**d**ual-enzyme **S**ingle **M**olecule **F**ootprinting), utilized an additional CpG-specific 5mC methyltransferase (M.SssI) to map promoter states in *Drosophila* cells in finer detail. Such an approach is applicable in the *Drosophila* context thanks to the absence of endogenous CpG methylation in flies.

Here, we utilize the m^6A methyltransferase EcoGII[19] as an alternative/addition to CpG/GpC methylation, and use nanopore sequencing to generate single-molecule readouts of accessibility states over many kilobases (Figure 1A). Briefly, isolated nuclei are treated with methyltransferase(s), preferentially modifying DNA within accessible chromatin. High-molecular weight (HMW) DNA is then isolated and subjected to long-read single-molecule sequencing using the Oxford Nanopore platform. Using the ability of nanopore sequencing to directly read modified DNA bases[20,21], we then obtain methylation maps for individual DNA molecules on a multikilobase scale (see the Methods section for details), which we then interpret in terms of chromatin accessibility.

The addition of m^6A methylation is of key importance to making SMAC-seq a high-resolution broadly applicable assay. Many eukaryote genomes are endogenously methylated at 5mC positions[23,24]. This is usually in a CpG context, in particular in metazoans, but it is not always limited to CpG dinucleotides. For example, in plants methylation in CHG and CHH contexts is also a common occurrence[25], thus even results obtained with M.CviPI alone are confounded by endogenous methylation. In addition, CpG and GpC dinucleotides are relatively rare in the genome. The average resolution achieved by the combination of CpG and GpC methyltransferases is >10 bp in *Drosophila* and ∼15 bp in yeast; it is ∼25bp for GpC alone in mammalian genomes. However, these are genome-wide theoretical averages; in practice many individual regions of interest either completely lack informative positions or contain so few of them that the signal-to-noise ratio is too high for reliable conclusions to be obtained. The addition of m^6A dramatically increases the resolution of SMAC-seq, down to a theoretical limit of ∼3 bp in all main model organisms, and ensures proper coverage of the genome over all individual loci (Supplementary Figures 1-3 and 16-49).

We initially developed and optimized the method in the budding yeast *Saccharomyces cerevisiae* under normal growth conditions, as *Saccharomyces cerevisiae* has no endogenous DNA methylation (allowing the simultaneous use of all three enzymes) and has a small genome (∼12 Mbp), making it possible to achieve very high depth of nanopore sequencing coverage. To verify the specificity and efficiency of the enzymatic treatments, we carried out both dSMF experiments (i.e M.CviPI + M.SssI treatment) on yeast chromatin and M.CviPI + M.SssI + EcoGII reactions on naked yeast DNA (as well as untreated naked DNA controls), and subjected the resulting material to whole-genome bisulfite sequencing (see the Methods section for

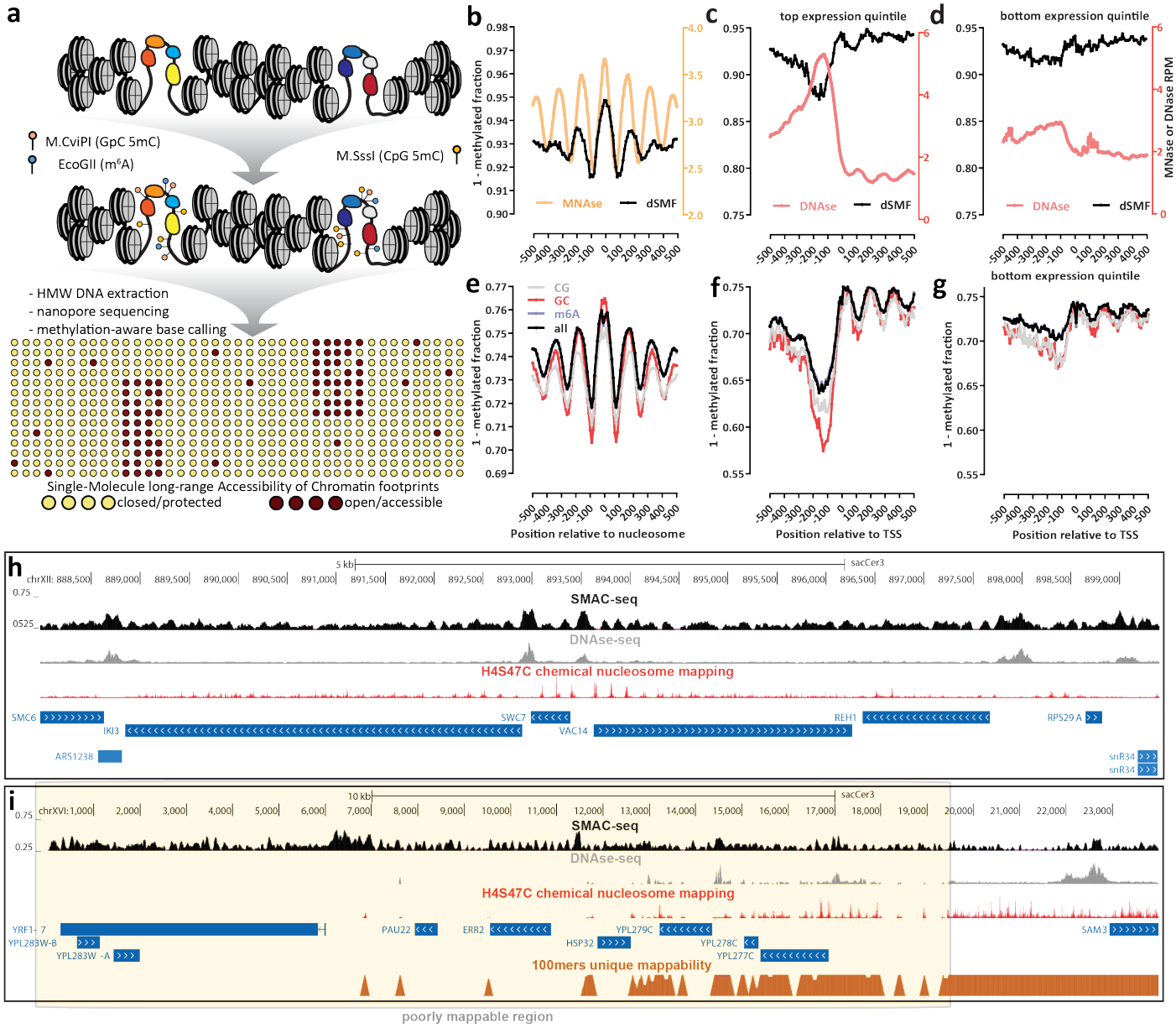


Figure 1: The SMAC-seq assay for profiling chromatin accessibility and nucleosome positioning at the multikilobase scale. (a) Outline of the SMAC-seq assay. Intact chromatin is treated with m^6A and CpG and GpC 5mC methyltransferases, which preferentially methylate DNA bases in open chromatin regions. HMW DNA is then isolated, subjected to nanopore sequencing and methylated bases are used to reconstruct the open chromatin state within individual molecules. (b-h) SMAC-seq faithfully captures chromatin accessibility around promoters and positioned nucleosomes in *S. cerevisiae*; (b) MNAse-seq and dSMF profiles around chemically mapped positioned nucleosome dyads; (c) DNAse-seq and dSMF profiles around the top 20% highly expressed genes in *S. cerevisiae*; (d) DNAse-seq and dSMF profiles around the bottom 20% expressed genes in *S. cerevisiae*; (e) Average SMAC-seq profile around chemically mapped positioned nucleosomes dyads (shown is the "diamide 0 min rep2" sample); (f) Average SMAC-seq profile around the top 20% highly expressed genes in *S. cerevisiae*; (g) Average SMAC-seq profile around the bottom 20% expressed genes in *S. cerevisiae*; (h) SMAC-seq correlates closely with both DNAse-seq and nucleosome occupancy profiling at the level of individual loci, and provides a combined readout of accessibility and nucleosome positioning. Shown is aggregate SMAC-seq signal along the genome (aggregated over 50-bp windows sliding every 5 bp; see Methods for details) together with DNAse-seq, nucleosome chemical mapping data, and transcriptional activity (measured by PRO-seq and PRO-cap). Large aggregate SMAC-seq signal enrichments match closely with DNAse accessibility peaks, while smaller aggregate SMAC-seq peaks are inversely correlated with positioned nucleosomes; (i) SMAC-seq profiles chromatin accessibility in repetitive regions of the genomes that are "invisible" to short reads. Shown is the telomeric region of chrXVI

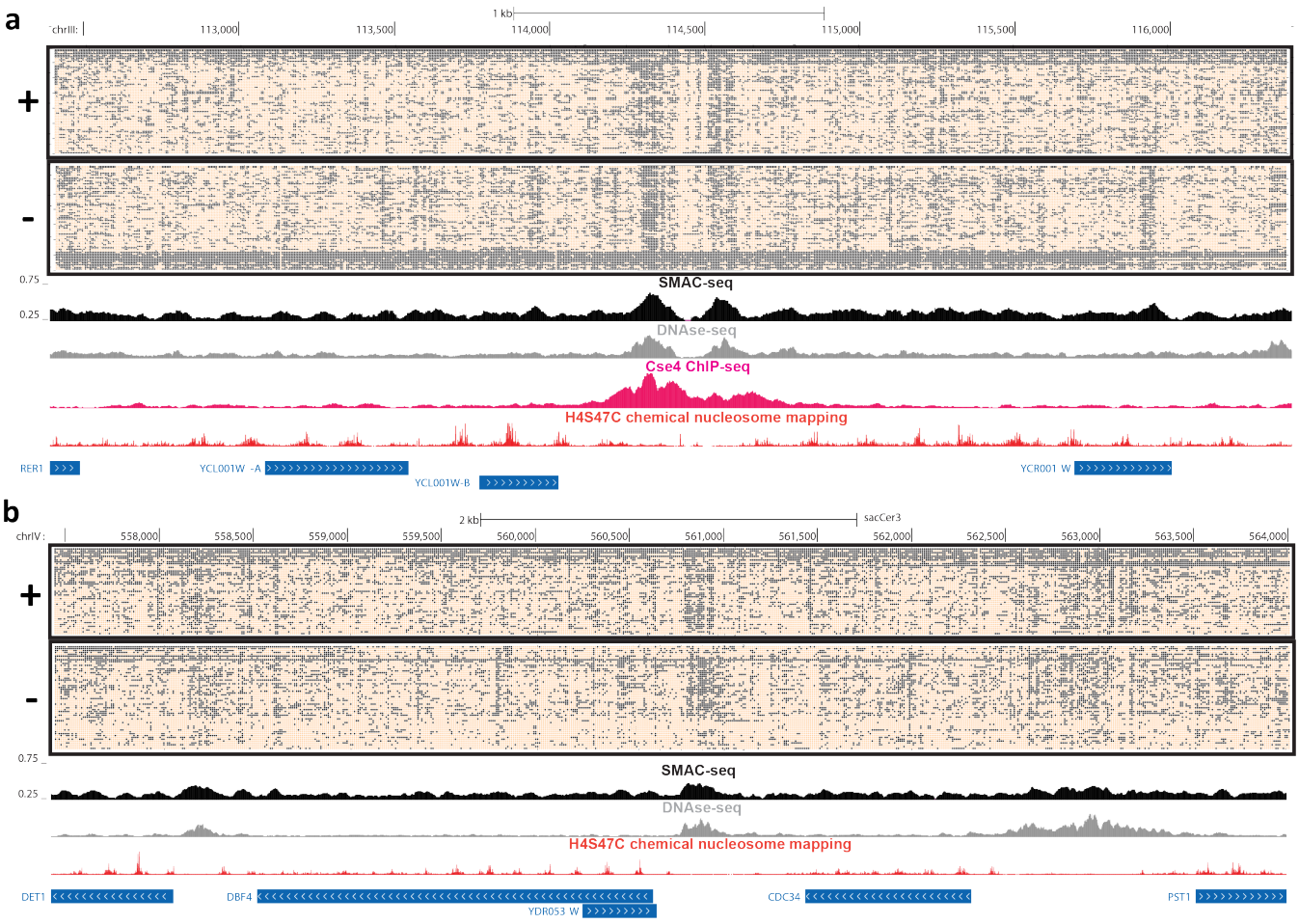


Figure 2: SMAC-seq provides a single-molecule linked-read view of the chromatin landscape. (a) Unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrIII ("aggregate" signal "Sample 1"). (b) Unfiltered nanopore reads fully spanning a 6.6-kilobase neighborhood encompassing several genes on chrIV ("aggregate" signal from "Sample 1"). In both cases, accessibility is shown at 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the average SMAC-seq track.

more details). We observe ≥95% methylation in both the CpG and GpC contexts when starting with naked DNA, ≤10% when working with chromatin, and nearly 0% on untreated naked DNA (Supplementary Figure 6). Comparing dSMF profiles to DNase-seq and MNase-seq profiles around transcription start sites (TSSs) and positioned nucleosomes (obtained from previous H4S47C chemical mapping studies[26]) revealed the expected nucleosome depletion and nucleosomal patterns, respectively (Figure 1b-d).

The efficiency of methylation by EcoGII is more difficult to estimate as fully methylated templates are notoriously difficult to sequence on the Oxford Nanopore platform. We did not obtain many reads using genomic DNA treated with a high dose of EcoGII (Supplementary Table)1; the available reads exhibited ∼50% methylation levels (Supplementary Figure 7). We also obtained a ∼40% methylation levels (Supplementary Figure 8) using EcoGII-treated λ DNA. However, these are likely underestimates given the aforementioned sequencing bias against fully methylated DNA. The original description of the EcoGII enzyme reported 50% methylation of naked DNA after 5 minutes of treatment, increasing to ≥85% after an hour of incubation; the efficiency of methylation of accessible bases within chromatin probably lies somewhere in between[19].

We then carried out a SMAC-seq experiment on unsynchronized *S. cerevisiae* cells using all three enzymes (experimental details described in the Methods section). We isolated HMW DNA and performed nanopore sequencing on the MinION platform. To call methylated bases, we first used Albacore for raw base calling, and then applied the Tombo[27] algorithm in "de novo" mode (running on top of the Minimap aligner[28]) for "resquiggling" of raw nanopore signal to the genomic sequence, and calling of methylated bases. After mapping, we obtained reads with a median

length of ∼1.5 kbp from this initial experiment ("Sample 1"; Supplementary Table 1), which allows the capture of multiple promoter regions per fragment for much of the yeast genome (Supplementary Figures 4 and 5). We also analyzed our initial dataset with Nanopolish[20], an alternative algorithm for calling methylated bases, which is capable of identifying 5mC events in CpG and GpC context.

Unlike bisulfite-based conversion followed by sequencing-by-synthesis, nanopore-based direct measurement of nucleotide modifications does not provide unambiguous binary calls for methylated bases. Instead, methylation probabilities are obtained for each base. We therefore examined multiple strategies for binarizing these methylation calls within each read. We observed that while per-base methylation probabilities are skewed towards the two ends of the $[0, 1]$ interval, a substantial number of bases have probabilities in between those extremes (Supplementary Figure 9). We find that binarization at $p = 0.5$ delivers the most optimal results (Supplementary Figures 10 and 11) in terms of signal-to-noise ratio. Using a simple average of these binarized per-base methylation values, we compared SMAC-seq to dSMF, MNase-seq, and DNase-seq profiles, as well as signal from ChIP-seq for RNA Polymerase (Pol2) and transcription initiation factors around known chromatin features such as active promoters and well positioned nucleosomes (Figure 1e-g, Supplementary Figures 50 and 53). SMAC-seq faithfully reproduces nucleosomal positioning throughout the genome and the extensively previously documented[29] nucleosome depletion around promoters, and, strikingly, has a larger observed dynamic range than dSMF data (possibly due to the higher mapping efficiency of long nanopore reads). Using comparisons with dSMF data and sequencing of untreated genomic DNA we estimate the false positive rate of methylation base calling to be on the order of 20% for Tombo and around 10-15% for Nanopolish (Supplementary Figures 15 and 7). We also examined potential sequence biases inherent to the combination of methylation enzymes and base calling algorithm. We find modest differences in methylation levels for different k-mers in the genome (less than two-fold for $k = 6$; Supplementary Figures 13 and 14).

In practice, the biologically relevant length scale of accessibility measurements is usually larger than an individual base. Furthermore, given noise intrinsic to single molecule sequencing methods, we reasoned that sharing methylation information between adjacent bases should improve the reliability of overall accessibility measurements. We therefore developed a simple Bayesian procedure to aggregate the methylation probabilities at individual bases and derive accessibility calls at the level of single reads over windows of arbitrary size (described in detail in the Methods section and thereafter referred to as "aggregate" signal as opposed to "average" signal, which refers to simple probability averaging and binarization). This approache decreases the assay resolution from the theoretical maximum while providing more reliable accessibility estimates over slightly larger windows.

We also note that sometimes we observe a subpopulation of reads that appear to be either entirely highly methylated or highly methylated over large segments of their length (Supplementary Figure 12 and Figure 2a). We interpret these as originating from naked DNA molecules most likely deriving from dead cells. As such reads can confound many analyses, in particular when measuring coaccessibility within single reads, we devised a procedure for filtering them out (the resulting sets of reads are referred to as "filtered reads"; (Supplementary Figures 57 and 58). However, as discussed below, in certain situations chromatin is indeed largely nucleosome-free over specific regions in vivo, and in such cases filtering out highly methylated reads removes real biological signal. For these unique special case loci, we do not eliminate reads based on a very high fraction of accessibility (see Methods).

We then compared average SMAC-seq profiles against chemical maps of positioned nucleosomes (generated using H4S47C substitutions and copper-induced cleavage[26]), DNase-seq, and maps of transcriptional activity at the level of individual loci in the genome (Figure 1h). Qualitatively, we observe that large peaks in SMAC-seq signal profiles match very closely with DNase-seq peaks, while smaller "bumps" in the SMAC-seq signal profile are inversely correlate with positioned nucleosomes, consistent with labeling of linker DNA. We also observe positive correlation between average SMAC-seq methylation levels and DNase-seq and ATAC-seq coverage over promoter regions (Supplementary Figure 52). Thus SMAC-seq simultaneously probes both regions of "open" chromatin, as well as the position of nucleosomes.

The long reads of nanopore sequencing allow SMAC-seq to provide accessibility maps for the whole yeast genome, not just for the portions of it that are uniquely mappable with short reads (Supplementary Figure 56). For example, SMAC-seq maps chromatin and nucleosomes in the otherwise not uniquely mappable telomere of chrXVI (Figure 1i), which is revealed to contain several active promoters and numerous well positioned nucleosomes. We also used SMAC-seq to characterize chromatin accessibility around multiple transposable elements (Supplementary Figures 55), for several of which we observe open chromatin peaks around their promoters.

SMAC-seq provides single-molecule accessibility profiles on individual chromatin fibers

Unlike short-read methods for probing chromatin, SMAC-seq allows the profiling of nucleosome positioning and open chromatin within individual long molecules. To demonstrate this capacity, we investigated all reads spanning the 4-kb neighborhood around the centromere of chrIII (Figure 2a). Centromeres in *S. cerevisiae* are specified by a precisely defined sequence element, are occupied by a single strongly positioned nucleosome containing the H3 histone variant Cse4, and are also associated with the CBF3 com-

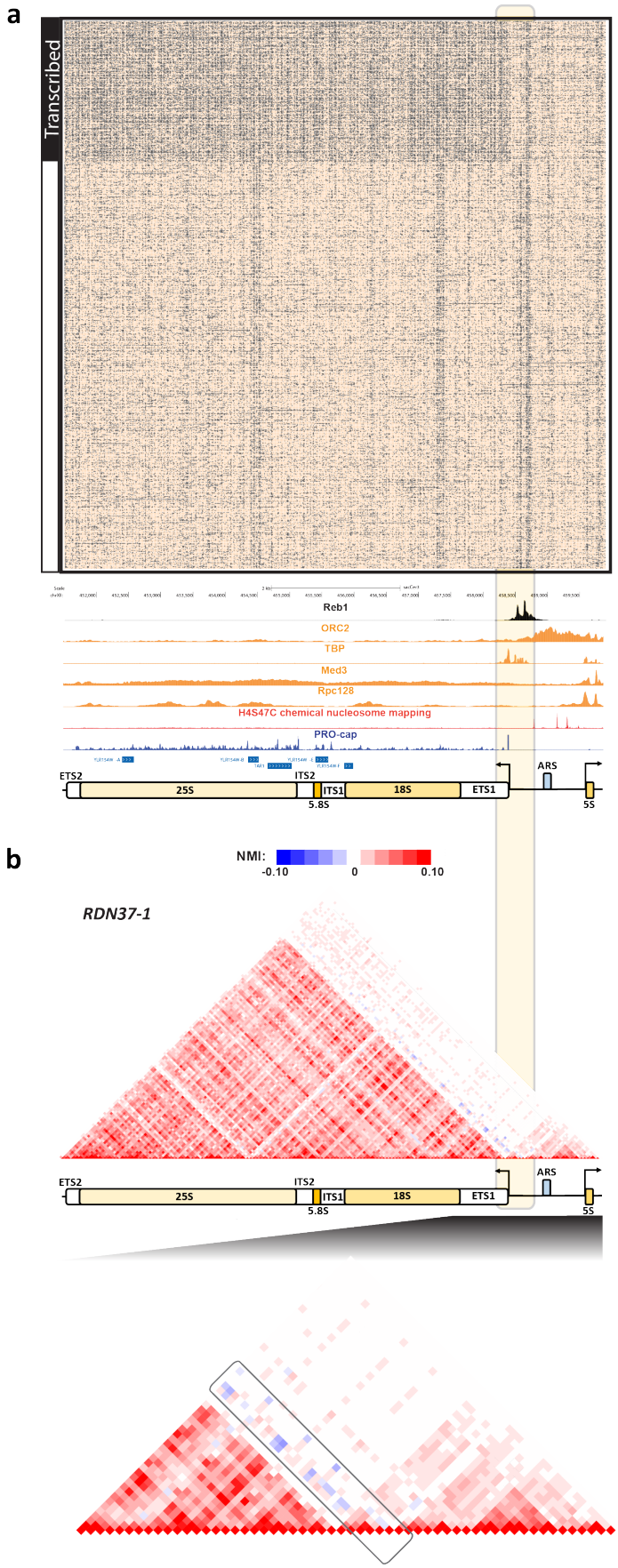


Figure 3: SMAC-seq's single-molecule readout provides insights into the distribution and relationship between mutually exclusive chromatin yeast rDNA states. (a) SMAC-seq reveals the distribution of alternative chromatin states of rDNA arrays. Shown are all reads covering the *RDN37-1* array in the *RDN1* locus in the "diamide 30 min rep1" experiment (unfiltered reads, "aggregate" signal). See Supplementary Figures 68–71 for additional details. ChIP-seq and ChIP-exo tracks were generated by including and normalizing all multimappers rather than the usual unique-only policy (See the Methods section for more details). (b) Normalized mutual information profiles for the *RDN37-1* array show anti-correlation between the accessibility peaks immediately upstream of the *35S* TSS and the nucleosome-free state over the *35S* transcriptional unit. (c) High-resolution SMAC-seq profiles reveal regulatory protein footprints in the immediate vicinity of the 35S TSS and the Reb1 binding site in the rDNA NTS region (shown are 3000 randomly sampled reads using 10-bp aggregate SMAC-seq signal at 1-bp resolution).

plex, an essential component of the kinetochore[30]. Yeast centromeric nucleosomes are thought to be nearly perfectly positioned[30,31] and thus represent a good case system to study nucleosome positioning at the single-molecule level. We indeed observe strong nucleosomal positioning using SMAC-seq, with nearly all individual reads exhibiting the expected from the presence of a strongly positioned centromere nucleosomal pattern. We also find hints of substructure within the centromeric nucleosome in the form of accessibility traces inside the protected centromeric region and potential protection footprints in its immediate open chromatin vicinity. We find similarly strong positioning for most other centromeric nucleosomes (Supplementary Figures 59-66), but not all (chrX, chrXII and chrXIII appear to be exceptions to this general strong footprinting pattern).

We also illustrate the ability of SMAC-seq to capture accessibility at long-range multikilobase scales with an example from a more generic genomic neighborhood in Figure 2b). This ∼6.6-kb span of chrIX contains five genes and three open chromatin regions, one of them fairly large and diffuse. In contrast to the more localized accessibility observed elsewhere, this region exhibits considerable heterogeneity in its accessibility suggesting a complex landscape of protein occupancy.

We next asked if SMAC-seq could reveal binary states of chromatin accessibility. To approach this question we investigated SMAC-seq profiles at ribosomal DNA (rDNA). In yeast, rDNA is organized into multicopy arrays, each unit of which contains a copy of the 35S precursor pre-rRNA, transcribed by Pol I and later processed into mature 18S, 5.8S and 25S rRNAs, as well as a copy of the 5S RNA gene, transcribed by Pol III, and an origin of replication (ARS element in yeast) located in non-transcribed (NTS) regions of the array. Each array unit is ∼9.1 kb in length, and each cell's genome has an array of ∼150 copies of this unit[32]. However, this number can vary from cell to cell, and the widely used `sacCer3` *S. cerevisiae* genome assembly only contains a single locus with two array copies. Chromatin structure at the rDNA locus has long been known to adopt two distinct conformations[33–35], depending on whether or not an individual unit is being transcribed. The active state is thought to be largely devoid of nucleosomes due to the extremely high levels of active transcription[34]; the high-mobility group protein Hmo1 is proposed to replace nucleosomes[32,36]. However, other studies have alternatively suggested that nucleosomes are found over actively transcribed rDNA arrays[37]. Of note, rDNA indeed appears to be extremely accessible in short-read assays – around half of reads in a typical ATAC-seq dataset in *S. cerevisiae* are not uniquely mappable even though only a small fraction of the yeast genome consists of repetitive elements; these reads originate primarily from rDNA arrays (Supplementary Figure 67a-e). The same phenomenon is also observed in other yeast species, such as *Schizosaccharomyces pombe* (Supplementary Figure 67f). Active and inactive rDNA arrays are usually estimated to exist in roughly equal proportions in untreated, normally growing cells[37]. However, methods to precisely observe these alternative states at the population level and relate them to sequence in fine detail have not been available.

SMAC-seq reveals a striking picture of the two alternative mutually exclusive rDNA states at the single molecule level (Figure 3a). About a quarter of full-length molecules exhibit near-full accessibility in the region spanning the 35S transcript, but not in the non-transcribed sequence (NTS) between the 35S and the 5S gene. The rest of the molecules show a typical nucleosomal state with several clearly accessible regions. We observe a broadly similar picture in all of our high-quality SMAC-seq samples (Supplementary Figures 68–71). We note that while this picture is in contrast with the usually reported ∼50% of arrays being fully accessible, it is possible that the fully accessible and nucleosomal states fragment differentially during DNA isolation; the fully accessible fraction appears to be in the 50% range when a shorter window around the 35S promoter is examined (Figure 3c). We also observe a localized accessible region just upstream of the 35S transcriptional unit that is present in the nucleosomal subpopulation but is not in an open state in the fully accessible population (in addition to a nearby open chromatin region present in all molecules), suggesting the possibility of a regulatory switch associated with that element. Finally, we also observe at least two (previously unreported) accessible regions located within the 35S transcriptional unit that exhibit strong accessibility in the nucleosome-protected fraction (Figure 3a).

To quantify the extent of (anti-)correlation between chromatin states, we developed a modified normalized mutual information (NMI) metric for assessing the degree of correlation between segments of the genome (see the Methods section for further details). NMI analysis of the rDNA arrays confirmed our observations of the inverse correlation between the 35S open-chromatin state and the accessibility of the upstream element (Figure 3b).

What factors might be behind the observed chromatin state switch? Silencing of rDNA in yeast is thought to be mediated by a Sir2-containing complex called RENT[38], and a Reb1 binding site in NTS1 has been suggested to recruit corepressors to rDNA repeats[37]. We took a higher-resolution view of NTS1 by integrating SMAC-seq data with available ChIP-exo data for Reb1 and transcription factor motif maps in the region. We find a clear pattern of protection from methylation around the Reb1 motif, which is concordant with ChIP-exo data (Figure 3c), and we also observe signal consistent with footprinting from several other TF motifs. However, the anti-correlated accessibility profile seems to not be exclusively associated with Reb1 binding but rather with the region closer to the 35S TSS. It is likely that other proteins are responsible for establishing this state, but no currently annotated transcription factor recognition motifs are found in the underlying sequence.

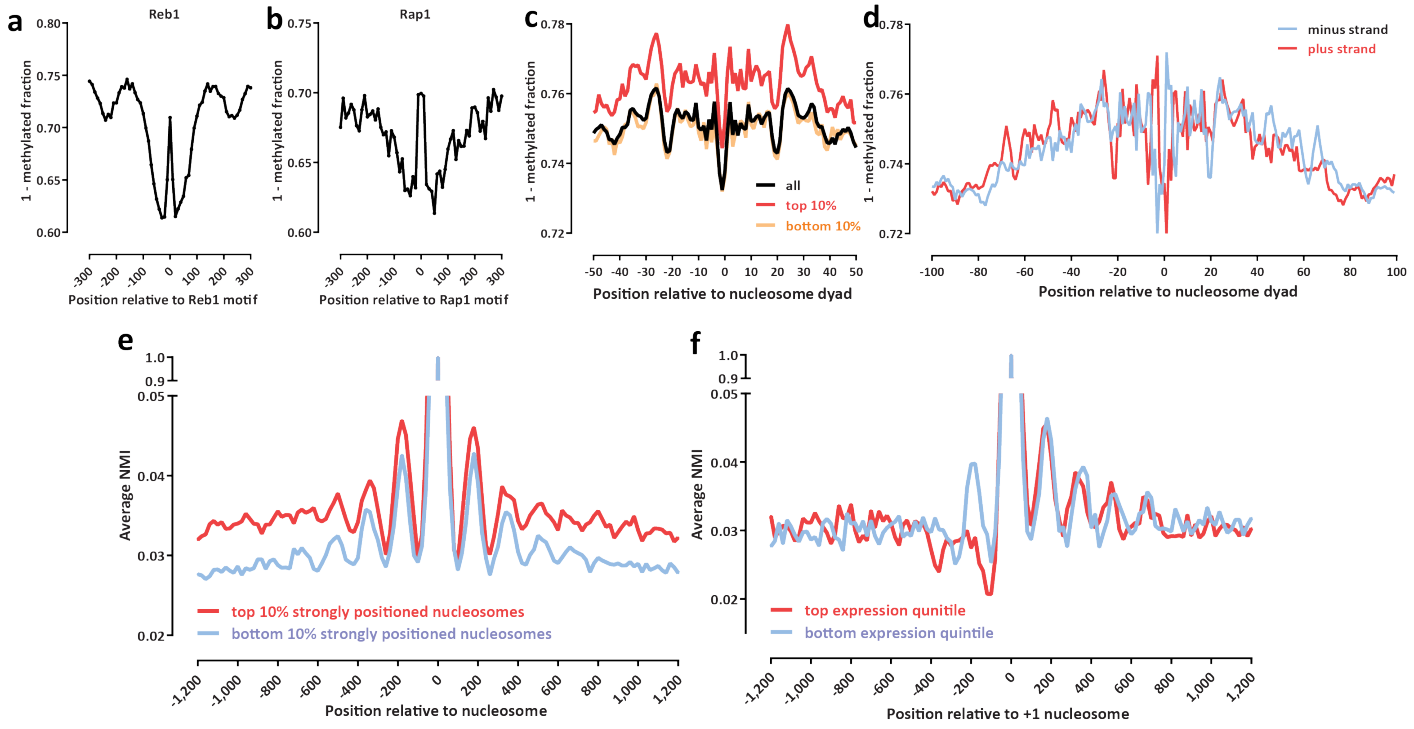


Figure 4: SMAC-seq provides a high-resolution strand-specific view of genomic occupancy by DNA-binding proteins and complexes. (a-b) SMAC-seq allows for footprinting of transcription factor binding events. Shown is aggregate genome-wide SMAC-seq signal around occupied (as measured by ChIP-exo) Reb1 (a), and Rap1 (b) sequence recognition motifs. (c) SMAC-seq profiles around positioned nucleosome dyads reveal increased accessibility in the dyad and increased protection at the points of contact with the nucleosome (see Supplementary Figure 53 for additional details.) (d) SMAC-seq provides a strand-specific view of nucleosome occupancy and reveals differential accessibility between the two DNA strands depending on their position on the nucleosomal particle. (e-f) Coordination between the positions of individual nucleosomes at the level of single chromatin fibers. (e) Shown is the average normalized mutual information between each strongly or poorly positioned nucleosome in the yeast genome and its immediate genomic neighborhood (measured for windows of 10 bp length tiling at every genomic position centered on the nucleosome dyad). (f) Shown is the average normalized mutual information between each +1 nucleosome and its immediate genomic neighborhood in highly expressed and in mostly silent genes (measured for windows of 10-bp length tiling at every genomic position centered on the +1 nucleosome dyad).

SMAC-seq provides a high-resolution strand-specific view of protein occupancy on DNA

We next asked if SMAC-seq can generally identify transcription factor footprints. We anticipated that the use of m^6A methylation ought to provide sufficient resolution to footprint many yeast TFs, which is confirmed by analysis of their consensus recognition motifs (Supplementary Figure 72). Averaging genome-wide SMAC-seq profiles over occupied motifs indeed revealed strong protection footprints for several factors, such as Reb1, Rap1 and ORC1 (Origin of Replication Complex) as shown in Figure 4b-c and Supplementary Figure 73. We also observed high concordance between the footprint profiles observed using DNase-seq, ATAC-seq and SMAC-seq over the footprints identified in the yeast genome by previous studes employing high-resolution DNAse-seq datasets[7] (Supplementary Figure 74). Examination of individual sites confirmed these observations (Supplementary Figures 75–83). We note that we did not observe strong footprinting for all TFs (e.g. Abf1 and Cbf1; Supplementary Figure 73) even though some of these TF do exhibit such footprints in DNase-seq datasets[7]. The likely explanation for this observation is that the different enzymes used to profile accessibility differ in their ability to access DNA in the context of protein occupancy; this explanation is supported by SMAC-seq profiles around positioned nucleosomes.

To further explore the limits of SMAC-seq's resolution, we studied methylation patterns around positioned nucleosomes in more detail (Figure 4e-f and Supplementary Figure 53). We find remarkably higher accessibility immediately at the dyad point, in contrast to the points of contact of the nucleosome with the DNA two helical turns away in each direction (Supplementary Figure 53). The same pattern was observed for all nucleosomes irrespective of the

strength of positioning (Figure 4e). We did not observe similar patterns in deeply sequenced DNAse-seq datasets (Supplementary Figure 53e).

Because SMAC-seq employs long-read single molecule sequencing and can therefore directly map accessibility independently on individual DNA strands, we next aimed to quantify strand-specificity in DNA accessibility within well-positioned nucleosomes. We observe a striking strand asymmetry in DNA accessibility around the nucleosome particle (Figure 4f), especially within the dyad and at the points two helical turns away from the dyad. The magnitude of these differences in average methylation levels are similar to those observed between nucleosomes and flanking linker regions. Thus SMAC-seq reveals significant heterogeneity in DNA's accessibility potential within the nucleosomal particle, which has important implications for understanding how transcription factors interact with the genome in vivo, in particular in light of recent studies demonstrating that certain classes of TFs may preferentially occupy nucleosomal DNA[39]. We also note that these patterns are most clearly observed when using m^6A positions into account, underlying the importance of its use in comparison to using CpG/GpC positions alone (Supplementary Figure 54).

SMAC-seq reveals distal co-accessibility patterns in the genome

We next examined co-accessibility patterns in the yeast genome. We first aimed to measure correlation between the positions of individual nucleosomes. Average NMI profiles centered on positioned nucleosomes reveal detectable correlation between nucleosome positions up to three to four nucleosomes away from an individual positioned nucleosome (Figure 4g; Supplementary Figure 84), with strongly positioned nucleosomal particles exhibiting stronger overall correlation over larger distances. These observations are consistent with a model whereby the restrictions on positioning that nucleosomes impose on each result in correlation between their protection footprints on short distances until random positional fluctuations of individual nucleosomes in the chromatin fiber eventually dephase this correlation signal on longer scales.

We then examined co-accessibility patterns in the vicinity of promoters (Figure 4h). Expressed yeast genes are characterized by a nucleosome-depleted/free region (NFR) upstream of the TSS and a well positioned +1 nucleosome. NMI profiles centered on the +1 nucleosome show significant differences between highly expressed and silent genes. While correlation patterns decay downstream of the TSS similarly for both groups of genes, highly expressed genes exhibit an inverse accessibility correlation pattern upstream of the TSS. The NMI profile for highly expressed genes also exhibits inverse accessibility correlation not just with the NFR but at a distance of at least one nucleosome beyond it.

Actively transcribed yeast genes often exist in a looped conformation, in which the promoter and termination regions are brought in physical proximity, potentially helping to enforce transcriptional directionality[40,41]. Given this physical coupling, we wondered if correlation also exists between chromatin accessibility around gene ends. SMAC-seq data reveals low levels of correlation between the NFR and the accessible region in the 3′ end of genes, and stronger correlation between positioned nucleosomes in these locations (Supplementary Figure 85). The correlation between the accessibility in the NFR and the 3′ end is increased for highly expressed genes and decreased for silent genes, suggesting that transcriptional activity and looping may help more strongly position nucleosomes at the beginning and end of transcribed regions; the decreased correlation between the nucleosome-depleted areas could be explained by transcriptional activity and dynamic regulatory occupancy leading to less stable protection patterns at each end that are accordingly less well correlated with one other.

We next assessed coordinated accessibility between yeast TSSs. To this end, we devised an explicit test of coordinated coaccessibility based on splitting reads into separate pieces, randomly reassembling them, then deriving an empirical coaccessibility distribution (see the Methods for details). Using this approach we identified 1,115 TSS pairs as significantly correlated out of 19,578 pairs covered with $\geq$100 reads in our initial sample (Supplementary Figure 86). Of these, 560 were located a distance $\geq$ 1 kb from each other. An example of significantly coordinated accessibility (between the *GAL10* and *GAL1* genes) is shown in (Supplementary Figure 87). As one possible mechanism for generating correlated accessibility is increased frequency of physical association in 3D space, we used publicly available Micro-C[42] data to assess whether promoters exhibiting coordinated accessibility are more often physically interacting with each other. Indeed, we observe that significantly coaccessible promoters interact more frequently than non-coaccessible promoters at a similar distance (Supplementary Figure 88).

SMAC-seq charts dynamic coordinated changes in chromatin accessibility in the course of the yeast stress response

To monitor chromatin accessibility at long-range scales during dynamic changes in gene regulatory activity brought about by environmental stimuli, we carried out SMAC-seq experiments during a time course of diamide treatment. Diamide oxidizes thiols in proteins, resulting in the generation of disulfides and activation of the stress response pathway, leading to changes in the expression of several hundred genes[43]. Stress response in yeast is mediated through the action of the Hsf1 transcription factor (with the Msn2/4 proteins also playing a key role). Hsf1 binds to its cognate genes when activated upon stress leading to their transcriptional activation[44].

We performed SMAC-seq at 0, 30 and 60 minutes after diamide treatment of yeast cells, as well as RNA-seq, ATAC-seq, and ChIP-seq for RNA Pol2, the elongating ver-

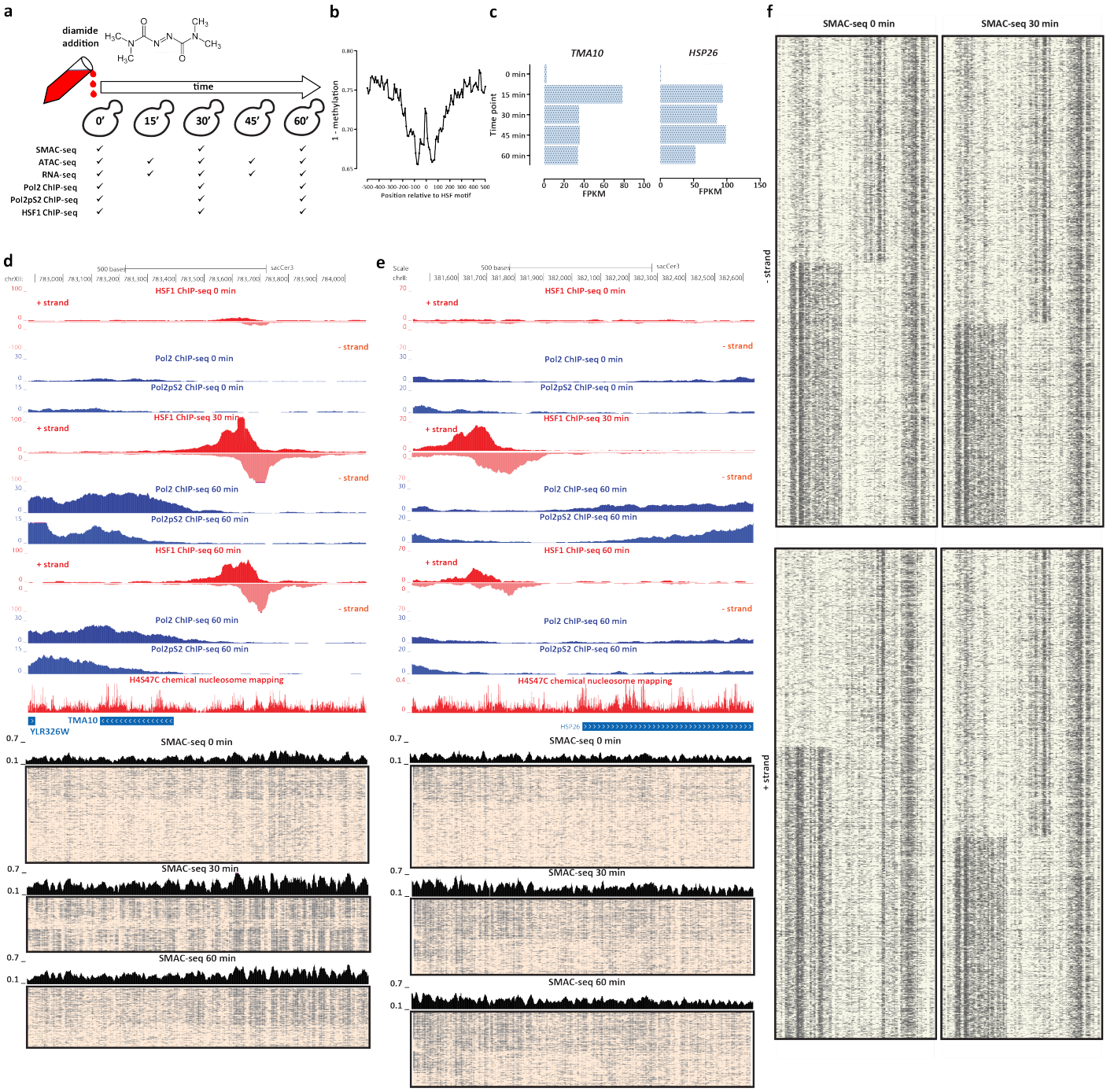


Figure 5: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. (a) Experimental outline. Yeast cells were treated with diamide, then SMAC-seq and other functional genomic assays where carried out at 15- or 30-minute intervals. (b) Sites occupied by the HSF1 transcription factor upon its activation by the stress response pathway exhibit strong footprints in SMAC-seq data. (c) Changes in the expression of the *TMA10* gene upon diamide treatment (d) Changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), and of chromatin accessibility at the single molecule level in the vicinity of the *TMA10* gene during the diamide time course. (e) Changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), and of chromatin accessibility at the single molecule level in the vicinity of the *HSP26* gene during the diamide time course. (f) Decrease in the fraction of transcribed rDNA arrays as a result of cellular response to diamide treatment.

sion of Pol2 (Pol2pS2), and a V5-tagged version of HSF1 (RNA-seq and ATAC-seq data was also collected at 15 and 45 minutes; Figure 5a). We observe several hundred genes exhibiting strong changes in gene expression during the time course (Supplementary Figure 89), and strong induction of Hsf1 occupancy at hundreds of sites in the genome (Supplementary Figure 90). SMAC-seq data at 30 minutes post-diamide treatment shows strong footprinting over Hsf1 motifs within induced Hsf1 binding sites.

We illustrate the dynamic patterns of chromatin accessibility that we observe upon diamide treatment using the *TMA10* and *HSP26* genes as examples in Figure 5d-e, and multiple others in Supplementary Figures 91-103. The *TMA10* and *HSP26* genes are strongly upregulated at 15 minutes after diamide treatment; TMA's expression subsequently declines somewhat and stabilizes (Figure 5c) while that of *HSP26* only declines at 60 minutes. SMAC-seq reveals a relatively modest level of accessibility usptream of these genes before diamide treatment. However, at 30 minutes and upon Hsf1 binding, dramatic changes are evident. Nearby nucleosomes appear to be evicted in many cells, and nucleosome depletion is also observed at increased levels within the gene bodies, where ChIP-seq data for Pol2 and Pol2pS2 shows highly active transcription. At 60 minutes we observe dampening of this response for *TMA10*, with the accessible fraction of reads decreasing; this effect is less pronounced for *HSP26*, whose expression remains relatively higher at later time points. Examination of NMI co-accessibility maps (Supplementary Figures 91-105) frequently shows loss of correlation between positioned nucleosomes within and upstream of activated gene bodies as a result of response to diamide treatment, consistent with an increased movement of nucleosomal particles due to the activities of transcribing polymerase molecules and chromatin remodelers. The aggregate observations derived from SMAC-seq were largely corroborated by the paired ATAC-seq measurements on the same samples (Supplementary Figures 106 and 107).

We also compared rDNA transcriptional states during the cellular response to diamide treatment. We observed a decrease in the fraction of transcribed rDNA arrays (Figure 5f), consistent with decreased overall rDNA transcription as a result of activation of the stress response program.

Discussion

SMAC-seq is a novel single-molecule method for profiling chromatin accessibility within individual chromatin fibers at a multikilobase resolution that leverages the power of nanopore sequencing to detect DNA modifications and the preference of DNA methyltransferases for chemically modifying accessible DNA. We show that SMAC-seq generates accessibility signals similar to those of widely used methods such as DNase-seq and ATAC-seq while also opening new windows into the long-range structure of eukaryotic chromatin. SMAC-seq enables the simultaneous profiling of nucleosome positioning and accessible chromatin on a truly genome-wide scale (including repetitive regions that are poorly mappable with short reads), the assessment of the absolute distribution of chromatin accessibility states within a population of cells, and the identification of pairs of genomic loci that exhibit significant correlations in co-accessibility.

Our initial work on SMAC-seq focused on applications in *S. cerevisiae* because of its modest genome size. Extending SMAC-seq to larger genomes will require significantly increased sequencing throughput, or a method for selective enrichment of a collection of individual loci. Fortunately, both of these approaches appear feasible in the near term, as nanopore sequencing throughput is increasing rapidly, while methods for selective enrichment of genomic regions for nanopore sequencing are also becoming available[45]. Read length increases will also be useful, in particular for correlating the activity of distal regulatory elements to their cognate promoters, which can often be located many tens of kilobases apart in mammalian genomes.

Improvements to base calling accuracy constitute another major area of potential future advances. At the time of our analysis, Tombo was the only readily available algorithm for calling m^6A, but, as discussed above, its base calls exhibit on the order of 20-25% error rates for all three bases combined. Results obtained with the other widely used methylation-aware base caller, Nanopolish, which is currently capable of calling only CpG and GpC methylation, show ∼15% error on CpG and GpC calls. The major barrier to base-calling improvements is the lack of sets of ground truth controls that can be used to train base calling algorithms. Pools of DNA templates with individual modifications in well-defined yet highly diverse base pair contexts are the ideal training set for developing accurate such models. Alternatively, the introduction of tags bulkier than a simple methyl group[46] may provide a much stronger modulation of electric current through the nanopore than does simple methylation, enabling much more reliable modified base calls and accessibility evaluation.

We also anticipate an increase in the diversity of the DNA modifying enzymes available to carry out variations of SMAC-seq and related assays. In mammalian systems, endogenous 5mC methylation occurs primarily in the CpG context, thus it is not possible to use M.SssI for mapping accessibility, leaving only GpC as an option for traditional NOME-seq, bisulfite-based assays. Fortunately, the bulk of SMAC-seq's increase in resolution is derived from m^6A (Supplementary Figure 1), so this will not be a significant obstacle to its widespread application in mammalian systems. The presence of endogenous CpG methylation does represent a potential concern regarding the use of m^6A in mammalian contexts, as the two modifications could in principle interfere with each other's detection. This is because nanopore sequencing as currently implemented works by reading multiple bases at a time, typically 3- to 6-mers. On the other hand, we note that the potential for interfer-

ence affects primarily the ability to read out endogenous methylation rather than accessibility, as positions nearby the generally widely spaced apart CpG dinucleotides can be simply filtered out of the analysis. To address this concern, we generated low-coverage (∼1.2×) SMAC-seq data for the GM12878 human lymphoblastoid cell line using only the EcoGII methyltransferase, and examined the aggregate profiles of the resulting "m^{6}A-SMAC" dataset around CTCF ChIP-seq peaks, DNAse hypersensitivity sites and TSSs. We recovered the expected features of chromatin accessibility (strong nucleosome positioning around CTCF occupanct sites[70], and accessibility peaks around all three types of features) and observed no significant difference between aggregate SMAC-seq profiles generated by filtering out A positions nearby CpGs, indicating that endogenous CpG methylation does not substantially interferes with m^{6}A-based accessibility measurements (Supplementary Figures 108-110).

However, there are also species where m^{6}A occurs endogenously and is strongly correlated with patterns of accessibility and nucleosome positioning (e.g. *Chlamydomonas*[47] and *Tetrahymena*[48,49]). Alternative methylation strategies such as 4mC methyltransferases[50], cytidine deaminases[51], or the conversion of thymine to modified bases such as 5-hydroxymethyluracil (dhmU)[52] are areas of potential future exploration. We expect that such improvements in base calling accuracy and resolution will allow the scope of SMAC-seq application to be further expanded, e.g. into direct de novo identification of nucleosomes and TF occupancy footprints within single molecules and measuring correlations between distal TF footprints.

Finally, we note that there is considerable scope for integration of SMAC-seq-type approaches with other measurements of the physical genome and the epigenome, especially once the improvements in accuracy outlined above are achieved. We anticipate the potential for obtaining simultaneous measurements of accessibility and nucleosomal positioning together with endogenous DNA methylation, general and specific protein occupancy, chromatin interactions, DNA replication, and other features on a multikilobase scale and within single molecules. In principle, similar approaches may also be applicable to individual RNA molecules. We expect long-read single-molecule approaches to provide an important new class of tools for the study of the functional and physical organization of genomes in the coming years.

Materials and Methods

Except for when explicitly stated otherwise, all analyses were carried out using custom-written Python or R scripts.

Cell lines and cell culture

The BY4741 *S. cerevisiae* strain (a kind gift from Ji-Ping Wang and Xiaozhong Wang) was used for all experiments except for Hsf1 ChIP-seq experiments where MS143 (H4S47C_Hsf1-V5::HphMX6, this study) was used. MS143 was generated by PCR-based C-terminal tagging of Hsf1 with the V5 epitope. Hsf1-V5 tagging was confirmed by colony PCR and western blotting. For all experiments, except the initial one ("Sample 1"), cells were grown in YPD media (30 °C) to OD∼0.8 before collection.

Yeast SMAC-seq experiments

Enzymatic treatment of chromatin

We developed and optimized SMAC-seq using the equivalent of 1×10^6 human cells, which in the case of *S. cerevisiae* translates in to 2.5×10^8 (the size of the haploid human genome is $\sim 3 \times 10^9$bp while that of *S. cerevisiae* is 1.2×10^6bp). As yeast cells have a cell wall, we adapted the spheroplasting protocol previously used for carrying out ATAC-seq in yeast cells[53] for our SMAC-seq experiments.

Yeast cells in log phase ($OD_{660} \leq 1.0$) were first centrifuged at 13,000 rpm for 1 minute, then washed with 100 μL Sorbitol Buffer(1.4 M Sorbitol, 40 mM HEPES-KOH pH 7.5, 0.5 mM $MgCl_2$), and centrifuged again at 13,000 rpm for 1 minute. Cells were then spheroplasted by resuspending in 200 μL Sorbitol Buffer with DTT added at a final concentration of 10 mM and 0.5 mg/mL 100T Zymolase, followed by incubating for 5 minutes at 30 °C at 300 rpm in a Thermomixer. The pellet was centrifuged for 2 minutes at 5,000 rpm, washed in 100 μL Sorbitol Buffer, and centrifuged again at 5,000 rpm for 2 minutes.

Cells were then resuspended in 100 μL ice-cold Nuclei Lysis Buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.5% NP-40) and incubated on ice for 10 minutes. Nuclei were then centrifuged at 5000 rpm for 5 min at 4 °C, resuspended in 100 μL cold Nuclei Wash Buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA), and centrifuged again at 5,000 rpm for 5 min at 4 °C. Finally, nuclei were resuspended in 100 μL M.CviPI Reaction Buffer (50 mM Tris-HCl pH 8.5, 50 mM NaCl, 10 mM DTT).

Nuclei were then first treated with M.CviPI + EcoGII by adding 200 U of M.CviPI (NEB) and 200 units of EcoGII (NEB), SAM at 0.6 mM and sucrose at 300 mM, and incubating at 30 °C for 7.5 min. After this incubation, 128 pmol SAM and another 100 U of enzymes were added, and a further incubation at 30 °C for 7.5 min was carried out. Immediately after, M.SssI treatment followed, by adding 60 U of M.SssI (NEB), 128 pmol SAM, $MgCl_2$ at 10 mM and incubation at 30 °C for 7.5 min.

The reaction was stopped by adding an equal volume of Stop Buffer (20 mM Tris-HCl pH 8.5, 600 mM NaCl, 1% SDS, 10 mM EDTA).

High-molecular weight DNA isolation

HMW DNA was isolated using the MagAttract HMW DNA Kit (Qiagen; cat # 67563) following the manufacturer's instructions.

Enzymatic treatment of naked DNA

Naked DNA was treated under exactly the same conditions as chromatin except that the reaction volume and enzyme amounts were reduced in half. HMW DNA was purified as described above

GM12878 cell culture

The GM12878 human lymphoblastoid cell lines were grown in media containing RPMI1640-GlutaMAX (Life Technologies, Carlsbad, CA) supplemented with 10% fetal bovine serum.

GM12878 SMAC-seq experiments

Briefly $1{\times}10^6$ human GM12878 cells were washed with 1× PBS then resuspended in 200 μL ice-cold Nuclei Lysis Buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1 mM EDTA, 0.5% NP-40) and incubated on ice for 10 minutes. Nuclei were then centrifuged at 500 *g* for 5 min at 4 °C, resuspended in 200 μL cold Nuclei Wash Buffer (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1 mM EDTA), and centrifuged again at 500 *g* for 5 min at 4 °C. Finally, nuclei were resuspended in 200 μL reaction buffer (1× NEB CutSmart buffer, 0.3 M sucrose). Nuclei were then treated with EcoGII by adding 200 units of EcoGII (NEB) and SAM at 0.6 mM, and incubating at 37 °C for 10 min. The reaction was stopped by adding 0.2% SDS, and HMW DNA was immediately isolated as previously described.

SMAC-seq analysis

Nanopore sequencing

HMW DNA was converted into libraries using the Ligation Sequencing Kit 1D (Oxford Nanopore Technologies, SQK-LSK108) following the manufacturer's instructions. Nanopore sequencing was carried out on R9.4 MinION flowcells (Oxford Nanopore Technologies) for up to 48 hours.

Nanopore base calling

Nanopore events were converted to DNA sequence using Albacore (V2.3.3) using default settings. Reads were resquiggled using Tombo[27], version 1.3, using the `sacCer3` reference genome. Methylated bases were identified using Tombo in the "de novo model" mode.

Aggregation of accessibility information over multibasepair windows

Even with the addition of m^6A methylation, the resolution of SMAC-seq still does not cover every nucleotide in the genome, and it varies substantially between different locations depending on local sequence content differences. In addition to that, nanopore base calling is still far from being a fully resolved problem, and even more so in methylation-aware mode. For these reasons, for many of the analyses described in this study we aimed at assigning aggregate accessibility scores over windows, taking the totality of the available evidence into account, thus obtaining more reliable, if coarser-grained, views of accessibility patterns along the genome. We used a Bayesian approach to carry out aggregation, as follows.

For a given window of width w in the genome, specified by coordinates $c, i, i+w$ (where c denotes the chromosome, and i the leftmost coordinate of the window), and for all reads $r \in R_{c,i,i+w}$ fully spanning the window, we obtain all Tombo probabilities $p_{r,(c,j)}$ such that $j \in [i, i+w)$ for sequence contexts CpG, GpC and A on the corresponding genomic strand. We use a Beta prior $B(\alpha,\beta)$, with $\alpha = \beta = 10$, which we then updated based on each probability $p_{r,(c,j)}$ for all $j \in [i, i+w)$. The final binary accessibility score $p_{r,(c,i,i+w)}$ for read r and window $c, i, i+w$ is determined by the final state of the prior.

Read filtering

As discussed above, we sometimes observe a population of reads that are fully methylated across their whole length or over large segments of it. There reads most likely derive from dead cells, as our initial experiment, which was carried out on a very dense yeast population containing a substantial number of dead cells, exhibited much higher proportion of such reads compared to subsequent experiments using early log-phase cells. In order to remove such potentially artifactual reads, "filtered" sets of reads were obtained by removing all reads containing a ≥1-kbp stretch that is ≥75% methylated (while also filtering out reads shorter than 1 kb).

Read clustering

For most analyses presented in this manuscript, the `tglkmeans` package was used to cluster SMAC-seq reads (implemented in `R`, `https://bitbucket.org/tanaylab/tglkmeans`. In addition, the hierarchical clustering implementation in `scipy` was also used in certain cases.

Co-accessibility assessment using Normalized Mutual Information

To evaluate co-accessibility patterns along the genome, we applied a Normalized Mutual Information as follows. Each chromosome in the genome c was split into windows of size w. For each such window $(c, i, i+w)$, we identified the maximum range to the right of it, $(c, j, j+w)$ such that the span $(c, i, j+w)$ is covered by $\geq M$ reads. All reads spanning $(c, i, j+w)$ were then extracted and subsampled down to M reads (usually $M = 100$, unless specified otherwise). Accessibility scores were then aggregated and binarized as described above for all windows located in the span $(c, i, j+w)$, and for all M reads fully spanning it, resulting in a local

co-accessibility matrix LCM of size $M \times (j+w-i)/w$. We then calculated Normalized Mutual Information scores for each pair of columns LCM_k and LCM_l as follows:

$$\begin{aligned} MI(LCM_k, LCM_l) = & \; p(0,0) \log_2 \left(\frac{p(0,0)}{p_k(0)\, p_l(0)} \right) \\ & + p(1,1) \log_2 \left(\frac{p(1,1)}{p_k(1)\, p_l(1)} \right) \\ & + p(0,1) \log_2 \left(\frac{p(0,1)}{p_k(0)\, p_l(1)} \right) \\ & + p(1,0) \log_2 \left(\frac{p(1,0)}{p_k(1)\, p_l(0)} \right) \end{aligned} \tag{1}$$

MI scores were then normalized and rescaled in the interval $(-1, 1)$:

$$NMI(LCM_k, LCM_l) = \begin{cases} \dfrac{MI(LCM_k, LCM_l)}{\sqrt{H(LCM_k) H(LCM_l)}} & \text{for } p(0,0) + p(1,1) \geq 0.5 \\ -\dfrac{MI(LCM_k, LCM_l)}{\sqrt{H(LCM_k) H(LCM_l)}} & \text{for } p(0,0) + p(1,1) < 0.5 \end{cases} \tag{2}$$

Where H refers to the entropy of each individual distribution.

For computational efficiency, local NMI matrices were calculated for even-sized (50kb) evenly spaced (every 10kb) tiles of the genome. The entries of the general genome-wide NMI matrix were then calculated as the average of all local NMI matrices containing each entry.

Testing for coordinated accessibility

Coordinated accessibility was evaluated as follows. For each pair of locations $(c, i_1, i_1 + r_1)$ and $(c, i_2, i_2 + r_2)$ (usually $r_1 = r_2$), a minimum number of reads N was required that fully spans the $(c, i_1, i_2 + r)$ interval. All such reads were then obtained for each pair, and then subsampled multiple times down to N reads (in order not to introduce bias in coordinated accessibility tests arising due to differential read coverage between locations closer/further apart). For each subsampling, the fraction of accessible regions p_1 and p_2 was estimated for each of the two locations using the Bayesian procedure described above, as well as the distribution of joint accessibilities over the four states $(0,0)$, $(1,0)$, $(0,1)$, and $(1,1)$. The two halves of the reads were then virtually split in half and recombined for a total of 10^3 random combinations. The empirical distribution $\mathcal{N}(\mu, \sigma)$ of the four states was then estimated from these random combinations, where $\mu = N * (p_{(0,0)} + p_{(1,1)})$ if $p_{(0,0)} + p_{(1,1)} > 0.5$ and $\mu = N * (p_{(1,0)} + p_{(1,0)})$ if $p_{(0,0)} + p_{(1,1)} \leq 0.5$. Empirical coordinated accessibility p-values were then estimated based on the observed counts $|(0,0)|+|(1,1)|$ if $|(0,0)|+|(1,1)| > 0.5*N$ or $|(0,1)|+|(1,0)|$ if $|(0,0)| + |(1,1)| \leq 0.5 * N$. Bonferroni correction was applied to account for multiple hypothesis testing.

dSMF and Bisulfite sequencing

Illumina measurements of CpG and GpC methylation levels were carried out using the PBAT protocol[54] with modifications. HMW DNA (∼500 ng) was bisulfite converted using the EZ DNA Methylation-Lightning Kit (Zymo, Cat # D5030) by mixing 20 μL of purified DNA (∼500 ng) with 130 μL DNA Methylation Lightning Conversion reagent and incubating at 98 °C for 8 minutes and then at 64 °C for 60 minutes. Bisulfite converted DNA was then cleaned up using the EZ DNA Methylation-Lightning Kit following manufacturer's instructions.

First strand synthesis was carried out by mixing 20 μL bisulfite converted DNA, 19.75 μL H_2O, 5 μL 10× Blue Buffer (ThermoFisher), 1.25 μL 10 mM dNTP (NEB), and 4 μL custom-designed biotinylated adapter. Samples were then incubated at 94 °C for 5 minutes, and at 4 °C for 5 minutes, after which 1.5 μL Klenow ($3' \rightarrow 5'$ exo minus; MCLab) were added, and the reaction was incubated at 4 °C for 15 minutes, at 37 °C for 90 minutes, and at 70 °C for 5 minutes. First-strand reaction cleanup was carried out using 50 μL AMPure XP beads (Beckman Coulter); DNA was eluted 50 μL EB buffer.

Biotinylated DNA was captured on streptavidin beads. A total of 20 μL streptavidin Dybaneads M-280 (ThermoFisher) per sample were added to a PCR tube, separated on a magnet and then resuspended in 50 μL 2× BW(Li) buffer (6.3 g LiCl, 0.5 mL Tris-HCL pH 8.0, and 0.1 mL 500 mM EDTA for 50 mL total volume), to which the 50 μL of eluted first-strand reaction DNA was added. Beads were

then incubated at room temperature for 30 minutes, washed with 180 μL 2× BW(Li) buffer, twice with 0.1 N NaOH (by resuspending well and incubating at room temperature for 2 minutes), washed again with 180 μL 2× BW(Li) buffer, then with 180 μL 10 mM Tris-HCL pH 7.5.

Second-strand synthesis was carried out by resuspending streptavidin beads in the following reaction mix: 5 μL 10 × Blue Buffer, 1.25 μL 10 mM dNTPs, 39.75 μL H_2O, 4 μL custom-designed second-strand adapter. Samples were then incubated at 94 °C for 5 minutes, and at 4 °C for 5 minutes, after which 1.5 μL Klenow ($3' \rightarrow 5'$ exo minus) were added, followed by further incubation at 4 °C for 15 minutes, at 37 °C for 30 minutes, and at 70 °C for 5 minutes.

Beads were separated on magnet and the chase reaction was carried out by resuspending in a mix of 5 μL 10× Thermo Pol Buffer, 1.25 μL 10 mM dNTPs, 43.5 μL H_2O, and 1 μL *Bst* DNA Polymerase Large Fragment (NEB). Samples were incubated at 65 °C for 30 minutes, then again separated on magnet.

PCR was performed on beads in 50 μL reactions composed of 25 μL 2× NEB Next PCR Master Mix, 20 μL H_2O, 2.5 μL i7 and 2.5 μL i5 primers (both custom-designed), with initial extension at 72 °C for 3 min, denaturation at 98 °C for 30 sec, 15 cycles of 98 °C for 10 sec, 63 °C for 30 sec, and 72 °C for 30 sec, and final extension at 72 °C for 5 min. PCR reactions were cleaned up and size-selected using AMPure XP beads.

Libraries were sequenced on Illumina NextSeq or MiSeq instruments, as 2×75mers or 2×300mers, respectively.

dSMF data processing

Bisulfite reads were trimmed using cutadapt (version 0.16) and Trim Galore (version 0.4.4), using the following settings (taking into account that the bisulfite sequencing libraries are generated with the PBAT protocol): `--clip_R1 9 --clip_R2 9 --three_prime_clip_r1 6 --three_prime_clip_r2 6 --paired`. Trimmed reads were the mapped to the `sacCer3` version of the yeast genome using Bismark[55] (version 0.19.0) with the following settings: `--bowtie2 --pbat`. Methylation calls were extract using the `bismark_methylation_extractor` program within Bismark and the following settings: `-s --no_overlap --comprehensive --merge_non_CpG --cytosine_report --CX`.

ATAC-seq

ATAC-seq was carried out on the same nuclei isolated for SMAC-seq as described above (before resuspension in M.CviPI Reaction Buffer), by resuspending nuclei with 25 μL 2× TD buffer (20 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 20% Dimethyl Formamide), 2.5 μL transposase (custom produced) and 22.5 μL nuclease-free H_2O, and incubating at 37 °C for 30 min in a Thermomixer at 1000 RPM. Transposed DNA was isolated using the DNA Clean & Concentrator Kit (Zymo, cat # D4014) and PCR amplified as described before[56]. Libraries were then sequenced on a Illumina NextSeq instrument as 2×36mers or as 2×75mers.

ATAC-seq data processing

Demultipexed fastq files were mapped to the `sacCer3` assembly of the *S. cerevisiae* genome as 2×36mers using Bowtie[57] with the following settings: `-v 2 -k 2 -m 1 --best --strata`. Duplicate reads were removed using `picard-tools` (version 1.99).

ChIP-seq experiments

Cell lysis and ChIP reactions were performed as previously described[58] with minor modifications. Cells were fixed with 1% formaldehyde for 20 minutes (Rpb1-CTD and Rbp1-CTD-S2P ChIP) or 30 minutes (Hsf1-V5 ChIP) and quenched with 0.125 M glycine for 5 minutes. A total of ∼50 ODs of cells were used per Rpb1-CTD or Rpb1-CTD-S2P ChIP and ∼300 ODs per Hsf1-V5 ChIP. Fixed cell were washed 2× in cold 1× PBS, pelleted and stored at −80 °C. Pellets were lysed in 300 μL FA lysis buffer (50 mM HEPES–KOH pH 8.0, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM PMSF, Roche protease inhibitor) with ∼1 mL ceramic beads on a Fastprep-24 (MP Biomedicals). The entire lysate was then collected and adjusted to 1 mL with FA lysis buffer before sonication with a 1/8' microtip on a Q500 sonicator (Qsonica) for 14 minutes (10 seconds on, 20 seconds off). The sample tube was held in a −20 °C 80% ethanol bath throughout sonication to prevent sample heating. After sonication, cell debris was pelleted and the supernatant was retained for ChIP. For each ChIP reaction, 30 μL Protein G Dynabeads (Invitrogen) were blocked (PBS + 0.5% BSA), prebound with 5-10 μL antibody (8wG16 Rpb1-CTD, Abcam cat # ab817); 3E10 Rpb1-CTD-S2P, Milipore cat # 04-1571-1) or SV5-Pk1 (anti-V5, BioRad cat # MCA1360G)) and washed 1× with PBS before incubation with supernatant (4 °C, overnight). Dynabeads were then washed (5 minutes per wash) 3× in FA lysis buffer, 3× in high-salt FA lysis buffer (50 mM HepesKOH pH 8.0, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM PMSF), 1× in ChIP wash buffer (10 mM TrisHCl pH 7.5, 0.25 M LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA, 1 mM PMSF) and 1× in TE wash buffer (10 mM TrisHCl pH 7.5, 1 mM EDTA, 50 mM NaCl). DNA was eluted from the beads in ChIP elution buffer (50 mM TrisHCl pH 7.5; 10 mM EDTA; 1% SDS) at 65 °C for 20 min. Eluted DNA was incubated at 65 °C overnight to reverse crosslinks, before treatment with RNAse A (37 °C, 1 hour) and then Proteinase K (65 °C, 2 hours). DNA was purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Sequencing libraries were generated using the NEBNext Ultra II DNA Library Prep kit (NEB Cat # E7645) and sequenced on a Illumina NextSeq instrument as 2×36mers or as 2×75mers.

ChIP-seq data processing

Demultipexed fastq files were mapped to the `sacCer3` assembly of the *S. cerevisiae* genome as 2×36mers using Bowtie[57] with the following settings: `-v 2 -k 2 -m 1 --best --strata`. Duplicate reads were removed using `picard-tools` (version 1.99). Hsf1 peaks were called using MACS2[59] (version 2.1.0) with the following settings: `-g 12000000 -f BAMPE`.

Multiread-preserving alignment and normalization

Multiread-preserving alignment and track generation was carried out by mapping reads to the `sacCer3` assembly of the *S. cerevisiae* genome using Bowtie[57] with the following settings: `-v 2 -a --best --strata`. Each alignment was then given a weight inversely proportional to the number of locations that the read maps to i.e. each position's score was normalized to RPMs as follows:

$$S_{c,i} = \frac{\sum_{R \in R_{c,i}} \frac{1}{NH_R}}{\frac{|R|}{10^6}} \quad (3)$$

Where NH_R is the number of locations in the genome a read maps to.

RNA-seq experiments

Cells (1 mL) were pelleted and flash frozen in liquid N_2. Pellets were resuspended in 300 μL TRIzol and lysed with ∼1 mL ceramic beads on a Fastprep-24 (MP Biomedicals). Cell debris were pelleted and RNA was extracted from the supernatant using the Direct-Zol RNA Microprep Kit (Zymo Research). RNA-seq libraries were generated using the NEBNext Ultra Directional RNA Library Prep Kit (NEB Cat # E7420)

RNA-seq processing and gene expression quantification

RNA-seq reads were mapped to the yeast genome as 1×50mers (external datasets) or 2×75mers (diamide experiments) using TopHat version 2.0.8[60]. Gene-level quantifications in FPKMs (Fragments Per Kilobase per Million mapped reads) were generated using Cufflinks version 2.0.2[60]. The mean from all replicates was taken as the expression level for each gene for subsequent analyses.

External sequencing datasets

A number of previously published *S. cerevisiae* genomics datasets were used in this study. ChIP-exo reads and called peaks for Abf1, Cbf1, Rap1 and Reb1 were downloaded from GEO accessions GSE93662 and GSE72106. ChIP-seq data for centromeric proteins was downloaded from GEO accessions GSE31466 and GSE51949. PRO-seq and PRO-CAP data was obtained from GEO accession GSE76142. ORC ChIP-seq data was downloaded from GEO accession GSE16926. DNase-seq was downloaded from GEO accession GSE69651 while DGF data was downloaded from SRA accession SRP000620. MNase-seq data was obtained from GEO accessions GSE26493 and GSE29292, TBP ChIP-seq from GSE44200, Rpb1 ChIP-seq from GSE93190, Rpb3 ChIP-seq from GSE74787, RPC128 ChIP-seq from GSE39566, and Mediator subunits ChIP-seq from GSE95051. RNA-seq data from accession GSE85590 was also used. Except where otherwise stated, raw reads were aligned using Bowtie[61] with the following settings: `-v 2 --k 2 -m 1 --best --strata`, with the addition of `-X 1000` for paired-end reads. Paired-end reads were aligned as 2×25mers), while single-end reads were aligned as 1×36mers. PRO-seq and PRO-CAP data was aligned as 1×16mers.

Micro-C data and processing

Micro-C data was downloaded from GEO accession GSE68016 and processed as described in the original publication[42].

Transcription factor motif mapping

Transcription factor motif recognition sequences were mapped genome-wide using FIMO[62] (version 4.11.2 of the MEME-Suite[63] using the CIS-BP database[64] as a reference set of position weight matrices.

Gene annotation update

Publicly available gene models for *S. cerevisiae* do not contain TSS and TTS information for a major fraction of genes in the genome, only including the coding ("CDS") portions instead. As the omission of UTRs presents a problem for TSS- and TTS-centered analyses, we updated the existing gene models following the approach described previously[53] and the *S. cerevisiae* TIF-seq dataset from GEO accession GSE39128[65]. New TSS and TTS positions were assigned to each gene for which such information was available based on the median UTR length as measured by TIF-seq.

Nucleosome positioning information

H4S47C[26,66] chemical mapping data was downloaded form GEO accessions GSE59523 and GSE36063. H3Q85C[67] chemical mapping data was downloaded from GEO accession GSE97290. We used the nucleosome positioning calls obtained from the original 2012 Brogaard et al. study for our analyses, after transforming them from coordinates in the `sacCer2` version of the *S. cerevisiae* genome assembly to `sacCer3` using the `liftOver` function in the UCSC Genome Browser utilities toolkit.

Mappability tracks generation

To evaluate unique read mappability, the whole genome was tiled with reads of given length at every position. The reads were then mapped back to the genome using the same settings used to map single-end ChIP-seq reads. For every position coverage by mapped reads was calculated, and mappability was scored as the ratio between read coverage and the read length used to tile the genome.

Data availability

Short read datasets associated with this study are available through GEO accession GSE128290.

Author contributions

Z.S, G.K.M and N.A.S.A. conceived and designed the study. Z.S, G.K.M and N.A.S.A. performed initial experiments. Z.S., M.P.S. and G.K.M. performed diamide time course experiments. G.K.M. and Z.S. analyzed data. W.J.G. and A.K. supervised the study. G.K.M., Z.S., W.J.G. and A.K. wrote the manuscript with input from all authors.

Acknowledgments

This work was supported by NIH grants (P50HG007735, RO1 HG008140, U19AI057266 and UM1HG009442 to W.J.G., 1UM1HG009436 to W.J.G. and A.K., 1DP2OD022870-01 and 1U01HG009431 to A.K.), the Rita Allen Foundation (to W.J.G.), the Baxter Foundation Faculty Scholar Grant, and the Human Frontiers Science Program grant RGY006S (to W.J.G). W.J.G is a Chan Zuckerberg Biohub investigator and acknowledges grants 2017-174468 and 2018-182817 from the Chan Zuckerberg Initiative. Z.S. is supported by EMBO Long-Term Fellowship EMBO ALTF 1119-2016 and by Human Frontier Science Program Long-Term Fellowship HFSP LT 000835/2017-L G.K.M. is supported by the Stanford School of Medicine Dean's Fellowship. N.A.S.A. is funded by the Department of Defense through a National Defense Science and Engineering Grant and by a Stanford Graduate Fellowship. The authors would also like to thank members of the Greenleaf and Kundaje labs for their helpful suggestions and discussions on the subject over the course of the study.

References

1. Wu C. 1980. The 5′ ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* **286**(5776):854–860.
2. Keene MA, Corces V, Lowenhaupt K, Elgin SC. 1981. DNase I hypersensitive sites in Drosophila chromatin occur at the 5′ ends of regions of transcription. *Proc Natl Acad Sci U S A* **78**(1):143–146.
3. McGhee JD, Wood WI, Dolan M, Engel JD, Felsenfeld G. 1981. A 200 base pair region at the 5′ end of the chicken adult β-globin gene is accessible to nuclease digestion. *Cell* **27**(1 Pt 2):45–55.
4. Dorschner MO, Hawrylycz M, Humbert R, Wallace JC, Shafer A, Kawamoto J, Mack J, Hall R, Goldy J, Sabo PJ, Kohli A, Li Q, McArthur M, Stamatoyannopoulos JA. 2004. High-throughput localization of functional elements by quantitative chromatin profiling. *Nat Methods* **1**(3):219–225.
5. Sabo PJ, Kuehn MS, Thurman R, Johnson BE, Johnson EM, Cao H, Yu M, Rosenzweig E, Goldy J, Haydock A, Weaver M, Shafer A, Lee K, Neri F, Humbert R, Singer MA, Richmond TA, Dorschner MO, McArthur M, Hawrylycz M, Green RD, Navas PA, Noble WS, Stamatoyannopoulos JA. 2006. Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. *Nat Methods* **3**(7):511–518.
6. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**(2):311–322.
7. Hesselberth JR, Chen X, Zhang Z, Sabo PJ, Sandstrom R, Reynolds AP, Thurman RE, Neph S, Kuehn MS, Noble WS, Fields S, Stamatoyannopoulos JA. 2009. Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. *Nat Methods* **6**(4):283–289.
8. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**(12):1213–1218.
9. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**(5):887–898.
10. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**(5950):289–293.
11. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y. 2009. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* **462**(7269):58–64.

12. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, Chang HY. 2016. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* **13**(11):919–922.
13. Risca VI, Denny SK, Straight AF, Greenleaf WJ. 2017. Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping. *Nature* **541**(7636):237–241.
14. Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, Imakaev M, Mirny LA, Wu CT, Zhuang X. 2016. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**(7586):418–422.
15. Ou HD, Phan S, Deerinck TJ, Thor A, Ellisman MH, O'Shea CC. 2017. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* **357**(6349).
16. Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. 2012. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res* **22**(12):2497–2506.
17. Nabilsi NH, Deleyrolle LP, Darst RP, Riva A, Reynolds BA, Kladde MP. 2014. Multiplex mapping of chromatin accessibility and DNA methylation within targeted single molecules identifies epigenetic heterogeneity in neural stem cells and glioblastoma. *Genome Res* **24**(2):329–339.
18. Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. 2017. Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. *Mol Cell* **67**(3):411–422.e4.
19. Murray IA, Morgan RD, Luyten Y, Fomenkov A, Corrêa IR Jr, Dai N, Allaw MB, Zhang X, Cheng X, Roberts RJ. 2018. The non-specific adenine DNA methyltransferase M.EcoGII. *Nucleic Acids Res* **46**(2):840–848.
20. Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. 2017. Detecting DNA cytosine methylation using nanopore sequencing. *Nat Methods* **14**(4):407–410.
21. Rand AC, Jain M, Eizenga JM, Musselman-Brown A, Olsen HE, Akeson M, Paten B. 2017. Mapping DNA methylation with high-throughput nanopore sequencing. *Nat Methods* **14**(4):411–413.
22. Baldi S, Krebs S, Blum H, Becker PB. 2018. Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. *Nat Struct Mol Biol* **25**(9):894–901.
23. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, Ukomadu C, Sadler KC, Pradhan S, Pellegrini M, Jacobsen SE. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* **107**(19):8689–8694.
24. Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**(5980):916–919.
25. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**(3):523–536.
26. Brogaard K, Xi L, Wang JP, Widom J. 2012. A map of nucleosome positions in yeast at base-pair resolution. *Nature* **486**(7404):496–501.
27. Stoiber MH, Quick J, Egan R, Lee JE, Celniker SE, Neely R, Loman N, Pennacchio L, Brown JB. 2017. De novo Identification of DNA Modifications Enabled by Genome-Guided Nanopore Signal Processing. *bioRxiv* 094672
28. Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* **32**(14):2103–2110.
29. Yuan GC, Liu YJ, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. 2005. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* **309**(5734):626–630.
30. Cole HA, Howard BH, Clark DJ. 2011. The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proc Natl Acad Sci U S A* **108**(31):12687–12692.
31. Henikoff S, Ramachandran S, Krassovsky K, Bryson TD, Codomo CA, Brogaard K, Widom J, Wang JP, Henikoff JG. 2014. The budding yeast Centromere DNA Element II wraps a stable Cse4 hemisome in either orientation in vivo. *Elife* **3**:e01861.
32. Merz K, Hondele M, Goetze H, Gmelch K, Stoeckl U, Griesenbeck J. 2008. Actively transcribed rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules. *Genes Dev* **22**(9):1190–1204.
33. Conconi A, Widmer RM, Koller T, Sogo JM. 1989. Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. *Cell* **57**(5):753–761.
34. French SL, Osheim YN, Cioci F, Nomura M, Beyer AL. 2003. In exponentially growing *Saccharomyces cerevisiae* cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Mol Cell Biol* **23**(5):1558–1568.
35. Goetze H, Wittner M, Hamperl S, Hondele M, Merz K, Stoeckl U, Griesenbeck J. 2010. Alternative chromatin structures of the 35S rRNA genes in *Saccharomyces cerevisiae* provide a molecular basis for the selective recruitment of RNA polymerases I and II. *Mol Cell Biol* **30**(8):2028–2045.
36. Panday A, Grove A. 2016. Yeast HMO1: Linker Histone Reinvented. *Microbiol Mol Biol Rev* **81**(1). pii: e00037–16.
37. Jones HS, Kawauchi J, Braglia P, Alen CM, Kent NA,

Proudfoot NJ. 2007. RNA polymerase I in yeast transcribes dynamic nucleosomal rDNA. *Nat Struct Mol Biol* **14**(2):123–130.

38. Huang J, Moazed D. 2003. Association of the RENT complex with nontranscribed and coding regions of rDNA and a regional requirement for the replication fork block protein Fob1 in rDNA silencing. *Genes Dev* **17**(17):2162–2176.
39. Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, Dodonova SO, Nitta KR, Morgunova E, Taipale M, Cramer P, Taipale J. 2018. The interaction landscape between transcription factors and the nucleosome. *Nature* **562**(7725):76–81.
40. O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, Mellor J, Proudfoot NJ. 2004. Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* **36**(9):1014–1018.
41. Tan-Wong SM, Zaugg JB, Camblong J, Xu Z, Zhang DW, Mischo HE, Ansari AZ, Luscombe NM, Steinmetz LM, Proudfoot NJ. 2012. Gene loops enhance transcriptional directionality. *Science* **338**(6107):671–675.
42. Hsieh TH, Weiner A, Lajoie B, Dekker J, Friedman N, Rando OJ. 2015. Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* **162**(1):108–119.
43. Weiner A, Hsieh TH, Appleboim A, Chen HV, Rahat A, Amit I, Rando OJ, Friedman N. 2015. High-resolution chromatin dynamics during a yeast stress response. *Mol Cell* **58**(2):371–386.
44. Morano KA, Grant CM, Moye-Rowley WS. 2012. The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics* **190**(4):1157–1195.
45. Gabrieli T, Sharim H, Fridman D, Arbib N, Michaeli Y, Ebenstein Y. 2018. Selective nanopore sequencing of human BRCA1 by Cas9-assisted targeting of chromosome segments (CATCH). *Nucleic Acids Res* **46**(14):e87.
46. Staševskij Z, Gibas P, Gordevičius J, Kriukienė E, Klimašauskas S. 2017. Tethered Oligonucleotide-Primed Sequencing, TOP-Seq: A High-Resolution Economical Approach for DNA Epigenome Profiling. *Mol Cell* **65**(3):554–564.e6.
47. Fu Y, Luo GZ, Chen K, Deng X, Yu M, Han D, Hao Z, Liu J, Lu X, Doré LC, Weng X, Ji Q, Mets L, He C. 2015. N^6-methyldeoxyadenosine marks active transcription start sites in *Chlamydomonas*. *Cell* **161**(4):879–892.
48. Wang Y, Chen X, Sheng Y, Liu Y, Gao S. 2017. N^6-adenine DNA methylation is associated with the linker DNA of H2A.Z-containing well-positioned nucleosomes in Pol II-transcribed genes in *Tetrahymena*. *Nucleic Acids Res* **45**(20):11594–11606.
49. Luo GZ, Hao Z, Luo L, Shen M, Sparvoli D, Zheng Y, Zhang Z, Weng X, Chen K, Cui Q, Turkewitz AP, He C. 2018. N^6-methyldeoxyadenosine directs nucleosome positioning in *Tetrahymena* DNA. *Genome Biol* **19**(1):200.
50. Timinskas A, Butkus V, Janulaitis A. 1995. Sequence motifs characteristic for DNA [cytosine-N4] and DNA [adenine-N6] methyltransferases. Classification of all DNA methyltransferases. *Gene* **157**(1–2):3–11.
51. Salter JD, Bennett RP, Smith HC. 2016. The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends Biochem Sci* **41**(7):578–594.
52. Kawasaki F, Beraldi D, Hardisty RE, McInroy GR, van Delft P, Balasubramanian S. 2017. Genome-wide mapping of 5-hydroxymethyluracil in the eukaryote parasite *Leishmania*. *Genome Biol* **18**(1):23.
53. Schep AN, Buenrostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. 2015. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res* **25**(11):1757–1770.
54. Miura F, Enomoto Y, Dairiki R, Ito T. 2012. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res* **40**(17):e136.
55. Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**(11):1571–1572.
56. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, Kathiria A, Cho SW, Mumbach MR, Carter AC, Kasowski M, Orloff LA, Risca VI, Kundaje A, Khavari PA, Montine TJ, Greenleaf WJ, Chang HY. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**(10):959–962.
57. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**(3):R25.
58. Hu B, Petela N, Kurze A, Chan KL, Chapard C, Nasmyth K. 2015. Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. *Nucleic Acids Res* **43**(20):e132.
59. Feng J, Liu T, Qin B, Zhang Y, Liu XS. 2012. Identifying ChIP-seq enrichment using MACS. *Nat Protoc* **7**(9):1728–1740.
60. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**(3):562–578.
61. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**(3):R25.
62. Grant CE, Bailey TL, Noble WS. 2011. FIMO: scan-

ning for occurrences of a given motif. *Bioinformatics* **27**(7):1017–1018.
63. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**(Web Server issue):W202–208.
64. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, Zheng H, Goity A, van Bakel H, Lozano JC, Galli M, Lewsey MG, Huang E, Mukherjee T, Chen X, Reece-Hoyes JS, Govindarajan S, Shaulsky G, Walhout AJ, Bouget FY, Ratsch G, Larrondo LF, Ecker JR, Hughes TR. 2014. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**(6):1431–1443.
65. Pelechano V, Wei W, Steinmetz LM. 2013. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* **497**(7447):127–131.
66. Ramachandran S, Zentner GE, Henikoff S. 2015. Asymmetric nucleosomes flank promoters in the budding yeast genome. *Genome Res* **25**(3):381–390.
67. Chereji RV, Ramachandran S, Bryson TD, Henikoff S. 2018. Precise genome-wide mapping of single nucleosomes and linkers in vivo. *Genome Biol* **19**(1):19.
68. Liang Z, Shen L, Cui X, Bao S, Geng Y, Yu G, Liang F, Xie S, Lu T, Gu X, Yu H. 2018. DNA N^6-Adenine Methylation in *Arabidopsis thaliana*. *Dev Cell* **45**(3):406–416.e3.
69. Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, Aristizábal-Corrales D, Hsu CH, Aravind L, He C, Shi Y. 2015. DNA Methylation on N6-Adenine in *C. elegans*. *Cell* **161**(4):868–878.
70. Fu Y, Sinha M, Peterson CL, Weng Z. 2008. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet* **4**(7):e1000138.

Supplementary Materials

Supplementary Tables

Supplementary Table 1: Mapping statistics for SMAC-seq datasets in this study

sample	number reads	total bases	mean read length	median read length
Sample 1	847,262	3,475,258,237	4,102	1,474
diamide 0 min rep1	2,462,311	4,018,221,647	1,632	722
diamide 30 min rep1	1,127,937	3,376,196,280	2,993	1,254
diamide 60 min rep1	1,543,217	4,187,296,821	2,713	1,171
diamide 0 min rep2	2,605,209	7,815,151,882	3,000	1,383
diamide 30 min rep2	2,266,191	6,046,929,720	2,668	977
diamide 60 min rep2	1,397,628	5,335,378,307	3,817	1,919
positive control gDNA High EcoGII	553	4,345,539	7,858	3,706
positive control gDNA Low EcoGII	11,265	137,959,334	12,247	6,383
negative control gDNA	740,001	3,437,777,966	4,646	1,161
negative control λ DNA	120,711	455,213,644	3,771	2,528
positive control λ DNA EcoGII	311,974	895,965,634	2,872	1,950
GM12878 m^6A-SMAC	1,054,719	5,385,628,670	5,106	2,768

Supplementary Table 2: Mapping and QC statistics for ATAC-seq datasets used in this study

Dataset	Complexity	TSS ratio	Read Length	uniquely mapped deduplicated reads	Raw fragments
L464 Diamide 0 min rep1	0.79	1.38	2×36	2,284,992	2,363,608
L465 Diamide 0 min rep2	0.79	1.36	2×36	2,383,094	2,409,446
L466 Diamide 15 min rep1	0.77	1.49	2×36	1,907,268	1,760,891
L467 Diamide 15 min rep2	0.75	1.42	2×36	3,415,058	3,058,834
L468 Diamide 30 min rep1	0.72	1.49	2×36	3,223,114	3,414,835
L469 Diamide 30 min rep2	0.74	1.41	2×36	3,081,970	2,846,193
L470 Diamide 45 min rep1	0.79	1.37	2×36	2,411,938	2,457,100
L471 Diamide 45 min rep2	0.75	1.40	2×36	3,135,316	2,885,651
L472 Diamide 60 min rep1	0.74	1.45	2×36	3,205,726	3,120,294
L473 Diamide 60 min rep2	0.76	1.43	2×36	2,387,398	2,244,141

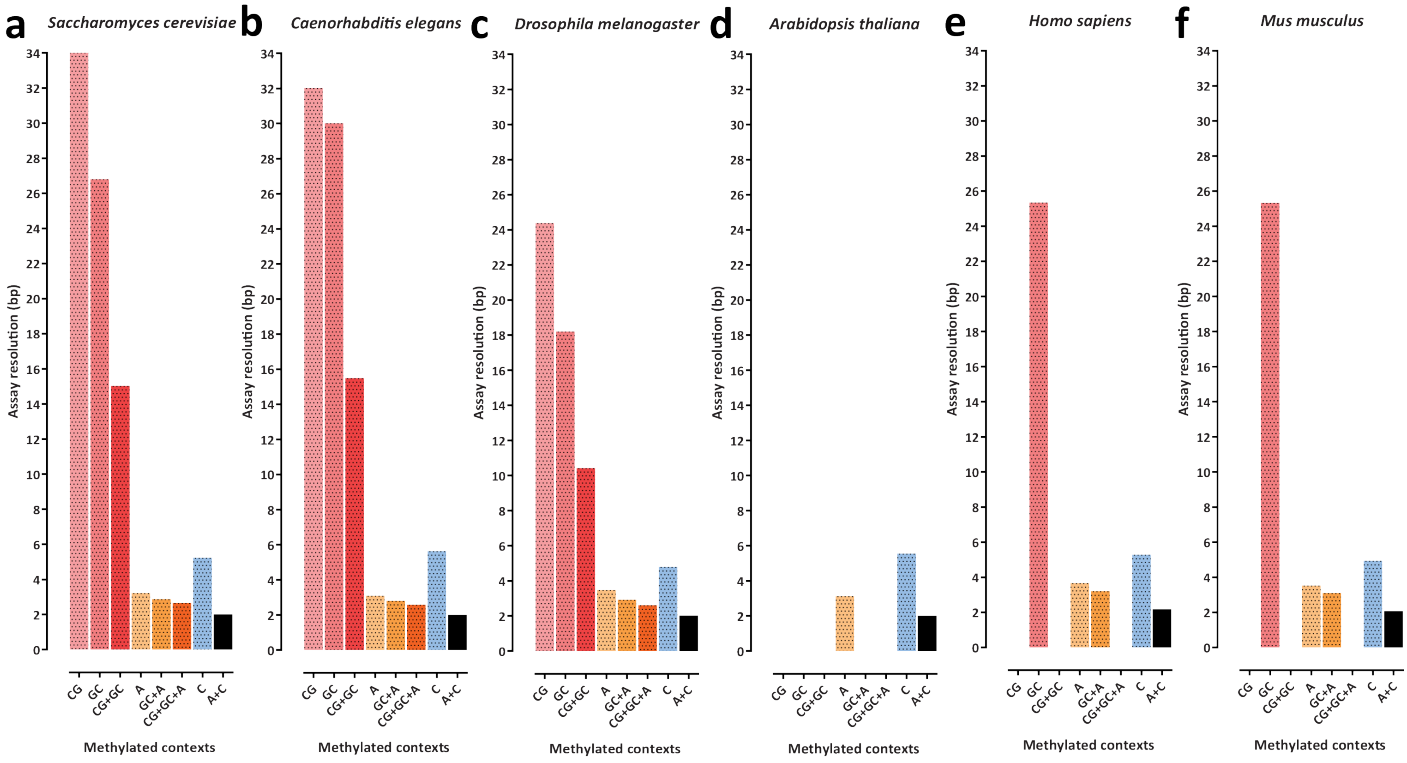
Supplementary Table 3: Mapping and QC statistics for ChIP-seq datasets used in this study

Dataset	Complexity	Read Length	Uniquely mapped deduplicated reads	Raw fragments
L482 Diamide 0 min Input	0.88	2 × 36	8,776,410	5,488,322
L483 Diamide 0 min Pol2	0.87	2 × 36	4,944,766	2,894,414
L484 Diamide 0 min Pol2pS2	0.80	2 × 36	5,834,864	3,434,723
L485 Diamide 0 min HSF1-V5 Input	0.90	2 × 36	6,089,572	3,850,484
L486 Diamide 0 min HSF1-V5	0.93	2 × 36	2,540,876	1,726,223
L487 Diamide 30 min Input	0.86	2 × 36	10,052,178	6,212,240
L488 Diamide 30 min Pol2	0.84	2 × 36	6,763,128	3,961,384
L489 Diamide 30 min Pol2pS2	0.84	2 × 36	6,332,462	3,901,689
L490 Diamide 30 min HSF1-V5 Input	0.92	2 × 36	4,587,466	2,831,871
L491 Diamide 30 min HSF1-V5	0.89	2 × 36	3,324,498	2,054,869
L492 Diamide 60 min Input	0.90	2 × 36	5,812,736	3,539,630
L493 Diamide 60 min Pol2	0.90	2 × 36	3,774,106	2,204,126
L494 Diamide 60 min Pol2pS2	0.85	2 × 36	4,873,244	2,924,345
L495 Diamide 60 min HSF1-V5 Input	0.87	2 × 36	7,683,586	4,679,564
L496 Diamide 60 min HSF1-V5	0.91	2 × 36	1,094,048	698,664

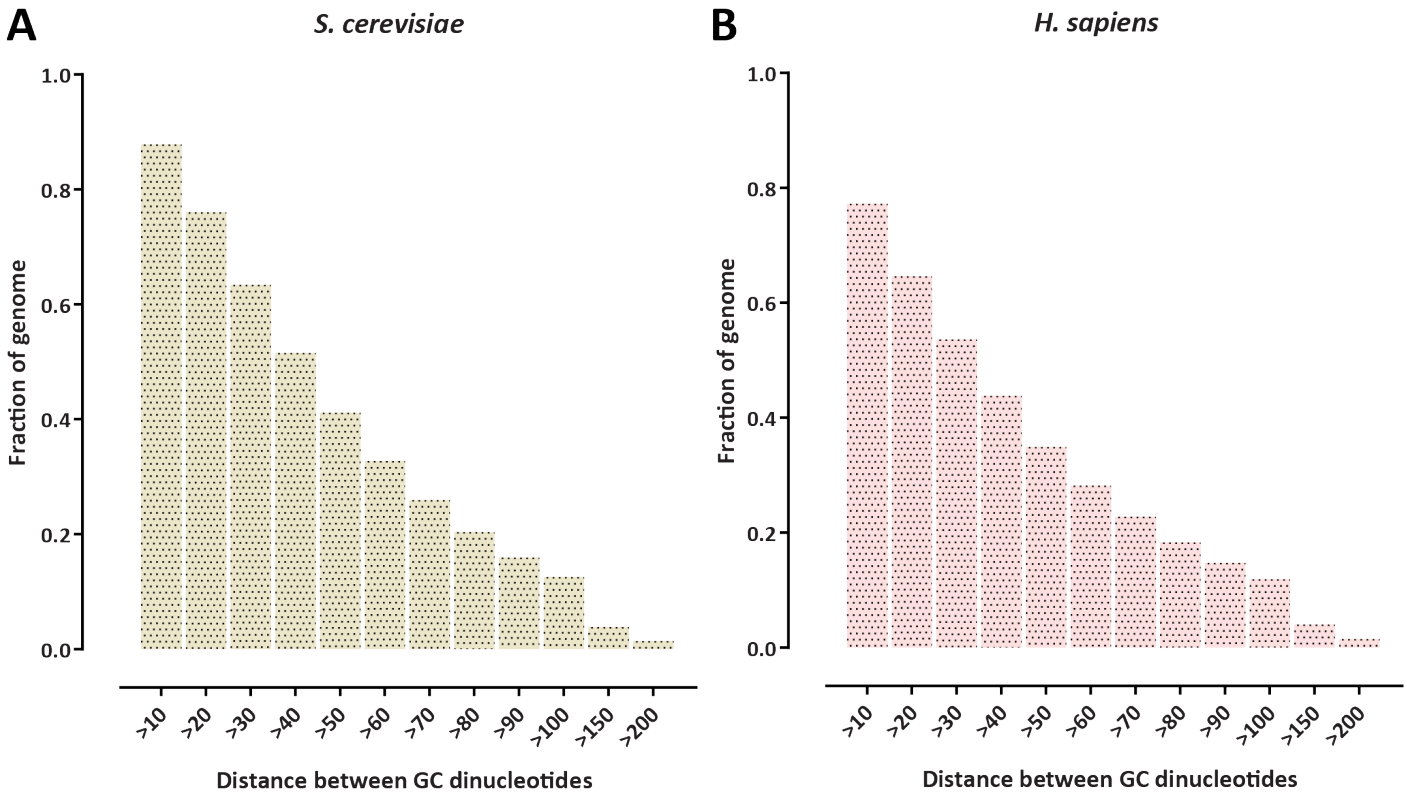
Supplementary Table 4: Mapping and QC statistics for RNA-seq datasets used in this study

Dataset	Complexity	Read Length	Unique	Unique Splices	Multi	Multi Splices	Raw fragments
Diamide 0 min	0.20	2 × 75	30,692,672	461,804	3,974,211	38,859	18,719,790
Diamide 15 min	0.21	2 × 75	26,991,043	182,208	3,217,924	18,788	15,960,277
Diamide 30 min	0.19	2 × 75	34,711,193	201,334	3,716,196	21,450	21,717,800
Diamide 45 min	0.22	2 × 75	28,668,901	116,773	3,103,601	19,301	17,832,222
Diamide 60 min	0.31	2 × 75	14,991,437	71,198	1,759,048	10,455	13,548,619
Diamide Hsf-V5 0min	0.19	2 × 75	32,981,363	519,649	4,320,132	35,513	21,535,487
Diamide Hsf-V5 30min	0.19	2 × 75	33,505,595	211,992	3,973,770	23,016	22,974,020
Diamide Hsf-V5 60min	0.19	2 × 75	40,290,524	196,082	4,698,437	29,628	28,894,301

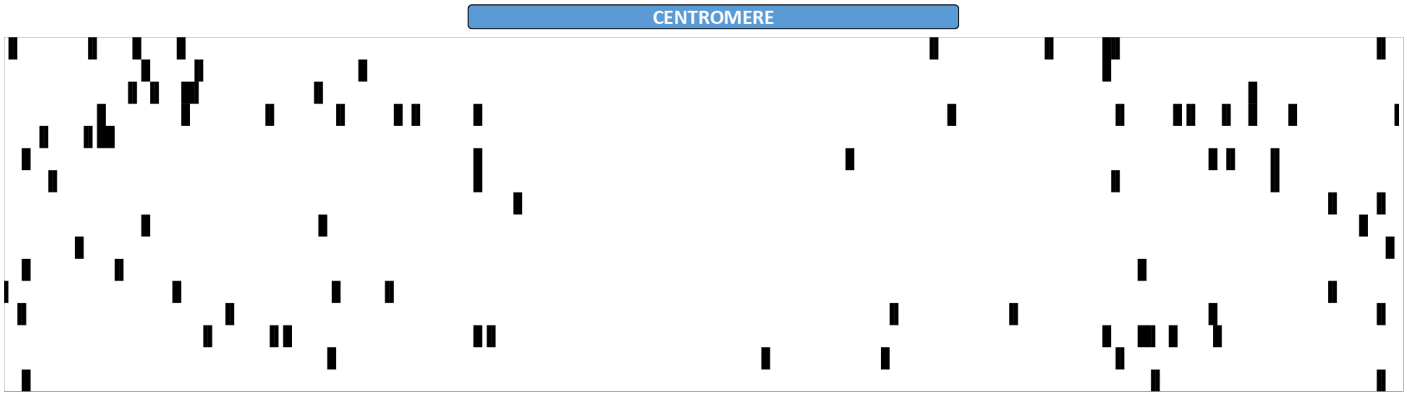
Supplementary Figures



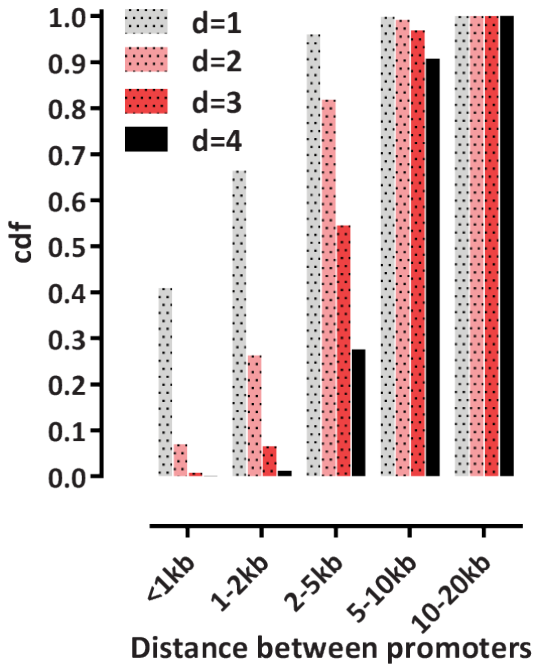
Supplementary Figure 1: Resolution of the SMAC-seq assay in its current form and some potential other versions of it in the main model organism systems. Note that where endogenous methylation confounds readout of accessibility, the corresponding combination of sequence contexts has been omitted from the plot. Also note that while m^6A has been reported to be present in the genomes of *Arabidopsis*[68] and *C. elegans*[69], it is generally found at low levels ($\leq 1\%$), and is not as strongly correlated with open chromatin and nucleosome positioning as it is in some other eukaryotes such as *Chlamydomonas*[47], thus its utility for accessibility profiling is not altered significantly. Nevertheless, a universally applicable version of SMAC-seq that is minimally confounded by endogenous methylation status in all eukaryotes will probably require the use of different methyltransferases (once they become available as efficient recombinant enzymes), for example, ones depositing the 4mC mark, which is what the "C" sequence context shown here corresponds to. (a) *Saccharomyces cerevisiae* (complete absence of endogenous methylation); (b) *Caenorhabditis elegans* (no endogenous 5mC, small amounts of endogenous m^6A); (c) *Drosophila melanogaster* (no significant endogenous 5mC or m^6A methylation); (d) *Arabidopsis thaliana* (endogenous 5mC in CpG, CHG and CHH contexts, small amounts of endogenous m^6A); (e) *Homo sapiens* (endogenous 5mC in CpG contexts, small amounts of endogenous 5mC in CHG and CHH contexts); (f) *Mus musculus* (endogenous 5mC in CpG contexts, small amounts of endogenous 5mC in CHG and CHH contexts).



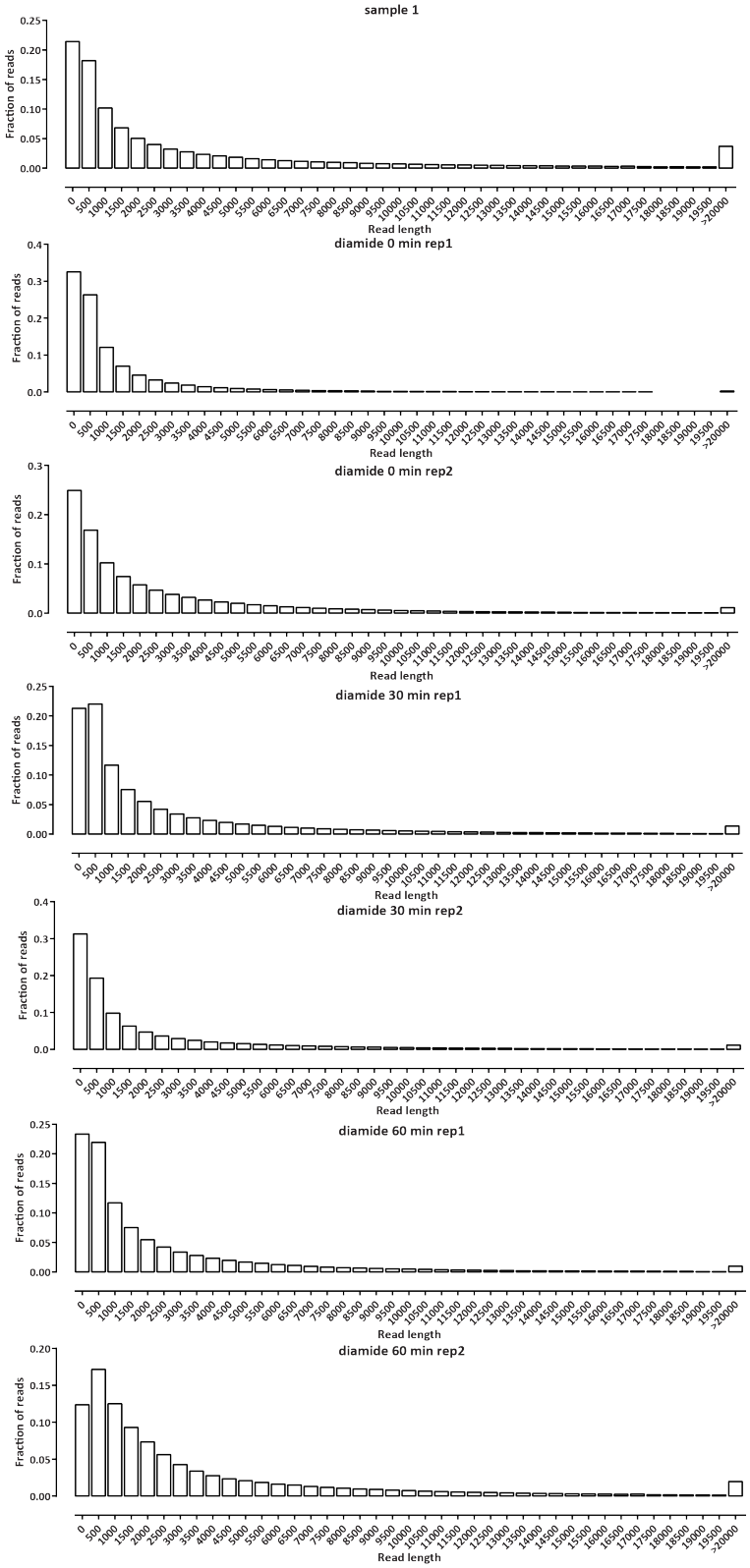
Supplementary Figure 2: GC positions alone are insufficient to provide proper coverage of the genome in the context of a methylation-based assay for profiling chromatin accessibility. Shown is the distribution of the fraction of the genome that contains no GC dinucleotides closer to each other than the indicated distance. The mitochondrial genome was excluded from the calculation. (a) *Saccharomyces cerevisiae*; (b) *Homo sapiens*. More than 50% of each genome consists of regions with GC dinucleotides spaced at least 30 bp apart (and 40 bp in the case of *S. cerevisiae*), i.e longer than a typical nucleosome linker.



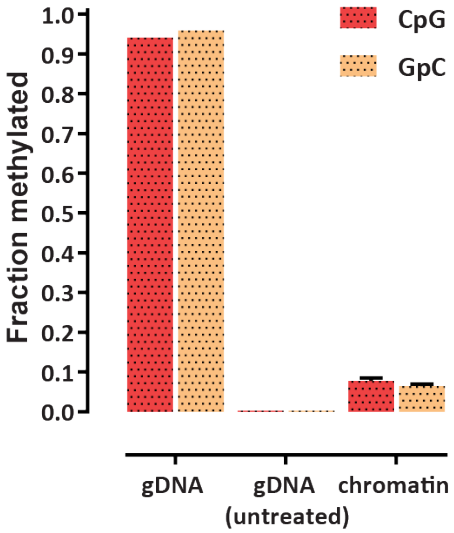
Supplementary Figure 3: Important functional elements in the yeast genome, such as, in this example, centromeres, are almost completely devoid of GC/GC dinucleotides. Shown is the distribution of GC dinucleotides (black rectangles) around *S. cerevisiae* centromeres.



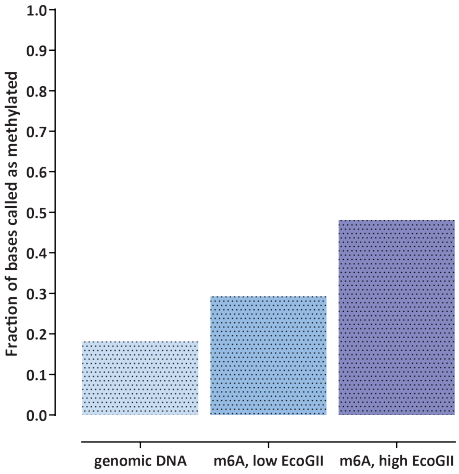
Supplementary Figure 4: Distance between regulatory elements (i.e. promoters in the case of *S. cerevisiae*).



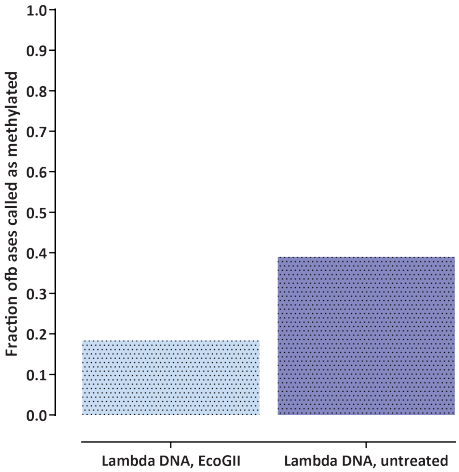
Supplementary Figure 5: Distribution of nanopore read lengths obtained for the experiments described in this study.



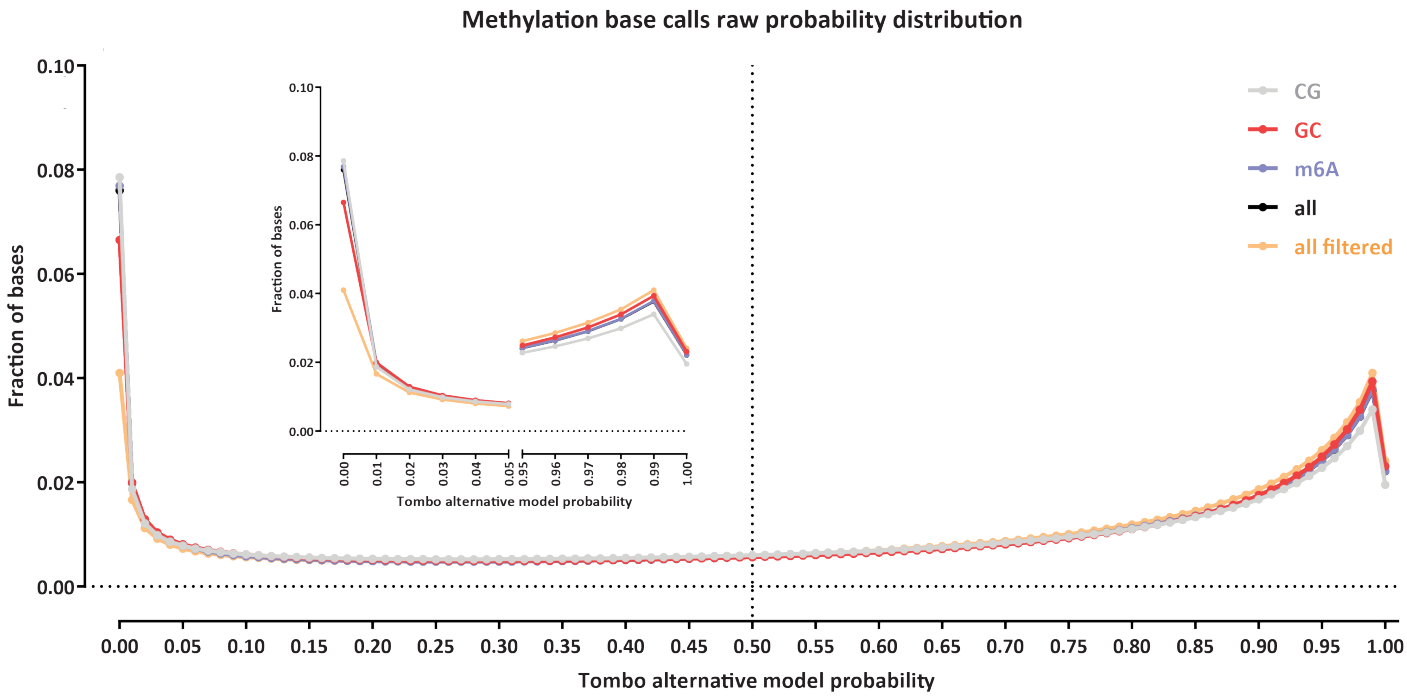
Supplementary Figure 6: Global methylation levels in yeast dSMF experiments and in positive and negative controls (measured by bisulfite sequencing). DNA from SMAC-seq experiments was subjected to Illumina bisulfite sequencing using the PBAT protocol. In parallel, naked genomic DNA (gDNA) was either treated with all three enzymes under the same conditions as in the SMAC-seq protocol or it was left untreated. These samples were also subjected to Illumina PBAT bisulfite sequencing. Shown are the global methylation levels for each sample.



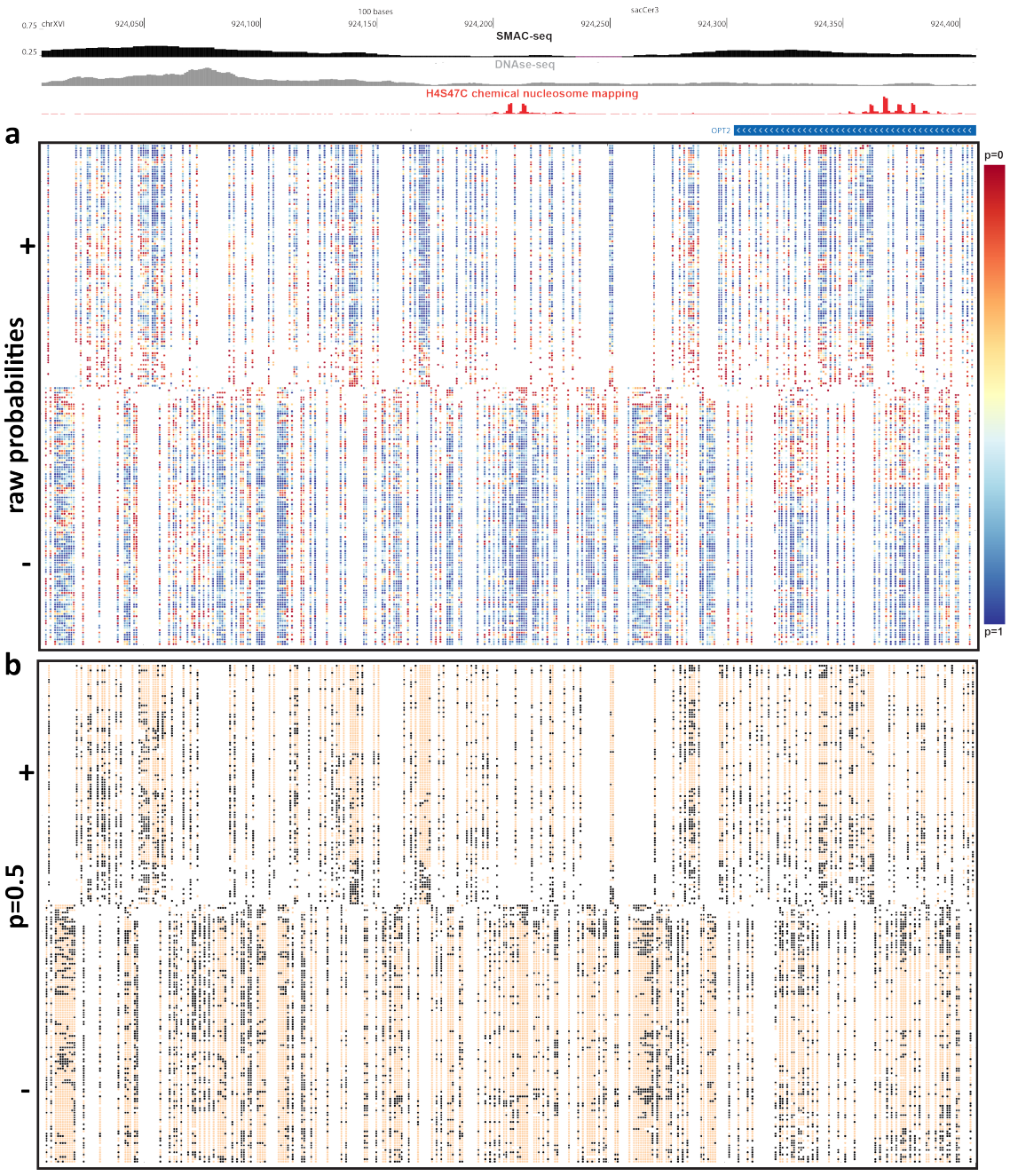
Supplementary Figure 7: Global m^{6}A methylation levels in positive and negative controls (measured by nanopore sequencing). Yeast genomic DNA was incubated with no or a low/high amount of the EcoGII enzyme and then subjected to nanopore sequencing. Shown are the average methylation levels for A nucleotides. Note that completely methylated DNA molecules are not sequenced well on the nanopore platform (See also Supplementary Table 1), thus the methylation estimates in the treated samples are most likely underestimates.

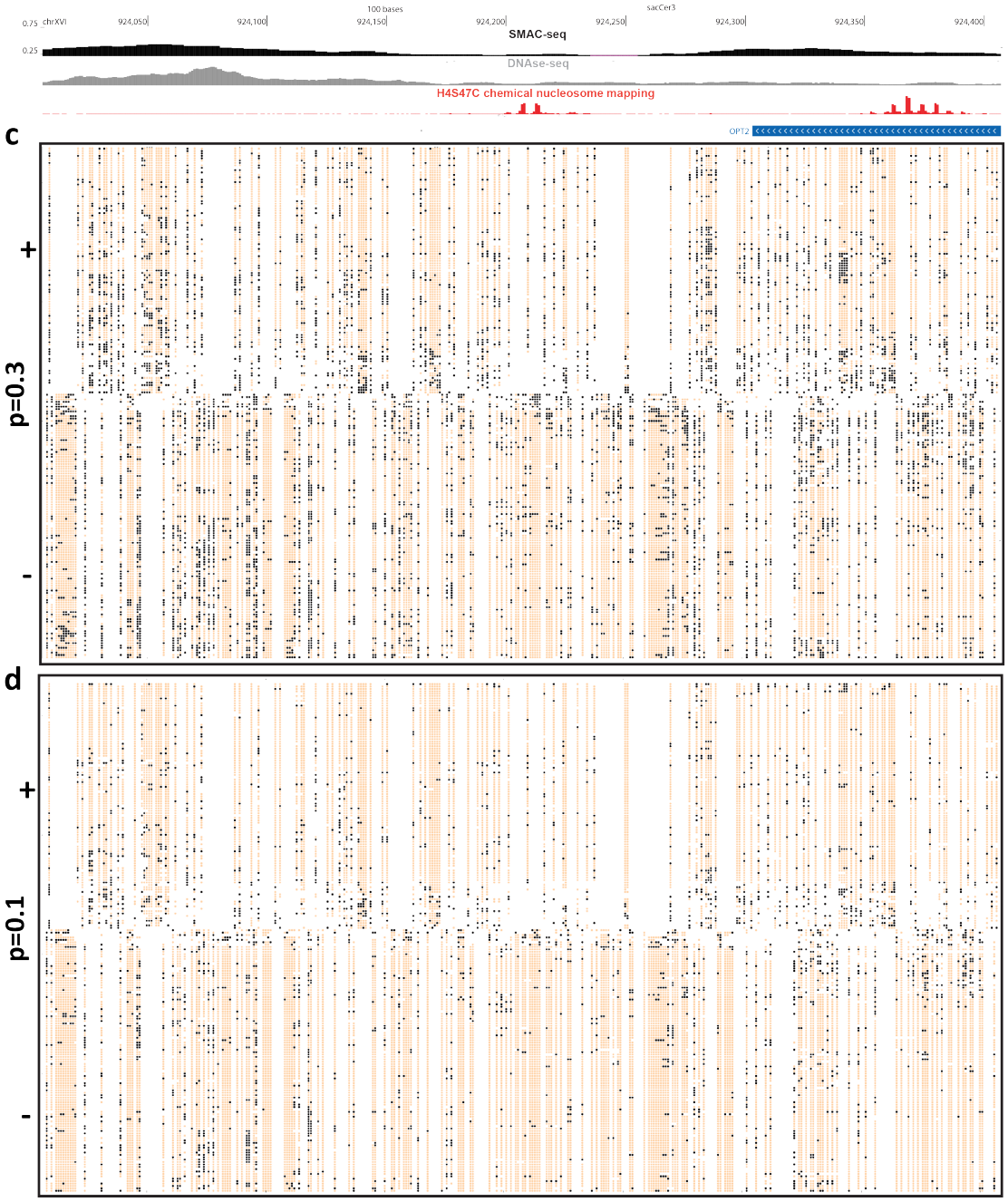


Supplementary Figure 8: Global m^6A methylation levels in positive and negative controls (measured by nanopore sequencing). Unmethylated Lambda DNA was incubated with the EcoGII enzyme and then subjected to nanopore sequencing. Shown are the average methylation levels for A nucleotides. Note that completely methylated DNA molecules are not sequenced well on the nanopore platform, thus the methylation estimates in the treated sample are most likely underestimates.

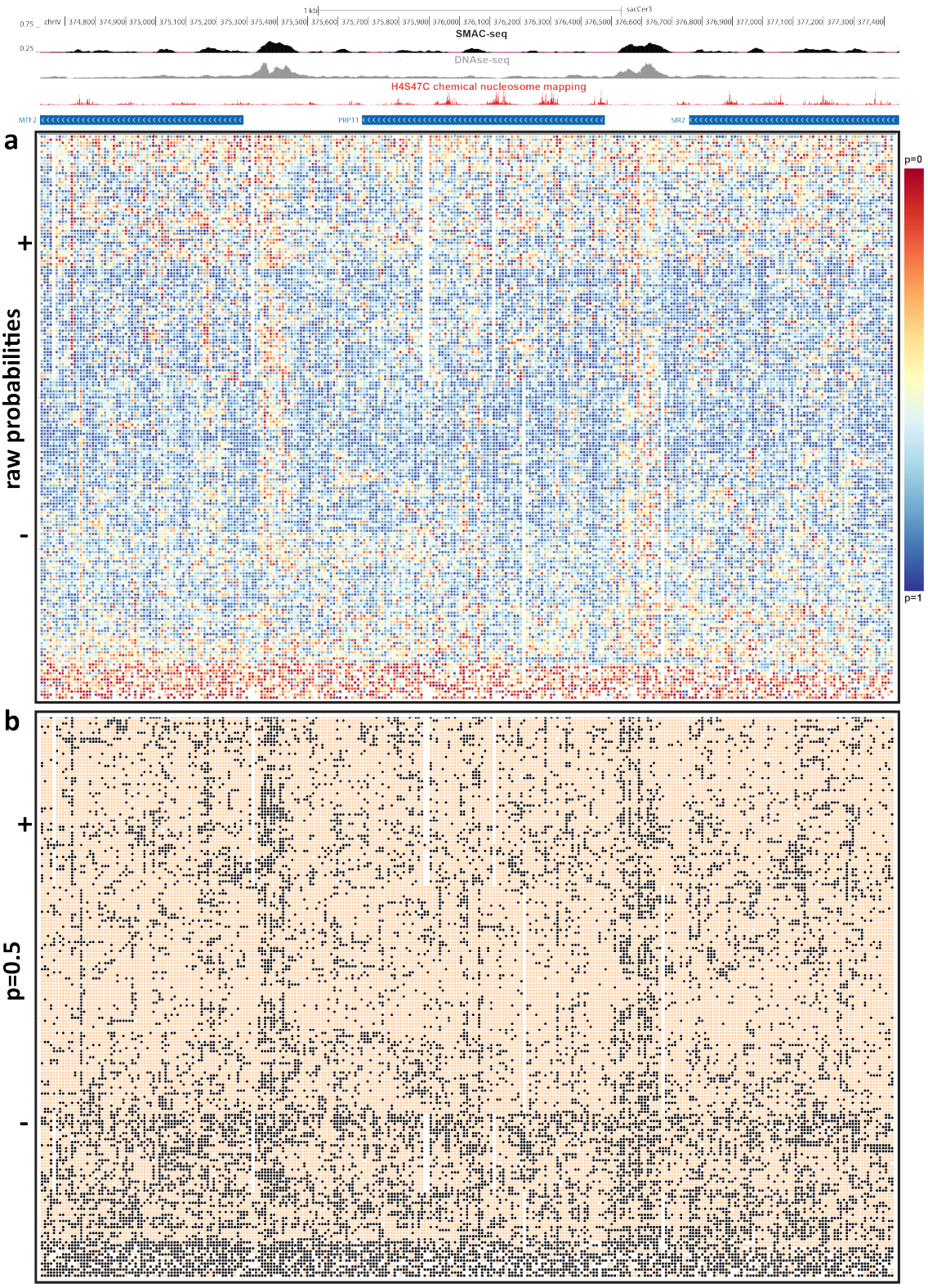


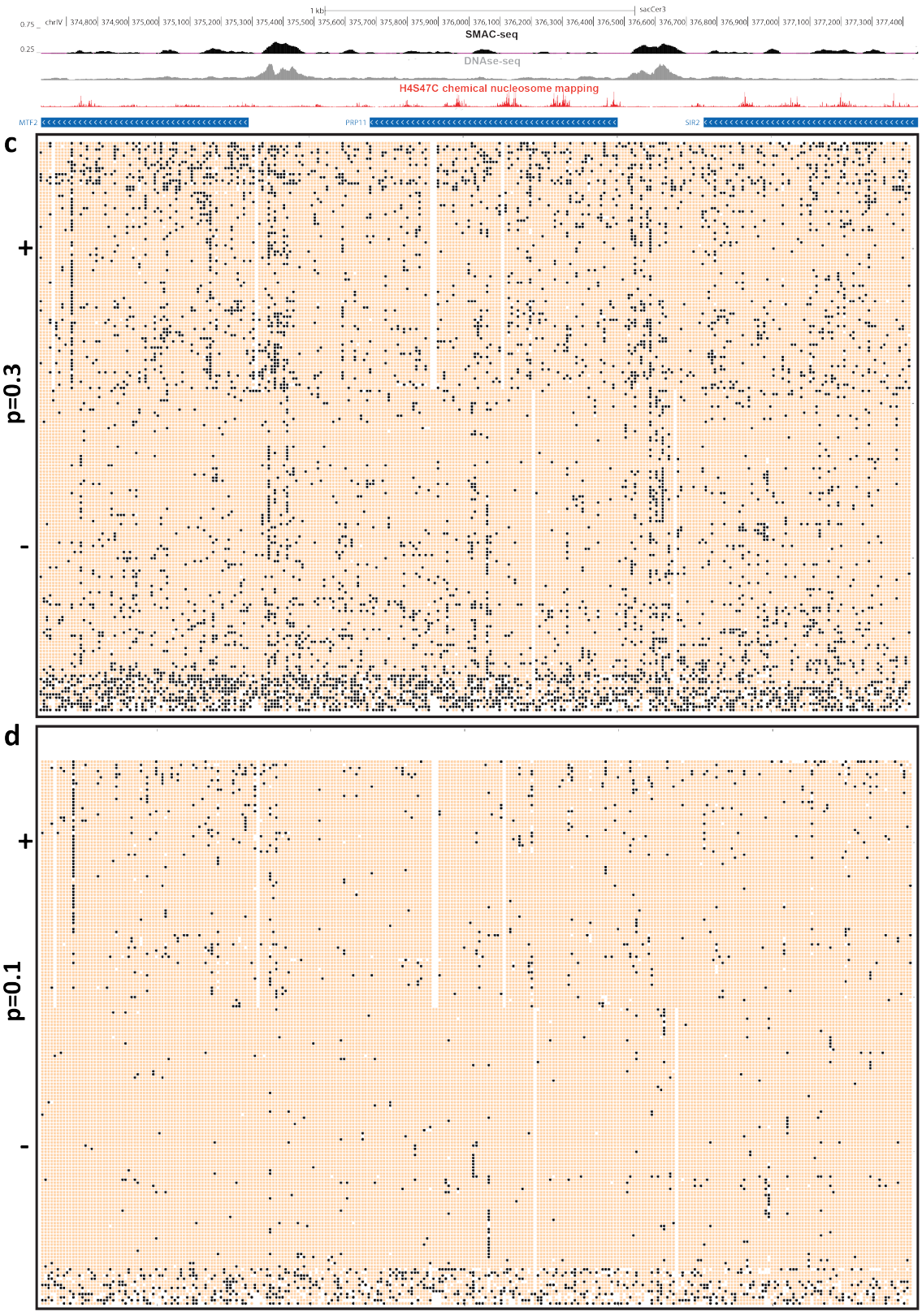
Supplementary Figure 9: Distribution of methylation call probabilities. Shown is the distribution of Tombo "alternative model" probabilities for all positions, and each of the three sequence contexts, as well as the distribution after filtering potential poor quality/non-chromatinized reads (see further below for more details).


100 bases
sacCer3
0.75 chrXVI
924,050
924,100
924,150
924,200
924,250
924,300
924,350
924,400
SMAC-seq
0.25
DNAse-seq
H4S47C chemical nucleosome mapping
OPT2
a
raw probabilities
+
-
p=0
p=1
b
p=0.5
+
-

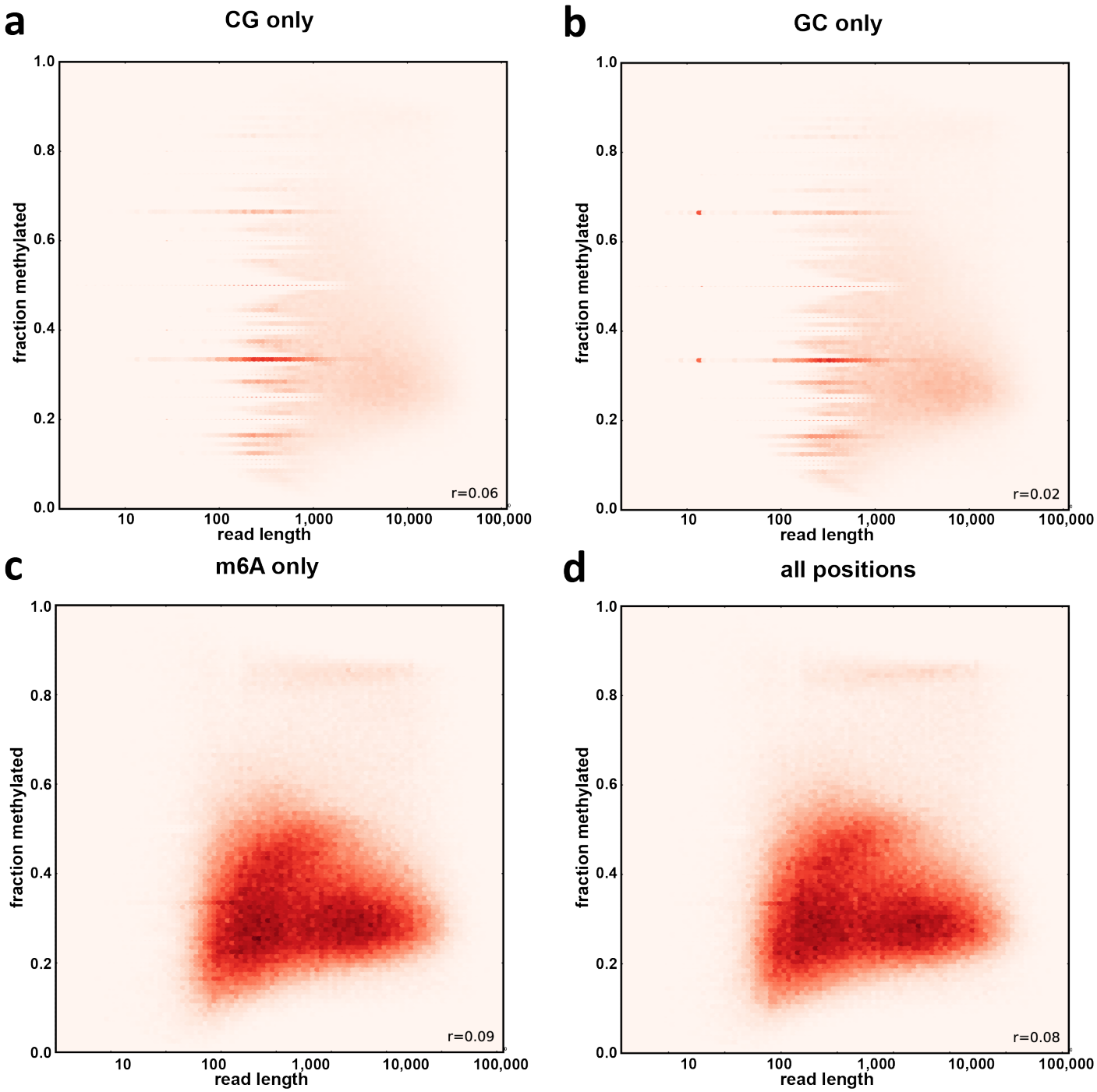


Supplementary Figure 10: Transformation of raw methylation probabilities into binary methylation calls. Shown is raw unfiltered SMAC-seq single-molecule data over a strongly positioned nucleosome on chrXVI (1-bp resolution). White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (a) raw Tombo alternative model methylation probabilities; (b) $p < 0.5$ thresholding; (d) $p < 0.3$ thresholding; (d) $p < 0.1$ thresholding.

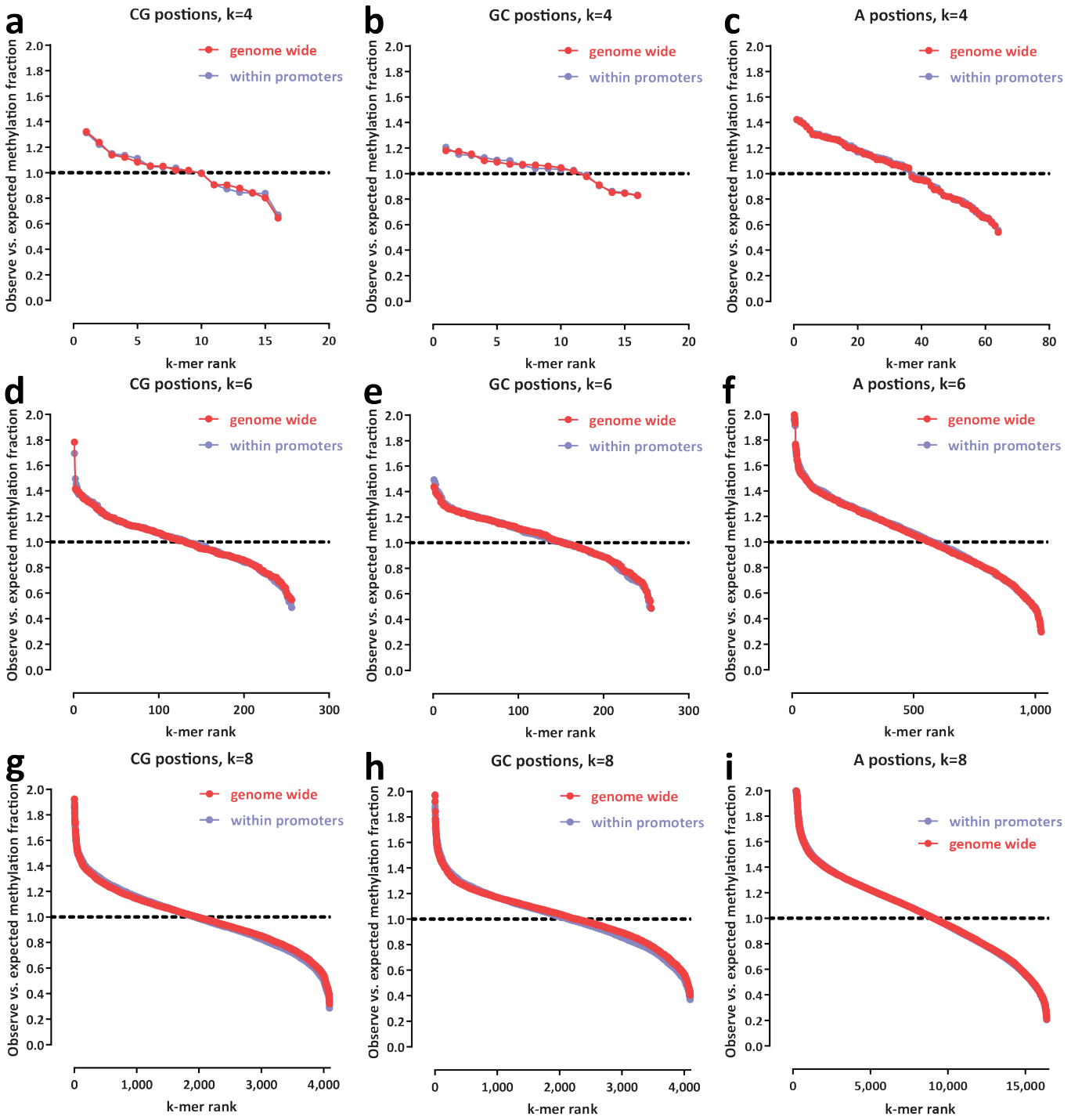

1 kb
sacCer3
chrIV 374,800 374,900 375,000 375,100 375,200 375,300 375,400 375,500 375,600 375,700 375,800 375,900 376,000 376,100 376,200 376,300 376,400 376,500 376,600 376,700 376,800 376,900 377,000 377,100 377,200 377,300 377,400
0.75
0.25
SMAC-seq
DNAse-seq
H4S47C chemical nucleosome mapping
MTF2
PRP11
SIR2
a
raw probabilities
+
-
p=0
p=1
b
p=0.5
+
-



Supplementary Figure 11: Transformation of raw methylation probabilities into binary methylation calls. Shown is raw unfiltered SMAC-seq single-molecule data over a ∼2.8kb locus on chrIV (10-bp average for all panels). White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (a) raw Tombo alternative model methylation probabilities; (b) $p < 0.5$ thresholding; (d) $p < 0.3$ thresholding; (d) $p < 0.1$ thresholding.



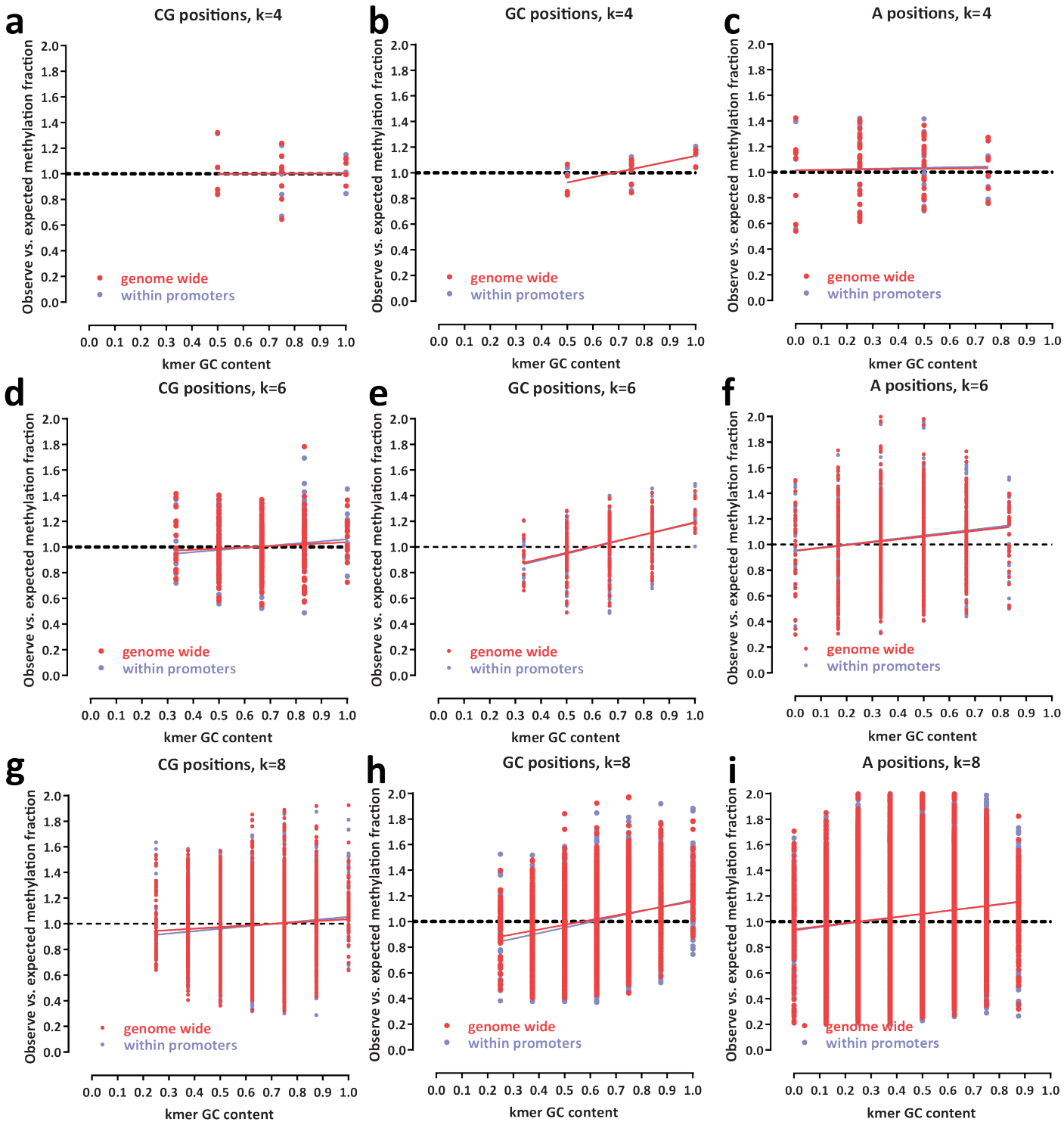
Supplementary Figure 12: Absence of strong correlation between nanopore sequencing read length and methylation status. Shown is the fraction of bases within each read that is scored as methylated. (a) CG positions only. (b) GC positions only. (c) m^6A positions only. (d) All positions.



Supplementary Figure 13: Examination of enzymatic/methylated base calling bias. Shown is the ratio of observed versus expected fraction of methylated bases for each sequence context of size k, calculated as follows:

$$f_{\text{obs/exp},k} = \frac{k_m/k_u}{\sum_k k_m / \sum_k k_u}$$

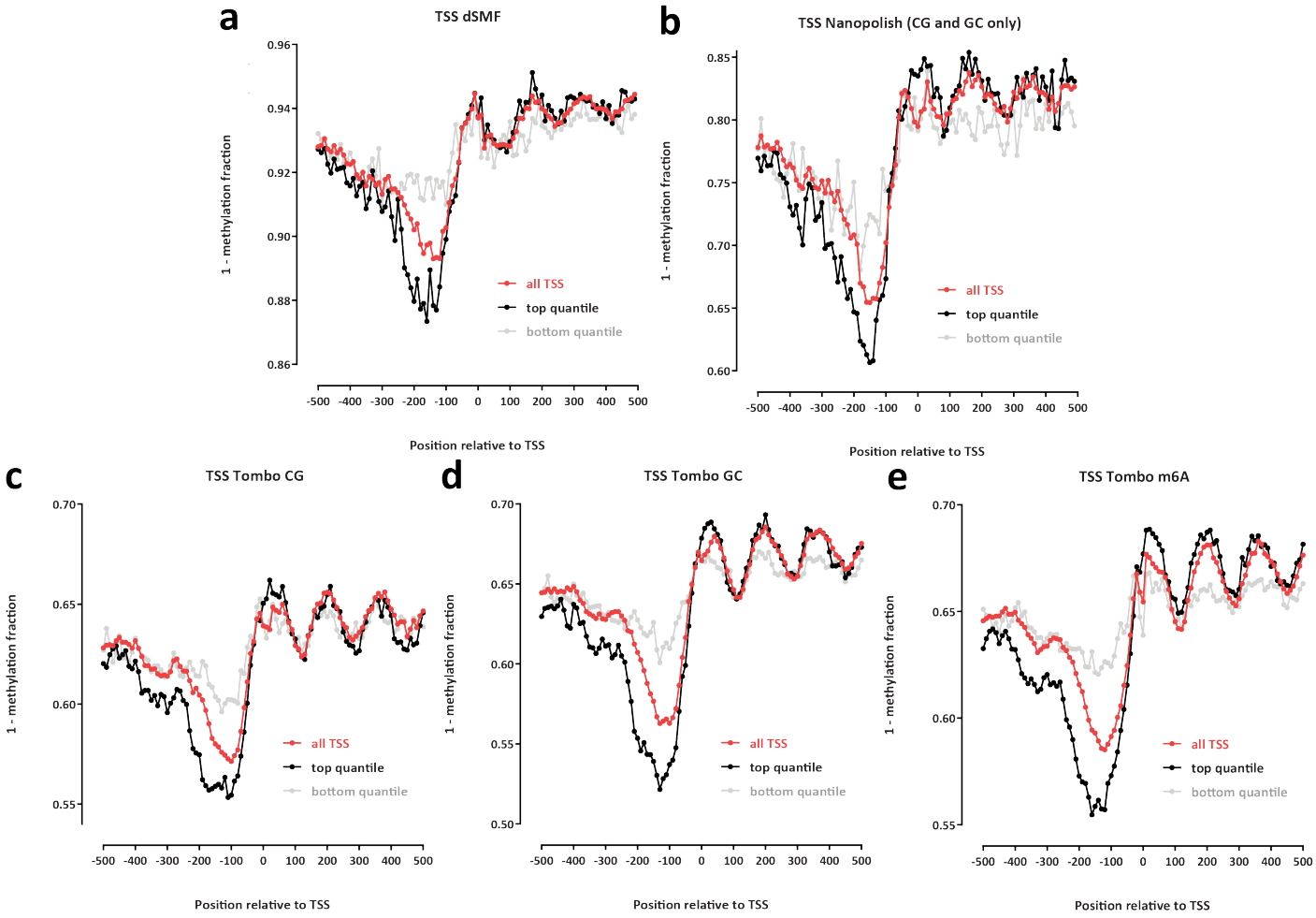
where k_m refers to the number of bases called as methylated across all reads and k_m refers to the number of bases called as unmethylated . (a) CG positions only, $k = 4$. (b) GC positions only, $k = 4$. (c) A positions only, $k = 4$. (d) CG positions only, $k = 6$. (e) GC positions only, $k = 6$. (f) A positions only, $k = 6$. (g) CG positions only, $k = 8$. (h) GC positions only, $k = 8$. (i) A positions only, $k = 8$.



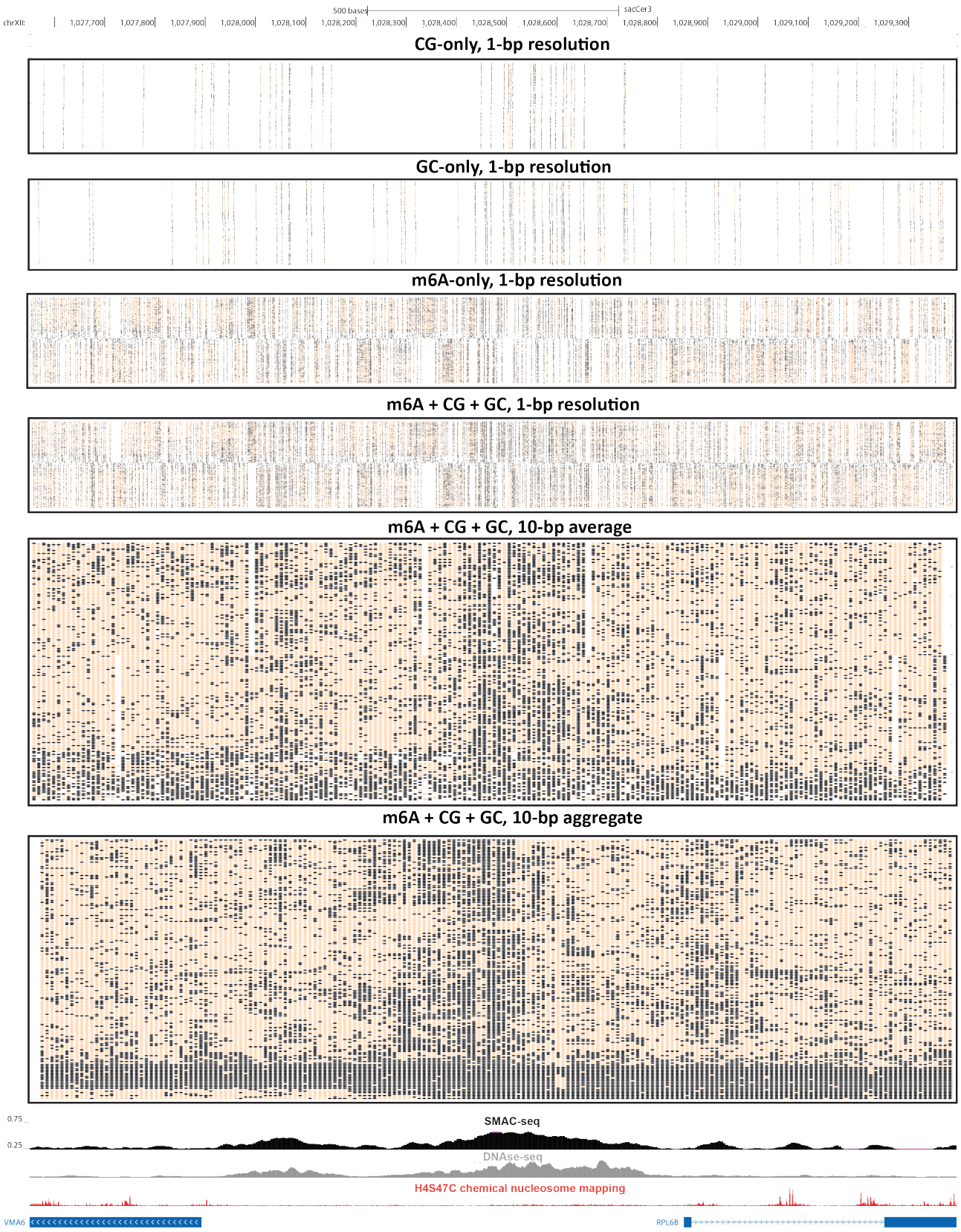
Supplementary Figure 14: Relationship between local GC content and enzymatic/methylated base calling bias. Shown is the ratio of observed versus expected fraction of methylated bases for each sequence context of size k, calculated as follows:

$$f_{\text{obs/exp},k} = \frac{k_m/k_u}{\sum_k k_m / \sum_k k_u}$$

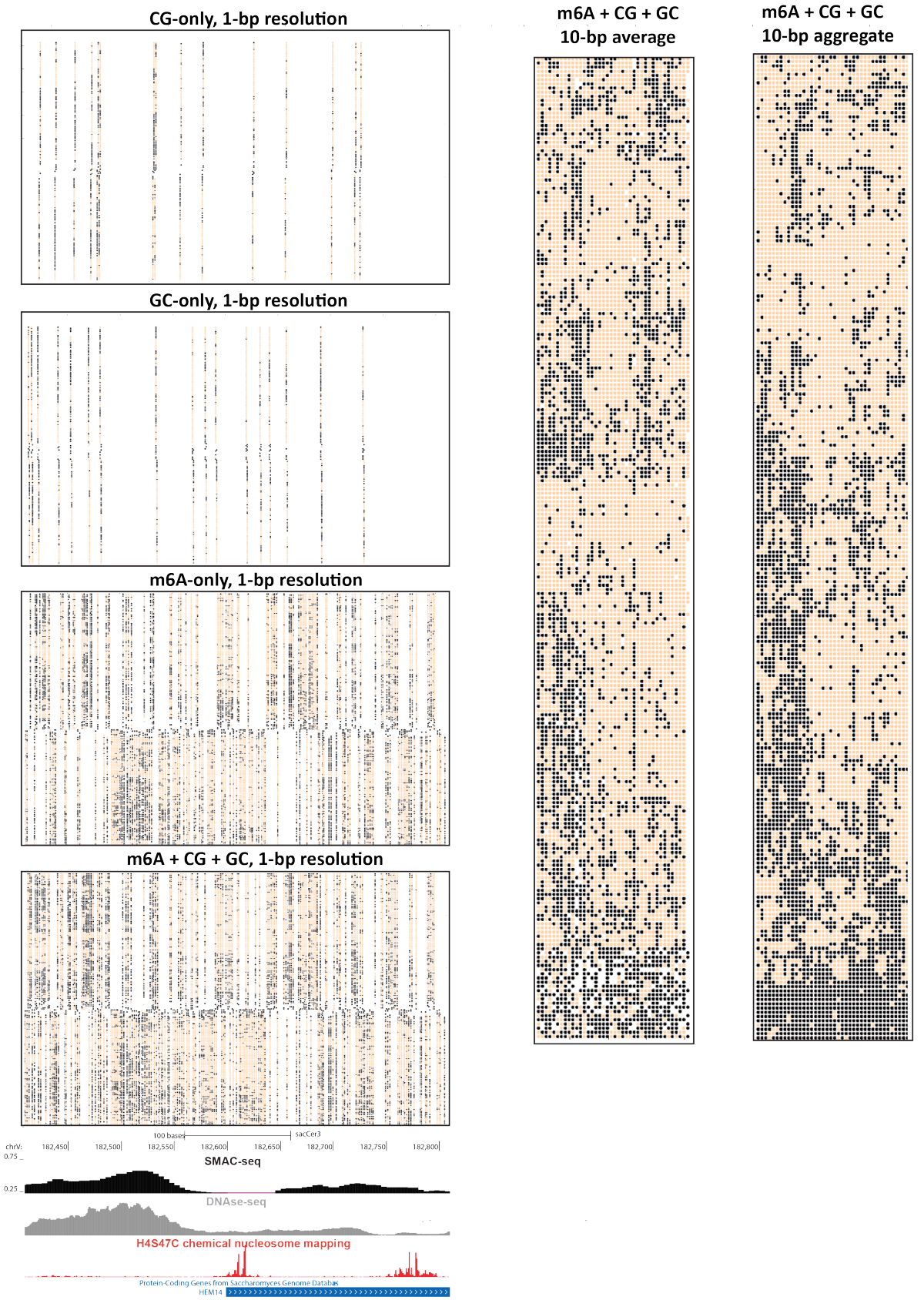
where k_m refers to the number of bases called as methylated across all reads and k_m refers to the number of bases called as unmethylated . (a) CG positions only, $k = 4$. (b) GC positions only, $k = 4$. (c) A positions only, $k = 4$. (d) CG positions only, $k = 6$. (e) GC positions only, $k = 6$. (f) A positions only, $k = 6$. (g) CG positions only, $k = 8$. (h) GC positions only, $k = 8$. (i) A positions only, $k = 8$.



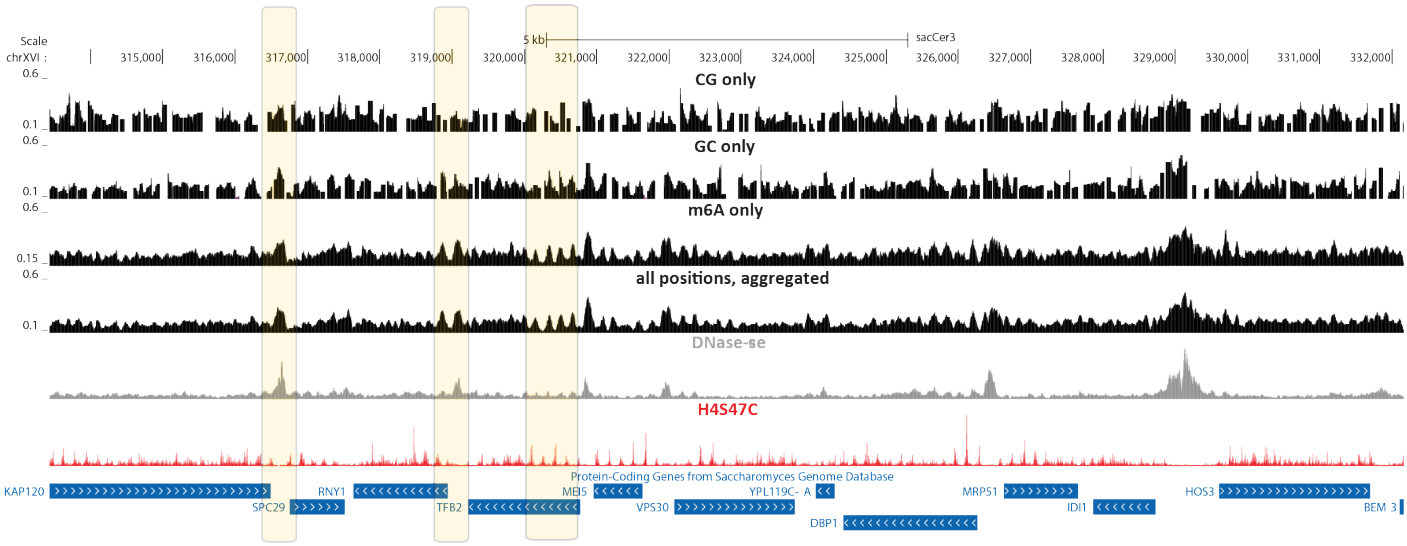
Supplementary Figure 15: Comparison of dSMF results and different approaches to methylation-aware base calling on SMAC-seq data. Shown is the inverse of the methylated fraction of nucleotides around TSSs of all, highly express (top quantile) and low expression-level (bottom quantiles) yeast genes (unfiltered "Sample 1" dataset). Note that the different panels are not drawn to the same scale. (a) dSMF; (b) SMAC-seq data using Nanopolish methylation base-calling on CG and GC nucleotides; (c) SMAC-seq data using Tombo methylation base-calling, CG positions only; (d) SMAC-seq data using Tombo methylation base-calling, GC positions only; (e) SMAC-seq data using Tombo methylation base-calling, m6A positions only.



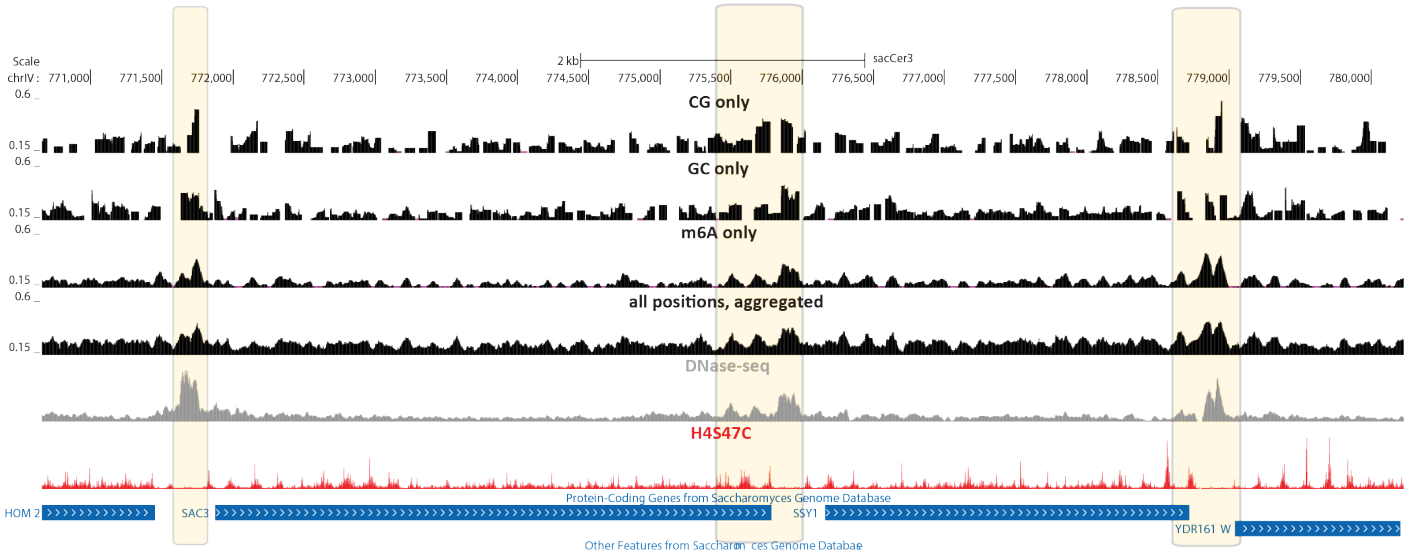
Supplementary Figure 16: Impact of the addition of m^6A on assay resolution. Shown is the raw unfiltered nanopore read coverage around a divergent promoter region on chrXII, considering only CG, only GC, only m^6A, and all bases at 1-bp resolution as well as all bases at averaged and aggregated 10-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A).



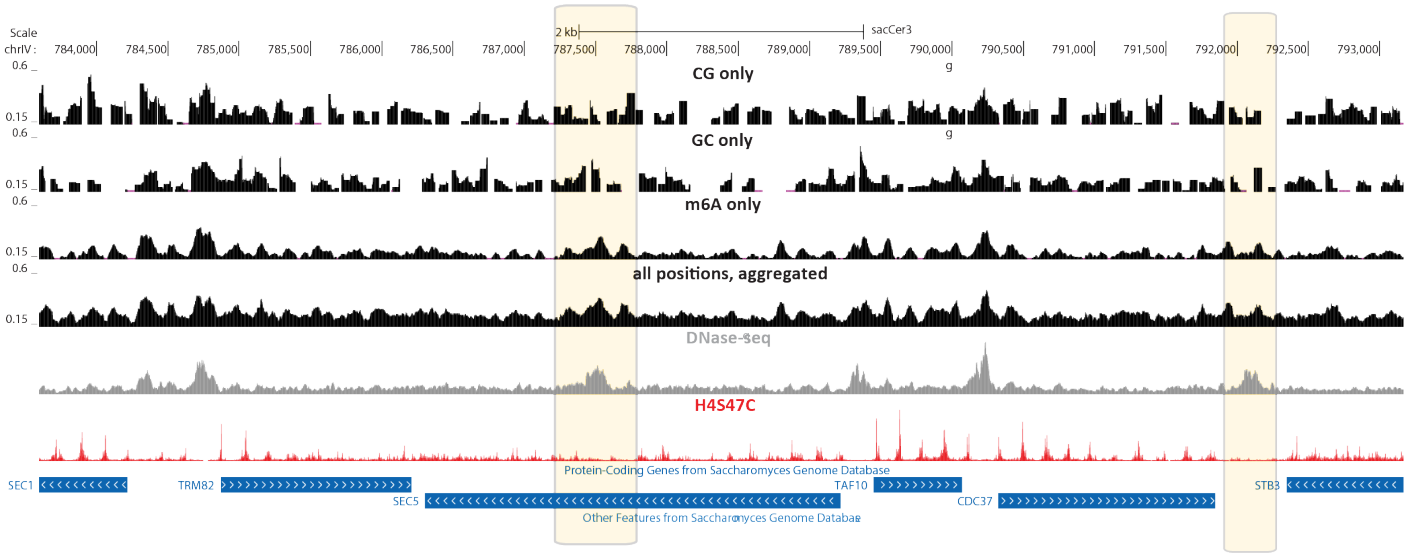
Supplementary Figure 17: Impact of the addition of m^{6}A on assay resolution. Shown is the raw unfiltered nanopore read coverage around a strongly positioned +1 nucleosome, considering only CG, only GC, only m^{6}A, and all bases at 1-bp resolution as well as all bases at averaged and aggregated 10-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A).



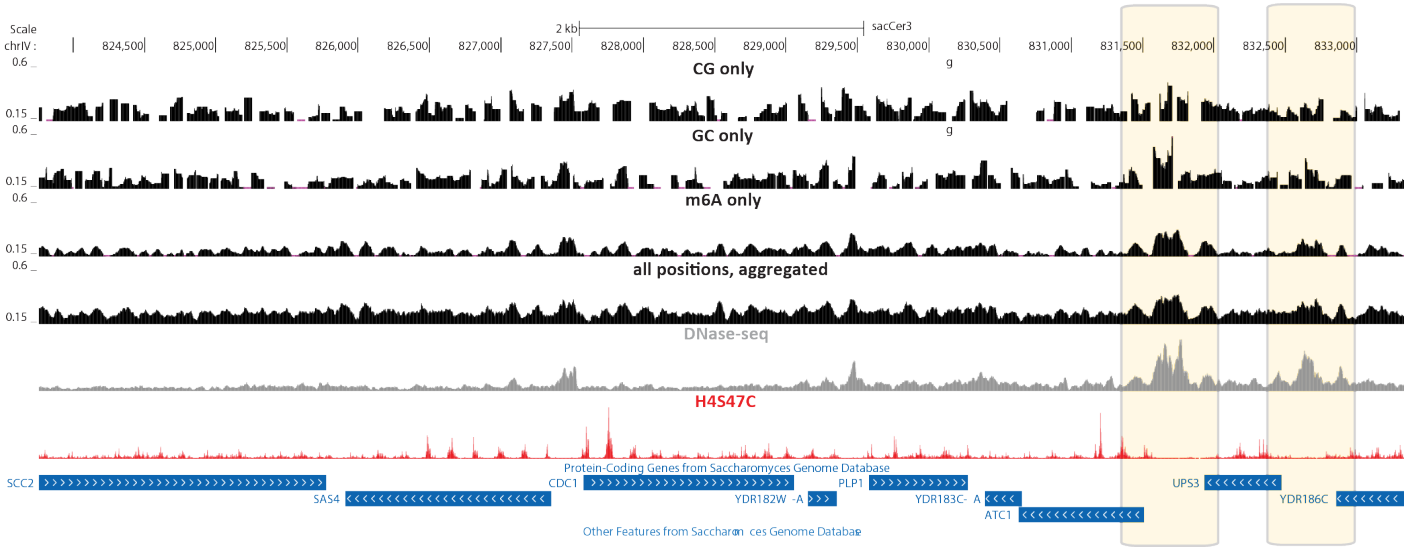
Supplementary Figure 18: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



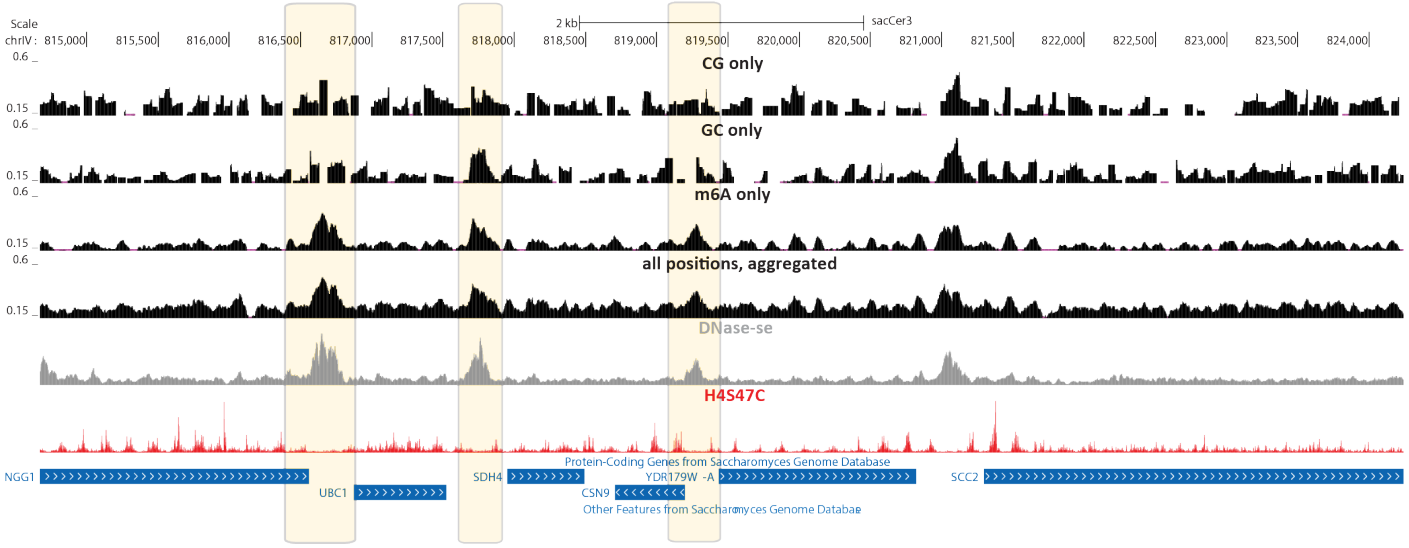
Supplementary Figure 19: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



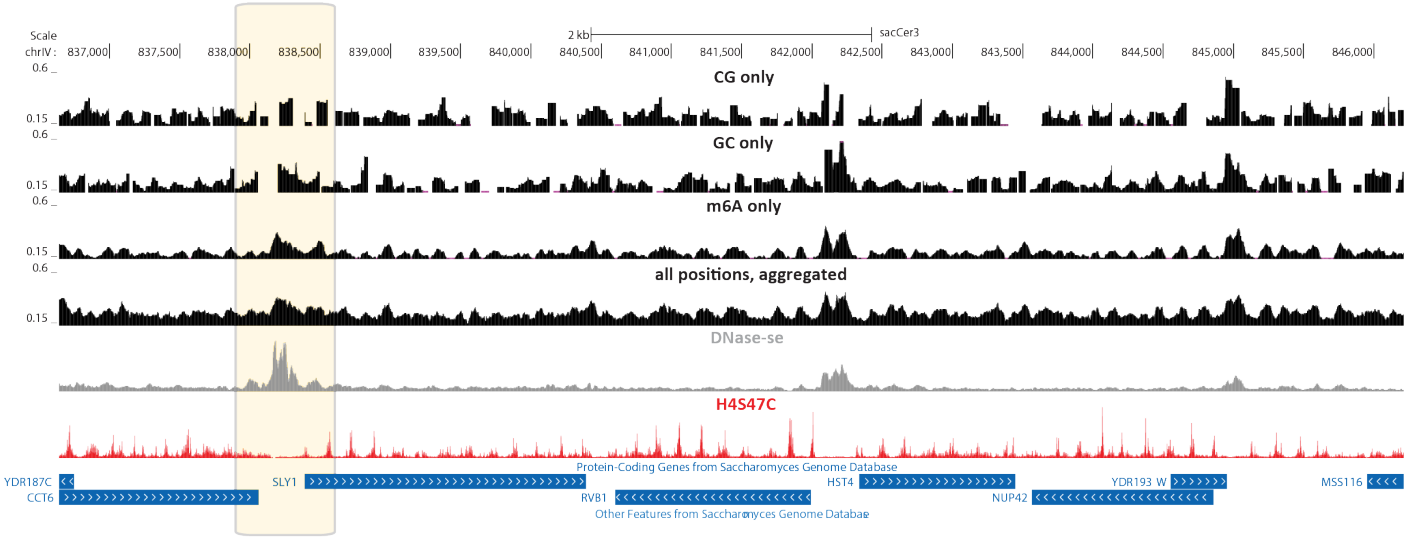
Supplementary Figure 20: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



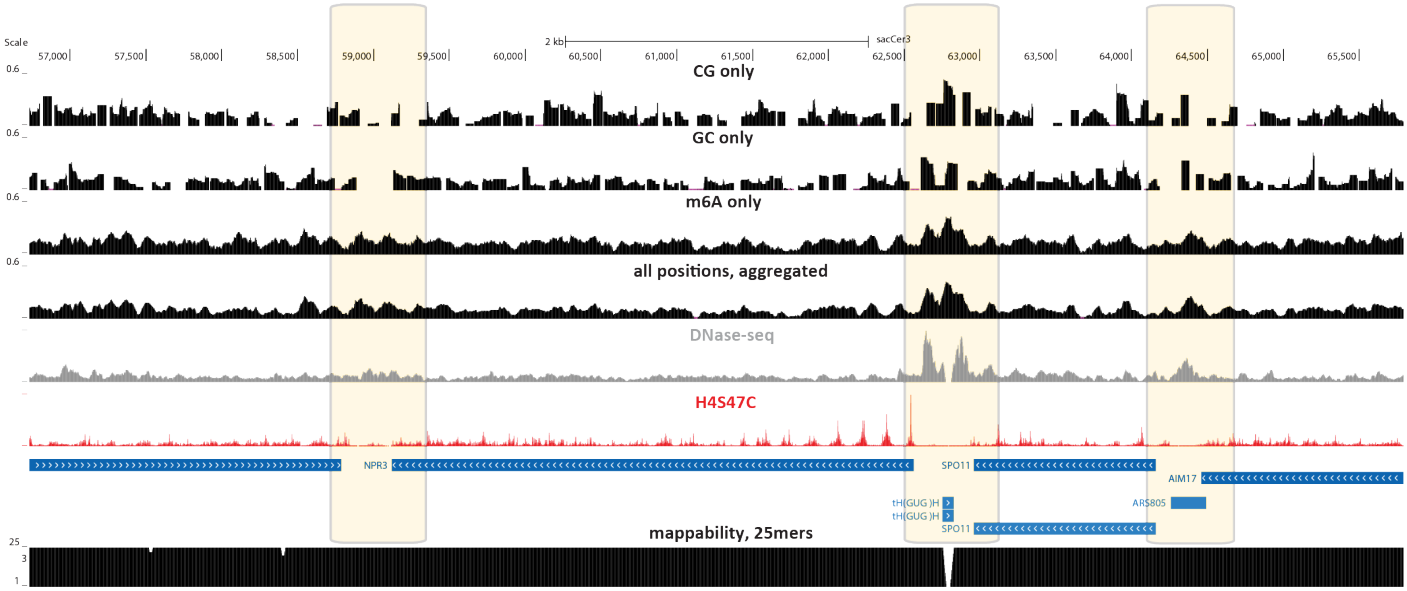
Supplementary Figure 21: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



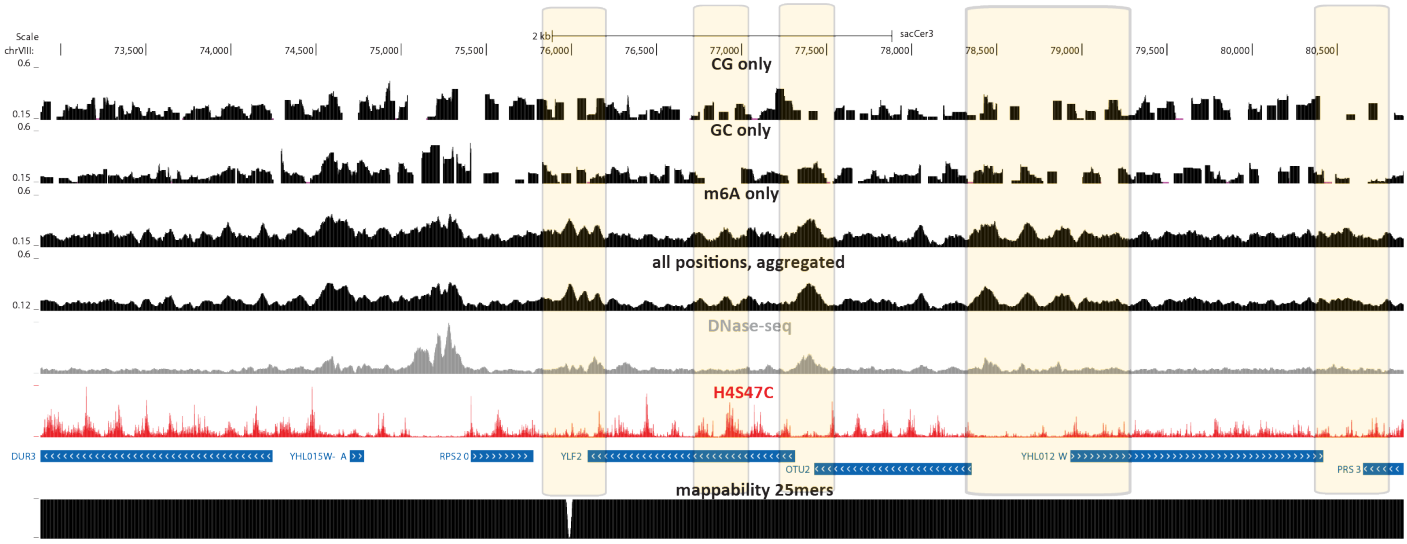
Supplementary Figure 22: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



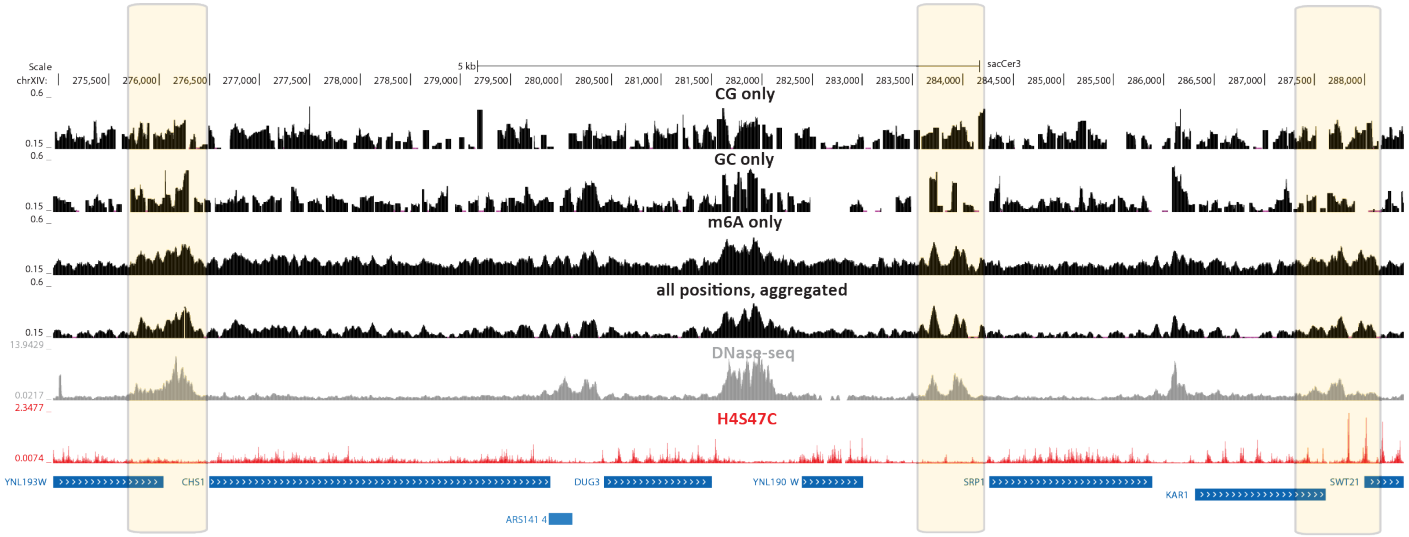
Supplementary Figure 23: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



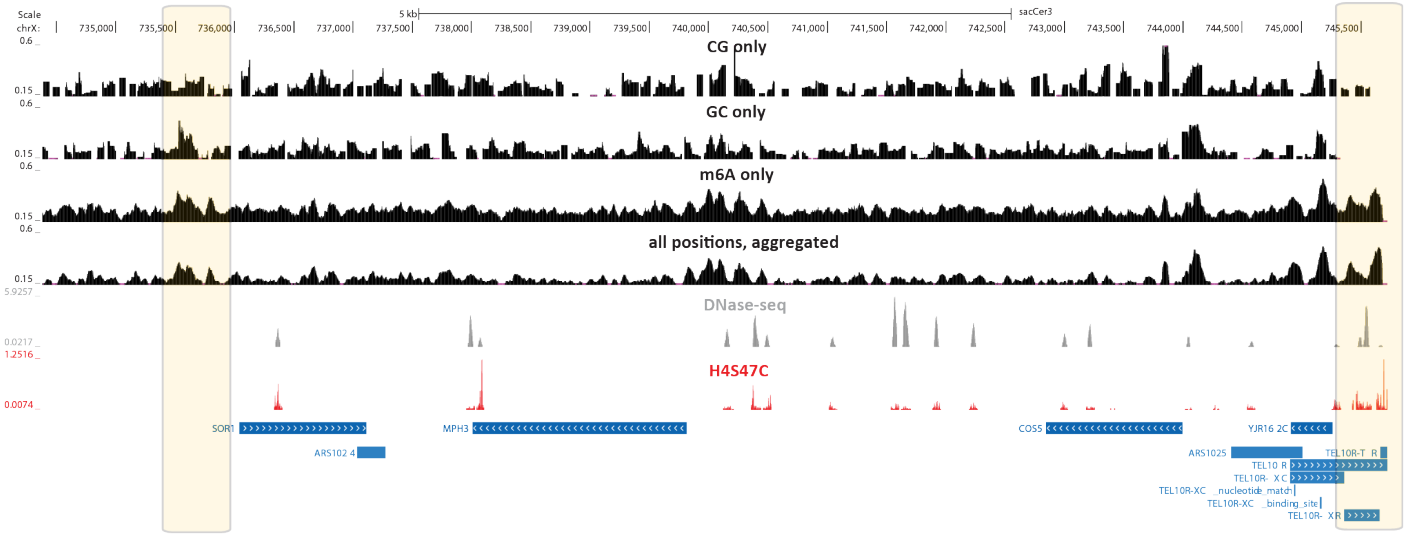
Supplementary Figure 24: Impact of the addition of m^6A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^6A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



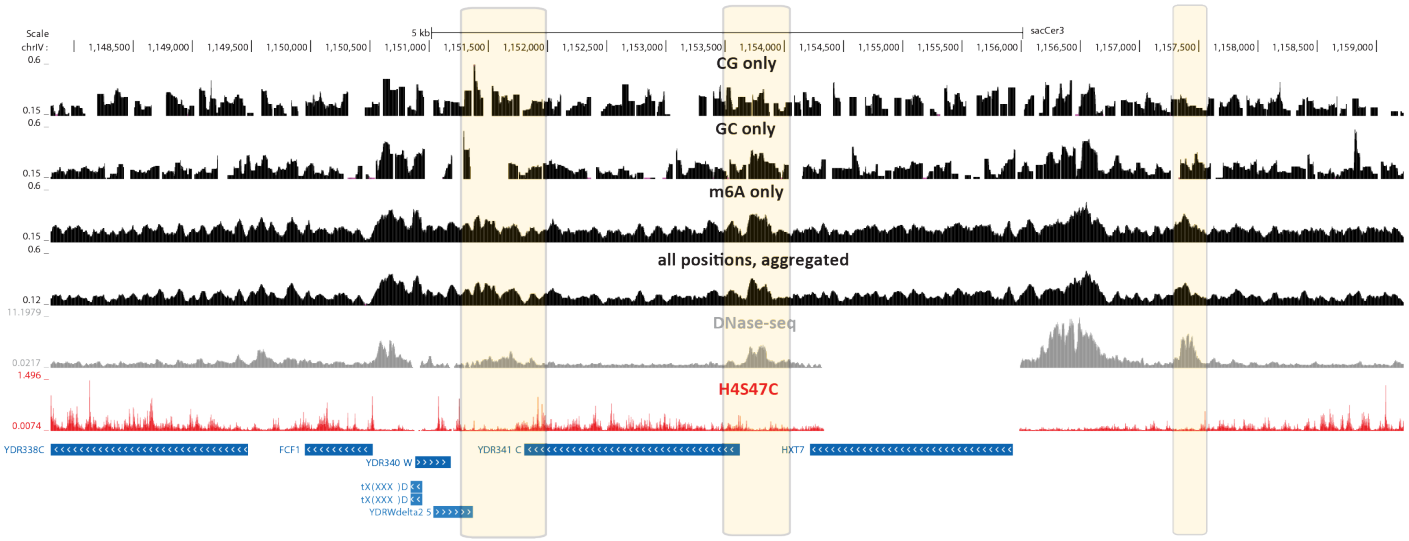
Supplementary Figure 25: Impact of the addition of m^6A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^6A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



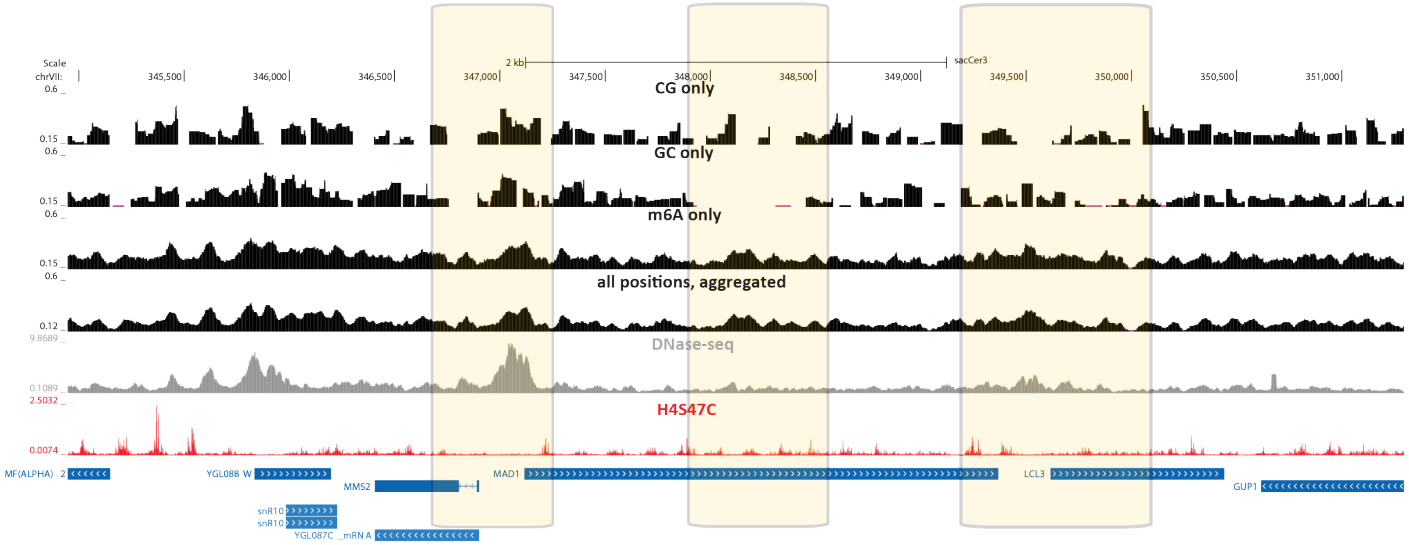
Supplementary Figure 26: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



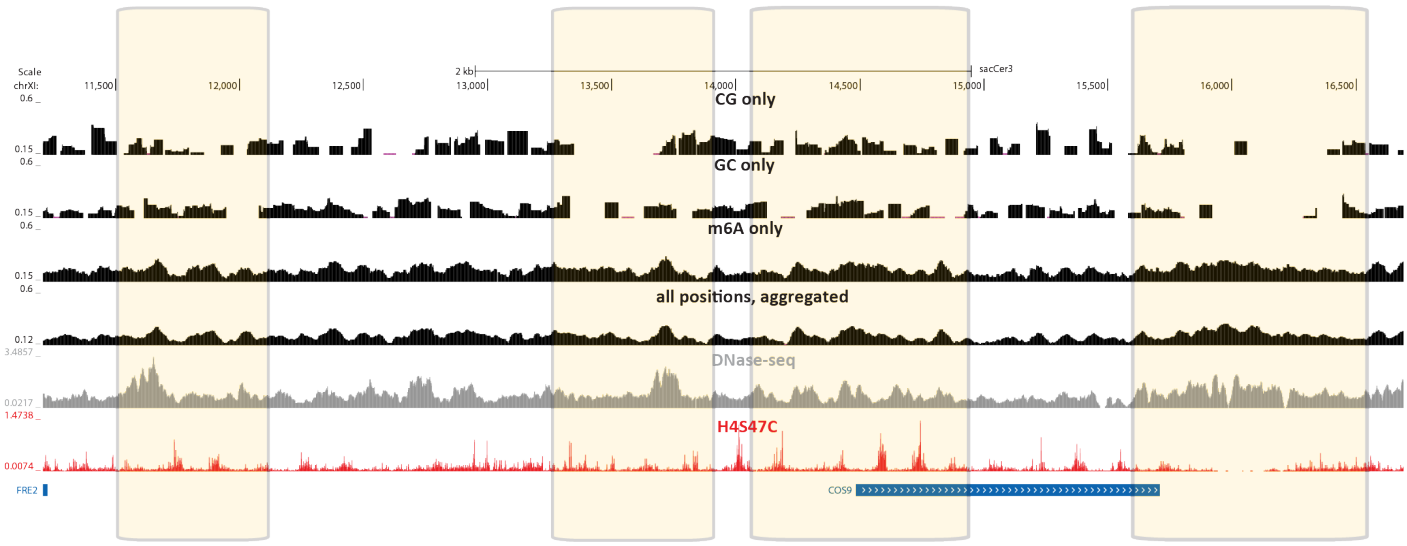
Supplementary Figure 27: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



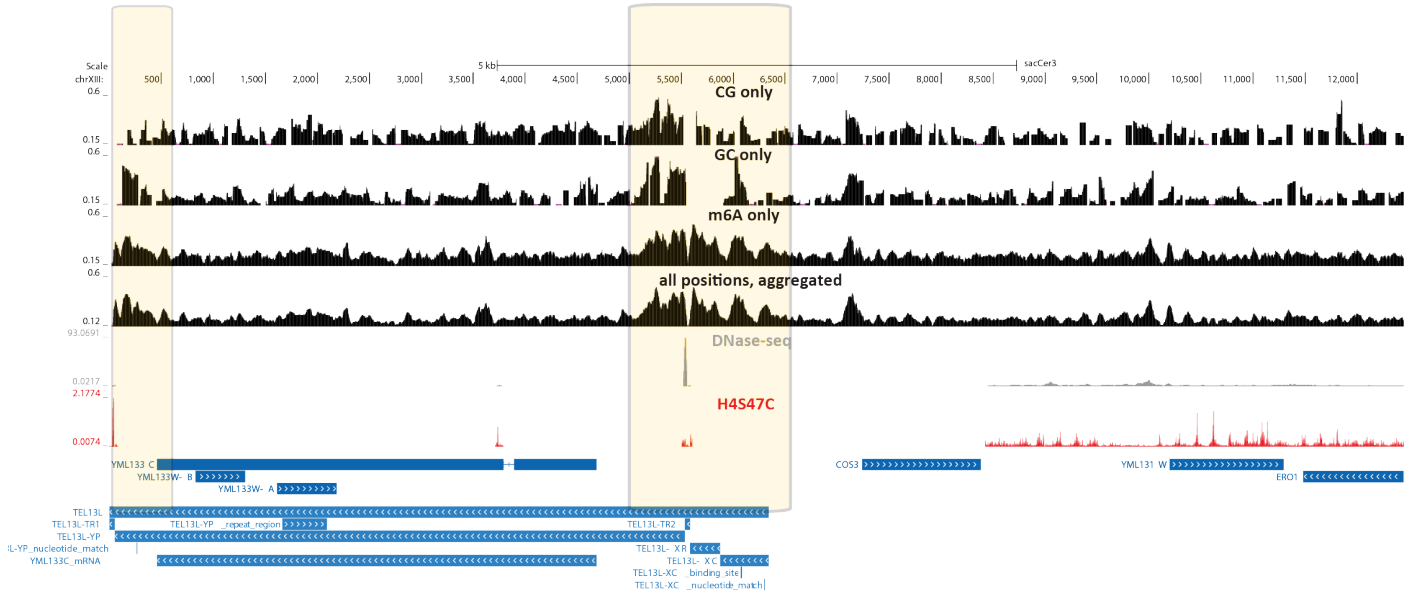
Supplementary Figure 28: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



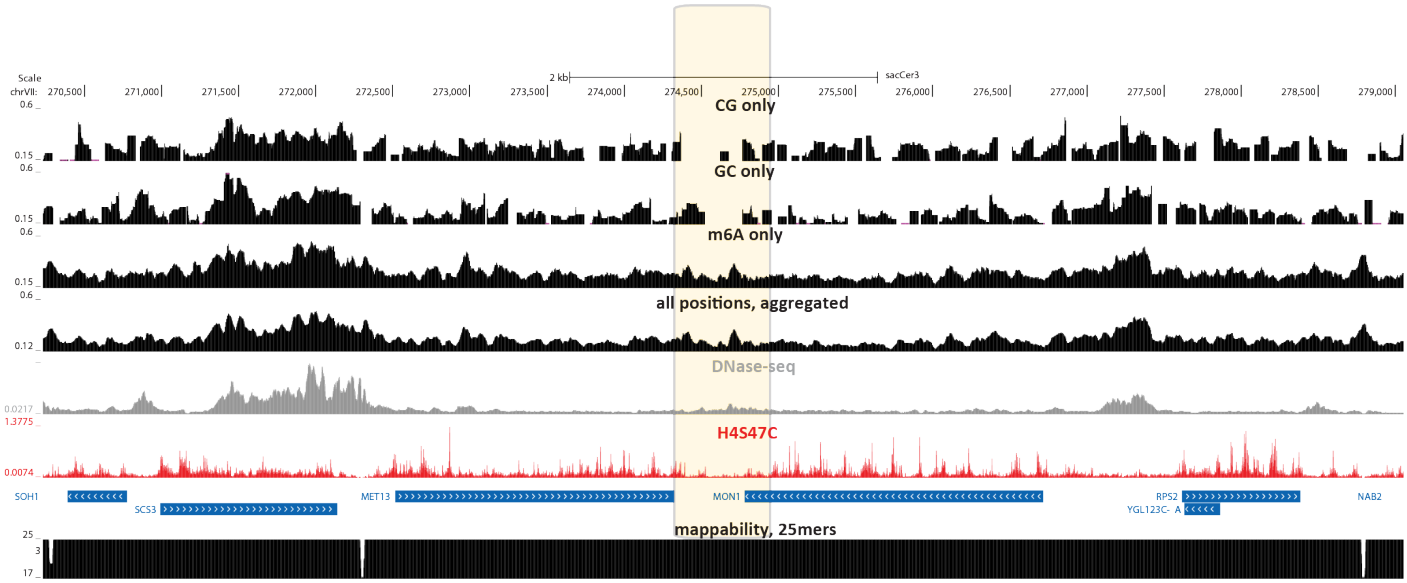
Supplementary Figure 29: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



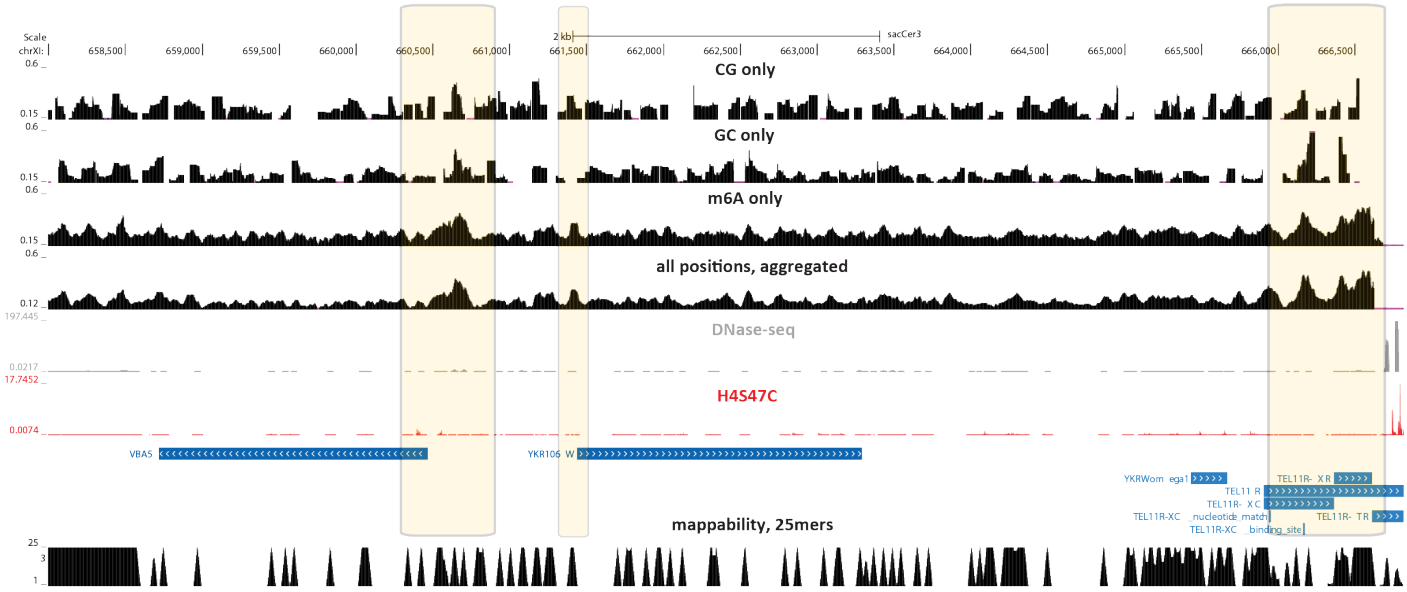
Supplementary Figure 30: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



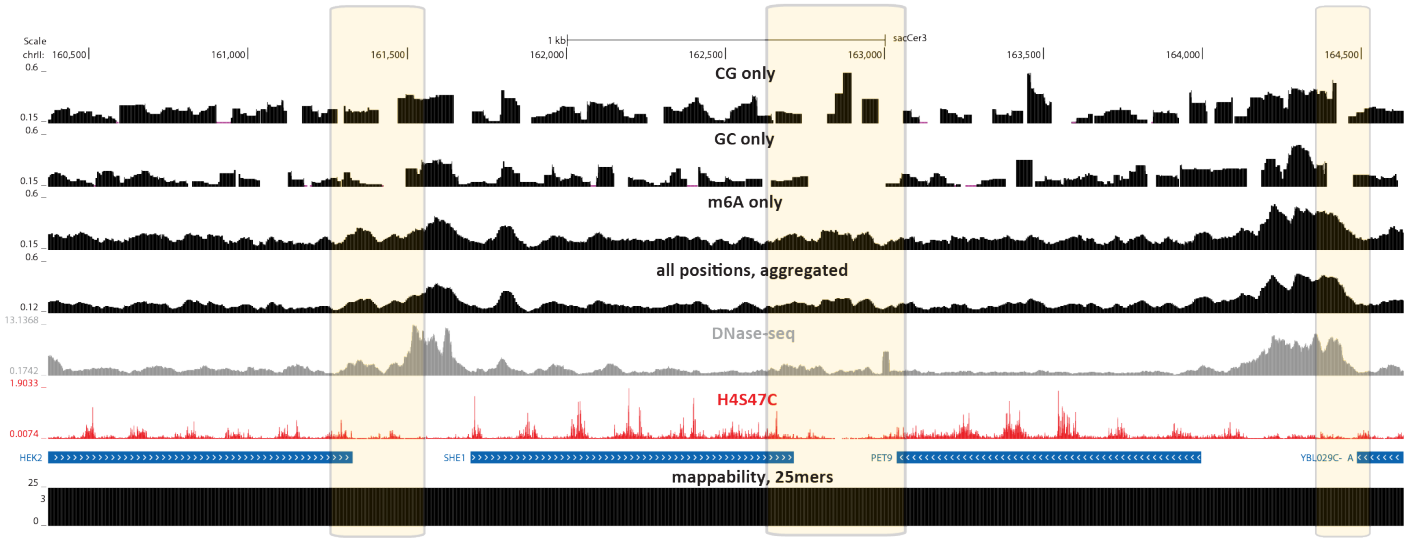
Supplementary Figure 31: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



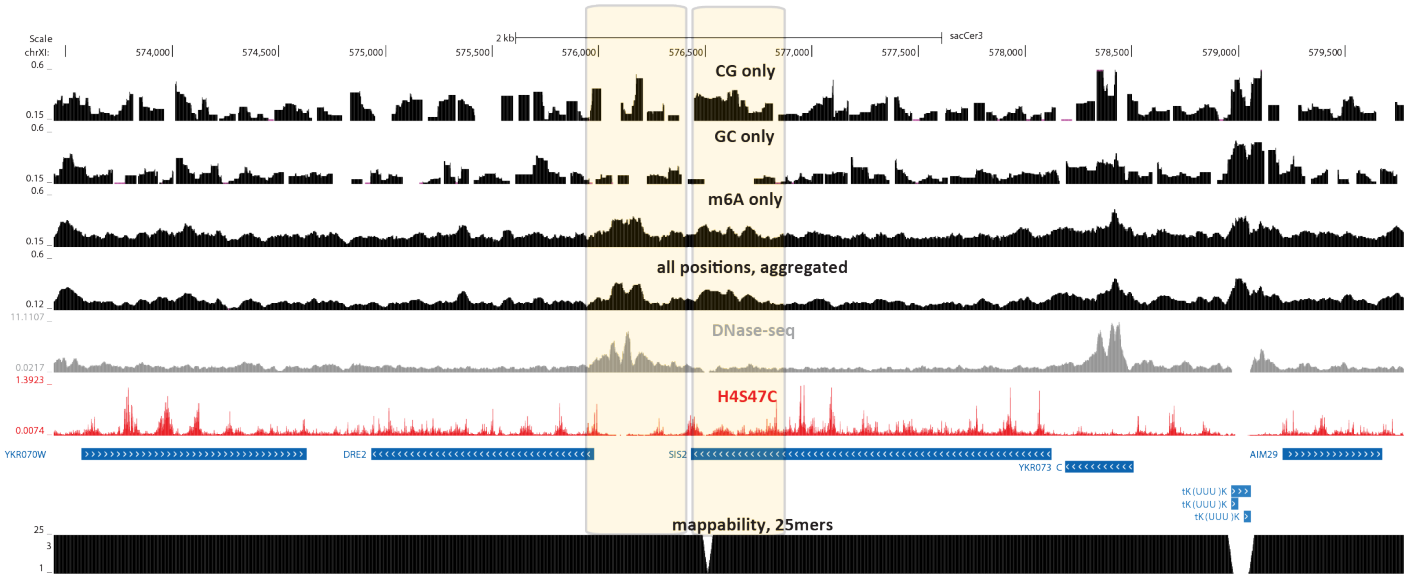
Supplementary Figure 32: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



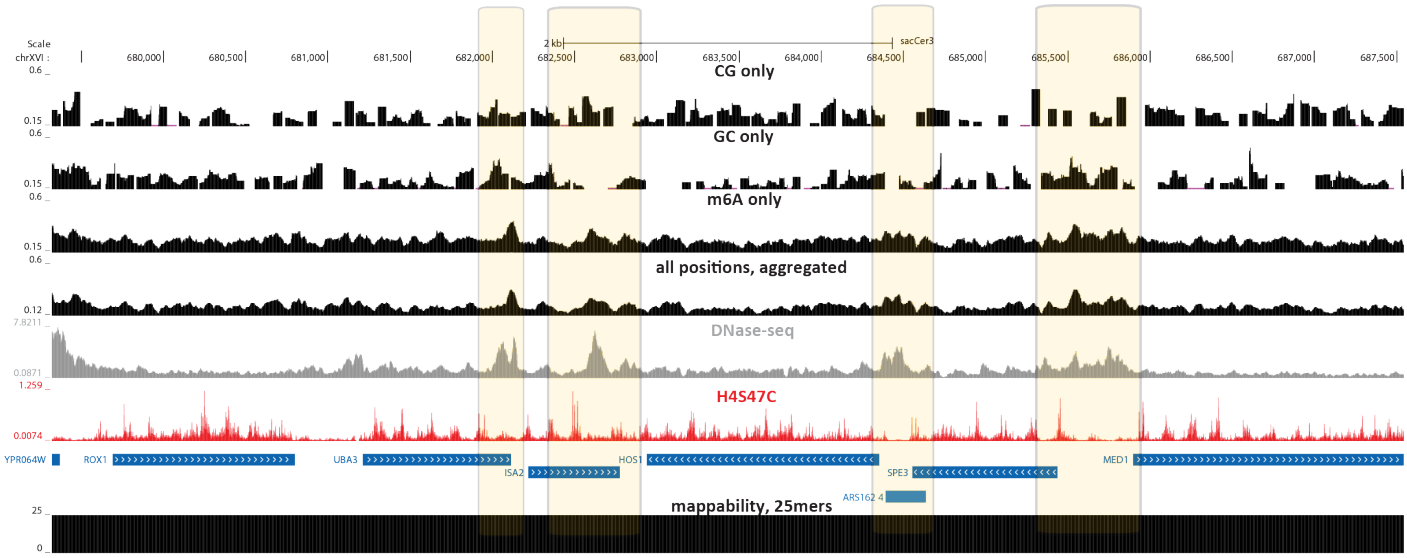
Supplementary Figure 33: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



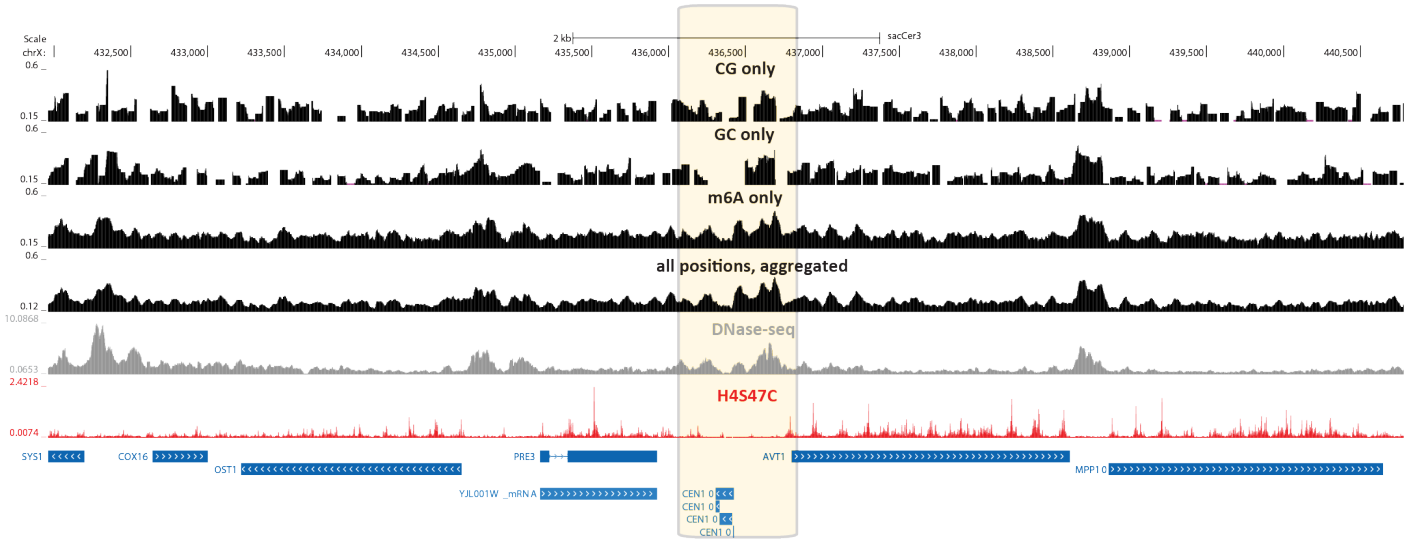
Supplementary Figure 34: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



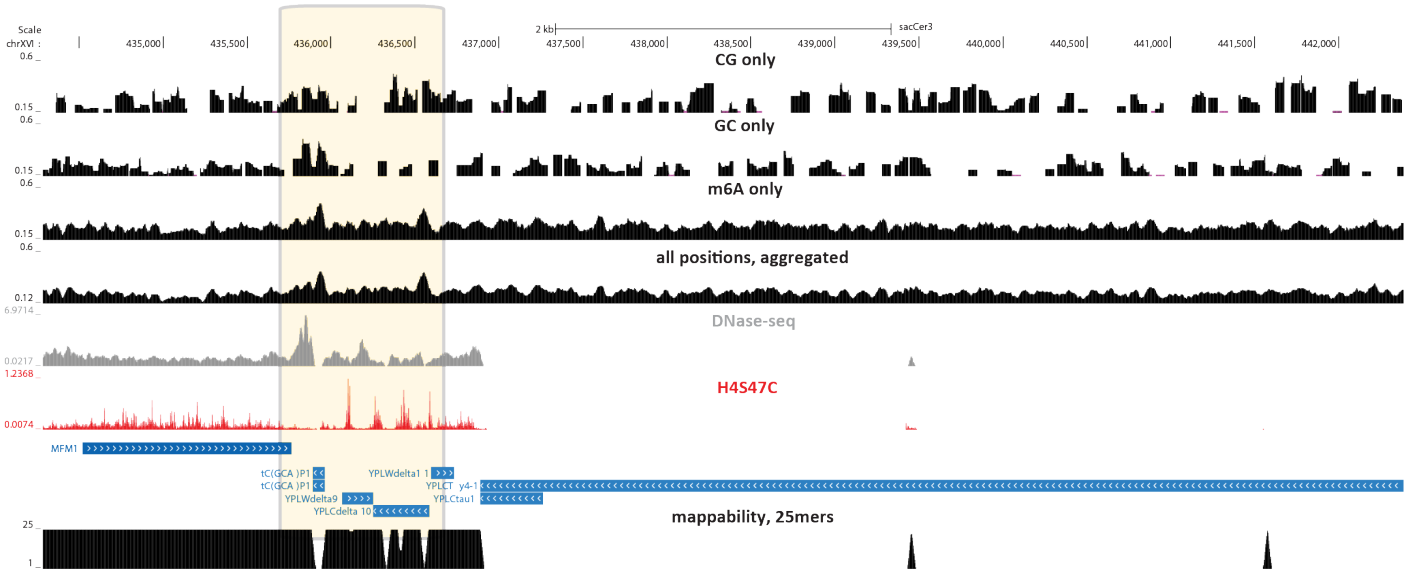
Supplementary Figure 35: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



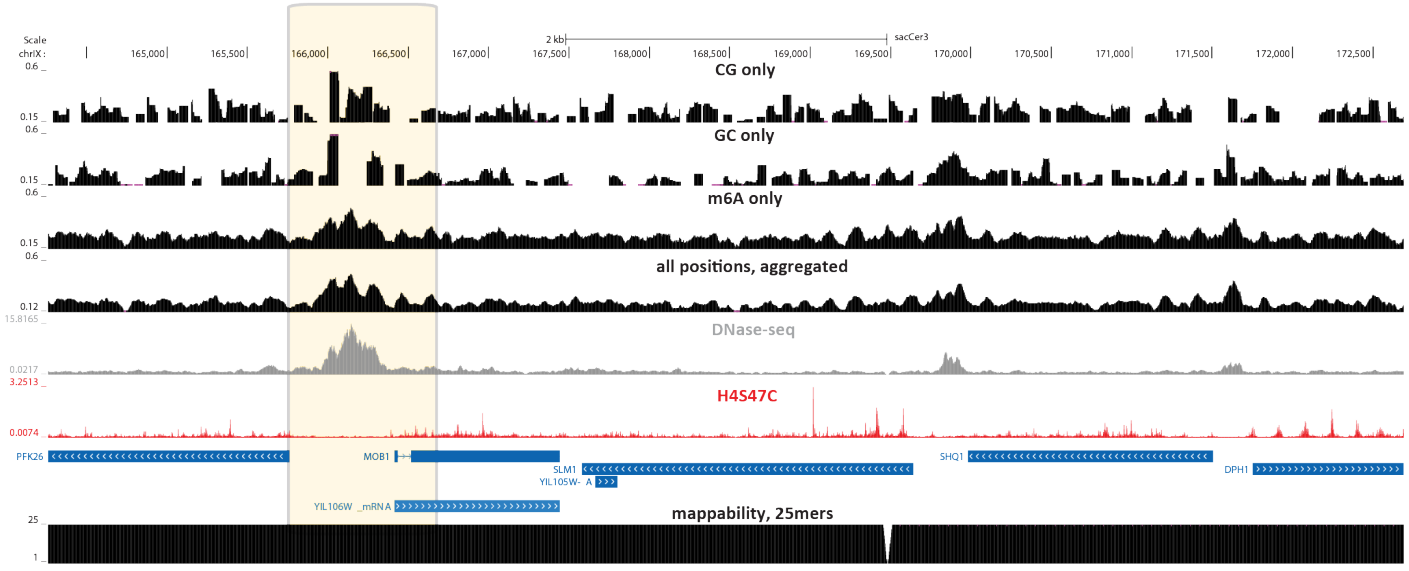
Supplementary Figure 36: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



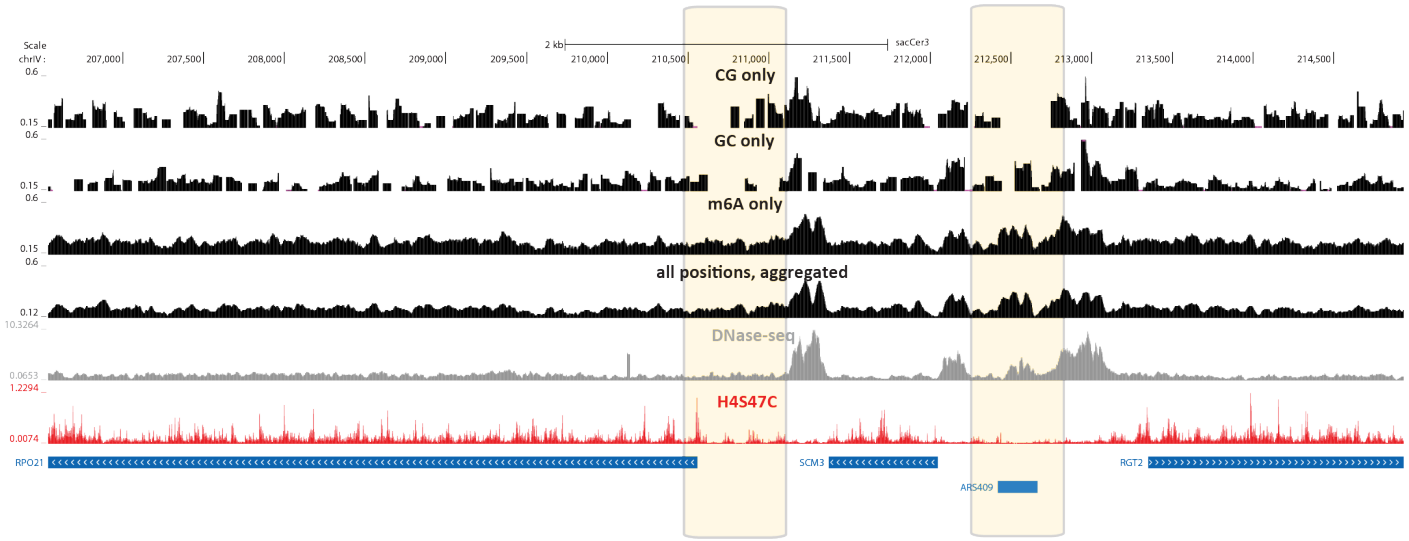
Supplementary Figure 37: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



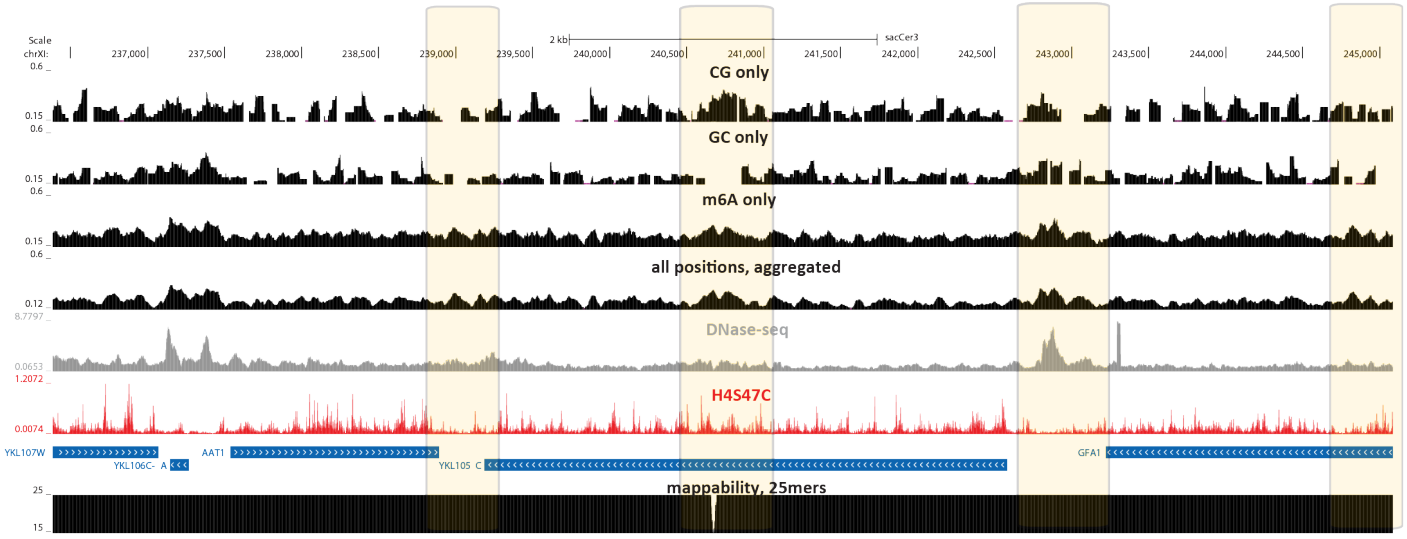
Supplementary Figure 38: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



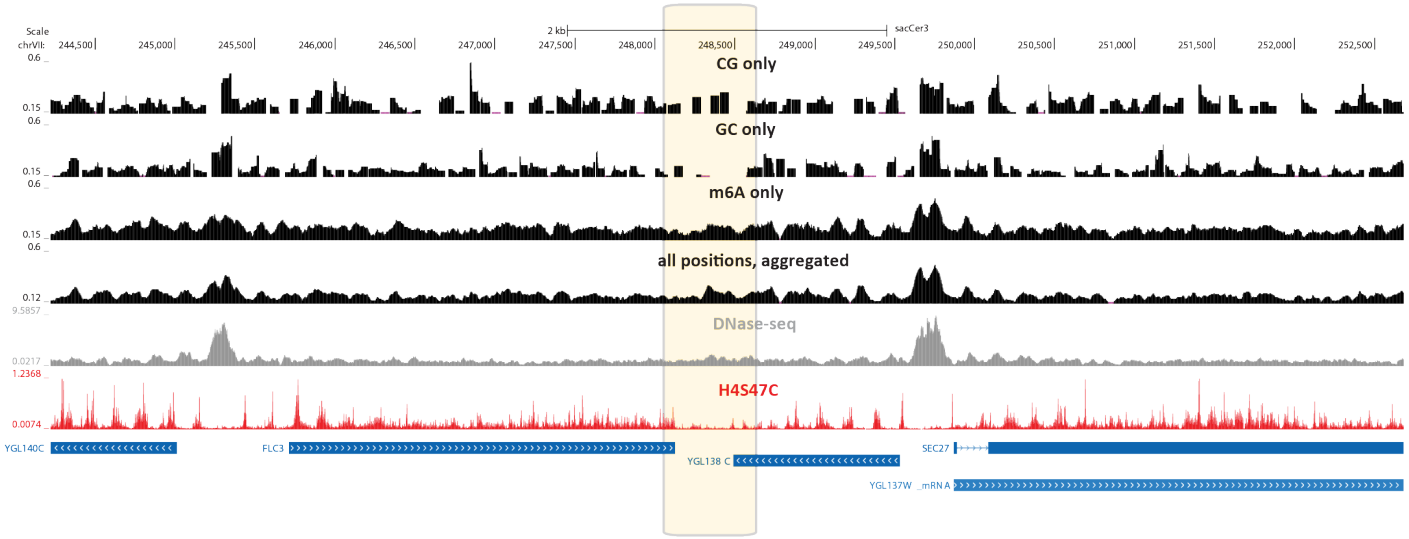
Supplementary Figure 39: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



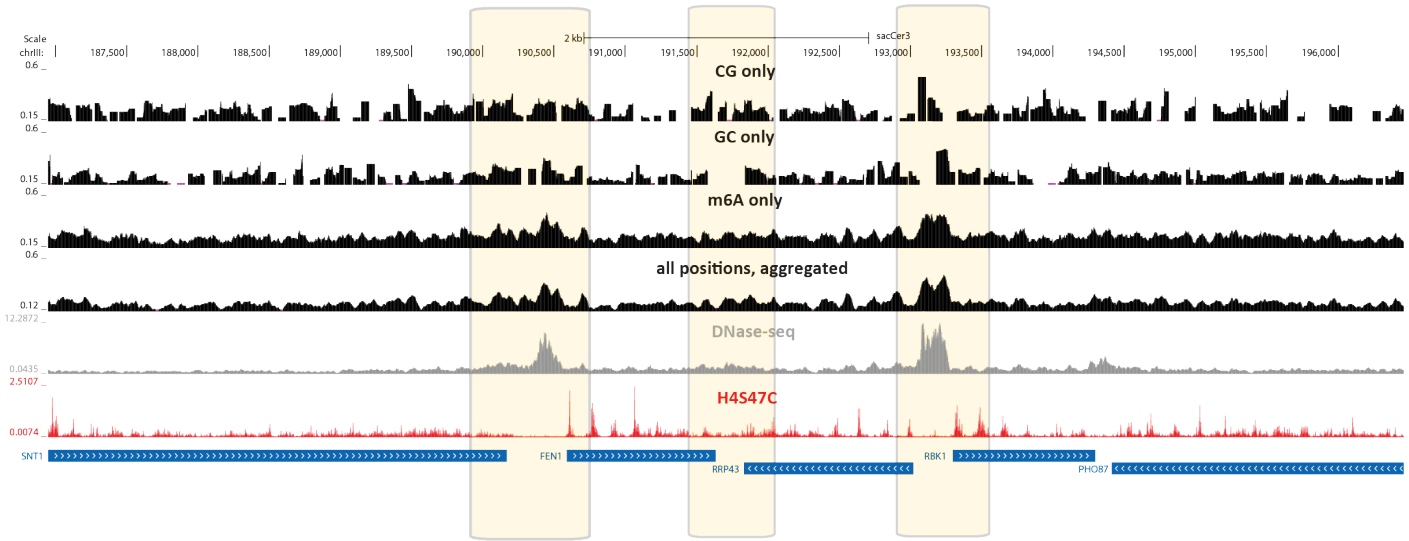
Supplementary Figure 40: Impact of the addition of m^6A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^6A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



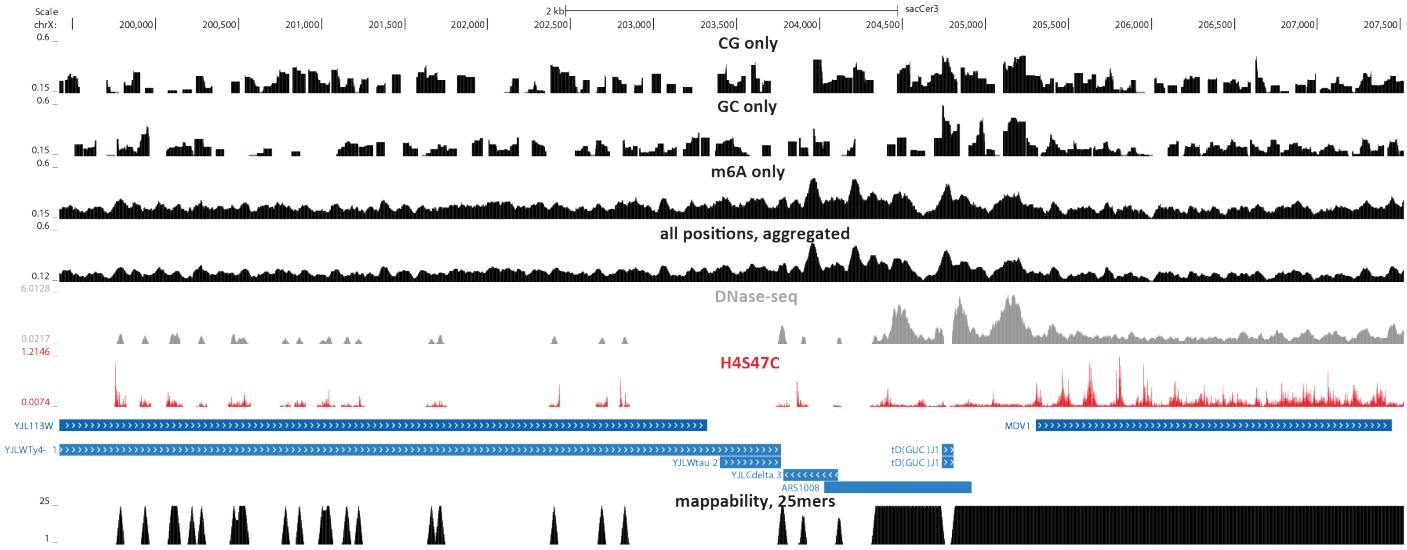
Supplementary Figure 41: Impact of the addition of m^6A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^6A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



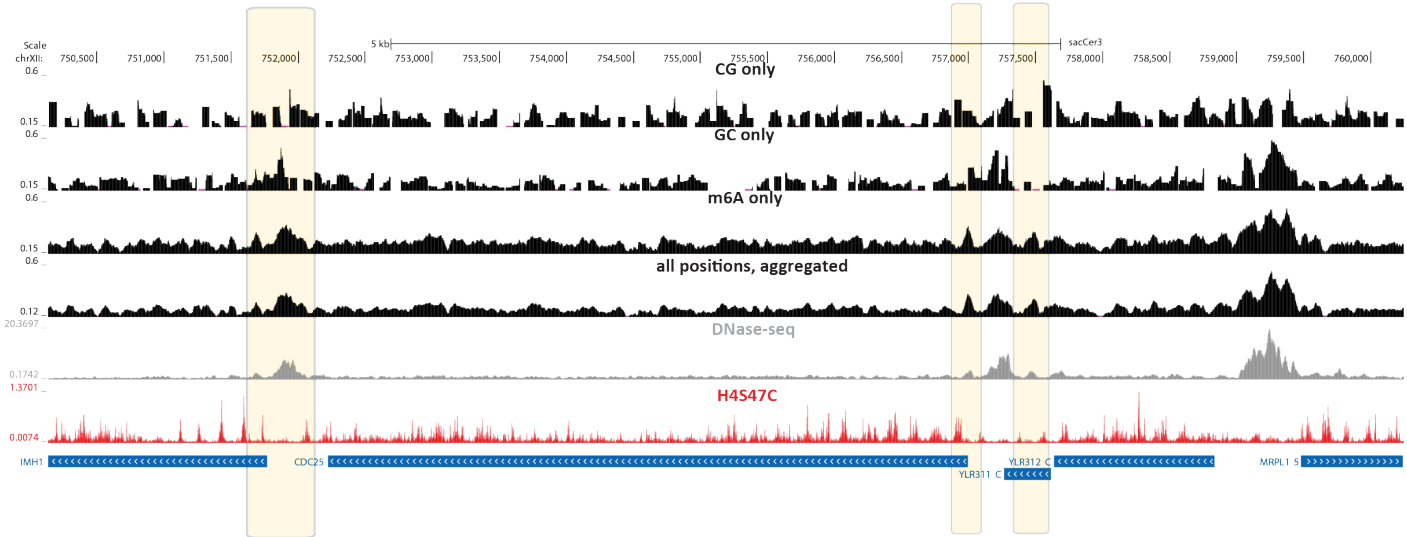
Supplementary Figure 42: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



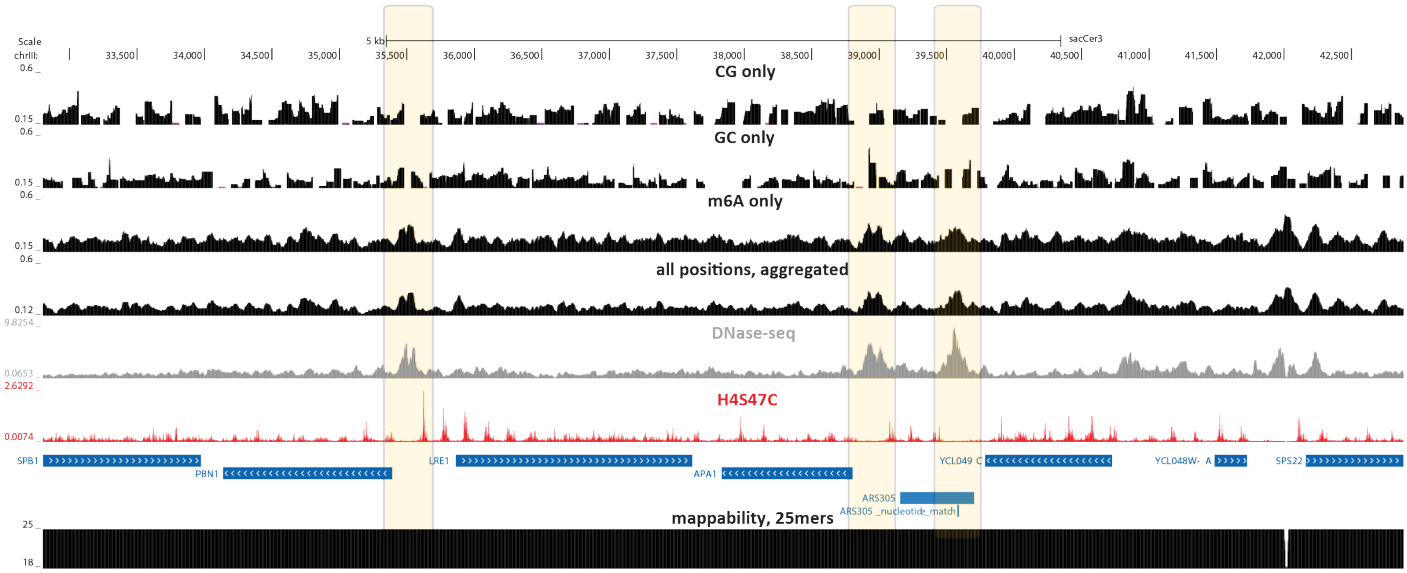
Supplementary Figure 43: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



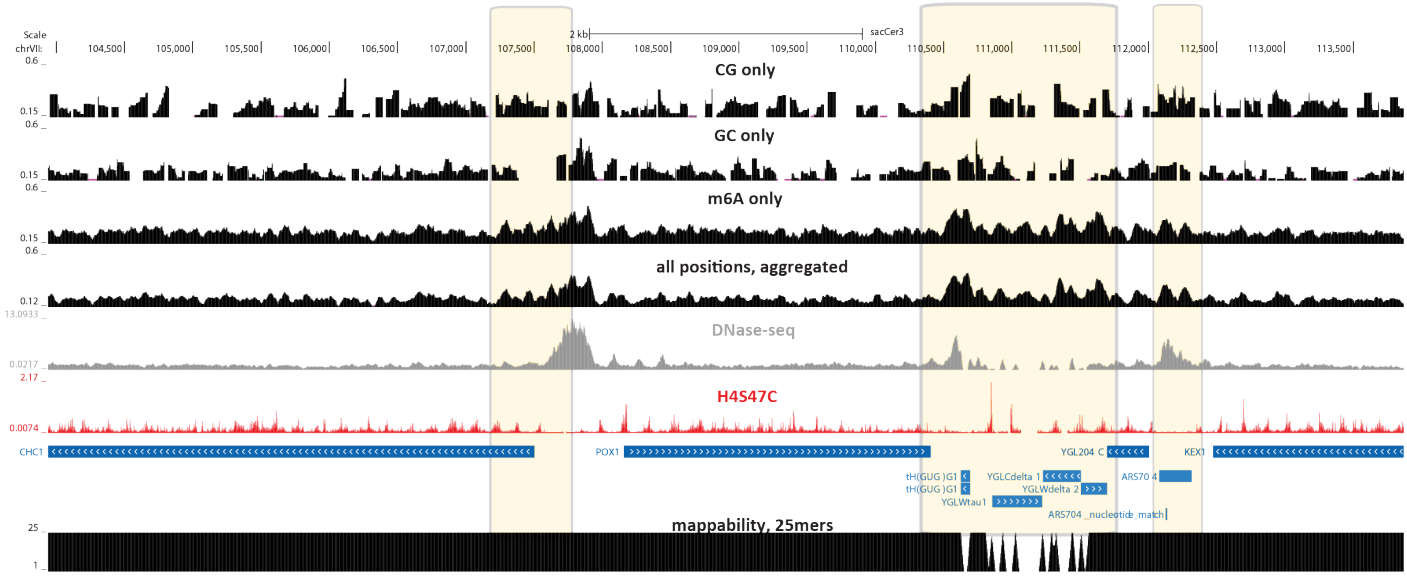
Supplementary Figure 44: Impact of the addition of m^6A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^6A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



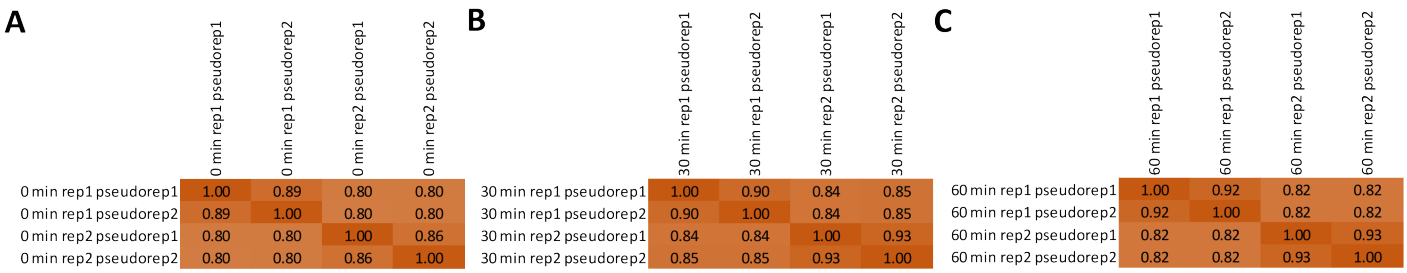
Supplementary Figure 45: Impact of the addition of m^6A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^6A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



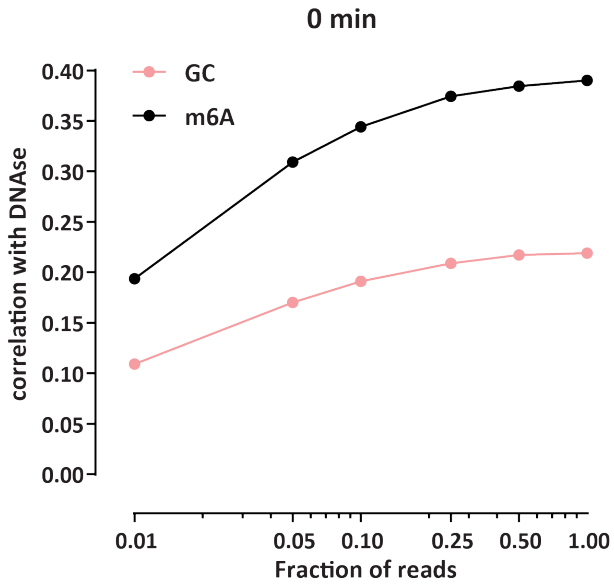
Supplementary Figure 46: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



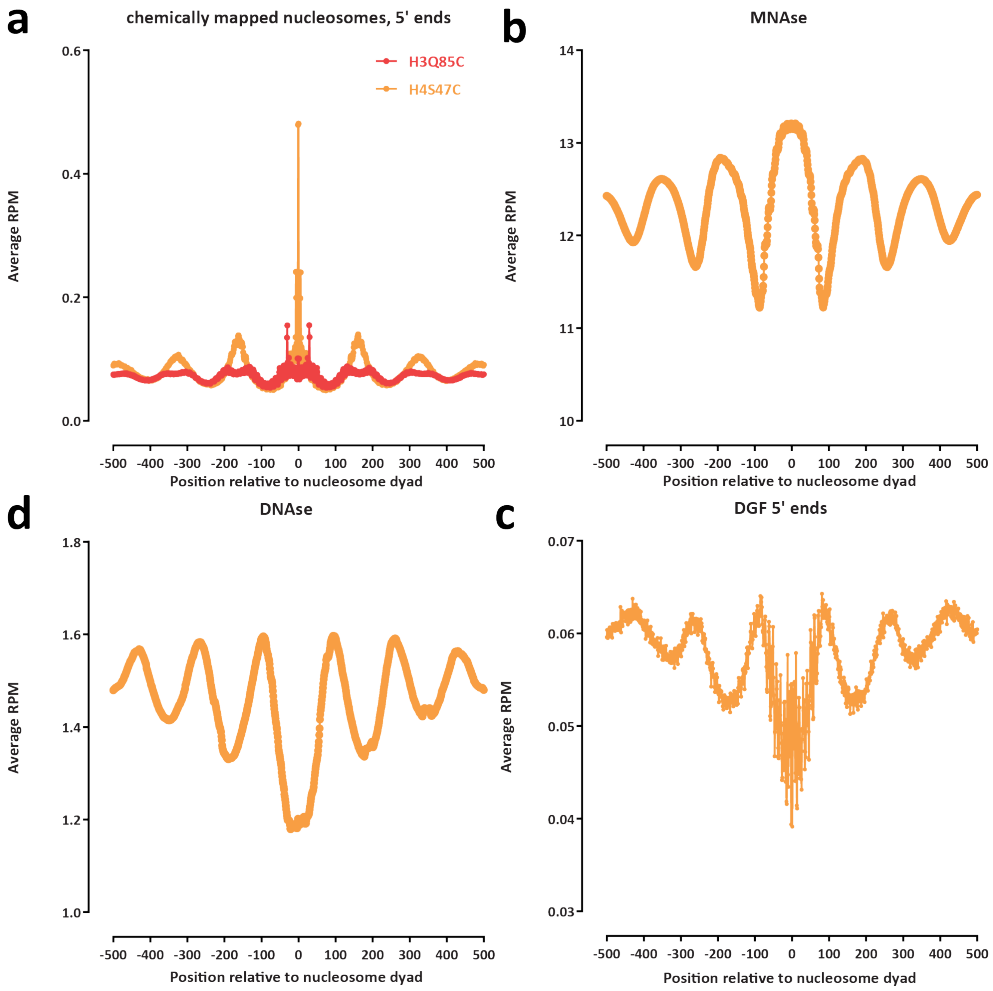
Supplementary Figure 47: Impact of the addition of m^{6}A on assay resolution. Shown is the average accessibility status measured by CG, GC or m^{6}A modifications alone and with all positions aggregated together (calculated over 50bp windows, with a step size of 5, as in Figure 1). The sparseness of CG and GC dinucleotides in the genome results in numerous positions where data is completely missing and to low resolution not allowing the identification of numerous positioned nucleosomes and even entire accessibility peaks.



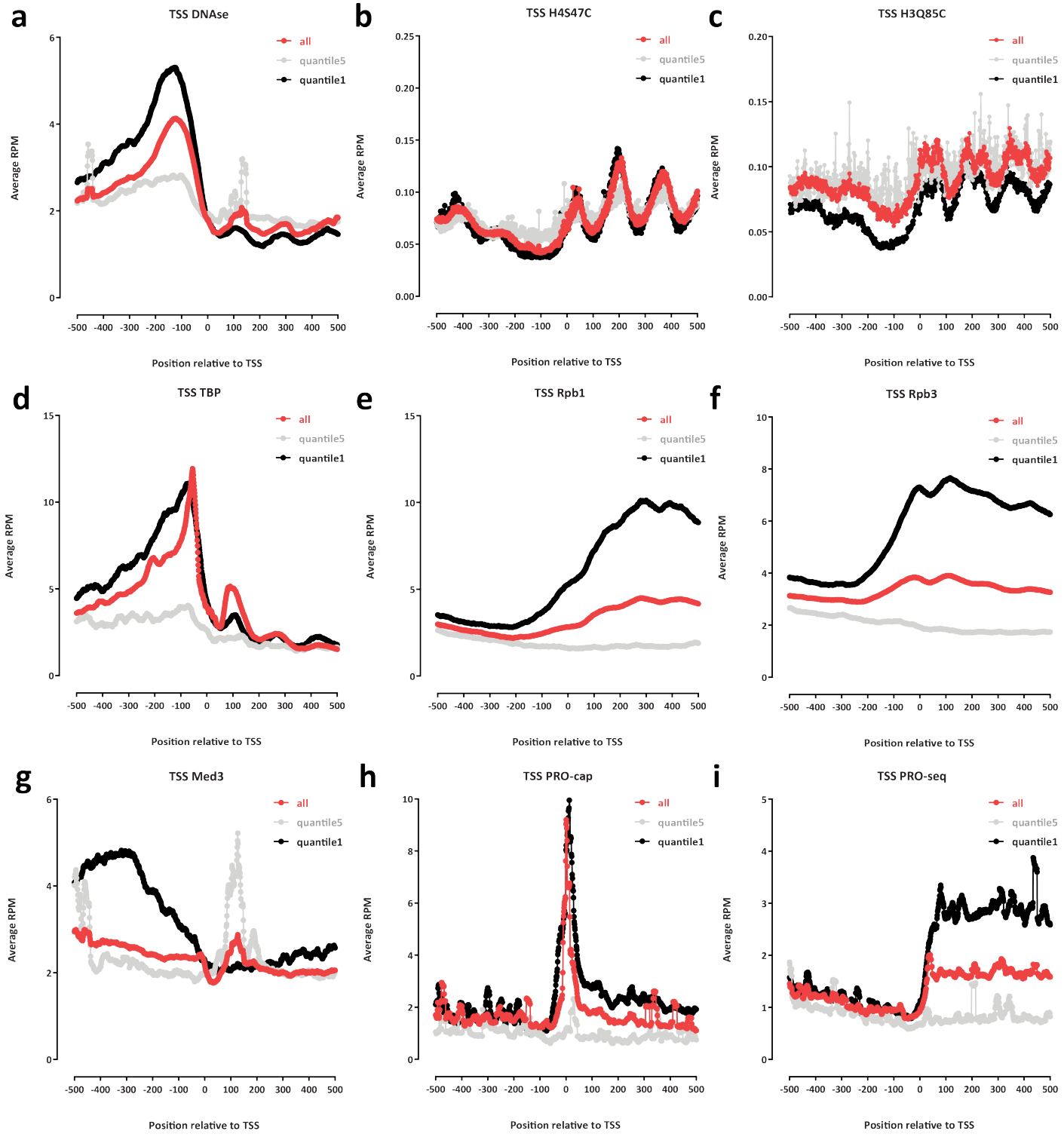
Supplementary Figure 48: Correlation at the base pair level between replicates and pseudoreplicates. Shown are the Pearson r^2 values for the average methylation calls for each position in the yeast genome between pseudoreplicates (generated by randomly splitting reads in two halves) of the same and different biological replicates (tracks generated as shown in Figure 1).



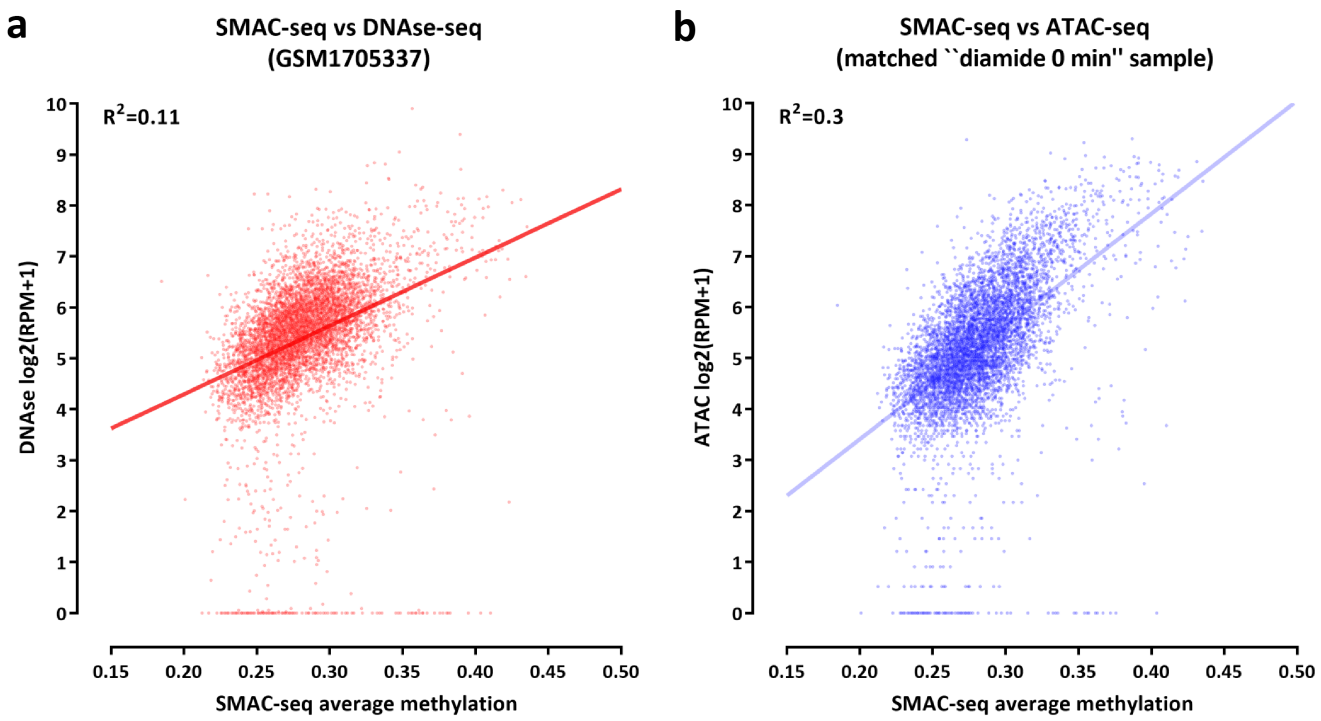
Supplementary Figure 49: Impact of the addition of m^{6}A on assay resolution. Shown are Spearman correlation values between average methylation calls and smoothed (over 10bp) DNAse-seq tracks for each position in the yeast genome (without filtering out positions that are not uniquely mappable) for different subsamplings of SMAC-seq reads. Due to the sparseness of GC dinucleotides in the genome results, using GC methylation alone captures the accessibility signal much more poorly than what is enabled by the dense coverage provided by m^{6}A, as also shown above in Supplementary Figures 19–47



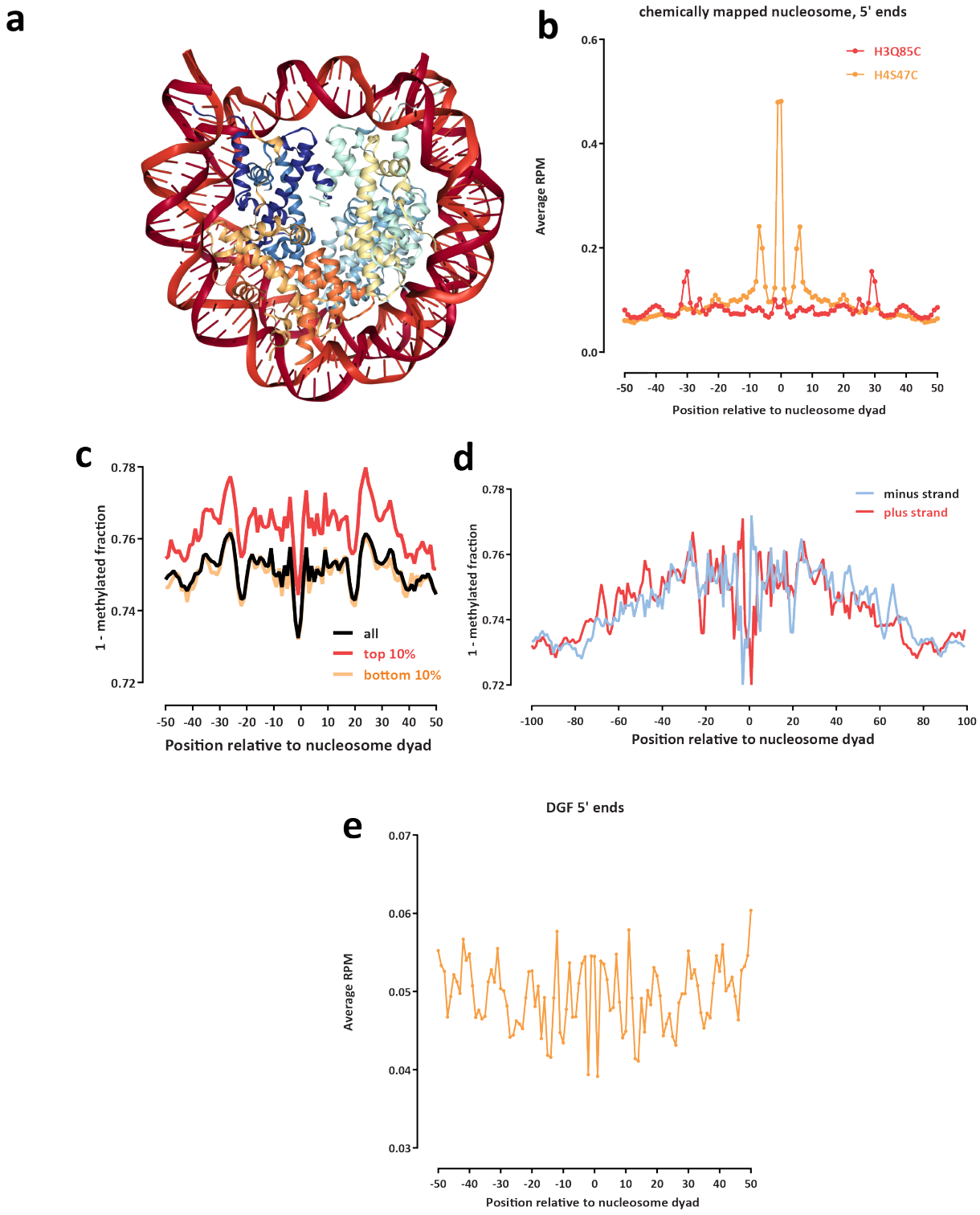
Supplementary Figure 50: Correlation of SMAC-seq (Figure 1e) with other measures of chromatin structure around the dyad centers of positioned nucleosomes in the *S. cerevisiae* genome. (a) H4S47C and H3Q85C nucleosome chemical mapping; (b) MNaseq-seq; (c) DNase-seq; (d) Digital Genomic Footprinting (DGF, 5′ ends of deeply sequenced DNAseq-seq data).



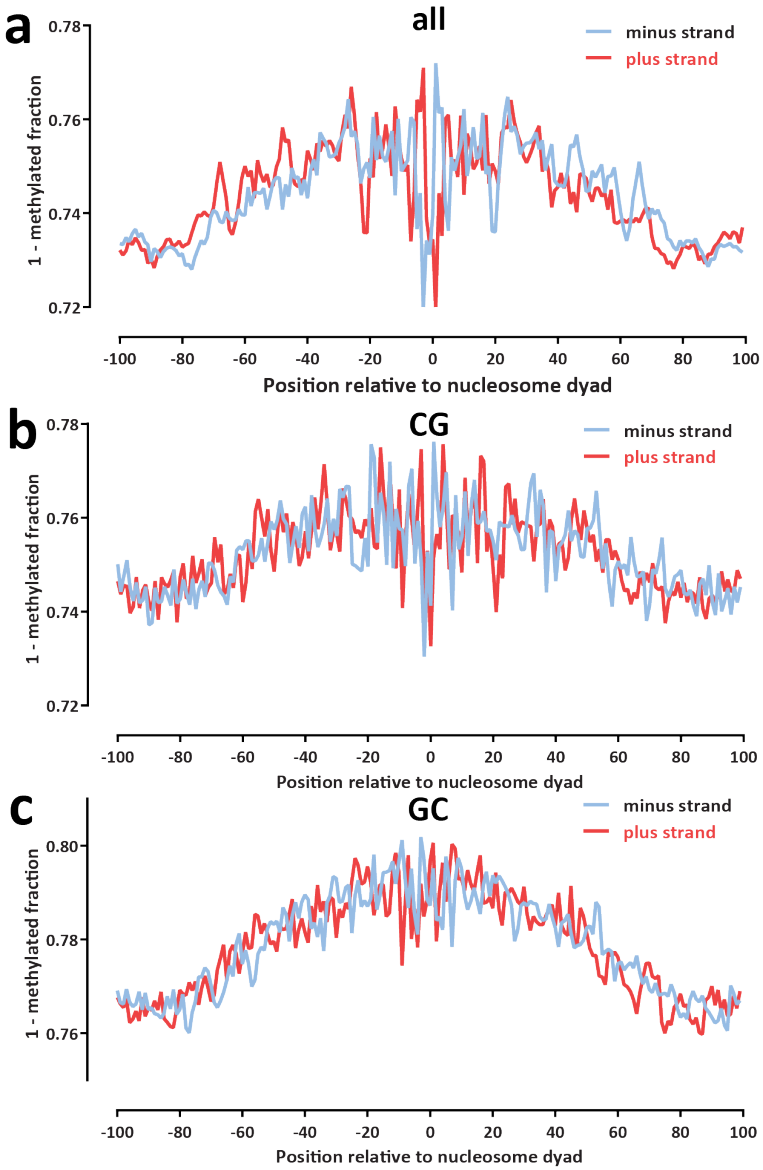
Supplementary Figure 51: Correlation of SMAC-seq (Figure 1f and g) with other types of functional genomic measurements of chromatin structure, protein occupancy and transcriptional activity around TSSs. Shown is average coverage over all *S. cerevisiae* genes, for the most highly expressed 20% of genes ("quantile1"), and for the bottom 20% of genes ("quantile5"). (a) DNase-seq; (b) H4S47C nucleosome chemical mapping; (c) H3Q85C nucleosome chemical mapping; (d) TBP ChIP-seq; (e) Rpb1 ChIP-seq; (f) Rpb3 ChIP-seq; (g) Med3 ChIP-seq; (h) PRO-cap; (i) PRO-seq.



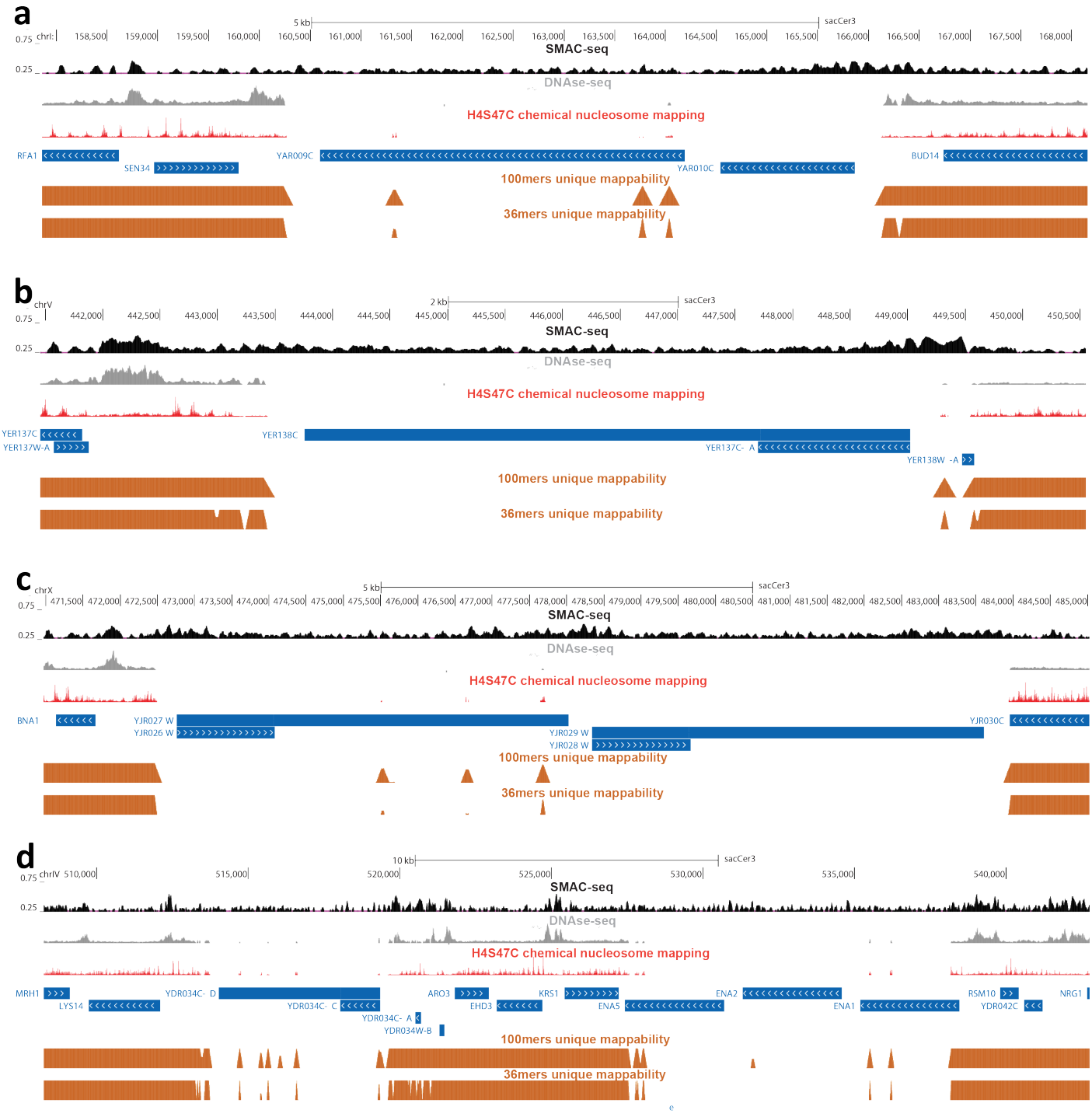
Supplementary Figure 52: Correlation of SMAC-seq signal with ATAC-seq and DNA-seq signal over yeast promoters. Shown is the average methylation over the TSS ± 200 bp for SMAC-seq and RPM (Reads Per Million mapped reads) values for DNase-seq (a) and ATAC-seq (b). Note that the DNase-seq dataset is obtained from an external study while the SMAC-seq and ATAC-seq ones are from the same sample ("diamide 0 min rep1").



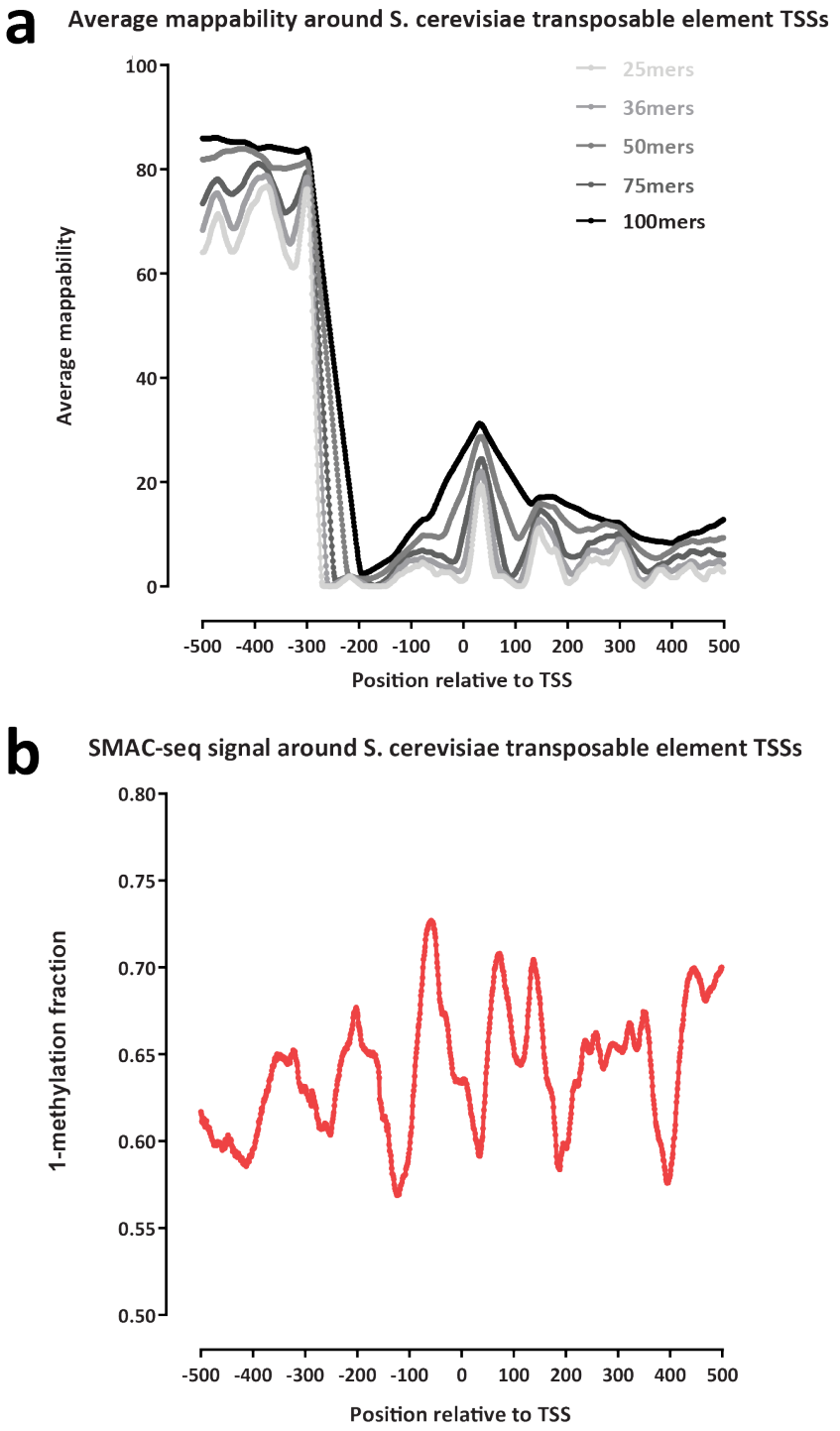
Supplementary Figure 53: SMAC-seq data displays higher methylation propensity in more exposed parts of the nucleosomal particle. (a) Structure of the eukaryotic nucleosome; (b) High-resolution (50-bp radius) view of chemical nucleosome mapping data relative to nucleosome dyads; (c) High-resolution (50-bp radius) view of SMAC-seq data relative to nucleosome dyads; (d) Strand-specific (100-bp radius) view of SMAC-seq data relative to nucleosome dyads; (e) High-resolution (50-bp radius) view of DGF cleavage profiles relative to nucleosome dyads.



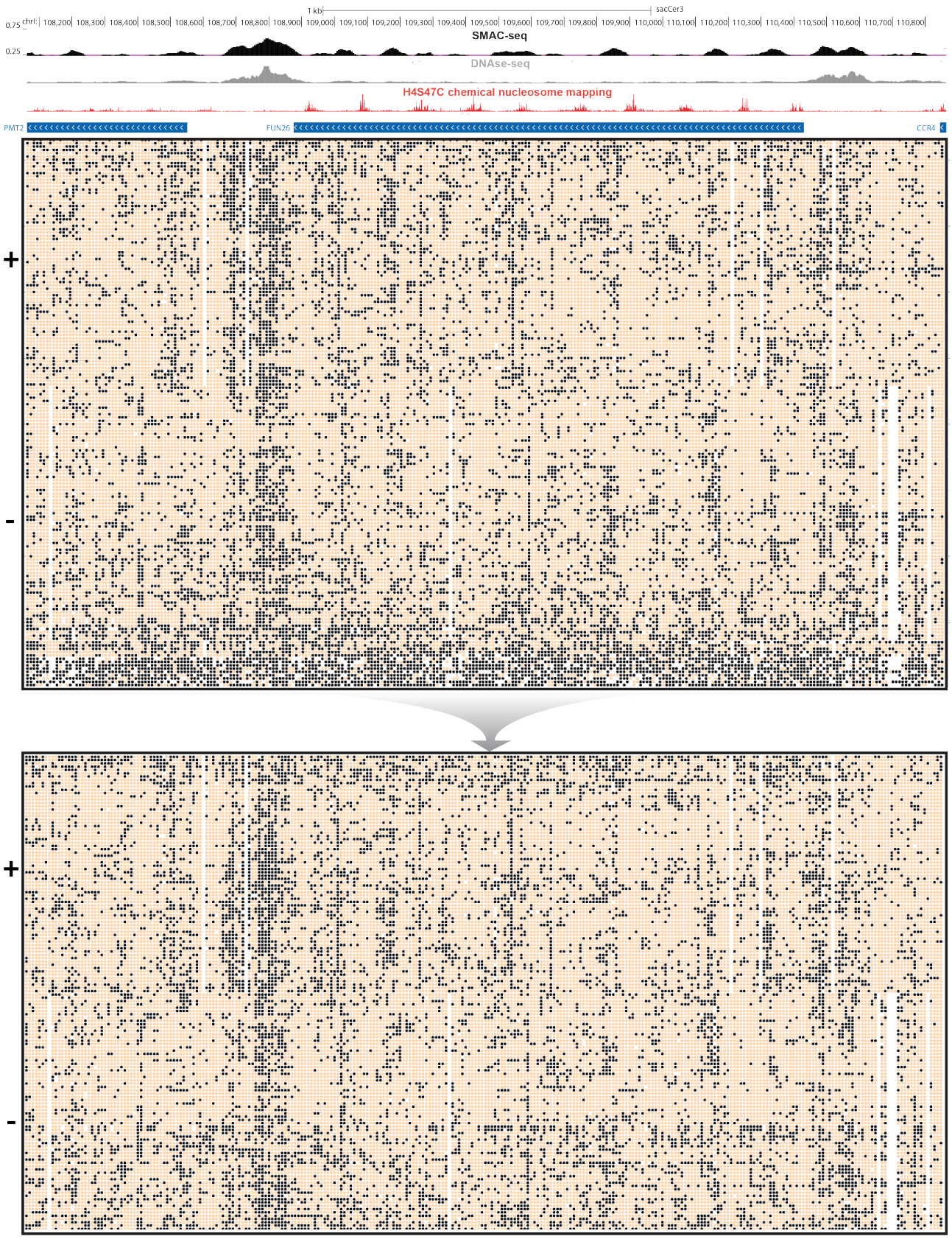
Supplementary Figure 54: Strand-specific nucleosome accessibility/occupancy is most clearly revealed by m^{6}A methylation using EcoGII. (a) Strand-specific (100-bp radius) view of SMAC-seq data relative to nucleosome dyads using all positions; (b) Strand-specific (100-bp radius) view of SMAC-seq data relative to nucleosome dyads using CG positions; (c) Strand-specific (100-bp radius) view of SMAC-seq data relative to nucleosome dyads using GC positions.



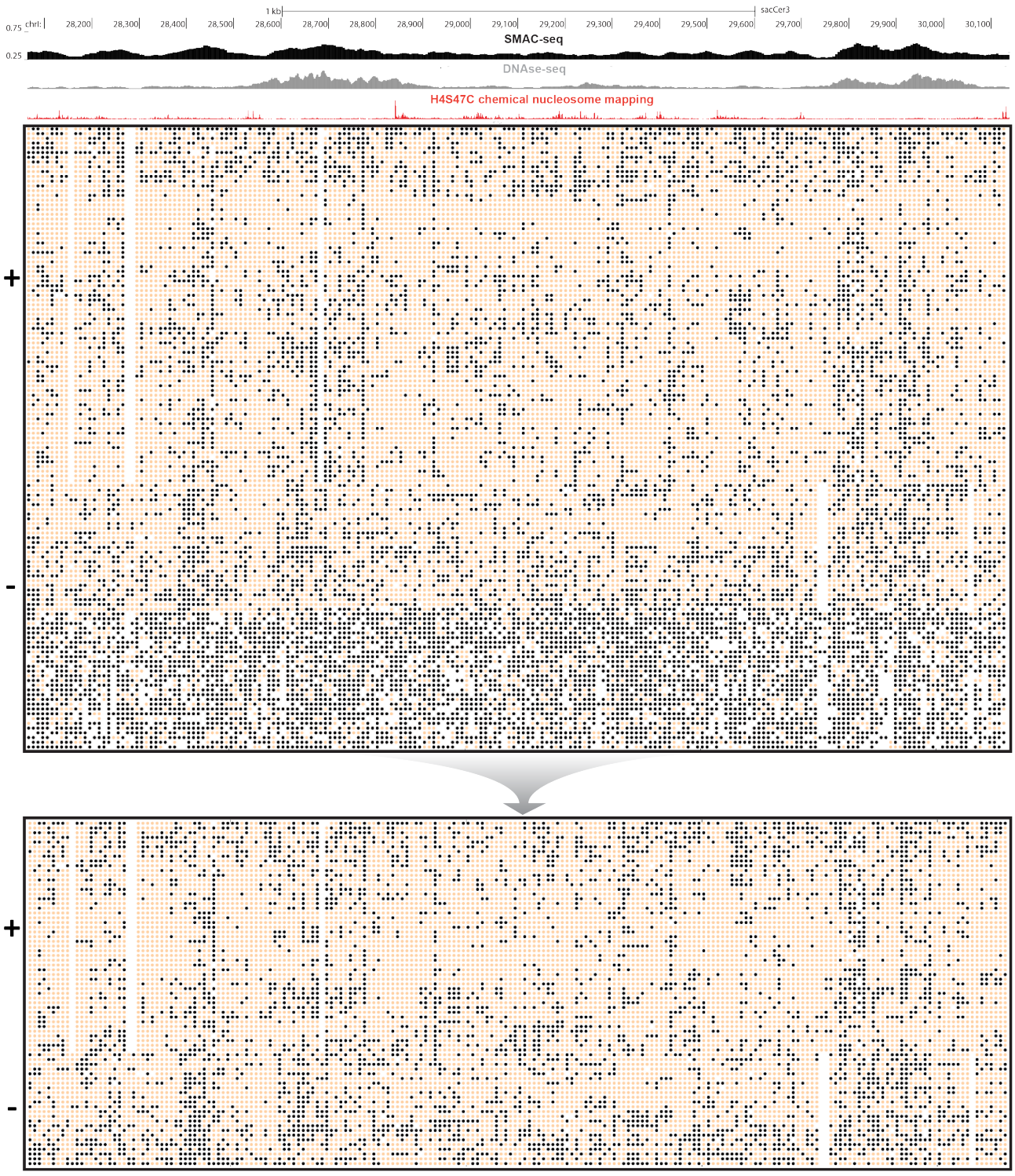
Supplementary Figure 55: SMAC-seq provides coverage of areas of the genome that cannot be uniquely mapped using short reads. Four different repetitive regions are shown (a, b, c and d) together with short read coverage for DNase-seq and chemical nucleosome mapping data and unique mappability tracks for 36mers and 100mers.



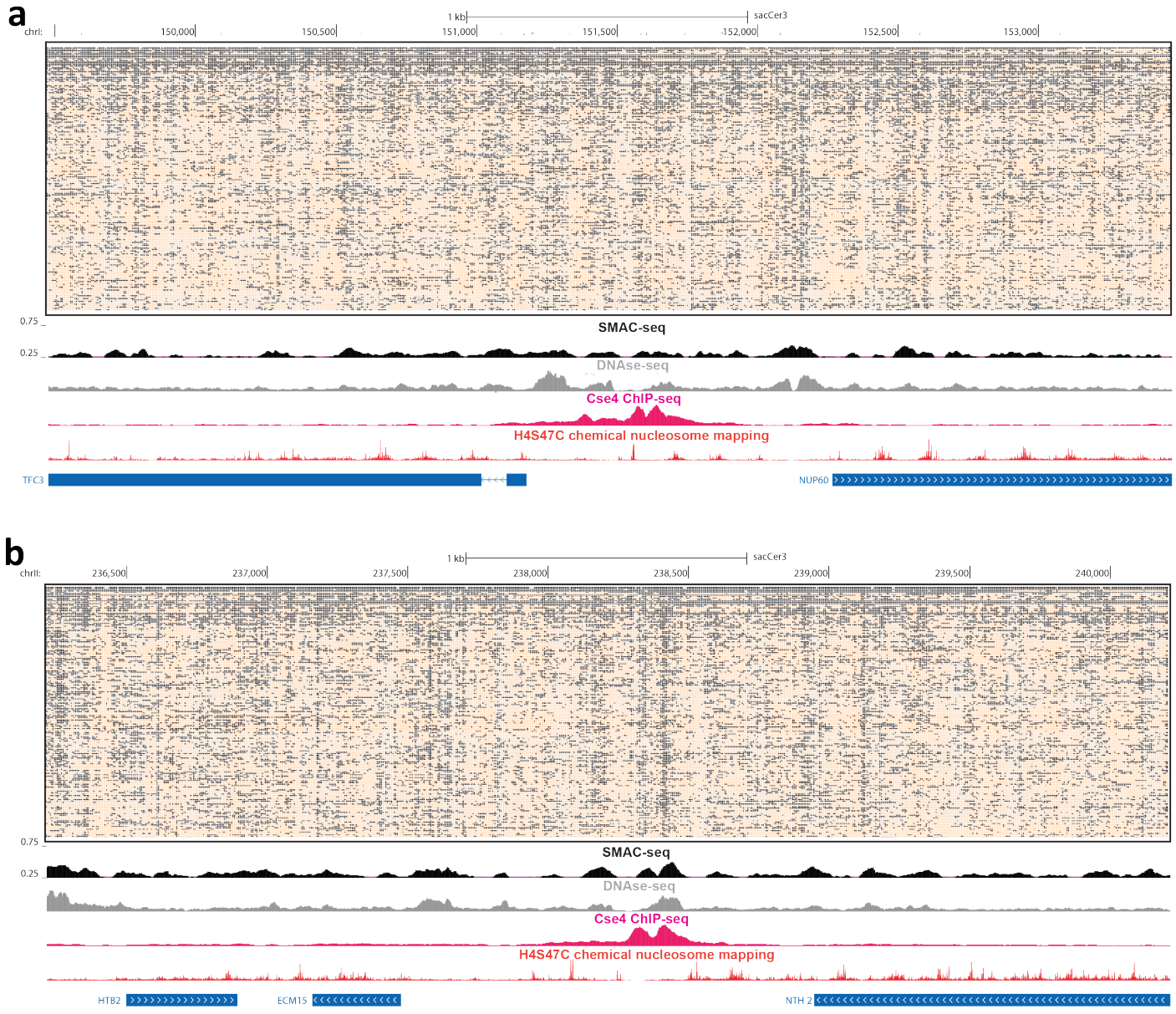
Supplementary Figure 56: SMAC-seq provides coverage of areas of the genome that cannot be uniquely mapped using short reads. (a) Average short-read mappability around TSSs of annotated transposable elements in the *S. cerevisiae* genome; (b) SMAC-seq signal around TSSs of annotated transposable elements in the *S. cerevisiae* genome.



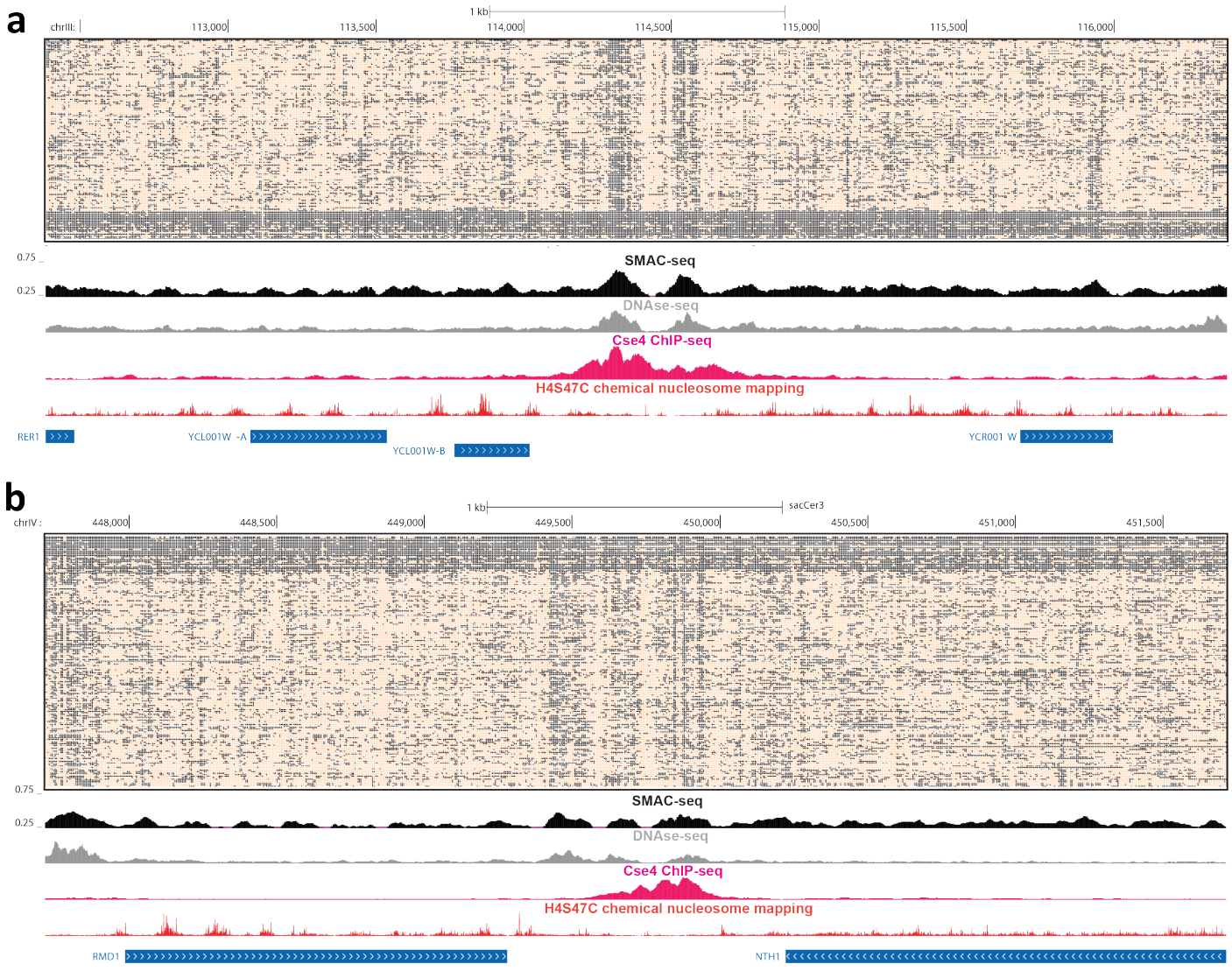
Supplementary Figure 57: Removal of potentially artifactual high-methylation reads. Shown is unfiltered SMAC-seq data and the same locus after removal of all reads containing a 1-kb ≥75% methylated stretch (average accessibility over 10 bp).



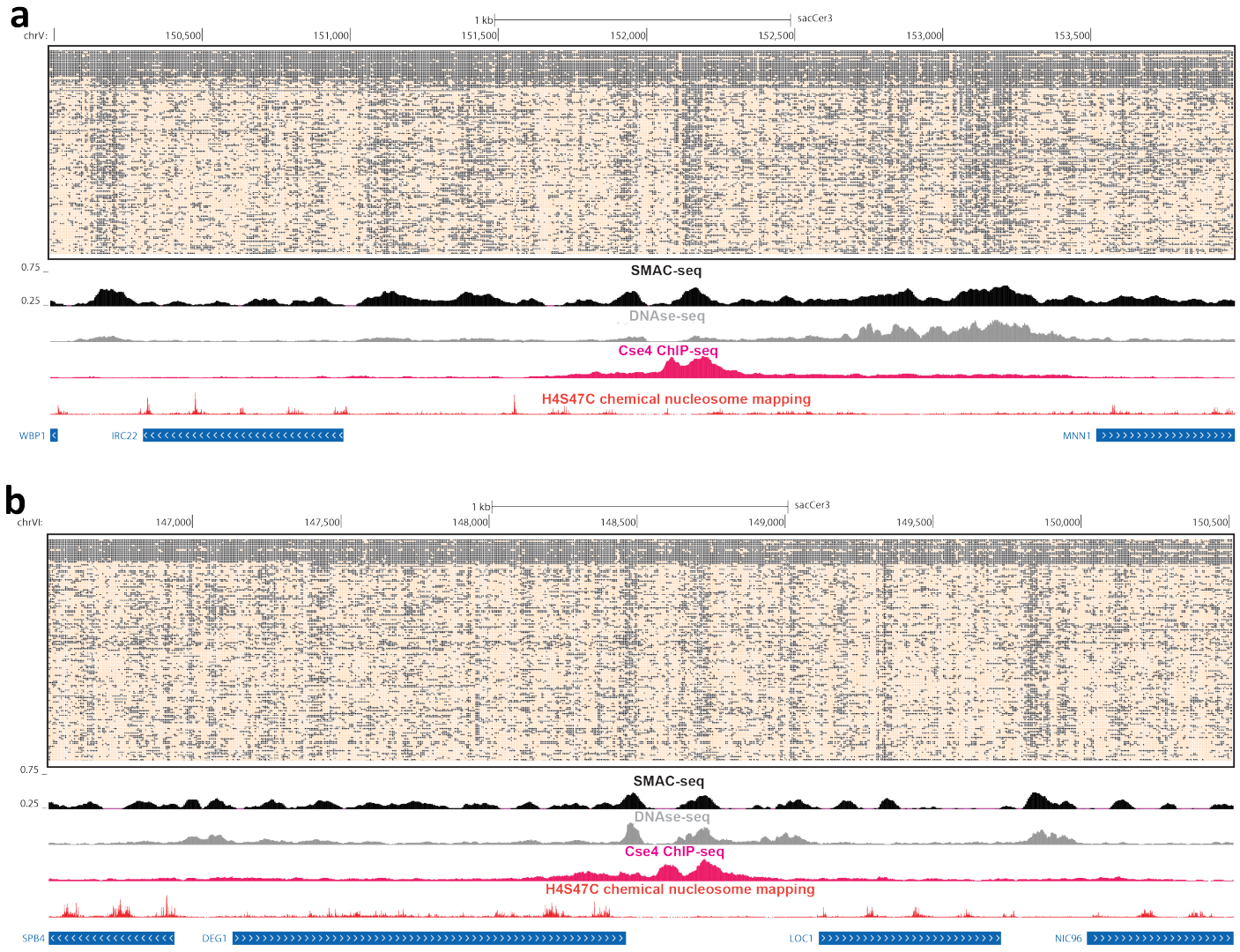
Supplementary Figure 58: Removal of potentially artifactual high-methylation reads. Shown is unfiltered SMAC-seq data and the same locus after removal of all reads containing a 1-kb ≥75% methylated stretch (average accessibility over 10 bp).



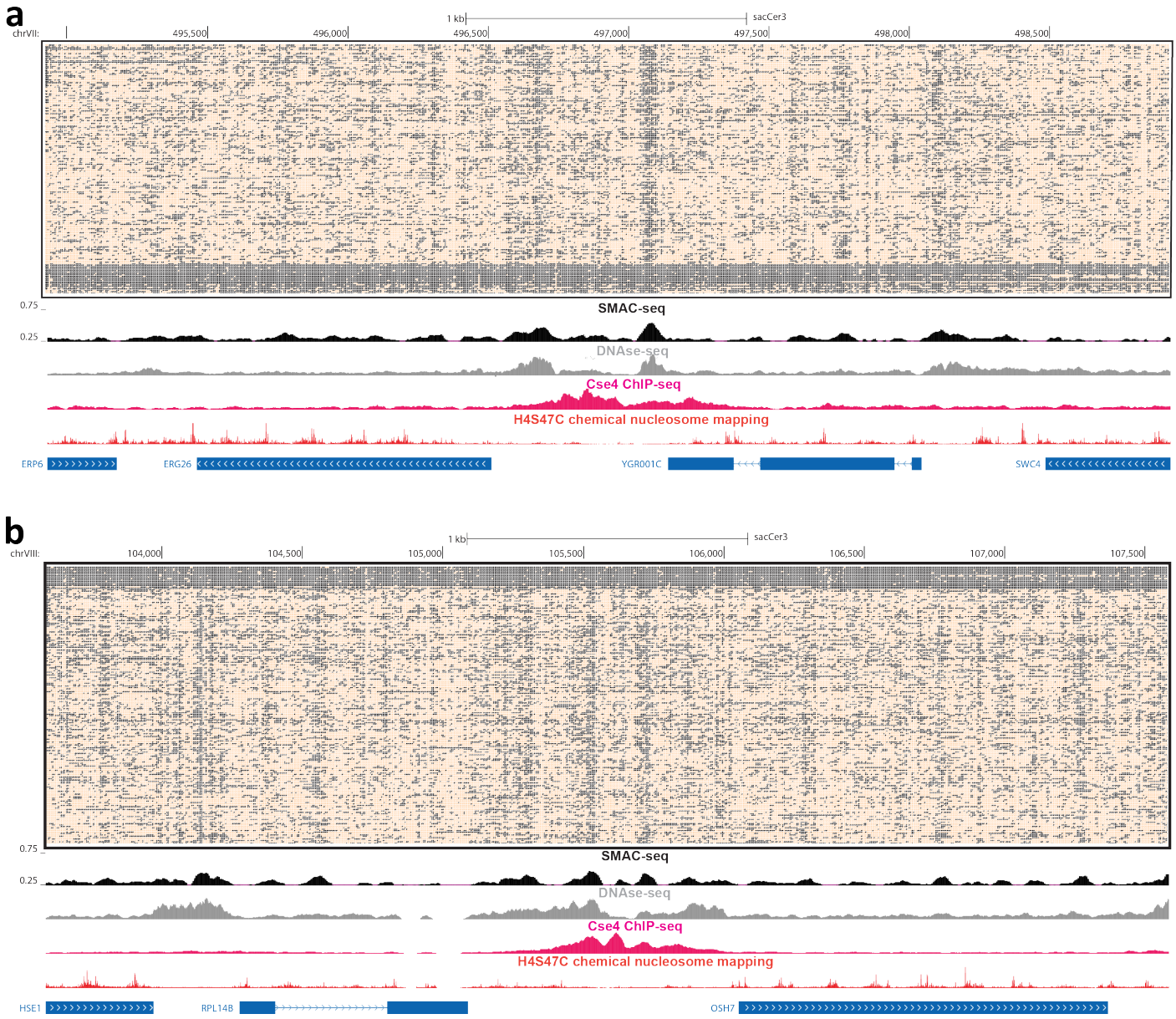
Supplementary Figure 59: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrI; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrII. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



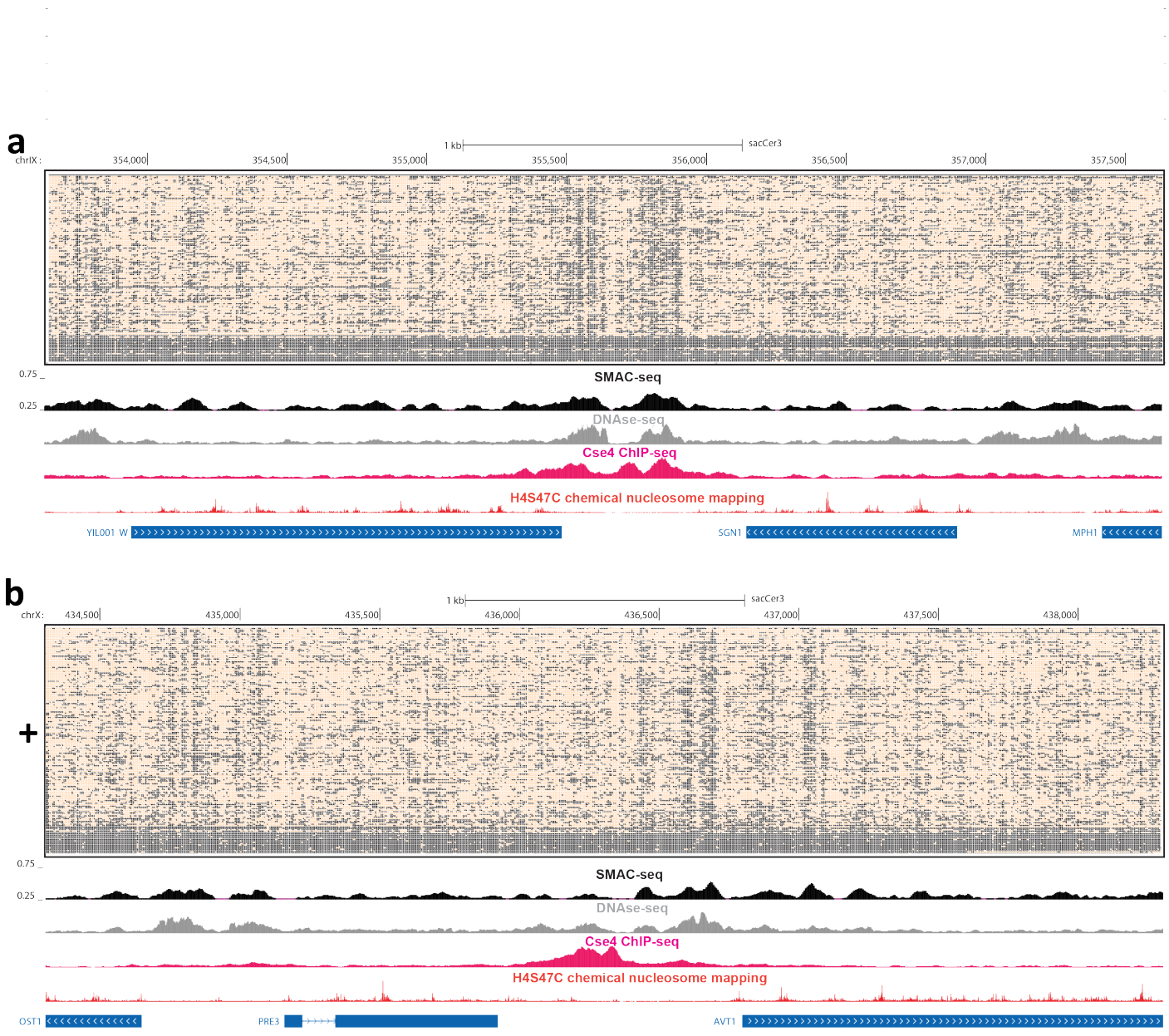
Supplementary Figure 60: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrIII; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrIV. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



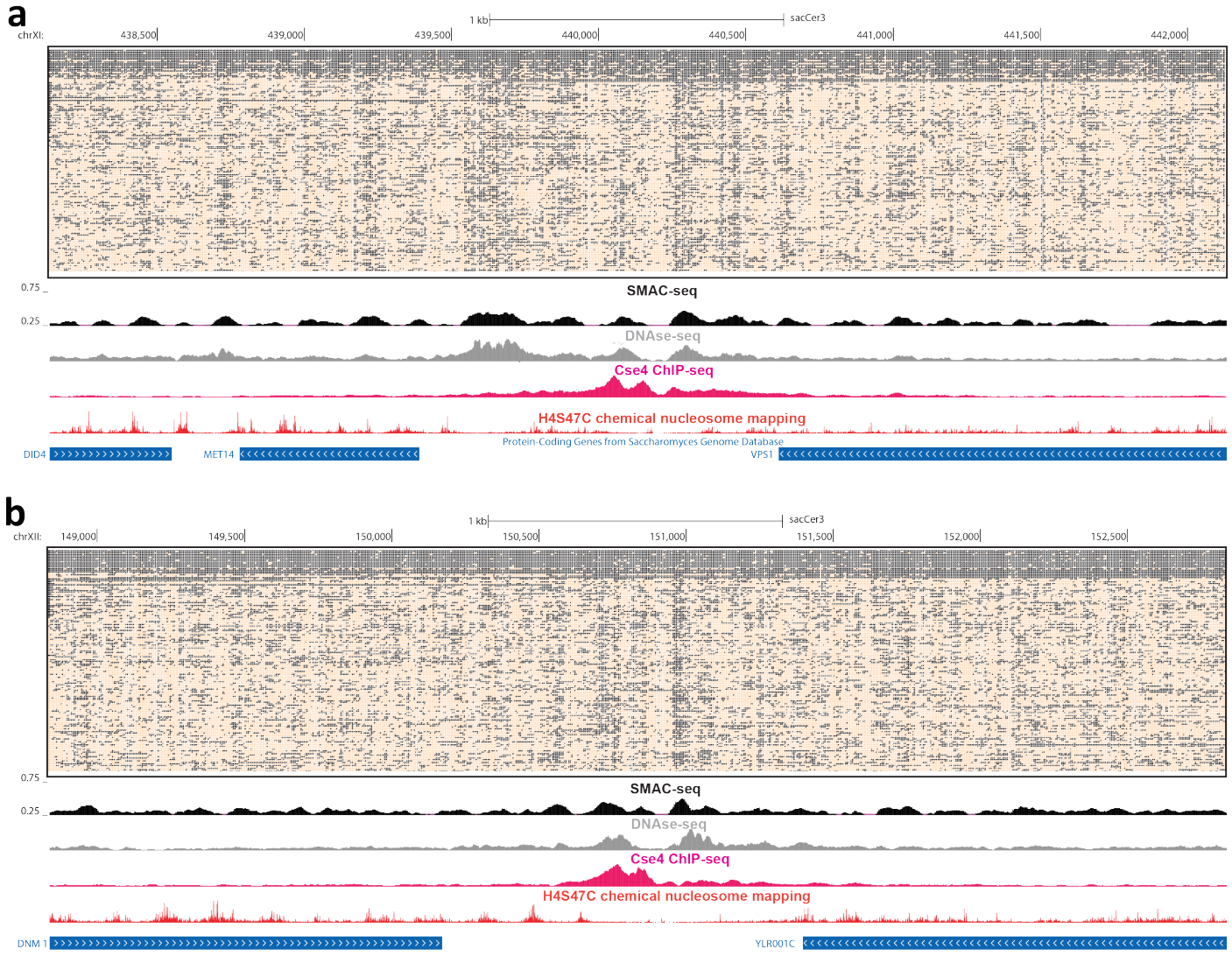
Supplementary Figure 61: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrV; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrVI. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



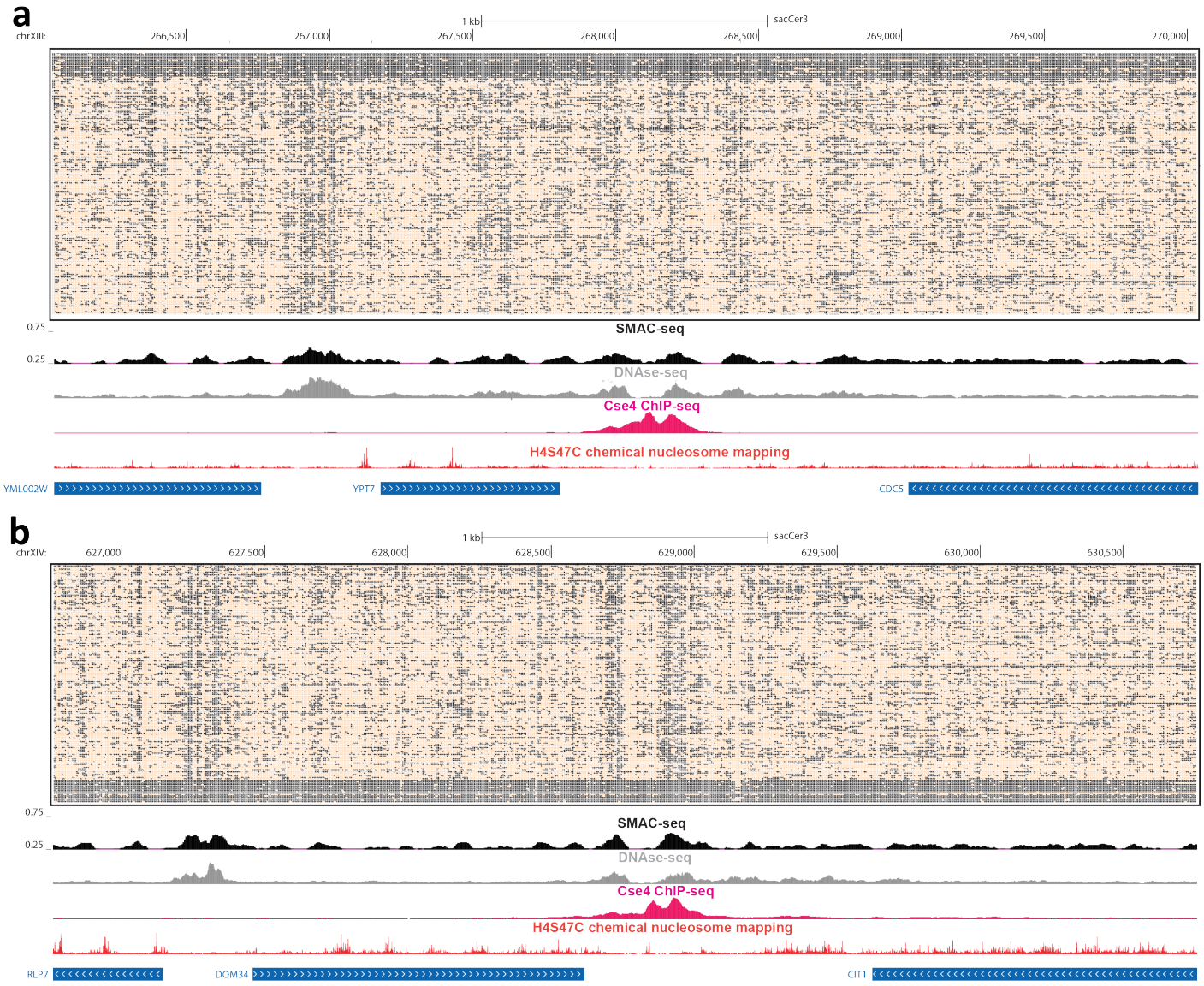
Supplementary Figure 62: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrVII; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrVIII. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



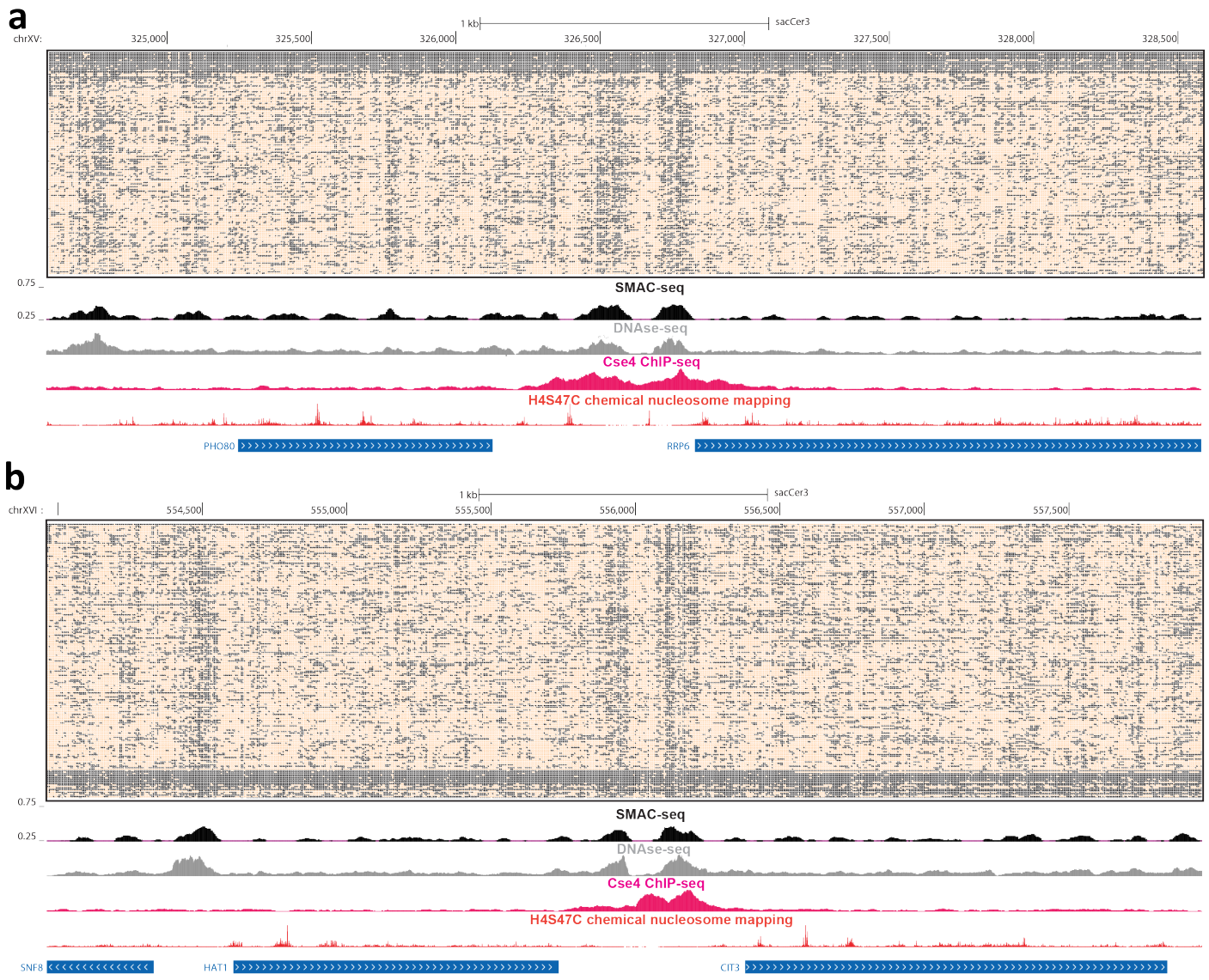
Supplementary Figure 63: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrIX; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrX. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



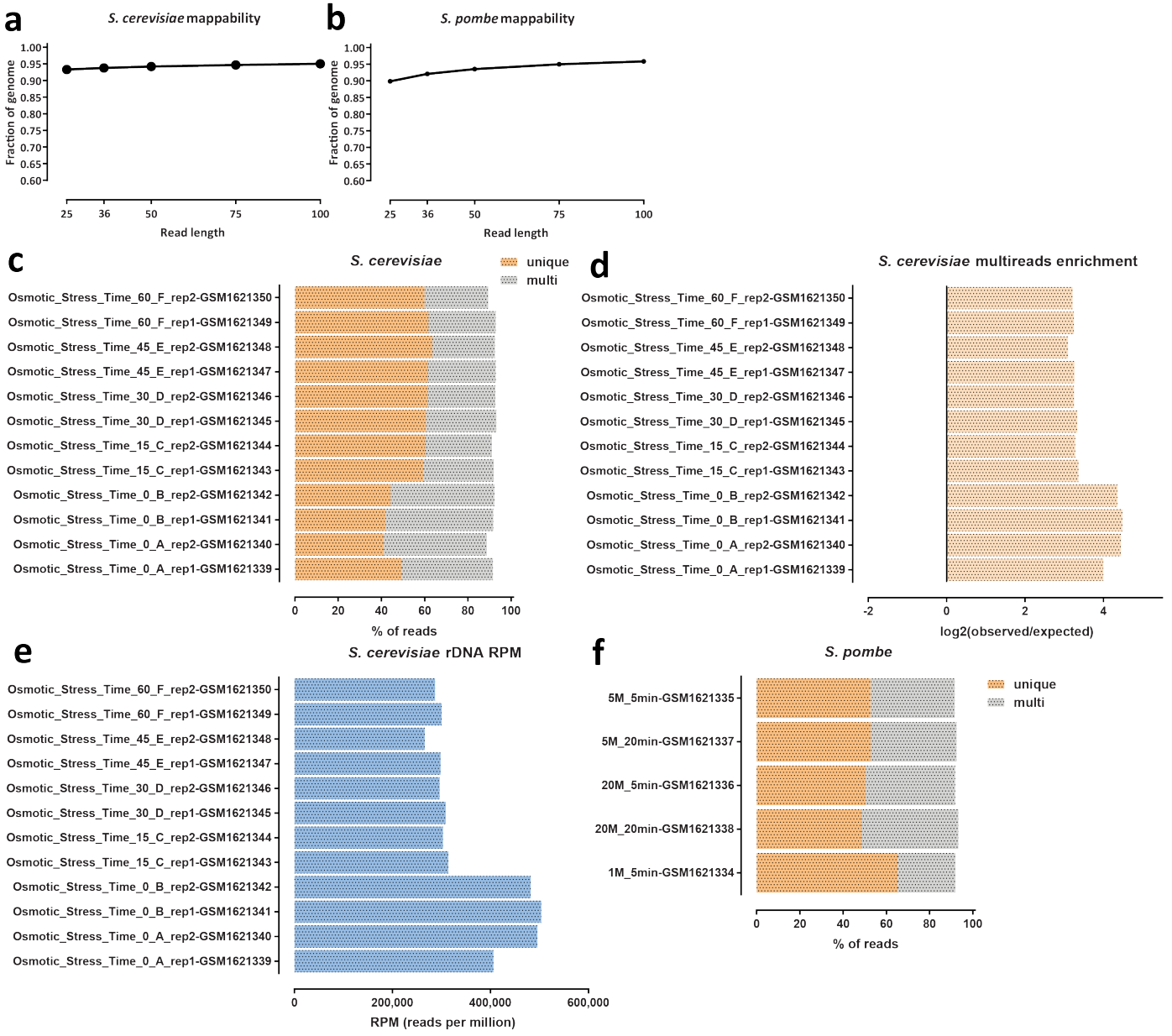
Supplementary Figure 64: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrXI; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrXII. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



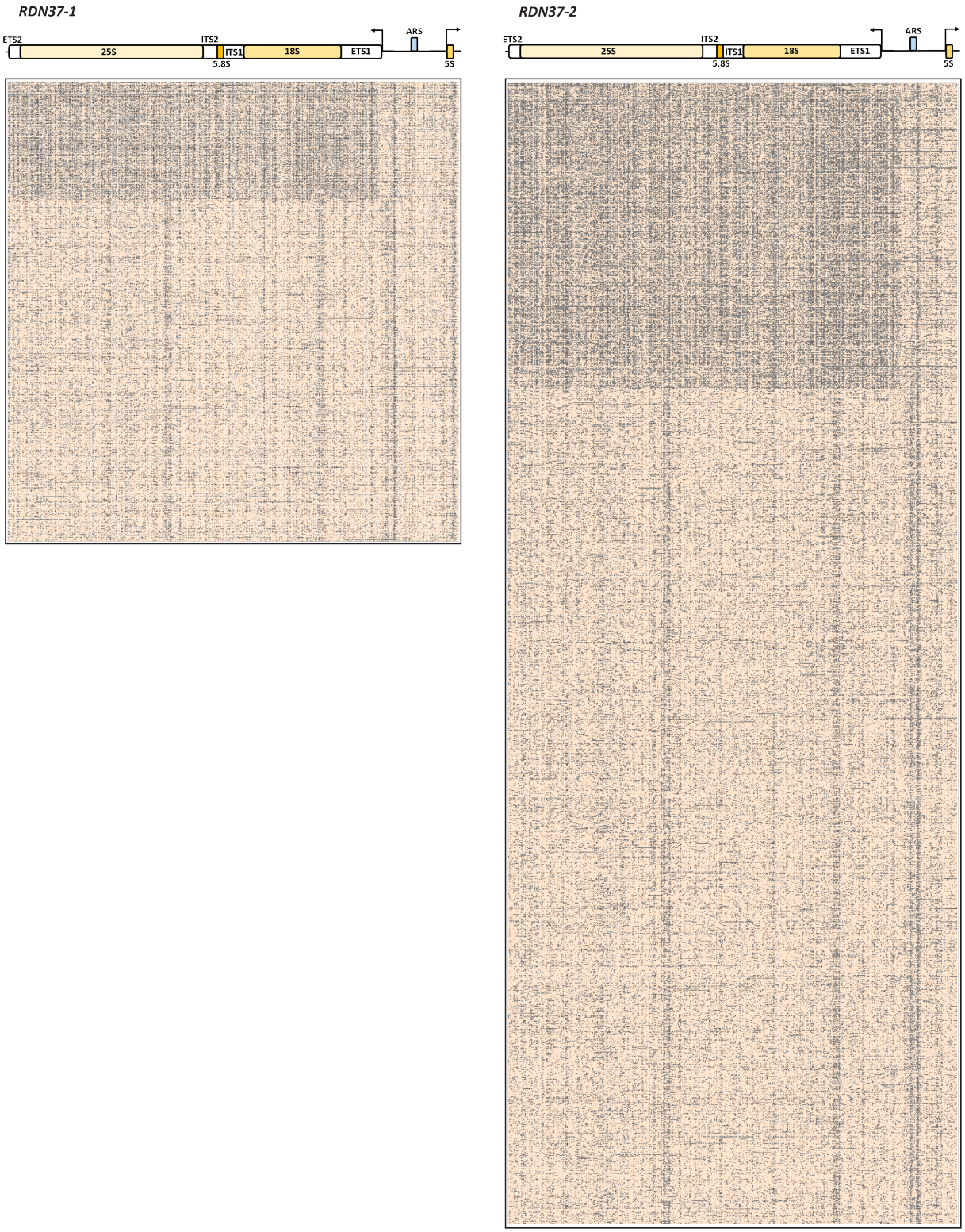
Supplementary Figure 65: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrXIII; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrXIV. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.



Supplementary Figure 66: Single-molecule long-read accessibility around well positioned centromeres. (a) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrXV; (b) Raw unfiltered nanopore reads fully spanning the 4-kilobase neighborhood of the centromere of *S. cerevisiae* chrXVI. In both cases, accessibility is shown at aggregated 10-bp resolution (see Methods section for details) for the single-molecule display, and aggregated over sliding (every 5 bases) 50-bp windows for the genome browser SMAC-seq track.

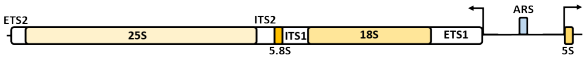


Supplementary Figure 67: Ribosomal DNA arrays are highly enriched for chromatin accessibility as measured by ATAC-seq. (a) Unique read mappability of the *Saccharomyces cerevisiae* genome as a function of read length (b) Unique read mappability of the *Schizosaccharomyces pombe* genome as a function of read length (c and d) Enrichment of multimapping reads in *Saccharomyces cerevisiae* ATAC-seq datasets (e) ATAC-seq multireads are highly enriched for rDNA-mapping reads (f) Enrichment of multimapping reads in *Schizosaccharomyces pombe* ATAC-seq datasets ATAC-seq datasets were obtained from Schep et al. 2015 [53]

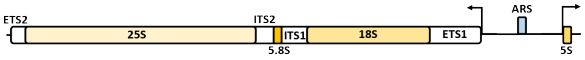


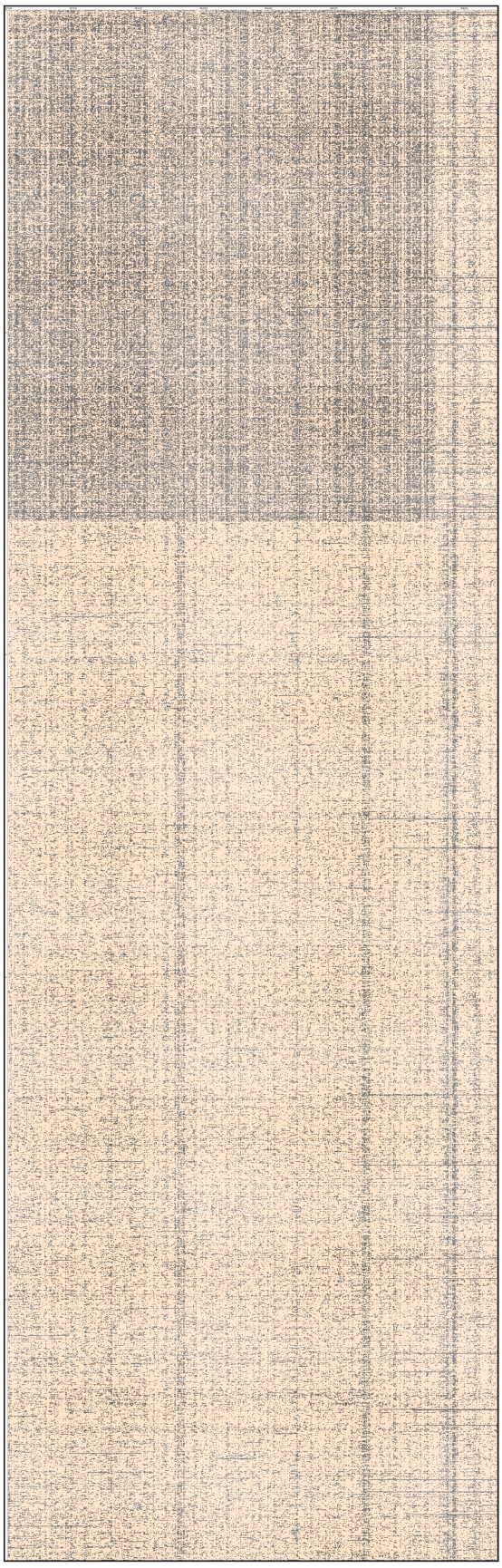
Supplementary Figure 68: SMAC-seq reveals the distribution of alternative chromatin states of rDNA arrays. Shown are all reads covering the *RDN37-1* and *RDN37-2* arrays in the *RDN1* locus in the "diamide 30 min rep1" experiment (unfiltered reads, "aggregate" signal).

RDN37-1

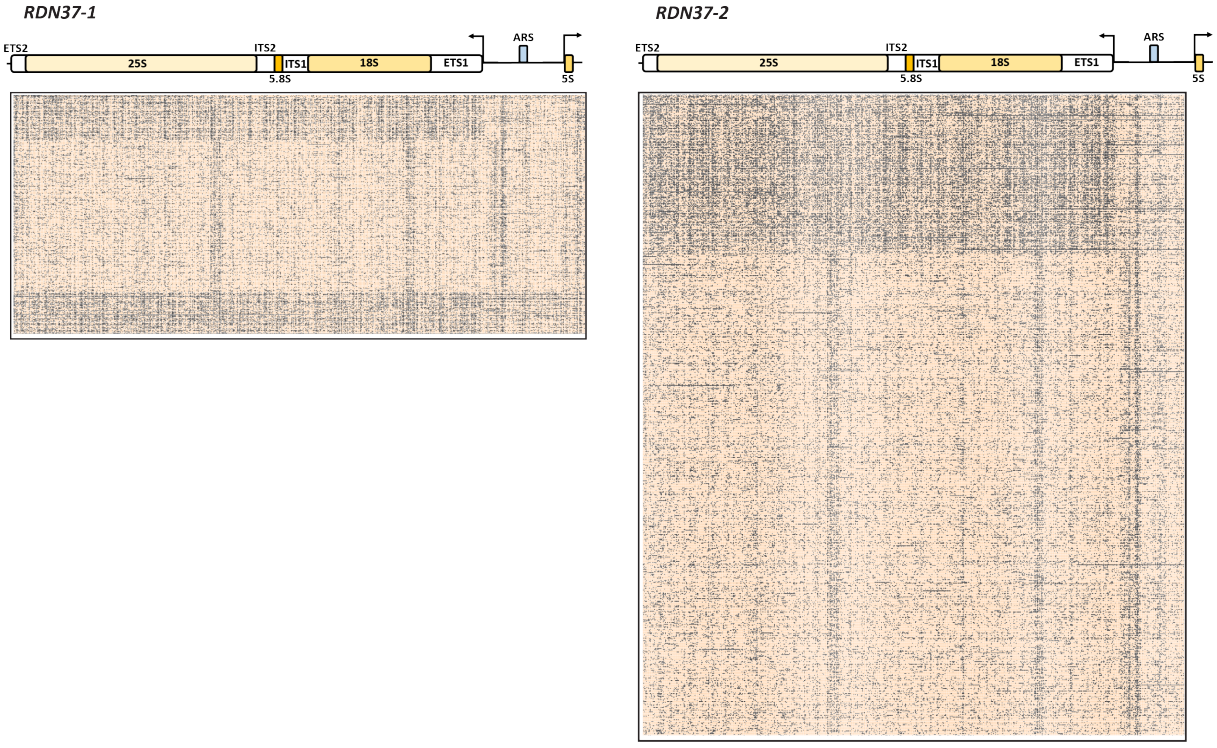


RDN37-2

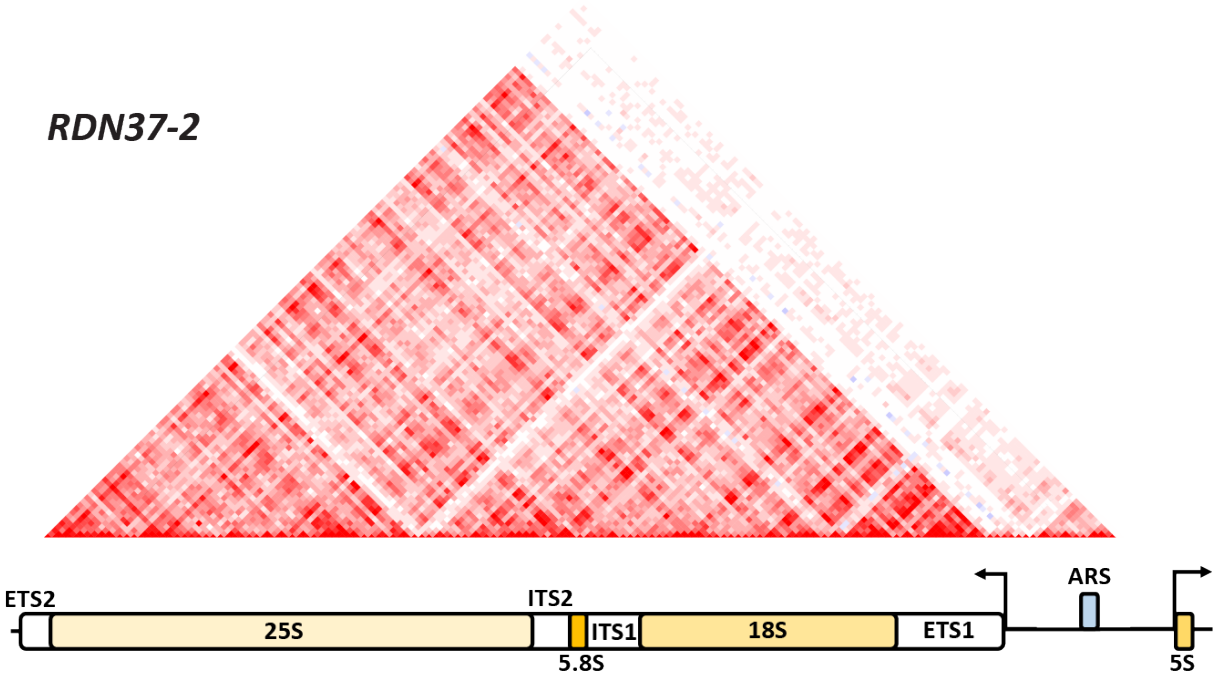




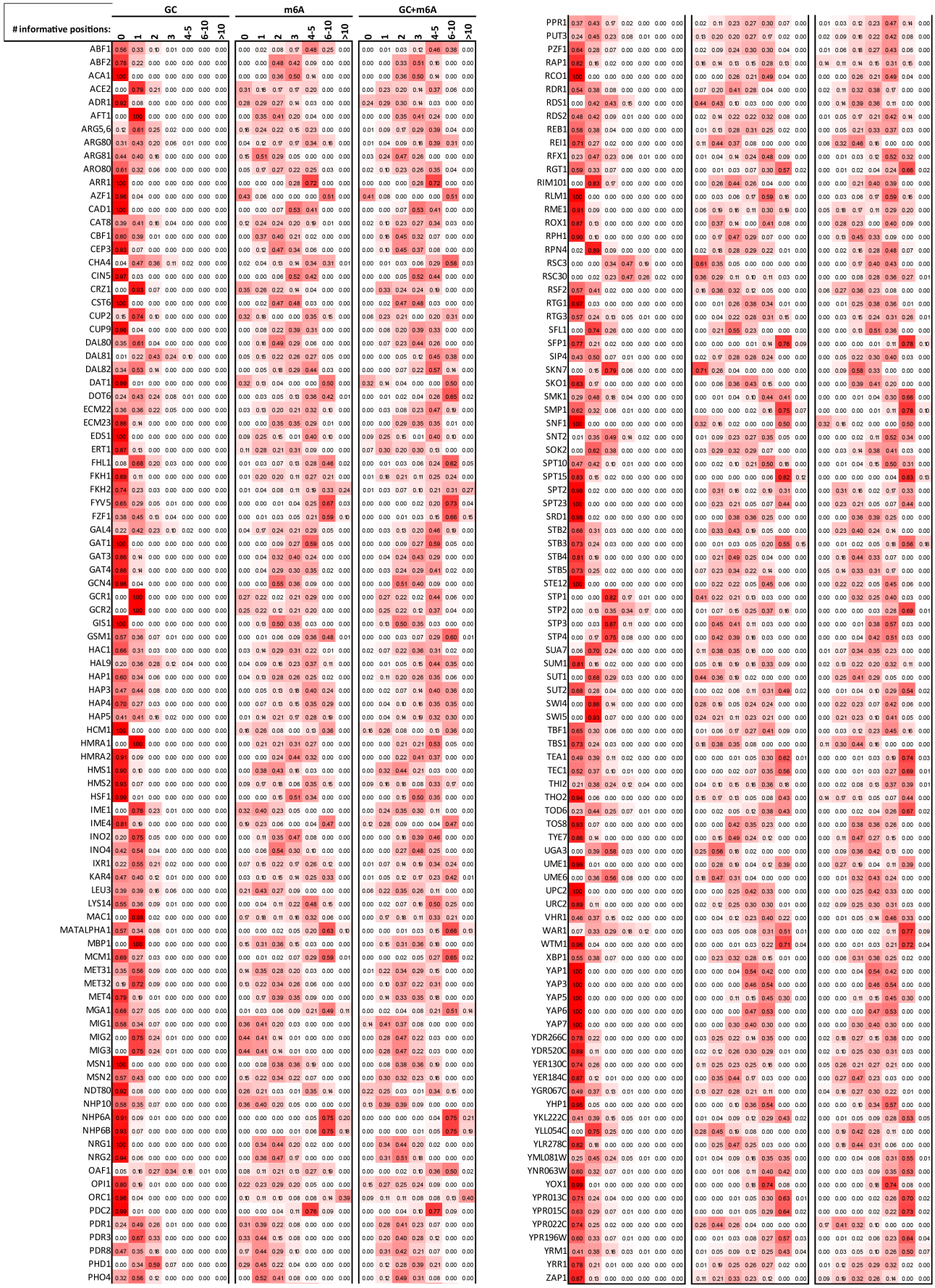
Supplementary Figure 69 ***(preceding page)*****: SMAC-seq reveals the distribution of alternative chromatin states of rDNA arrays. Shown are all reads covering the *RDN37-1* and *RDN37-2* arrays in the *RDN1* locus in the "diamide 60 min rep1" experiment (unfiltered reads, "aggregate" signal).**



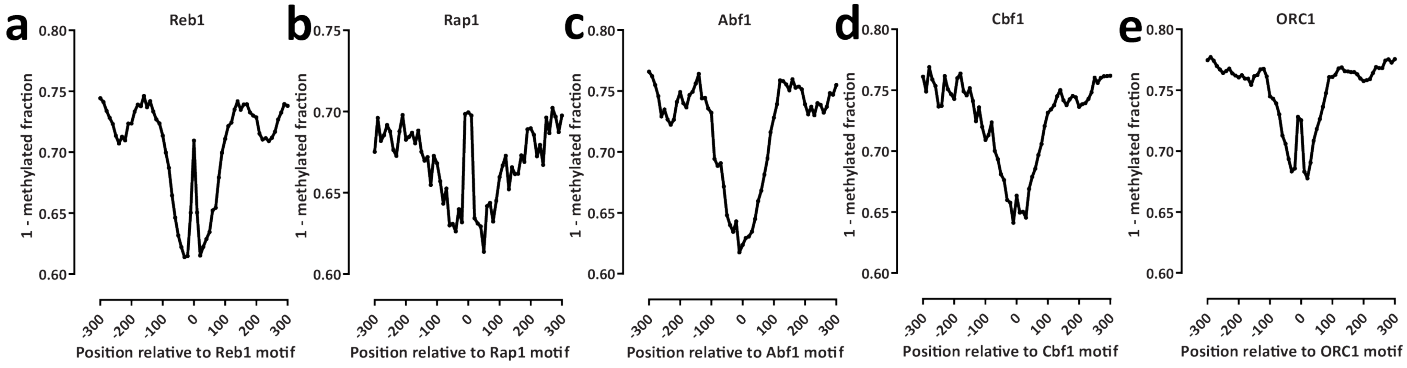
Supplementary Figure 70: SMAC-seq reveals the distribution of alternative chromatin states of rDNA arrays. Shown are all reads covering the *RDN37-1* and *RDN37-2* arrays in the *RDN1* locus in the "diamide 0 min rep1" experiment (unfiltered reads, "aggregate" signal).



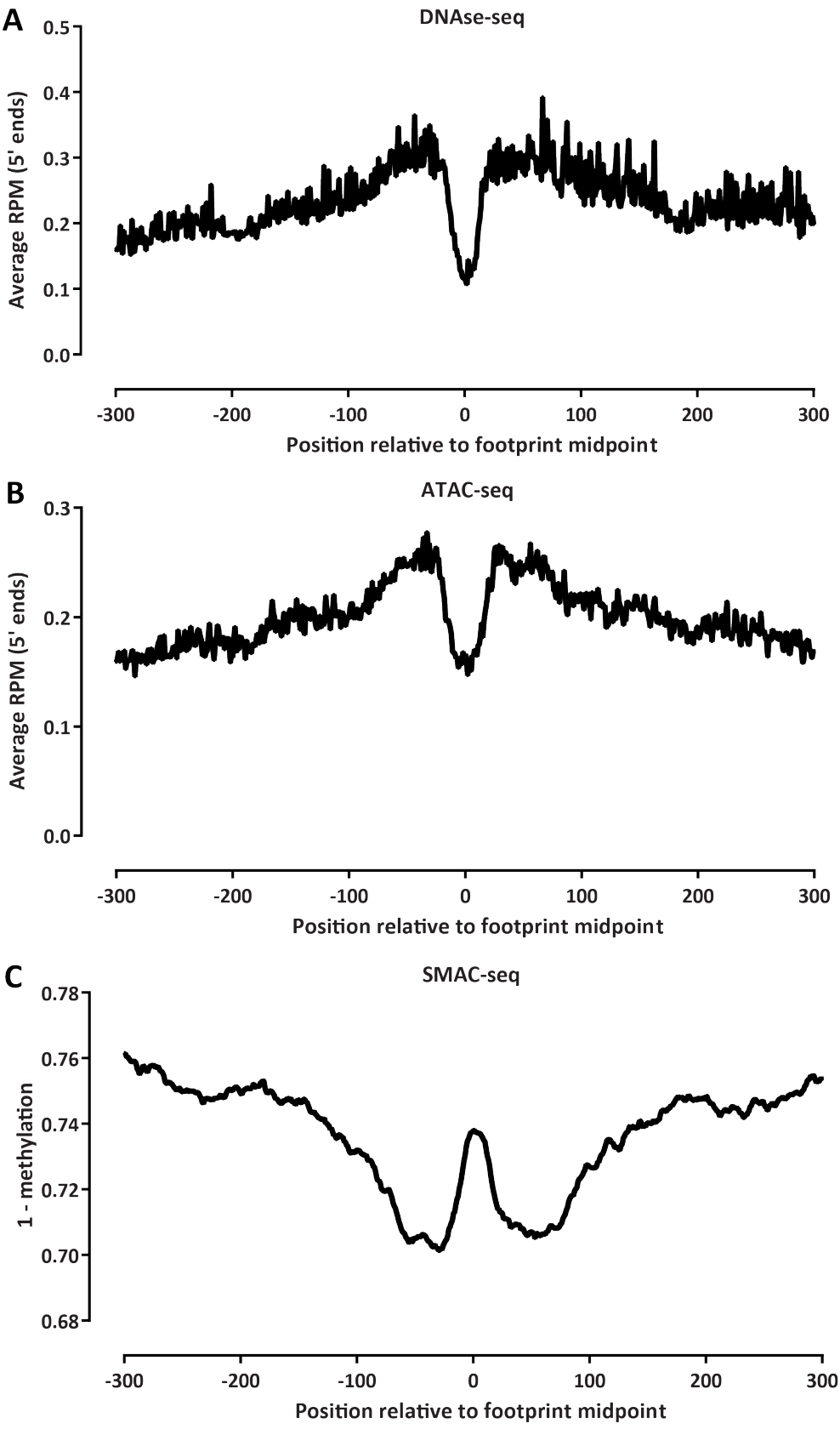
Supplementary Figure 71: NMI profile for the *RDN37-2* array, as in Figure 3b.


GC
m6A
GC+m6A
informative positions:
0 1 2 3 4-5 6-10 >10
ABF1 0.56 0.33 0.10 0.01 0.00 0.00 0.00 | 0.00 0.02 0.08 0.17 0.48 0.25 0.00 | 0.00 0.01 0.03 0.12 0.46 0.38 0.00
ABF2 0.78 0.22 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.48 0.42 0.09 0.00 0.00 | 0.00 0.00 0.33 0.51 0.16 0.00 0.00
ACA1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.36 0.50 0.14 0.00 0.00 | 0.00 0.00 0.36 0.50 0.14 0.00 0.00
ACE2 0.00 0.79 0.21 0.00 0.00 0.00 0.00 | 0.31 0.16 0.17 0.17 0.20 0.00 0.00 | 0.00 0.23 0.20 0.14 0.37 0.06 0.00
ADR1 0.92 0.08 0.00 0.00 0.00 0.00 0.00 | 0.28 0.29 0.27 0.14 0.03 0.00 0.00 | 0.24 0.29 0.30 0.14 0.03 0.00 0.00
AFT1 0.00 1.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.35 0.41 0.20 0.04 0.00 0.00 | 0.00 0.00 0.35 0.41 0.24 0.00 0.00
ARG5,6 0.12 0.61 0.25 0.02 0.00 0.00 0.00 | 0.16 0.24 0.22 0.15 0.23 0.00 0.00 | 0.01 0.09 0.17 0.29 0.39 0.04 0.00
ARG80 0.31 0.43 0.20 0.06 0.01 0.00 0.00 | 0.04 0.12 0.17 0.17 0.34 0.16 0.00 | 0.01 0.04 0.09 0.16 0.39 0.31 0.00
ARG81 0.44 0.40 0.16 0.00 0.00 0.00 0.00 | 0.15 0.51 0.29 0.05 0.00 0.00 0.00 | 0.03 0.24 0.47 0.26 0.00 0.00 0.00
ARO80 0.61 0.32 0.06 0.00 0.00 0.00 0.00 | 0.05 0.17 0.27 0.22 0.25 0.03 0.00 | 0.02 0.10 0.23 0.26 0.35 0.04 0.00
ARR1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.28 0.72 0.00 0.00 | 0.00 0.00 0.00 0.28 0.72 0.00 0.00
AZF1 0.96 0.04 0.00 0.00 0.00 0.00 0.00 | 0.43 0.06 0.00 0.00 0.00 0.51 0.00 | 0.41 0.08 0.00 0.00 0.00 0.51 0.00
CAD1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.07 0.53 0.41 0.00 0.00 | 0.00 0.00 0.07 0.53 0.41 0.00 0.00
CAT8 0.39 0.41 0.16 0.04 0.00 0.00 0.00 | 0.12 0.24 0.24 0.20 0.19 0.01 0.00 | 0.02 0.10 0.23 0.27 0.34 0.03 0.00
CBF1 0.60 0.39 0.01 0.00 0.00 0.00 0.00 | 0.00 0.37 0.40 0.21 0.02 0.00 0.00 | 0.00 0.16 0.45 0.32 0.07 0.00 0.00
CEP3 0.93 0.07 0.00 0.00 0.00 0.00 0.00 | 0.00 0.12 0.47 0.34 0.06 0.00 0.00 | 0.00 0.10 0.45 0.37 0.08 0.00 0.00
CHA4 0.04 0.47 0.36 0.11 0.02 0.00 0.00 | 0.02 0.04 0.13 0.14 0.34 0.31 0.01 | 0.00 0.01 0.03 0.06 0.29 0.58 0.03
CIN5 0.97 0.03 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.06 0.52 0.42 0.00 0.00 | 0.00 0.00 0.05 0.52 0.44 0.00 0.00
CRZ1 0.00 0.93 0.07 0.00 0.00 0.00 0.00 | 0.35 0.26 0.22 0.14 0.04 0.00 0.00 | 0.00 0.33 0.24 0.24 0.19 0.00 0.00
CST6 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.02 0.47 0.48 0.03 0.00 0.00 | 0.00 0.02 0.47 0.48 0.03 0.00 0.00
CUP2 0.15 0.74 0.10 0.00 0.00 0.00 0.00 | 0.32 0.18 0.00 0.00 0.35 0.15 0.00 | 0.06 0.23 0.21 0.00 0.20 0.31 0.00
CUP9 0.96 0.04 0.00 0.00 0.00 0.00 0.00 | 0.00 0.08 0.22 0.39 0.31 0.00 0.00 | 0.00 0.08 0.20 0.39 0.33 0.00 0.00
DAL80 0.35 0.61 0.04 0.00 0.00 0.00 0.00 | 0.00 0.16 0.49 0.29 0.06 0.00 0.00 | 0.00 0.07 0.23 0.44 0.26 0.00 0.00
DAL81 0.01 0.22 0.43 0.24 0.10 0.00 0.00 | 0.05 0.15 0.22 0.26 0.27 0.05 0.00 | 0.00 0.00 0.05 0.12 0.45 0.38 0.00
DAL82 0.34 0.53 0.14 0.00 0.00 0.00 0.00 | 0.00 0.05 0.18 0.29 0.44 0.03 0.00 | 0.00 0.00 0.07 0.22 0.57 0.14 0.00
DAT1 0.99 0.01 0.00 0.00 0.00 0.00 0.00 | 0.32 0.13 0.04 0.00 0.00 0.50 0.00 | 0.32 0.14 0.04 0.00 0.00 0.50 0.00
DOT6 0.24 0.43 0.24 0.08 0.01 0.00 0.00 | 0.00 0.03 0.05 0.13 0.36 0.42 0.01 | 0.00 0.01 0.02 0.04 0.26 0.65 0.02
ECM22 0.36 0.36 0.22 0.05 0.00 0.00 0.00 | 0.03 0.13 0.20 0.21 0.32 0.10 0.00 | 0.00 0.03 0.08 0.23 0.47 0.19 0.00
ECM23 0.86 0.14 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.35 0.35 0.29 0.01 0.00 | 0.00 0.00 0.29 0.35 0.35 0.01 0.00
EDS1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.09 0.25 0.15 0.01 0.40 0.10 0.00 | 0.09 0.25 0.15 0.01 0.40 0.10 0.00
ERT1 0.87 0.13 0.00 0.00 0.00 0.00 0.00 | 0.11 0.28 0.21 0.31 0.09 0.00 0.00 | 0.07 0.30 0.20 0.30 0.13 0.00 0.00
FHL1 0.08 0.68 0.20 0.03 0.00 0.00 0.00 | 0.01 0.03 0.07 0.13 0.28 0.48 0.02 | 0.00 0.01 0.03 0.06 0.24 0.62 0.05
FKH1 0.89 0.11 0.00 0.00 0.00 0.00 0.00 | 0.01 0.20 0.20 0.11 0.27 0.21 0.00 | 0.01 0.18 0.21 0.11 0.25 0.24 0.00
FKH2 0.74 0.23 0.03 0.00 0.00 0.00 0.00 | 0.01 0.04 0.08 0.11 0.19 0.33 0.24 | 0.01 0.03 0.07 0.10 0.21 0.31 0.27
FYV5 0.65 0.29 0.05 0.01 0.00 0.00 0.00 | 0.00 0.00 0.01 0.04 0.25 0.67 0.03 | 0.00 0.00 0.00 0.02 0.20 0.73 0.04
FZF1 0.38 0.45 0.13 0.04 0.00 0.00 0.00 | 0.00 0.01 0.03 0.05 0.21 0.59 0.10 | 0.00 0.00 0.01 0.03 0.15 0.66 0.15
GAL4 0.22 0.42 0.23 0.10 0.02 0.00 0.00 | 0.04 0.17 0.24 0.21 0.29 0.05 0.00 | 0.00 0.03 0.13 0.20 0.46 0.19 0.00
GAT1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.09 0.27 0.59 0.05 0.00 | 0.00 0.00 0.09 0.27 0.59 0.05 0.00
GAT3 0.86 0.14 0.00 0.00 0.00 0.00 0.00 | 0.00 0.04 0.32 0.40 0.24 0.00 0.00 | 0.00 0.04 0.24 0.43 0.29 0.00 0.00
GAT4 0.86 0.14 0.00 0.00 0.00 0.00 0.00 | 0.00 0.04 0.29 0.30 0.35 0.02 0.00 | 0.00 0.03 0.24 0.29 0.41 0.02 0.00
GCN4 0.96 0.04 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.55 0.36 0.09 0.00 0.00 | 0.00 0.00 0.51 0.40 0.09 0.00 0.00
GCR1 0.00 1.00 0.00 0.00 0.00 0.00 0.00 | 0.27 0.22 0.02 0.21 0.29 0.00 0.00 | 0.00 0.27 0.22 0.02 0.44 0.06 0.00
GCR2 0.00 1.00 0.00 0.00 0.00 0.00 0.00 | 0.25 0.22 0.12 0.21 0.20 0.00 0.00 | 0.00 0.25 0.22 0.12 0.37 0.04 0.00
GIS1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.13 0.50 0.35 0.03 0.00 0.00 | 0.00 0.13 0.50 0.35 0.03 0.00 0.00
GSM1 0.57 0.36 0.07 0.01 0.00 0.00 0.00 | 0.00 0.01 0.06 0.09 0.36 0.48 0.01 | 0.00 0.00 0.03 0.07 0.29 0.60 0.01
HAC1 0.66 0.31 0.03 0.00 0.00 0.00 0.00 | 0.03 0.14 0.29 0.31 0.22 0.01 0.00 | 0.02 0.07 0.22 0.36 0.31 0.02 0.00
HAL9 0.20 0.36 0.28 0.12 0.04 0.00 0.00 | 0.04 0.09 0.18 0.23 0.37 0.11 0.00 | 0.00 0.01 0.05 0.15 0.44 0.35 0.00
HAP1 0.60 0.34 0.06 0.00 0.00 0.00 0.00 | 0.04 0.13 0.28 0.26 0.25 0.02 0.00 | 0.02 0.11 0.20 0.26 0.35 0.06 0.00
HAP3 0.47 0.44 0.08 0.00 0.00 0.00 0.00 | 0.00 0.05 0.13 0.18 0.40 0.24 0.00 | 0.00 0.02 0.07 0.14 0.40 0.36 0.00
HAP4 0.70 0.27 0.03 0.00 0.00 0.00 0.00 | 0.00 0.07 0.13 0.15 0.35 0.29 0.00 | 0.00 0.04 0.10 0.16 0.35 0.35 0.00
HAP5 0.41 0.41 0.16 0.02 0.00 0.00 0.00 | 0.01 0.14 0.21 0.17 0.28 0.19 0.00 | 0.00 0.04 0.14 0.19 0.32 0.30 0.00
HCM1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.16 0.26 0.08 0.00 0.13 0.36 0.00 | 0.16 0.26 0.08 0.00 0.13 0.36 0.00
HMRA1 0.00 1.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.21 0.21 0.31 0.27 0.00 0.00 | 0.00 0.00 0.21 0.21 0.53 0.05 0.00
HMRA2 0.91 0.09 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.24 0.44 0.32 0.00 0.00 | 0.00 0.00 0.22 0.41 0.37 0.00 0.00
HMS1 0.90 0.10 0.00 0.00 0.00 0.00 0.00 | 0.00 0.38 0.43 0.16 0.03 0.00 0.00 | 0.00 0.32 0.44 0.21 0.03 0.00 0.00
HMS2 0.93 0.07 0.00 0.00 0.00 0.00 0.00 | 0.09 0.17 0.18 0.06 0.35 0.15 0.00 | 0.09 0.16 0.18 0.08 0.33 0.17 0.00
HSF1 0.99 0.01 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.15 0.51 0.34 0.00 0.00 | 0.00 0.00 0.15 0.50 0.35 0.00 0.00
IME1 0.00 0.76 0.23 0.01 0.00 0.00 0.00 | 0.32 0.40 0.23 0.05 0.00 0.00 0.00 | 0.00 0.24 0.35 0.30 0.11 0.00 0.00
IME4 0.81 0.19 0.00 0.00 0.00 0.00 0.00 | 0.19 0.23 0.06 0.00 0.04 0.47 0.00 | 0.12 0.28 0.09 0.00 0.04 0.47 0.00
INO2 0.20 0.75 0.05 0.00 0.00 0.00 0.00 | 0.00 0.11 0.35 0.47 0.08 0.00 0.00 | 0.00 0.00 0.16 0.39 0.46 0.00 0.00
INO4 0.42 0.54 0.04 0.00 0.00 0.00 0.00 | 0.00 0.06 0.54 0.30 0.10 0.00 0.00 | 0.00 0.00 0.27 0.48 0.25 0.00 0.00
IXR1 0.22 0.55 0.21 0.02 0.00 0.00 0.00 | 0.07 0.15 0.22 0.17 0.26 0.12 0.00 | 0.01 0.07 0.14 0.19 0.34 0.24 0.00
KAR4 0.47 0.40 0.12 0.01 0.00 0.00 0.00 | 0.03 0.10 0.15 0.14 0.25 0.33 0.00 | 0.01 0.05 0.12 0.17 0.23 0.42 0.01
LEU3 0.39 0.39 0.16 0.06 0.00 0.00 0.00 | 0.21 0.43 0.27 0.09 0.00 0.00 0.00 | 0.06 0.22 0.35 0.26 0.11 0.00 0.00
LYS14 0.55 0.36 0.09 0.01 0.00 0.00 0.00 | 0.00 0.04 0.11 0.22 0.48 0.15 0.00 | 0.00 0.02 0.07 0.16 0.50 0.25 0.00
MAC1 0.00 0.98 0.02 0.00 0.00 0.00 0.00 | 0.17 0.18 0.11 0.18 0.32 0.06 0.00 | 0.00 0.17 0.18 0.11 0.33 0.21 0.00
MATALPHA1 0.57 0.34 0.08 0.01 0.00 0.00 0.00 | 0.00 0.00 0.02 0.05 0.20 0.63 0.10 | 0.00 0.00 0.01 0.03 0.15 0.68 0.13
MBP1 0.00 1.00 0.00 0.00 0.00 0.00 0.00 | 0.15 0.31 0.36 0.15 0.03 0.00 0.00 | 0.00 0.15 0.31 0.36 0.18 0.00 0.00
MCM1 0.69 0.27 0.03 0.00 0.00 0.00 0.00 | 0.00 0.01 0.02 0.07 0.29 0.59 0.01 | 0.00 0.00 0.02 0.05 0.27 0.65 0.02
MET31 0.35 0.56 0.09 0.00 0.00 0.00 0.00 | 0.14 0.35 0.28 0.20 0.03 0.00 0.00 | 0.01 0.22 0.34 0.29 0.15 0.00 0.00
MET32 0.19 0.72 0.09 0.00 0.00 0.00 0.00 | 0.13 0.22 0.34 0.26 0.06 0.00 0.00 | 0.00 0.10 0.37 0.22 0.31 0.00 0.00
MET4 0.79 0.19 0.01 0.00 0.00 0.00 0.00 | 0.00 0.17 0.39 0.35 0.09 0.00 0.00 | 0.00 0.14 0.33 0.35 0.18 0.00 0.00
MGA1 0.68 0.27 0.05 0.00 0.00 0.00 0.00 | 0.01 0.03 0.06 0.09 0.21 0.49 0.11 | 0.00 0.02 0.04 0.08 0.21 0.51 0.14
MIG1 0.58 0.34 0.07 0.00 0.00 0.00 0.00 | 0.36 0.41 0.20 0.03 0.00 0.00 0.00 | 0.14 0.41 0.37 0.08 0.00 0.00 0.00
MIG2 0.00 0.75 0.24 0.01 0.00 0.00 0.00 | 0.44 0.41 0.14 0.01 0.00 0.00 0.00 | 0.00 0.28 0.47 0.22 0.03 0.00 0.00
MIG3 0.00 0.75 0.24 0.01 0.00 0.00 0.00 | 0.44 0.41 0.14 0.01 0.00 0.00 0.00 | 0.00 0.28 0.47 0.22 0.03 0.00 0.00
MSN1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.08 0.38 0.36 0.19 0.00 0.00 | 0.00 0.08 0.38 0.36 0.19 0.00 0.00
MSN2 0.57 0.43 0.00 0.00 0.00 0.00 0.00 | 0.15 0.22 0.34 0.22 0.07 0.00 0.00 | 0.00 0.30 0.32 0.23 0.15 0.00 0.00
NDT80 0.92 0.08 0.00 0.00 0.00 0.00 0.00 | 0.26 0.21 0.03 0.01 0.35 0.14 0.00 | 0.22 0.25 0.03 0.01 0.34 0.15 0.00
NHP10 0.58 0.35 0.07 0.00 0.00 0.00 0.00 | 0.36 0.40 0.20 0.05 0.00 0.00 0.00 | 0.13 0.39 0.39 0.09 0.00 0.00 0.00
NHP6A 0.91 0.09 0.01 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.05 0.75 0.20 0.00 | 0.00 0.00 0.00 0.04 0.75 0.21 0.00
NHP6B 0.93 0.07 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.01 0.06 0.75 0.18 0.00 | 0.00 0.00 0.01 0.05 0.75 0.19 0.00
NRG1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.34 0.44 0.20 0.02 0.00 0.00 | 0.00 0.34 0.44 0.20 0.02 0.00 0.00
NRG2 0.94 0.06 0.00 0.00 0.00 0.00 0.00 | 0.00 0.36 0.47 0.17 0.00 0.00 0.00 | 0.00 0.31 0.51 0.18 0.00 0.00 0.00
OAF1 0.05 0.16 0.27 0.34 0.18 0.01 0.00 | 0.08 0.11 0.21 0.13 0.27 0.19 0.00 | 0.00 0.00 0.02 0.10 0.36 0.50 0.02
OPI1 0.80 0.19 0.01 0.00 0.00 0.00 0.00 | 0.22 0.23 0.29 0.20 0.05 0.00 0.00 | 0.15 0.27 0.25 0.24 0.09 0.00 0.00
ORC1 0.98 0.04 0.00 0.00 0.00 0.00 0.00 | 0.10 0.11 0.10 0.08 0.08 0.14 0.39 | 0.09 0.11 0.11 0.08 0.08 0.13 0.40
PDC2 0.99 0.01 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.04 0.11 0.76 0.09 0.00 | 0.00 0.00 0.04 0.10 0.77 0.09 0.00
PDR1 0.24 0.49 0.26 0.01 0.00 0.00 0.00 | 0.31 0.39 0.22 0.08 0.00 0.00 0.00 | 0.00 0.28 0.41 0.23 0.07 0.00 0.00
PDR3 0.00 0.67 0.33 0.00 0.00 0.00 0.00 | 0.33 0.44 0.15 0.08 0.00 0.00 0.00 | 0.00 0.20 0.40 0.28 0.12 0.00 0.00
PDR8 0.47 0.35 0.18 0.00 0.00 0.00 0.00 | 0.17 0.44 0.29 0.10 0.00 0.00 0.00 | 0.00 0.31 0.42 0.21 0.07 0.00 0.00
PHD1 0.00 0.34 0.59 0.07 0.00 0.00 0.00 | 0.29 0.45 0.22 0.04 0.00 0.00 0.00 | 0.00 0.12 0.28 0.39 0.21 0.00 0.00
PHO4 0.32 0.56 0.12 0.00 0.00 0.00 0.00 | 0.00 0.52 0.41 0.08 0.00 0.00 0.00 | 0.00 0.12 0.49 0.31 0.08 0.00 0.00
PPR1 0.37 0.43 0.17 0.02 0.00 0.00 0.00 | 0.02 0.11 0.23 0.27 0.30 0.07 0.00 | 0.01 0.03 0.12 0.23 0.47 0.14 0.00
PUT3 0.24 0.45 0.23 0.08 0.01 0.00 0.00 | 0.13 0.20 0.20 0.27 0.17 0.02 0.00 | 0.02 0.09 0.16 0.22 0.42 0.08 0.00
PZF1 0.64 0.28 0.07 0.00 0.00 0.00 0.00 | 0.02 0.09 0.24 0.30 0.30 0.04 0.00 | 0.01 0.06 0.18 0.27 0.43 0.06 0.00
RAP1 0.82 0.16 0.02 0.00 0.00 0.00 0.00 | 0.16 0.14 0.13 0.15 0.28 0.13 0.00 | 0.14 0.11 0.14 0.15 0.31 0.15 0.00
RCO1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.26 0.21 0.49 0.04 | 0.00 0.00 0.00 0.26 0.21 0.49 0.04
RDR1 0.54 0.38 0.08 0.00 0.00 0.00 0.00 | 0.07 0.20 0.41 0.28 0.04 0.00 0.00 | 0.02 0.11 0.32 0.38 0.17 0.00 0.00
RDS1 0.00 0.42 0.43 0.15 0.00 0.00 0.00 | 0.44 0.43 0.10 0.03 0.00 0.00 0.00 | 0.00 0.14 0.39 0.36 0.11 0.00 0.00
RDS2 0.48 0.42 0.09 0.01 0.00 0.00 0.00 | 0.02 0.14 0.22 0.22 0.32 0.08 0.00 | 0.01 0.05 0.17 0.21 0.42 0.14 0.00
REB1 0.58 0.38 0.04 0.00 0.00 0.00 0.00 | 0.01 0.13 0.28 0.31 0.25 0.02 0.00 | 0.00 0.05 0.21 0.33 0.37 0.03 0.00
REI1 0.71 0.27 0.02 0.00 0.00 0.00 0.00 | 0.11 0.44 0.37 0.08 0.00 0.00 0.00 | 0.06 0.32 0.46 0.16 0.00 0.00 0.00
RFX1 0.23 0.47 0.23 0.06 0.01 0.00 0.00 | 0.01 0.04 0.14 0.24 0.48 0.09 0.00 | 0.00 0.01 0.03 0.12 0.52 0.32 0.00
RGT1 0.59 0.33 0.07 0.00 0.00 0.00 0.00 | 0.00 0.01 0.03 0.07 0.30 0.57 0.02 | 0.00 0.01 0.02 0.04 0.24 0.66 0.02
RIM101 0.00 0.83 0.17 0.00 0.00 0.00 0.00 | 0.00 0.26 0.44 0.26 0.04 0.00 0.00 | 0.00 0.00 0.21 0.40 0.39 0.00 0.00
RLM1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.03 0.06 0.17 0.59 0.16 0.00 | 0.00 0.03 0.06 0.17 0.59 0.16 0.00
RME1 0.91 0.09 0.00 0.00 0.00 0.00 0.00 | 0.06 0.19 0.16 0.11 0.30 0.19 0.00 | 0.05 0.18 0.17 0.11 0.29 0.20 0.00
ROX1 0.87 0.13 0.00 0.00 0.00 0.00 0.00 | 0.00 0.37 0.14 0.00 0.41 0.08 0.00 | 0.00 0.28 0.23 0.00 0.40 0.09 0.00
RPH1 0.90 0.10 0.00 0.00 0.00 0.00 0.00 | 0.00 0.17 0.47 0.29 0.07 0.00 0.00 | 0.00 0.13 0.45 0.33 0.09 0.00 0.00
RPN4 0.02 0.89 0.09 0.00 0.00 0.00 0.00 | 0.02 0.18 0.28 0.29 0.22 0.01 0.00 | 0.00 0.02 0.16 0.28 0.48 0.07 0.00
RSC3 0.00 0.00 0.00 0.34 0.47 0.19 0.00 | 0.61 0.35 0.05 0.00 0.00 0.00 0.00 | 0.00 0.00 0.17 0.40 0.43 0.00 0.00
RSC30 0.00 0.00 0.22 0.23 0.47 0.26 0.02 | 0.36 0.29 0.11 0.10 0.11 0.03 0.00 | 0.00 0.00 0.08 0.28 0.36 0.27 0.01
RSF2 0.57 0.41 0.02 0.00 0.00 0.00 0.00 | 0.16 0.36 0.32 0.12 0.05 0.00 0.00 | 0.06 0.27 0.36 0.23 0.08 0.00 0.00
RTG1 0.97 0.03 0.00 0.00 0.00 0.00 0.00 | 0.00 0.01 0.26 0.38 0.34 0.01 0.00 | 0.00 0.01 0.25 0.38 0.36 0.01 0.00
RTG3 0.57 0.24 0.13 0.05 0.01 0.00 0.00 | 0.00 0.04 0.22 0.28 0.31 0.15 0.00 | 0.00 0.03 0.15 0.24 0.31 0.26 0.01
SFL1 0.00 0.74 0.26 0.00 0.00 0.00 0.00 | 0.00 0.21 0.55 0.23 0.00 0.00 0.00 | 0.00 0.00 0.13 0.51 0.36 0.00 0.00
SFP1 0.77 0.21 0.02 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.01 0.14 0.76 0.09 | 0.00 0.00 0.00 0.01 0.11 0.78 0.10
SIP4 0.43 0.50 0.07 0.01 0.00 0.00 0.00 | 0.02 0.17 0.28 0.28 0.24 0.00 0.00 | 0.00 0.05 0.22 0.30 0.40 0.03 0.00
SKN7 0.00 0.15 0.79 0.06 0.00 0.00 0.00 | 0.71 0.26 0.04 0.00 0.00 0.00 0.00 | 0.00 0.09 0.58 0.33 0.00 0.00 0.00
SKO1 0.83 0.17 0.00 0.00 0.00 0.00 0.00 | 0.00 0.06 0.36 0.43 0.15 0.00 0.00 | 0.00 0.00 0.39 0.41 0.20 0.00 0.00
SMK1 0.29 0.48 0.18 0.04 0.00 0.00 0.00 | 0.00 0.01 0.04 0.10 0.44 0.41 0.00 | 0.00 0.00 0.01 0.04 0.30 0.66 0.00
SMP1 0.62 0.32 0.06 0.01 0.00 0.00 0.00 | 0.00 0.00 0.00 0.02 0.18 0.75 0.07 | 0.00 0.00 0.00 0.01 0.11 0.76 0.10
SNF1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.32 0.16 0.02 0.00 0.00 0.50 0.00 | 0.32 0.16 0.02 0.00 0.00 0.50 0.00
SNT2 0.01 0.35 0.49 0.14 0.02 0.00 0.00 | 0.01 0.09 0.23 0.27 0.35 0.05 0.00 | 0.00 0.00 0.02 0.11 0.52 0.34 0.00
SOK2 0.00 0.62 0.38 0.00 0.00 0.00 0.00 | 0.03 0.29 0.32 0.29 0.07 0.00 0.00 | 0.00 0.03 0.14 0.38 0.41 0.03 0.00
SPT10 0.47 0.42 0.10 0.01 0.00 0.00 0.00 | 0.00 0.02 0.10 0.21 0.50 0.18 0.00 | 0.00 0.01 0.04 0.15 0.50 0.31 0.00
SPT15 0.83 0.15 0.02 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.00 0.06 0.82 0.12 | 0.00 0.00 0.00 0.00 0.04 0.83 0.13
SPT2 0.98 0.02 0.00 0.00 0.00 0.00 0.00 | 0.00 0.31 0.16 0.02 0.19 0.31 0.00 | 0.00 0.31 0.16 0.02 0.17 0.33 0.00
SPT23 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.23 0.21 0.05 0.07 0.44 0.00 | 0.00 0.23 0.21 0.05 0.07 0.44 0.00
SRD1 0.98 0.02 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.38 0.36 0.25 0.00 0.00 | 0.00 0.00 0.36 0.39 0.25 0.00 0.00
STB2 0.86 0.31 0.03 0.00 0.00 0.00 0.00 | 0.00 0.33 0.43 0.19 0.05 0.00 0.00 | 0.00 0.23 0.40 0.24 0.14 0.00 0.00
STB3 0.73 0.24 0.03 0.00 0.00 0.00 0.00 | 0.00 0.01 0.03 0.05 0.20 0.55 0.15 | 0.00 0.01 0.02 0.05 0.18 0.56 0.18
STB4 0.81 0.19 0.00 0.00 0.00 0.00 0.00 | 0.00 0.21 0.49 0.25 0.04 0.00 0.00 | 0.00 0.18 0.44 0.33 0.07 0.00 0.00
STB5 0.73 0.25 0.02 0.00 0.00 0.00 0.00 | 0.07 0.23 0.34 0.22 0.14 0.00 0.00 | 0.05 0.14 0.33 0.31 0.17 0.00 0.00
STE12 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.22 0.22 0.05 0.45 0.06 0.00 | 0.00 0.22 0.22 0.05 0.45 0.06 0.00
STP1 0.00 0.00 0.82 0.17 0.01 0.00 0.00 | 0.41 0.22 0.21 0.13 0.03 0.00 0.00 | 0.00 0.00 0.32 0.25 0.40 0.03 0.00
STP2 0.00 0.13 0.35 0.34 0.17 0.00 0.00 | 0.01 0.07 0.15 0.25 0.37 0.16 0.00 | 0.00 0.00 0.00 0.03 0.28 0.69 0.01
STP3 0.00 0.03 0.87 0.11 0.00 0.00 0.00 | 0.00 0.45 0.41 0.11 0.03 0.00 0.00 | 0.00 0.00 0.01 0.38 0.57 0.03 0.00
STP4 0.00 0.17 0.75 0.08 0.00 0.00 0.00 | 0.00 0.42 0.39 0.16 0.03 0.00 0.00 | 0.00 0.00 0.04 0.42 0.51 0.03 0.00
SUA7 0.06 0.70 0.24 0.00 0.00 0.00 0.00 | 0.11 0.38 0.35 0.16 0.00 0.00 0.00 | 0.01 0.07 0.34 0.35 0.23 0.00 0.00
SUM1 0.81 0.16 0.02 0.00 0.00 0.00 0.00 | 0.05 0.18 0.19 0.16 0.33 0.09 0.00 | 0.02 0.13 0.22 0.20 0.32 0.11 0.00
SUT1 0.00 0.68 0.29 0.03 0.00 0.00 0.00 | 0.44 0.36 0.19 0.02 0.00 0.00 0.00 | 0.00 0.25 0.41 0.29 0.05 0.00 0.00
SUT2 0.68 0.28 0.04 0.00 0.00 0.00 0.00 | 0.00 0.02 0.06 0.11 0.31 0.49 0.02 | 0.00 0.01 0.04 0.10 0.29 0.54 0.02
SWI4 0.00 0.86 0.14 0.00 0.00 0.00 0.00 | 0.28 0.19 0.05 0.24 0.24 0.00 0.00 | 0.00 0.22 0.23 0.07 0.42 0.06 0.00
SWI5 0.00 0.93 0.07 0.00 0.00 0.00 0.00 | 0.24 0.21 0.11 0.23 0.21 0.00 0.00 | 0.00 0.21 0.23 0.10 0.41 0.05 0.00
TBF1 0.65 0.30 0.06 0.00 0.00 0.00 0.00 | 0.01 0.06 0.17 0.27 0.41 0.09 0.00 | 0.00 0.03 0.12 0.23 0.45 0.16 0.00
TBS1 0.73 0.24 0.03 0.00 0.00 0.00 0.00 | 0.18 0.38 0.35 0.08 0.00 0.00 0.00 | 0.11 0.30 0.44 0.16 0.00 0.00 0.00
TEA1 0.49 0.39 0.11 0.02 0.00 0.00 0.00 | 0.00 0.00 0.01 0.05 0.30 0.62 0.01 | 0.00 0.00 0.01 0.03 0.19 0.74 0.03
TEC1 0.52 0.37 0.10 0.01 0.00 0.00 0.00 | 0.00 0.00 0.02 0.07 0.35 0.56 0.00 | 0.00 0.00 0.01 0.03 0.27 0.69 0.01
THI2 0.21 0.38 0.24 0.12 0.04 0.00 0.00 | 0.04 0.10 0.15 0.16 0.36 0.18 0.00 | 0.00 0.02 0.08 0.13 0.37 0.39 0.01
THO2 0.94 0.06 0.00 0.00 0.00 0.00 0.00 | 0.15 0.17 0.13 0.05 0.08 0.43 0.00 | 0.14 0.17 0.13 0.05 0.07 0.44 0.00
TOD6 0.23 0.44 0.25 0.07 0.01 0.00 0.00 | 0.00 0.02 0.05 0.12 0.38 0.43 0.00 | 0.00 0.00 0.02 0.04 0.26 0.67 0.02
TOS8 0.93 0.07 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.42 0.35 0.23 0.00 0.00 | 0.00 0.00 0.38 0.36 0.26 0.00 0.00
TYE7 0.86 0.14 0.00 0.00 0.00 0.00 0.00 | 0.00 0.15 0.49 0.24 0.12 0.00 0.00 | 0.00 0.11 0.47 0.27 0.15 0.00 0.00
UGA3 0.00 0.39 0.58 0.03 0.00 0.00 0.00 | 0.25 0.56 0.18 0.02 0.00 0.00 0.00 | 0.00 0.09 0.36 0.42 0.13 0.00 0.00
UME1 0.99 0.01 0.00 0.00 0.00 0.00 0.00 | 0.00 0.28 0.19 0.04 0.12 0.39 0.00 | 0.00 0.27 0.19 0.04 0.11 0.39 0.00
UME6 0.00 0.36 0.56 0.08 0.00 0.00 0.00 | 0.18 0.47 0.31 0.04 0.00 0.00 0.00 | 0.00 0.01 0.32 0.43 0.24 0.00 0.00
UPC2 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.25 0.42 0.33 0.00 0.00 | 0.00 0.00 0.25 0.42 0.33 0.00 0.00
URC2 0.89 0.11 0.00 0.00 0.00 0.00 0.00 | 0.02 0.12 0.25 0.30 0.30 0.01 0.00 | 0.02 0.09 0.25 0.30 0.31 0.03 0.00
VHR1 0.46 0.37 0.15 0.02 0.00 0.00 0.00 | 0.00 0.04 0.12 0.23 0.40 0.22 0.00 | 0.00 0.01 0.05 0.14 0.46 0.33 0.00
WAR1 0.07 0.33 0.29 0.18 0.12 0.00 0.00 | 0.00 0.03 0.05 0.08 0.31 0.51 0.01 | 0.00 0.00 0.00 0.02 0.11 0.77 0.09
WTM1 0.98 0.04 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.01 0.03 0.22 0.71 0.04 | 0.00 0.00 0.01 0.03 0.21 0.72 0.04
XBP1 0.55 0.38 0.07 0.00 0.00 0.00 0.00 | 0.00 0.23 0.32 0.28 0.15 0.01 0.00 | 0.00 0.06 0.31 0.36 0.25 0.02 0.00
YAP1 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.04 0.54 0.42 0.00 | 0.00 0.00 0.00 0.04 0.54 0.42 0.00
YAP3 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.00 0.46 0.54 0.00 | 0.00 0.00 0.00 0.00 0.46 0.54 0.00
YAP5 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.11 0.15 0.45 0.30 | 0.00 0.00 0.00 0.11 0.15 0.45 0.30
YAP6 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.00 0.47 0.53 0.00 | 0.00 0.00 0.00 0.00 0.47 0.53 0.00
YAP7 1.00 0.00 0.00 0.00 0.00 0.00 0.00 | 0.00 0.00 0.00 0.00 0.30 0.40 0.30 | 0.00 0.00 0.00 0.00 0.30 0.40 0.30
YDR266C 0.78 0.22 0.00 0.00 0.00 0.00 0.00 | 0.03 0.25 0.22 0.14 0.35 0.00 0.00 | 0.02 0.18 0.27 0.15 0.38 0.00 0.00
YDR520C 0.89 0.11 0.00 0.00 0.00 0.00 0.00 | 0.02 0.12 0.26 0.30 0.29 0.01 0.00 | 0.02 0.10 0.25 0.30 0.31 0.03 0.00
YER130C 0.74 0.26 0.

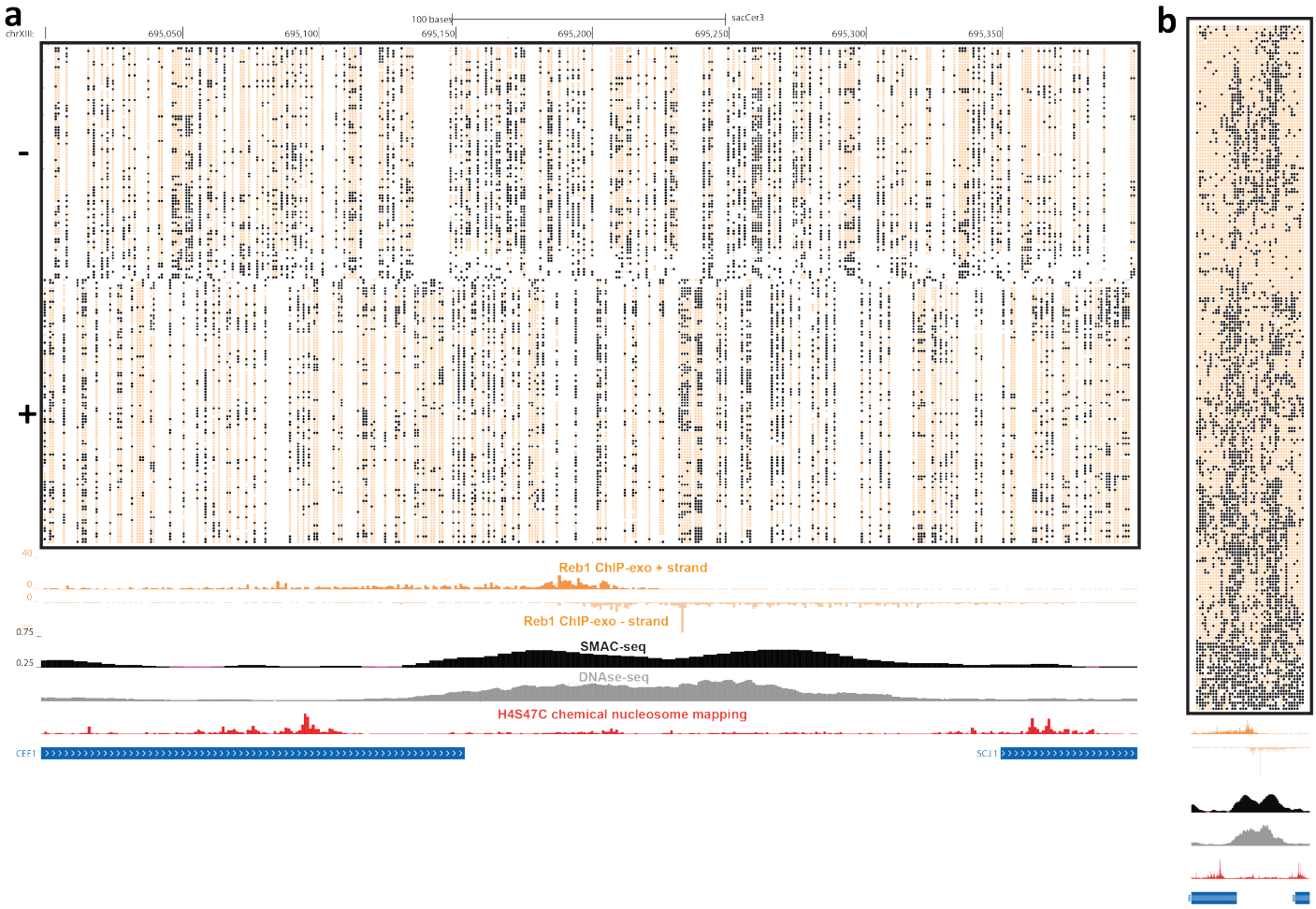
Supplementary Figure 72 *(preceding page)*: The impact of the addition of m6A to SMAC-seq on assay resolution and the potential ability to footprint individual transcription factors. Shown is the fraction of motifs in the genome for each transcription factor in the yeast genome containing the indicated number of informative positions using GC alone, m6A alone, and GC + m6A methyltransferase.



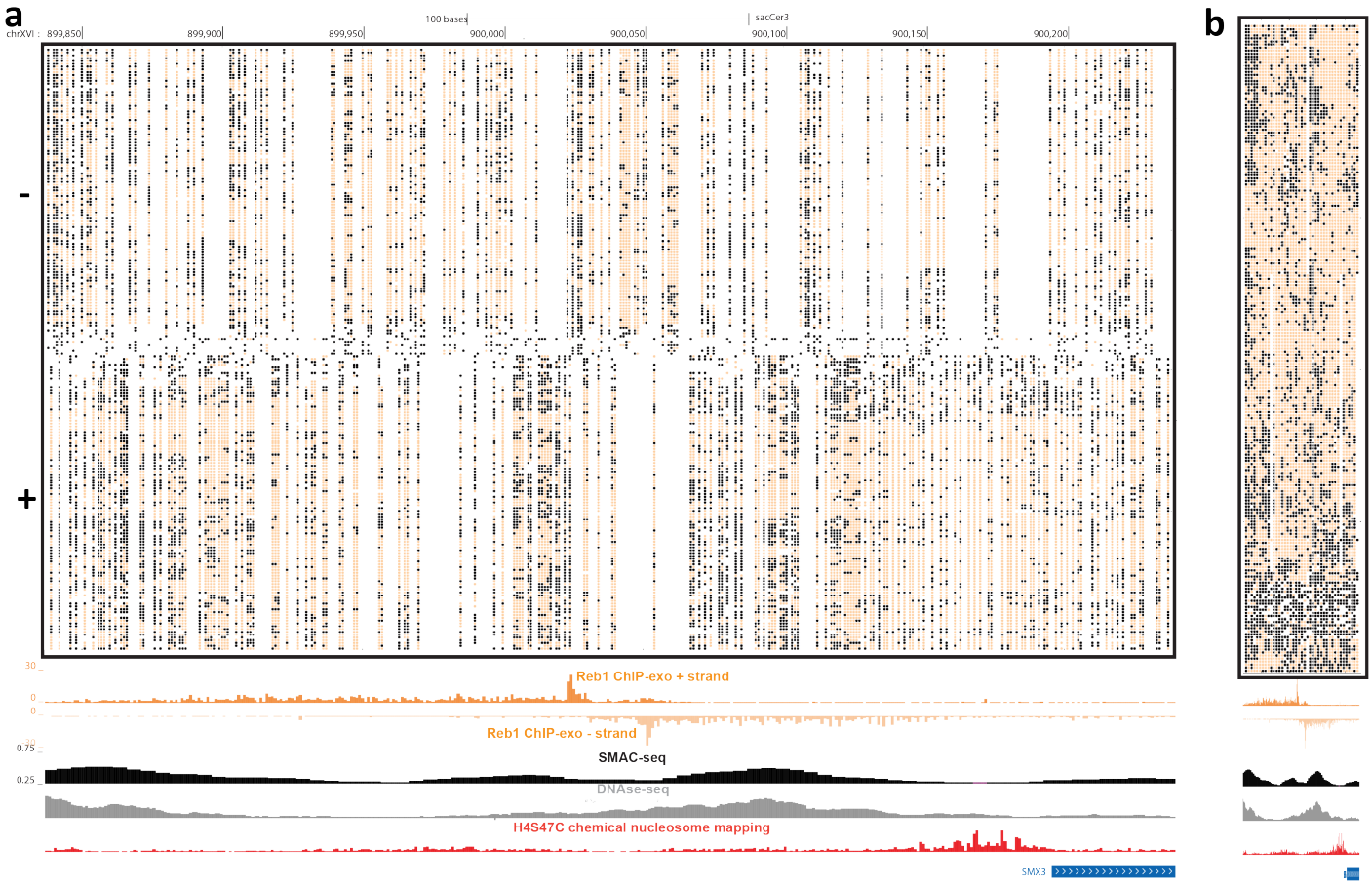
Supplementary Figure 73: Single-molecule footprinting by transcription factors. Shown is the average methylation status (averaged over 10bp) in the neighborhood of occupied (as measured by ChIP-exo or ChIP-seq) recognition motifs for several *S. cerevisiae* DNA binding proteins: (a) Reb1; (b) Rap1; (c) Abf1; (d) Cbf1. (e) ORC1. Strong footprinting is observed for Reb1, Rap1, and ORC1, while Abf1 and Cbf1 occupancy does not appear to be strongly protective against methylation.



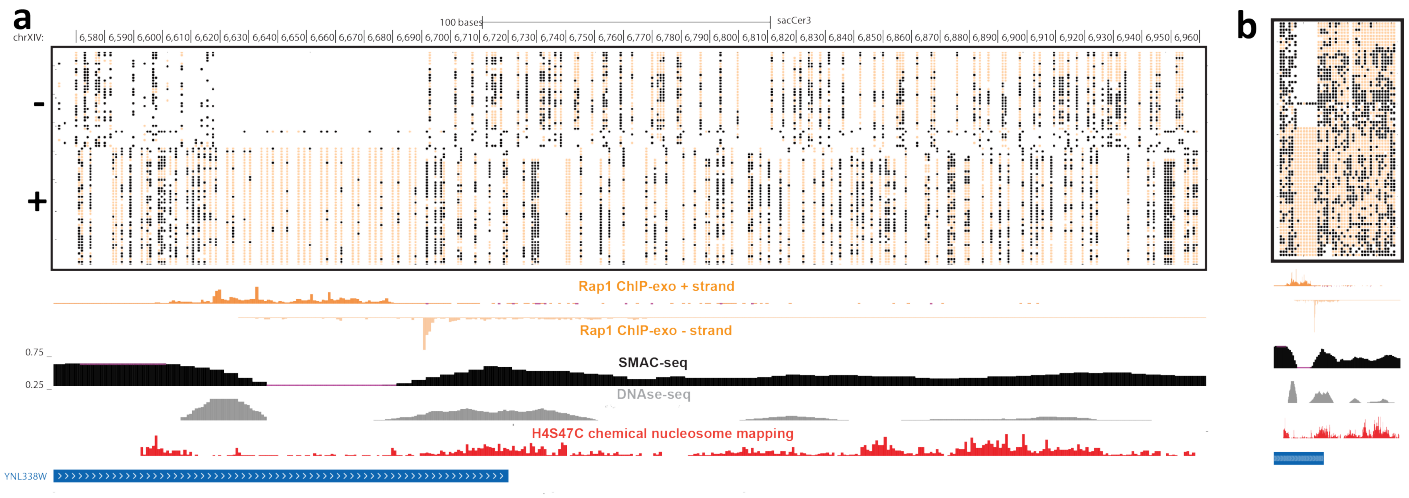
Supplementary Figure 74: Single-molecule footprinting by transcription factors. Shown is the average DNAse-seq (A) and ATAC-seq (B) cut profiles, and the SMAC-seq methylation profile (C) around DNAse footprints identified in yeast by Hesselbreth et al.[7].



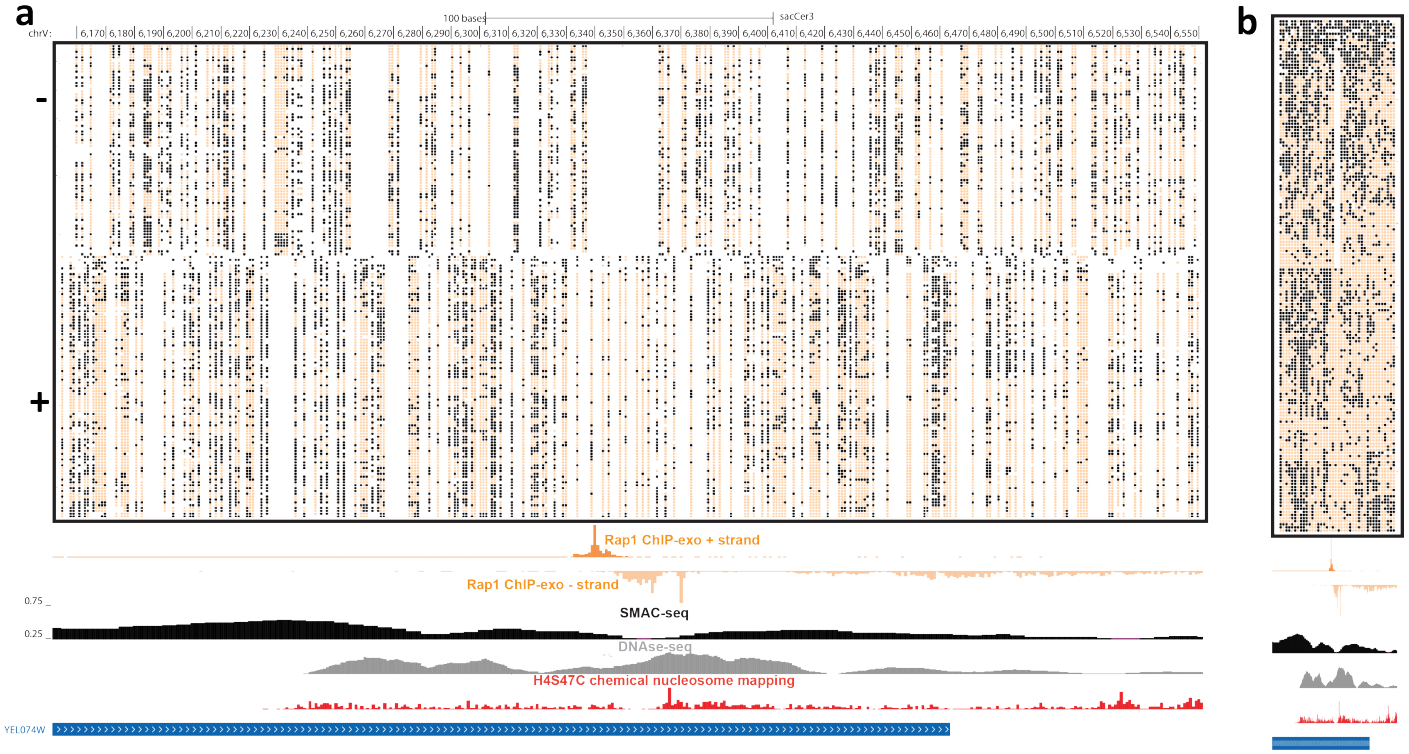
Supplementary Figure 75: Single-molecule footprinting by Reb1 binding sites. (a) Raw unfiltered nanopore reads fully spanning the 400-bp neighborhood of a Reb1 binding site on chrXVIII, at single-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (b) Same as in (a), but at aggregated 10-bp resolution.



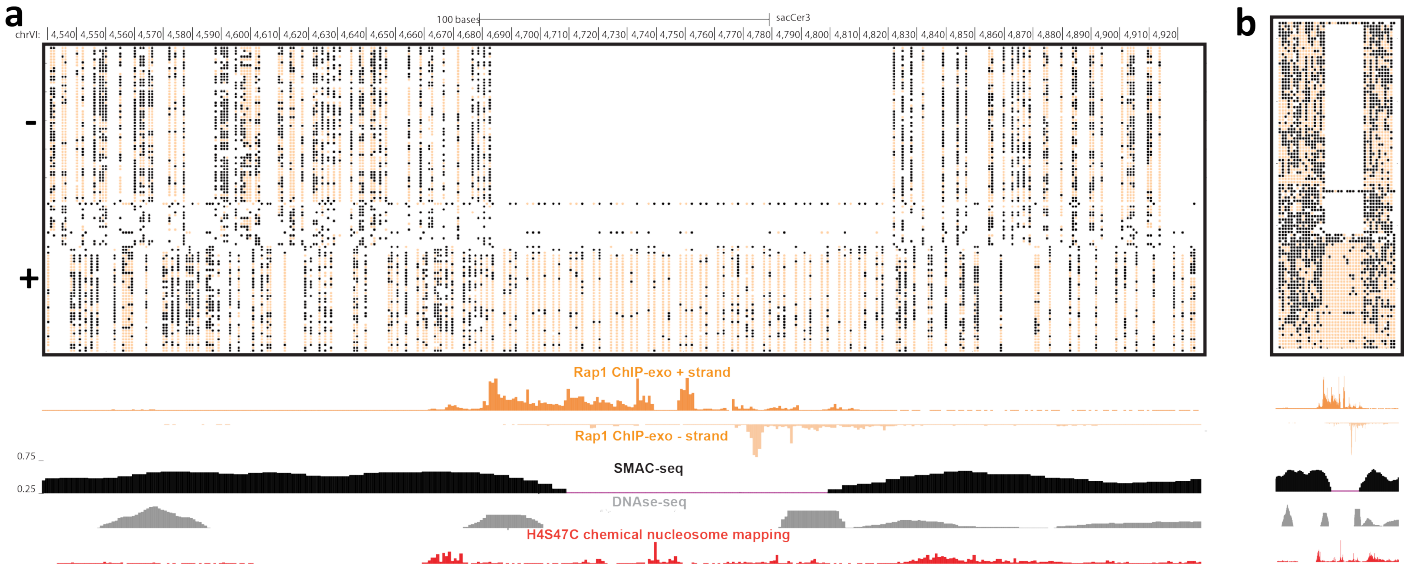
Supplementary Figure 76: Single-molecule footprinting by Reb1 binding sites. (a) Raw unfiltered nanopore reads fully spanning the 400-bp neighborhood of a Reb1 binding site on chrXVIII, at single-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (b) Same as in (a), but at aggregated 10-bp resolution.



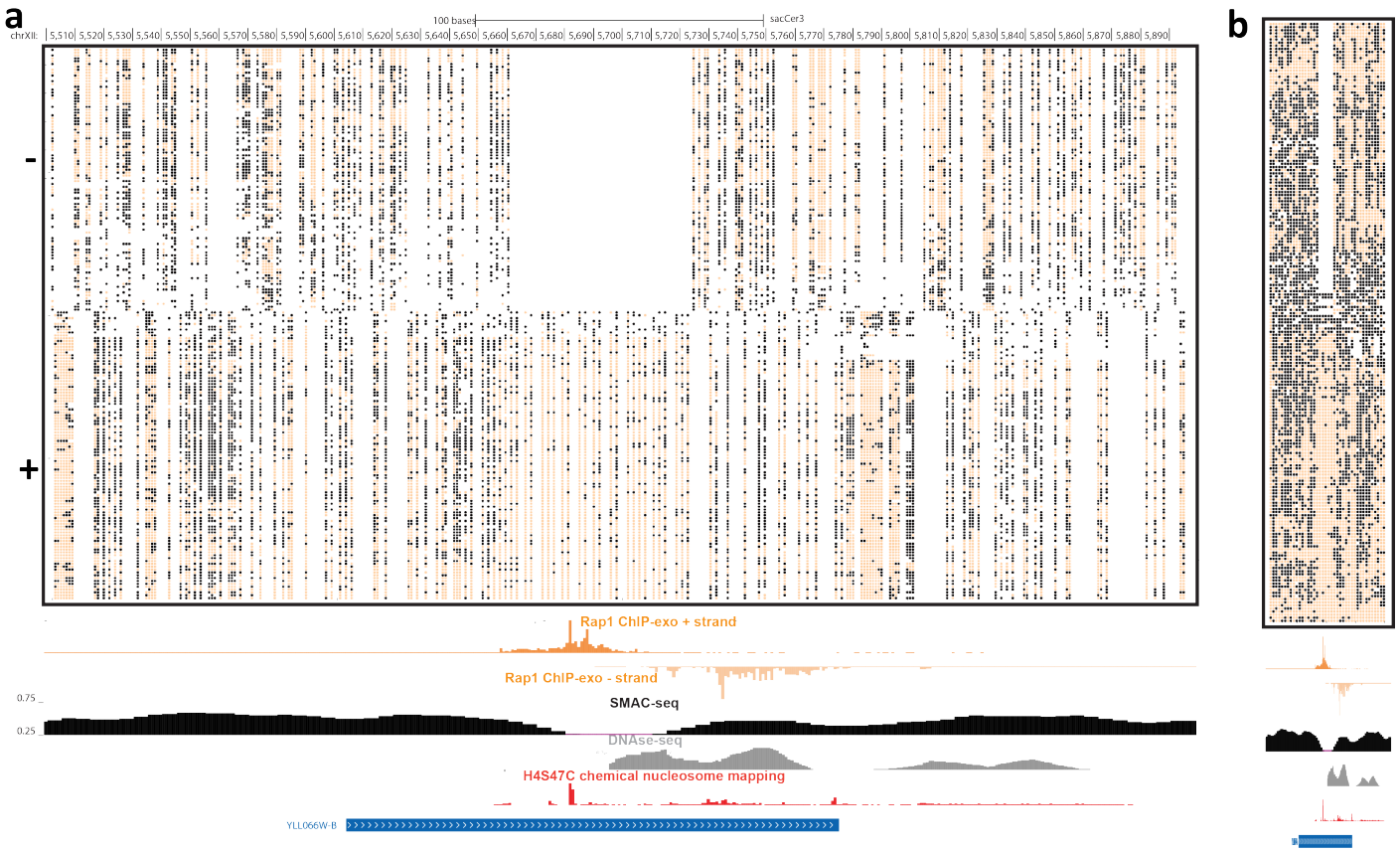
Supplementary Figure 77: Single-molecule footprinting associated with Rap1 occupancy. (a) Raw unfiltered nanopore reads fully spanning a 400-bp neighborhood of the subtelomeric region of chrXIV, at single-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (b) Same as in (a), but at aggregated 10-bp resolution.



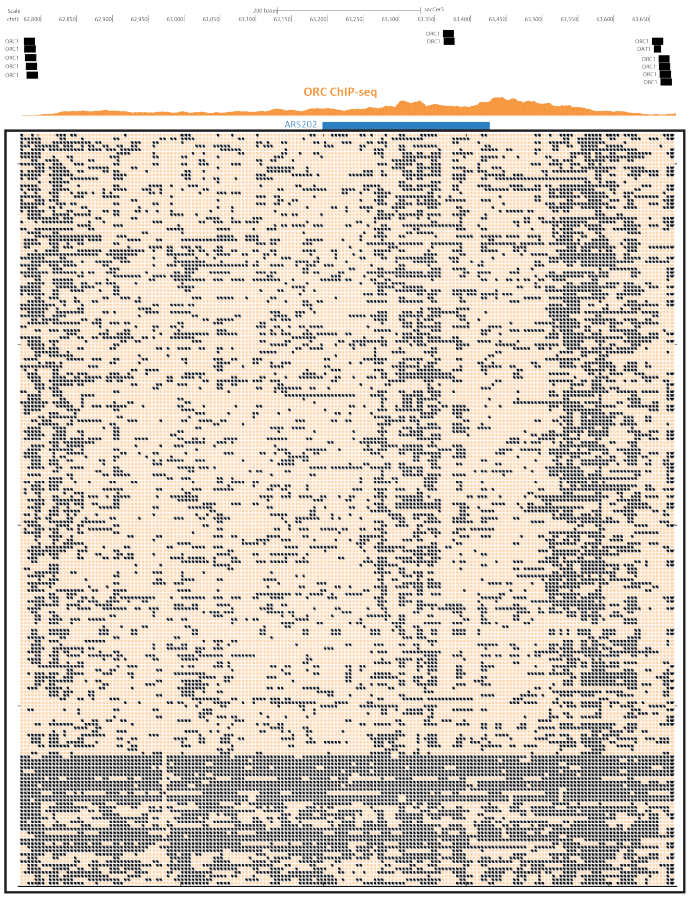
Supplementary Figure 78: Single-molecule footprinting associated with Rap1 occupancy. (a) Raw unfiltered nanopore reads fully spanning a 400-bp neighborhood of the subtelomeric region of chrXV, at single-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (b) Same as in (a), but at aggregated 10-bp resolution.



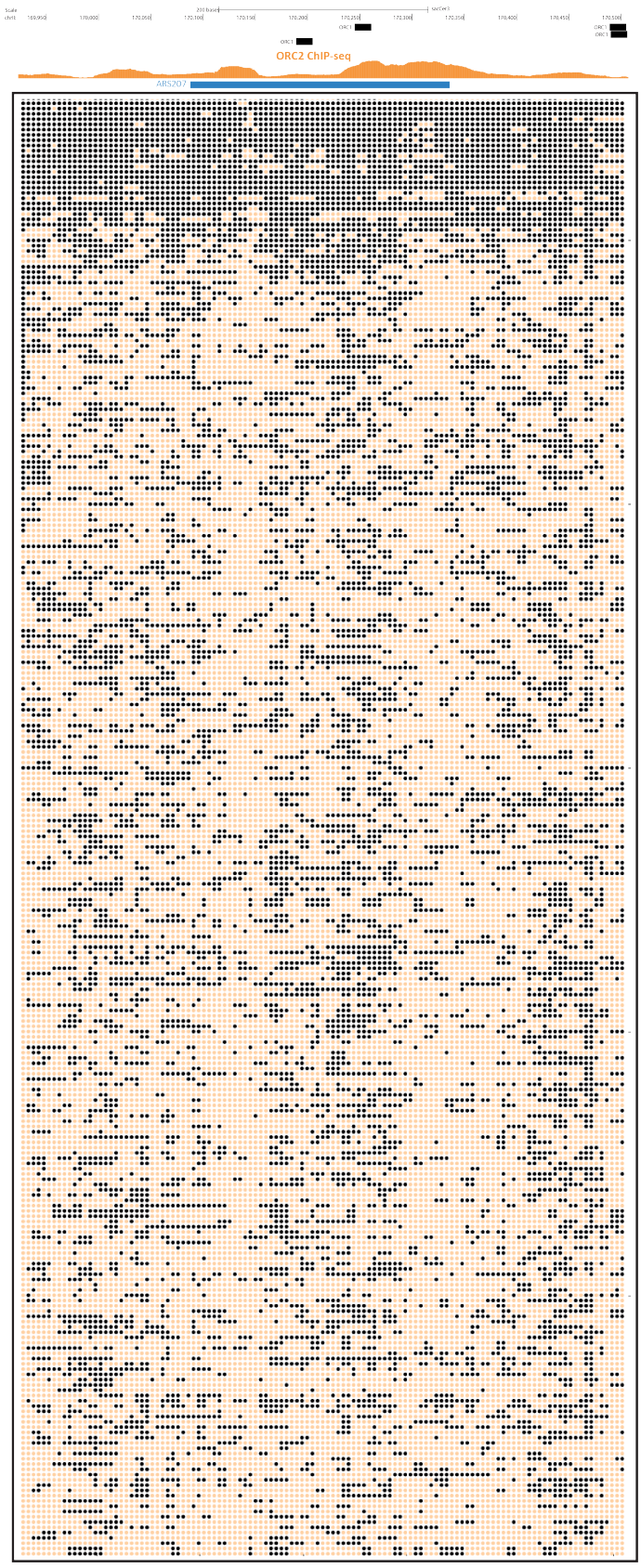
Supplementary Figure 79: Single-molecule footprinting associated with Rap1 occupancy. (a) Raw unfiltered nanopore reads fully spanning a 400-bp neighborhood of the subtelomeric region of chrVI, at single-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (b) Same as in (a), but at aggregated 10-bp resolution.



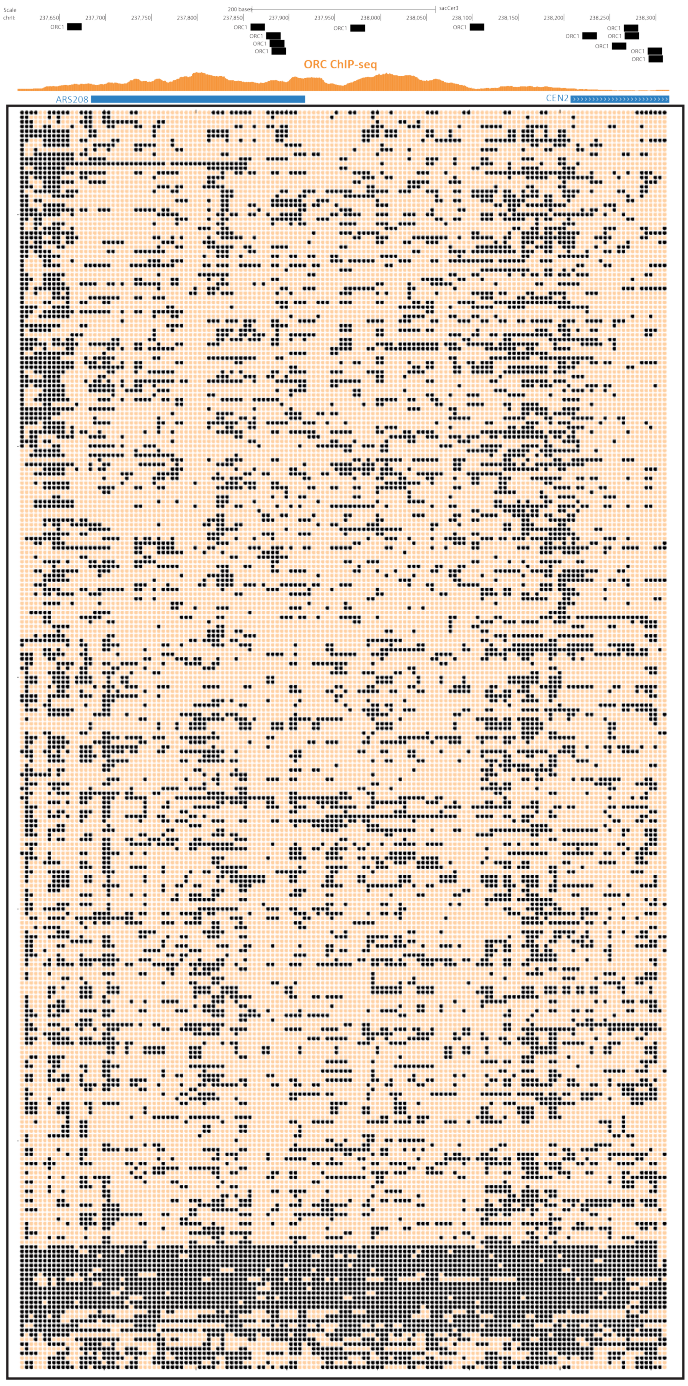
Supplementary Figure 80: Single-molecule footprinting associated with Rap1 occupancy. (a) Raw unfiltered nanopore reads fully spanning a 400-bp neighborhood of the subtelomeric region of chrXII, at single-bp resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A). (b) Same as in (a), but at aggregated 10-bp resolution.



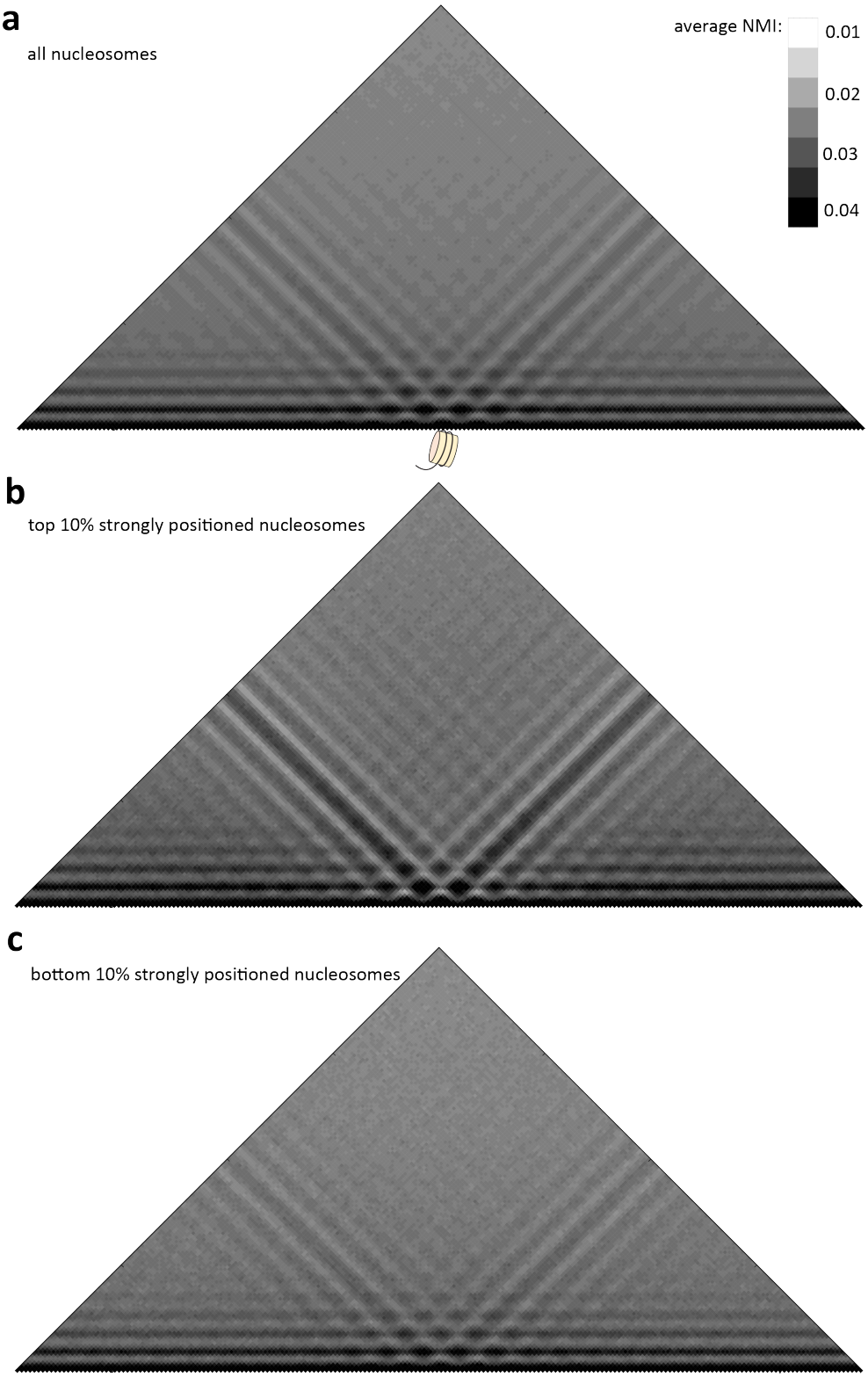
Supplementary Figure 81: Single-molecule footprinting associated with ORC occupancy. (a) Raw unfiltered nanopore reads fully spanning the neighborhood of an ARS site on chrII, at 5-bp aggregated resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A).



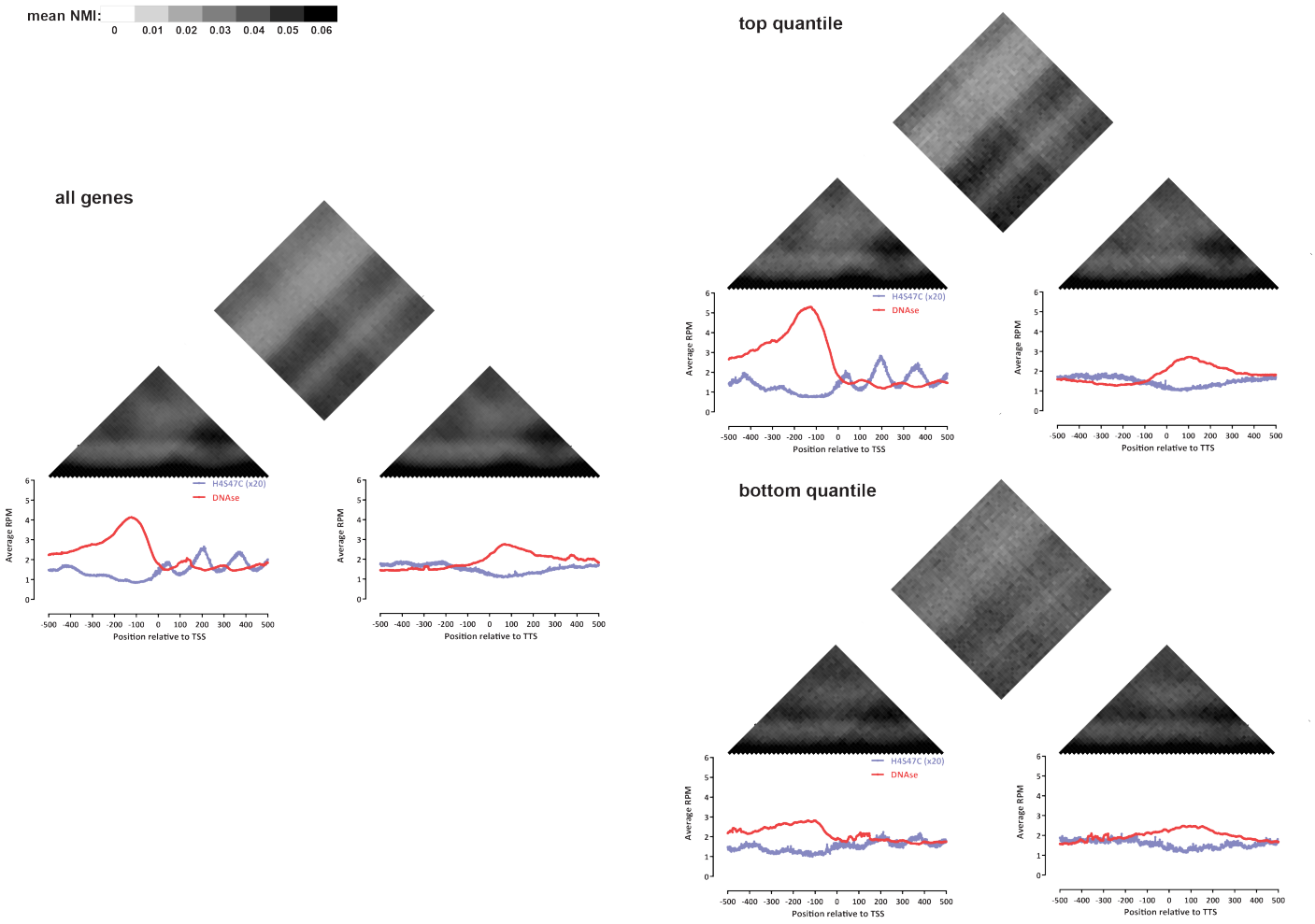
Supplementary Figure 82: Single-molecule footprinting associated with ORC occupancy. (a) Raw unfiltered nanopore reads fully spanning the neighborhood of an ARS site on chrII, at 5-bp aggregated resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A).



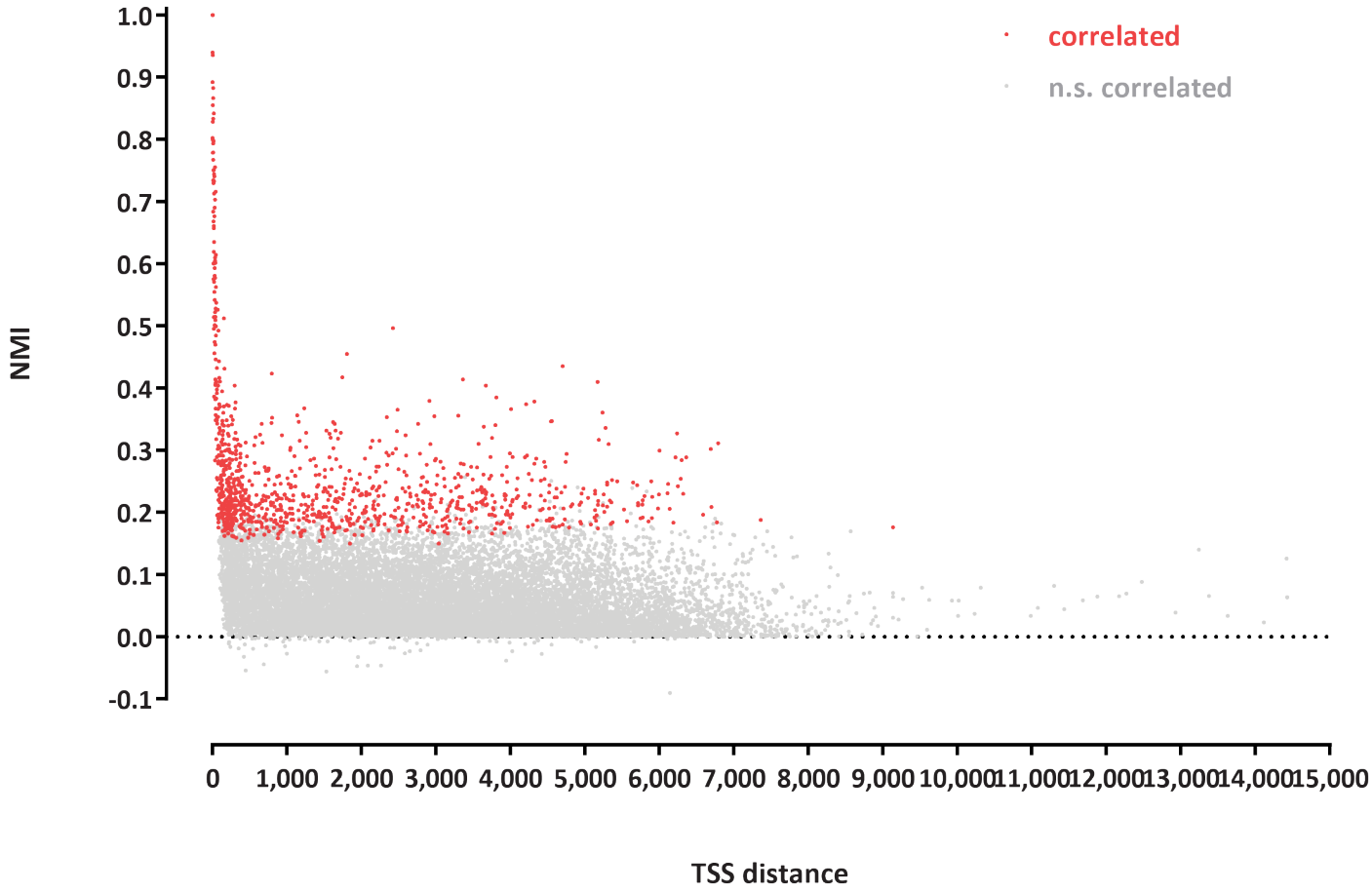
Supplementary Figure 83: Single-molecule footprinting associated with ORC occupancy. (a) Raw unfiltered nanopore reads fully spanning the neighborhood of an ARS site on chrII, at 5-bp aggregated resolution. White spaces indicate positions for which there is no data (i.e. no CG, GC or A).



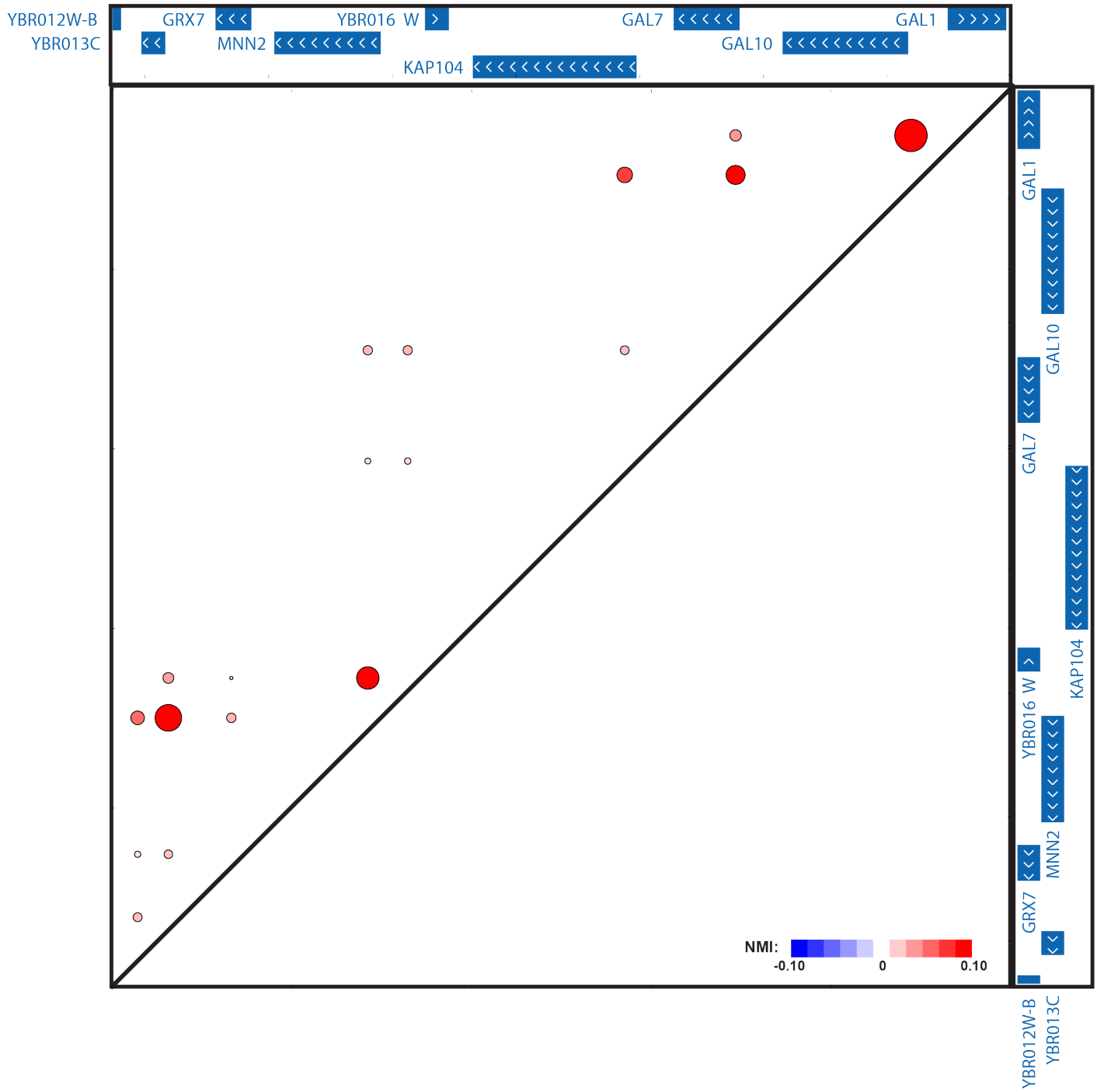
Supplementary Figure 84: Metanucleosome NMI profiles in the yeast genome. Shown are average NMI maps between all 20-bp segments centered on each positioned nucleosome in the genome (a), the top 10% strongly positioned nucleosomes (b), or the top 10% strongly positioned nucleosomes (c).



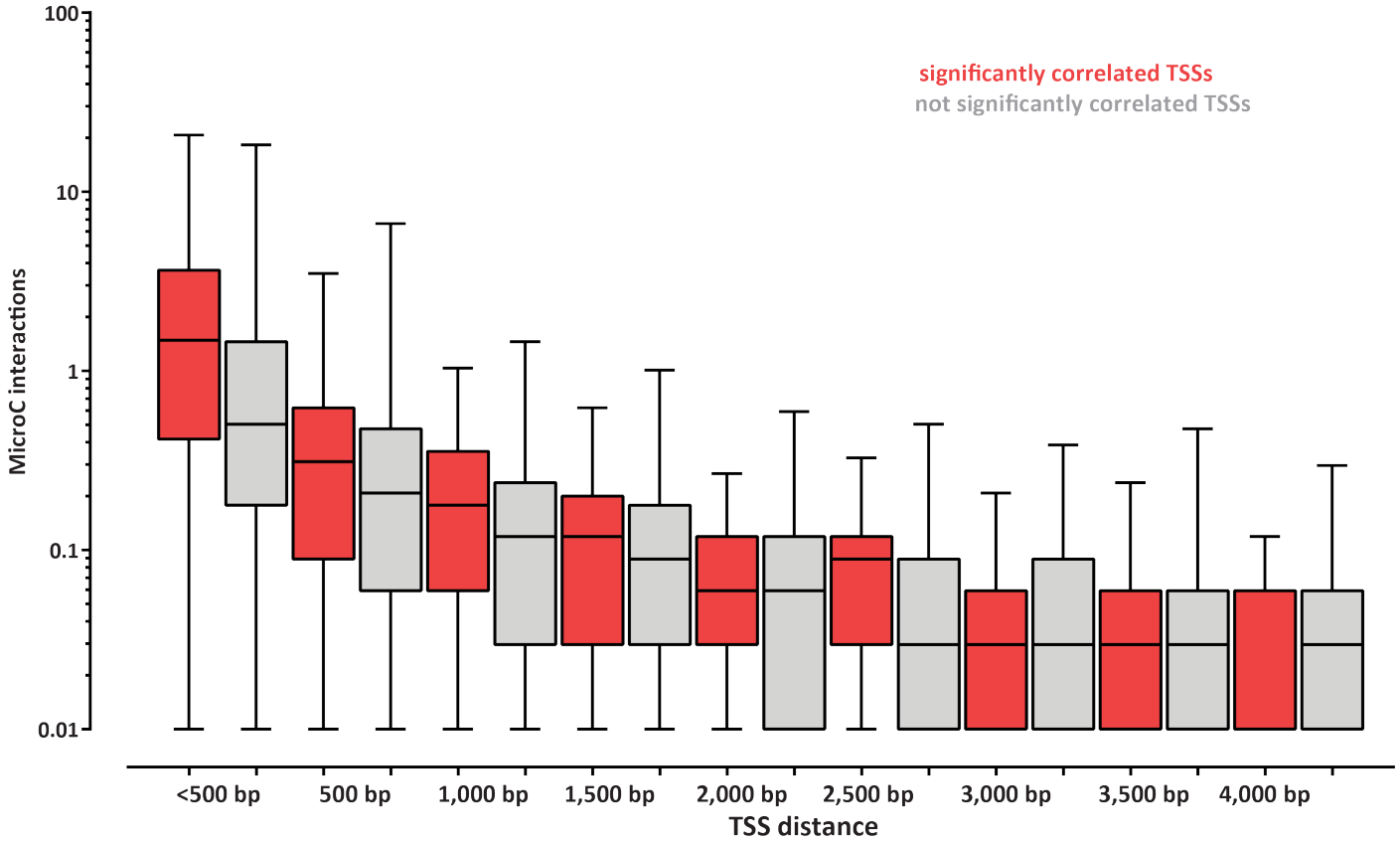
Supplementary Figure 85: Patterns of coaccessibility between the 5′ and 3′ ends of genes. Shown is the average NMI for the ±500bp regions in the 5′ and 3′ end of all yeast genes as well as the top and bottom 20% expression-ranked genes (calculated over 10-bp windows). Only genes ≥1000 bp in length are shown. Similar results are obtained using windows of size 20bp or 50bp (data not shown).



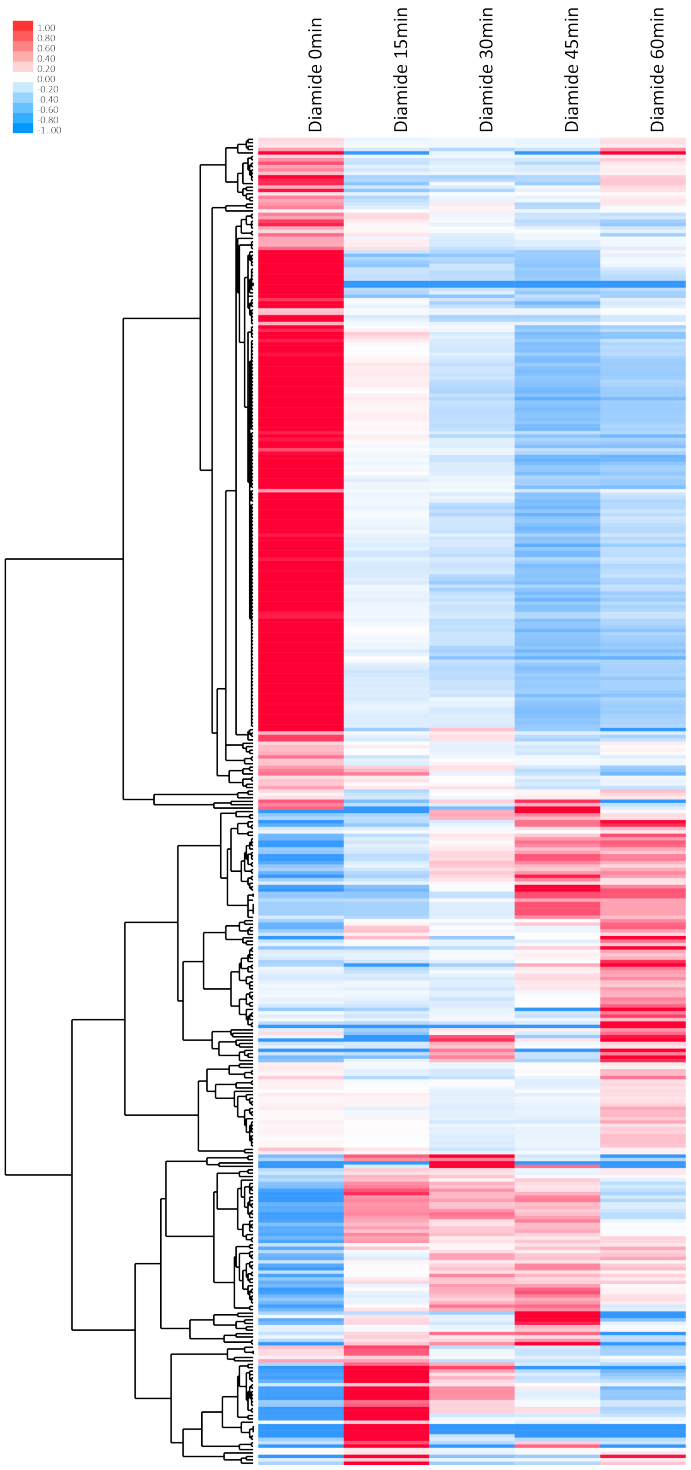
Supplementary Figure 86: Accessibility correlation between TSSs in the yeast genome. Shown are NMI values for each pair of significantly and non-significantly correlated TSSs (defined as the regions ±100 bp around the TSS).



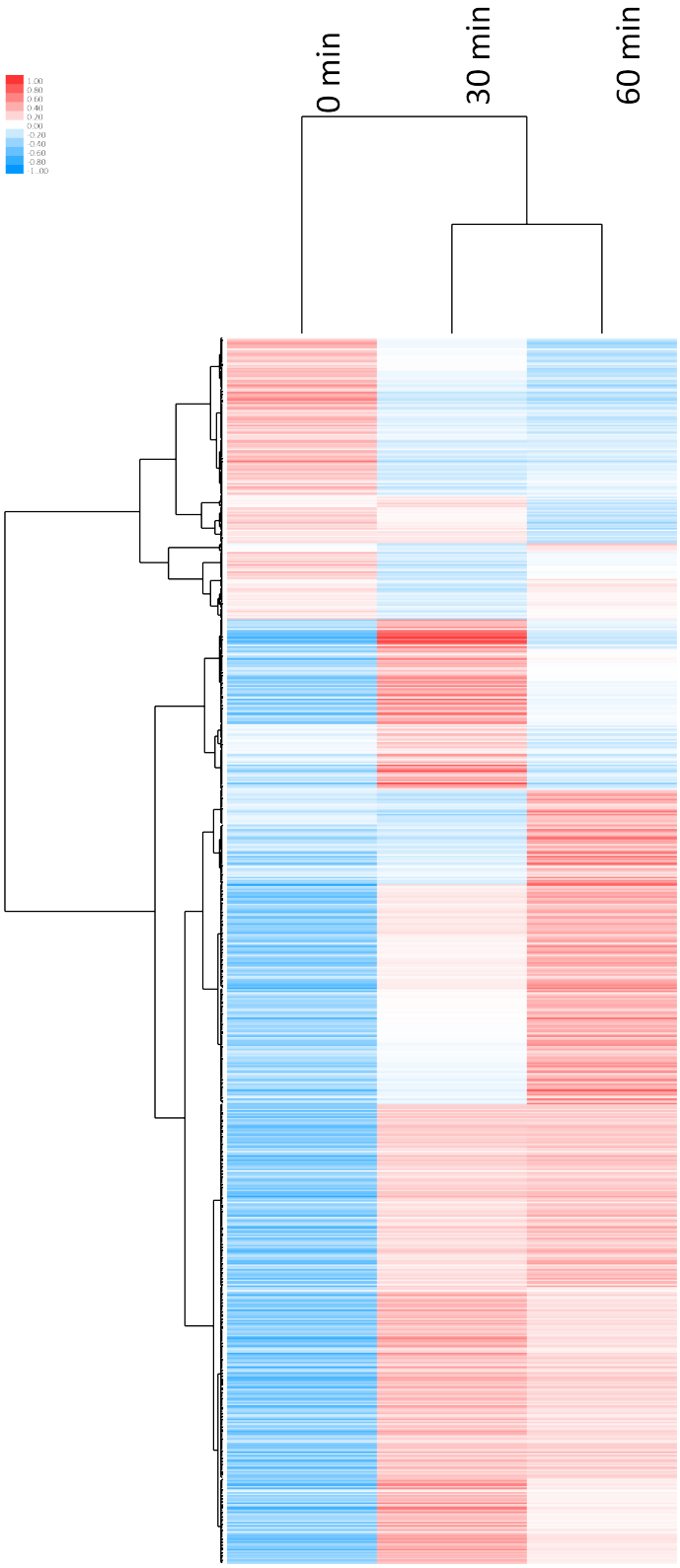
Supplementary Figure 87: Example of accessibility correlation between TSSs. Shown are accessibility correlations between all loci in the region with coordinates "chrII:265000-280000". The size and color of the dots corresponds to their NMI scores.



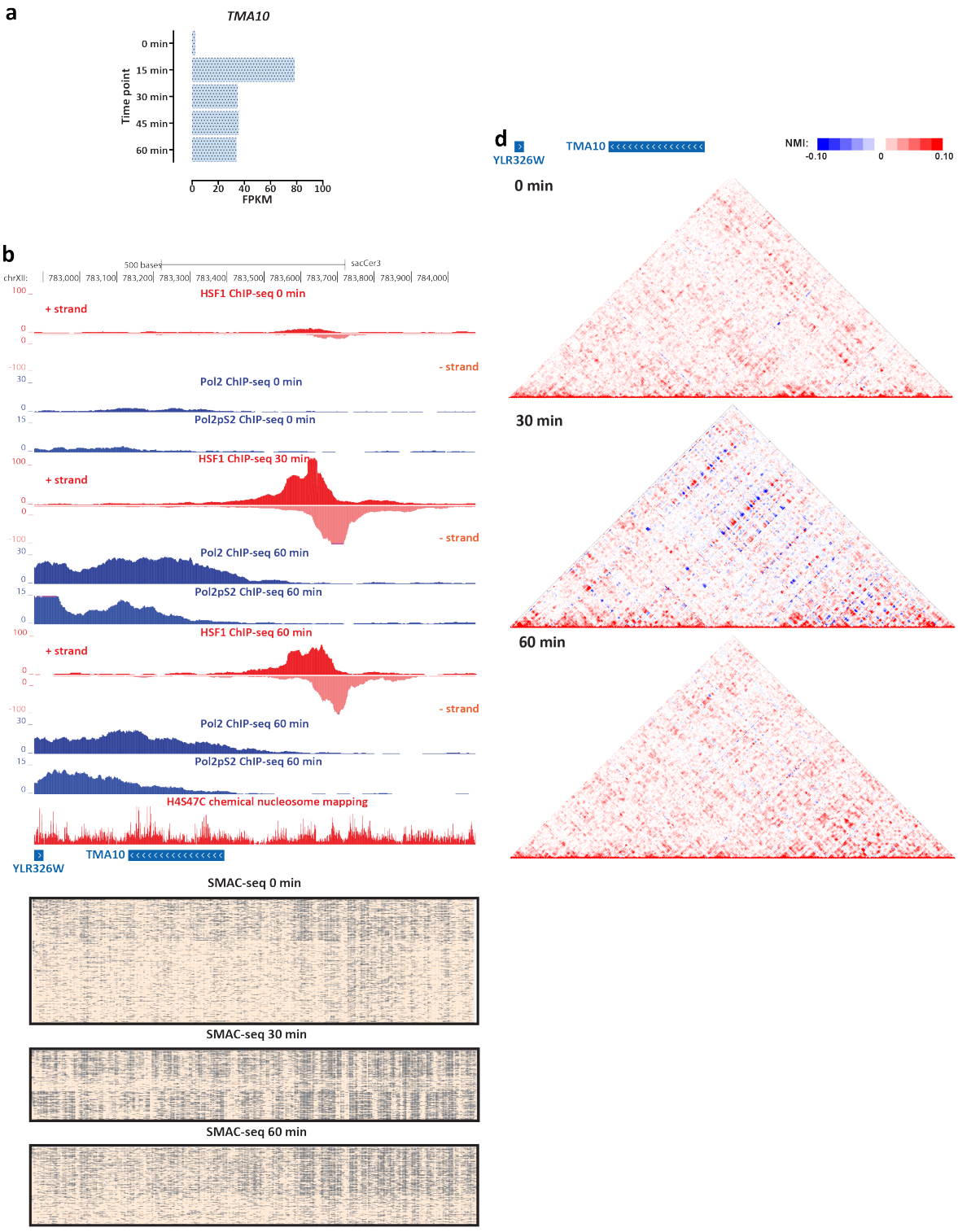
Supplementary Figure 88: Accessibility correlation between TSSs and 3D interactions. Shown are significantly and non-significantly correlated TSSs (defined as the regions ±100 bp around the TSS) split into distance bins and the number of 3D interactions between each group (measured by MicroC).



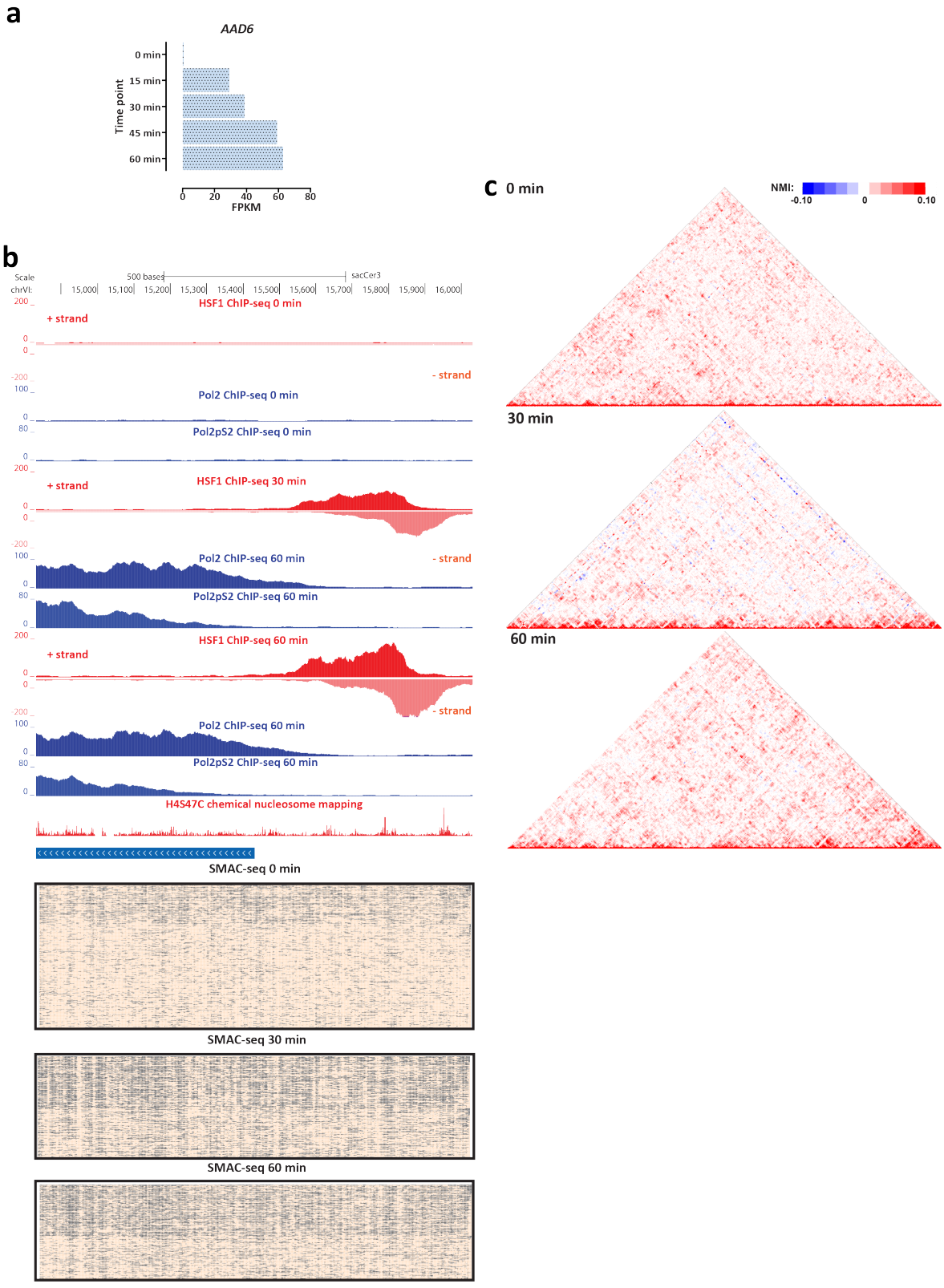
Supplementary Figure 89: Gene expression changes upon diamide treatment. Shown is RNA-seq data (mean and unit-variance normalized across time points) for all genes expressed at ≥50 FPKM in at least one time point.



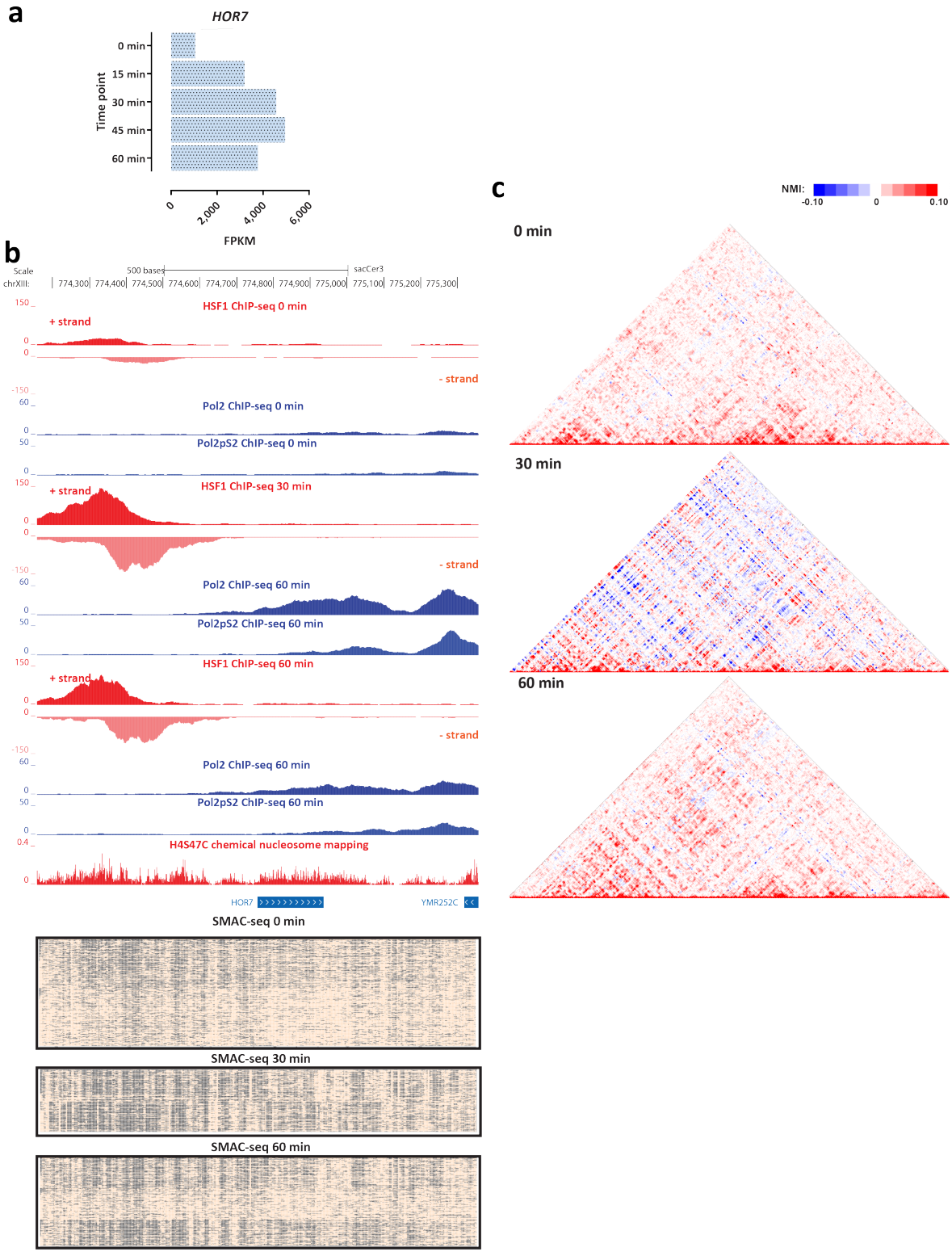
Supplementary Figure 90: Dynamic changes in HSF1 occupancy upon diamide treatment. Shown are ChIP-seq RPMs (mean and unit-variance normalized across time points) for all HSF1 peaks detected in least one time point.



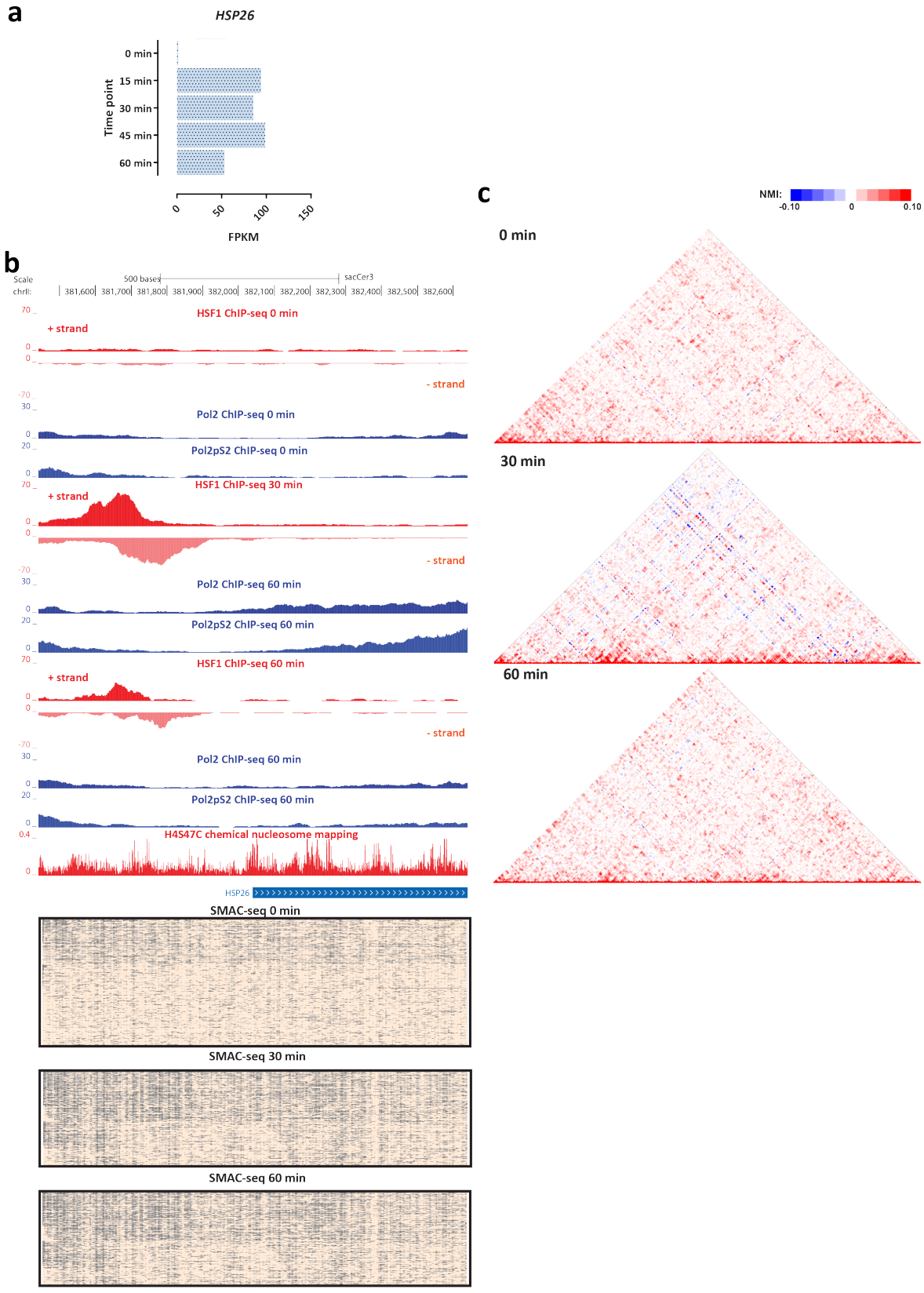
Supplementary Figure 91: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *TMA10* gene.



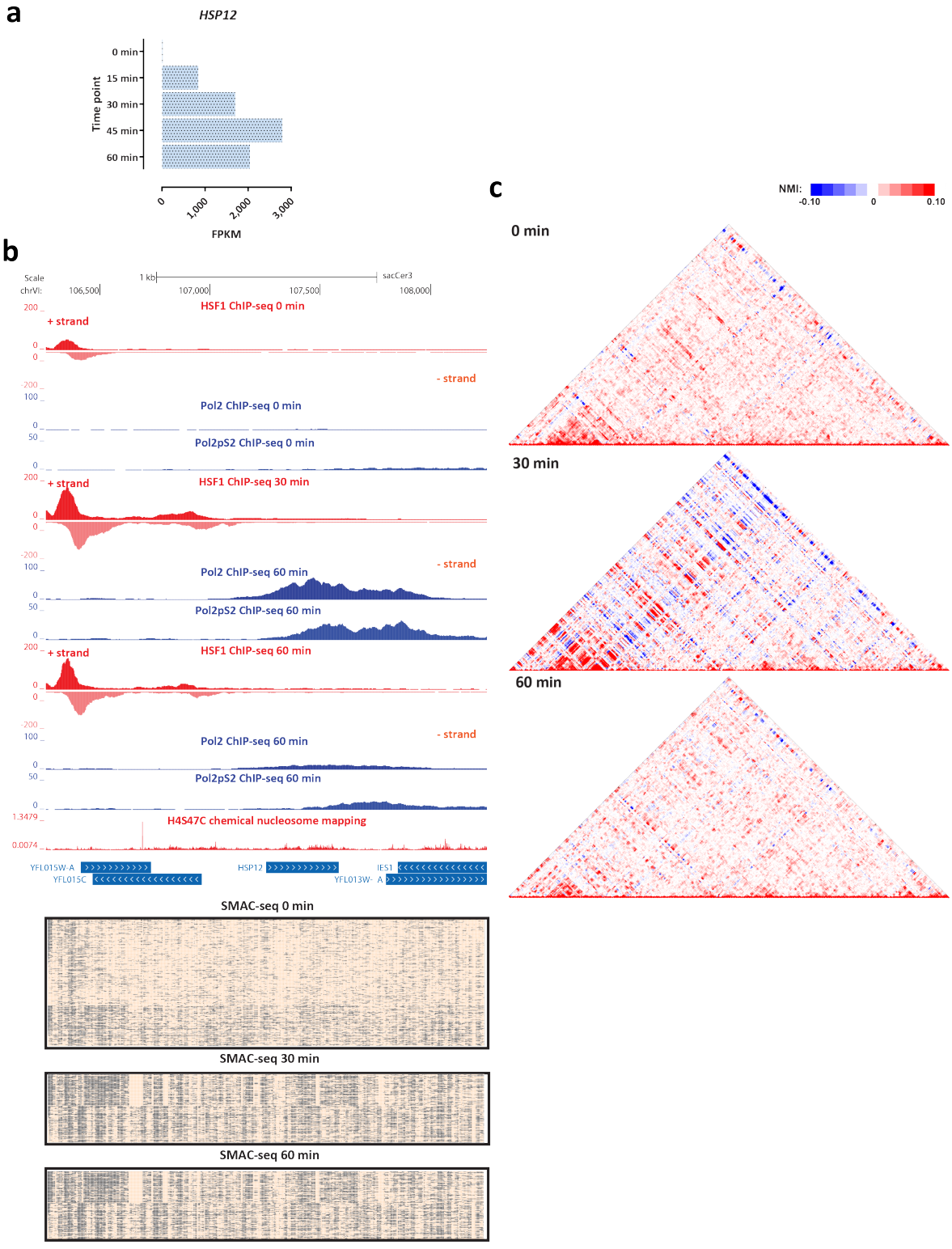
Supplementary Figure 92: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *AAD6* gene.



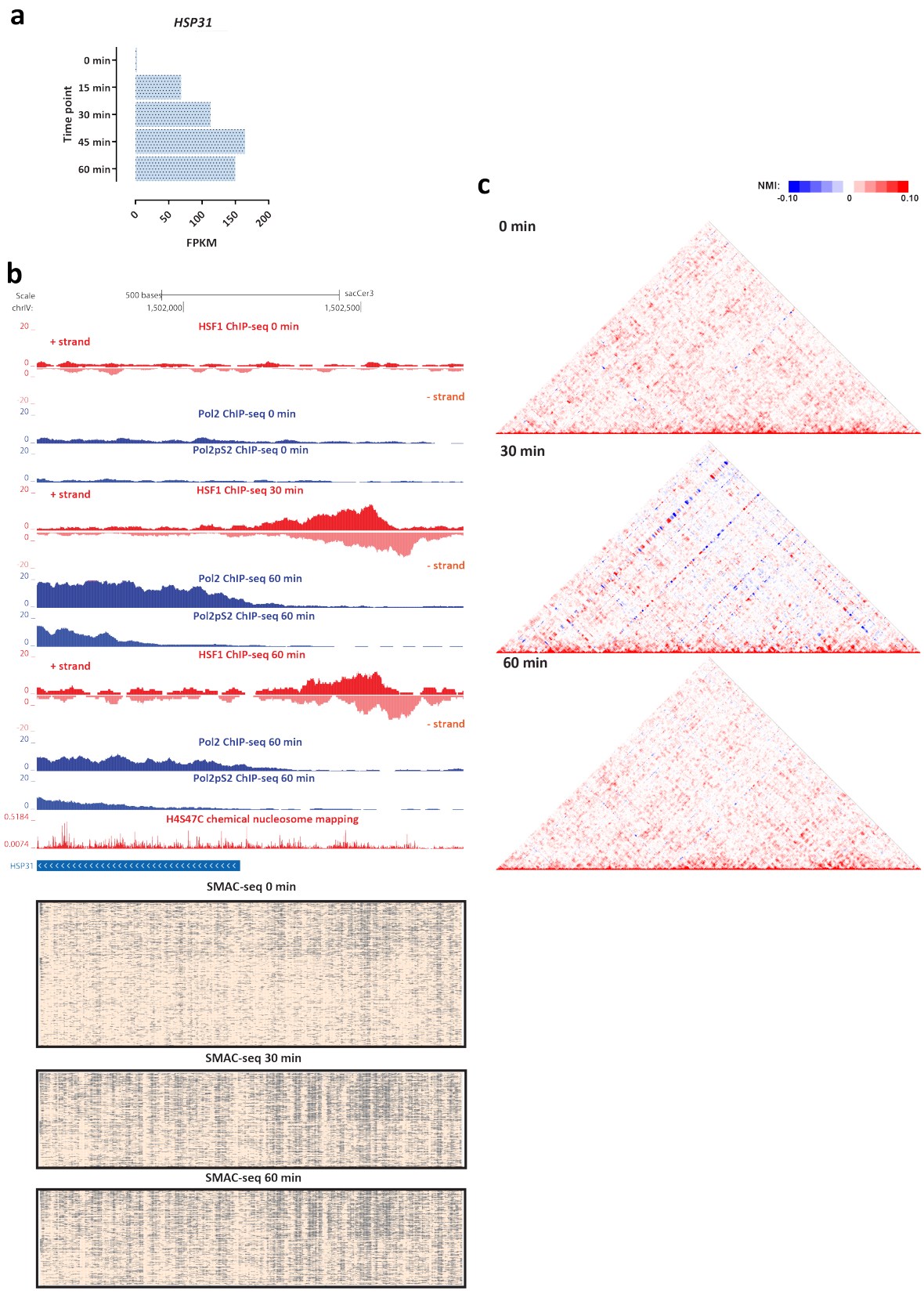
Supplementary Figure 93: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *HOR7* gene.



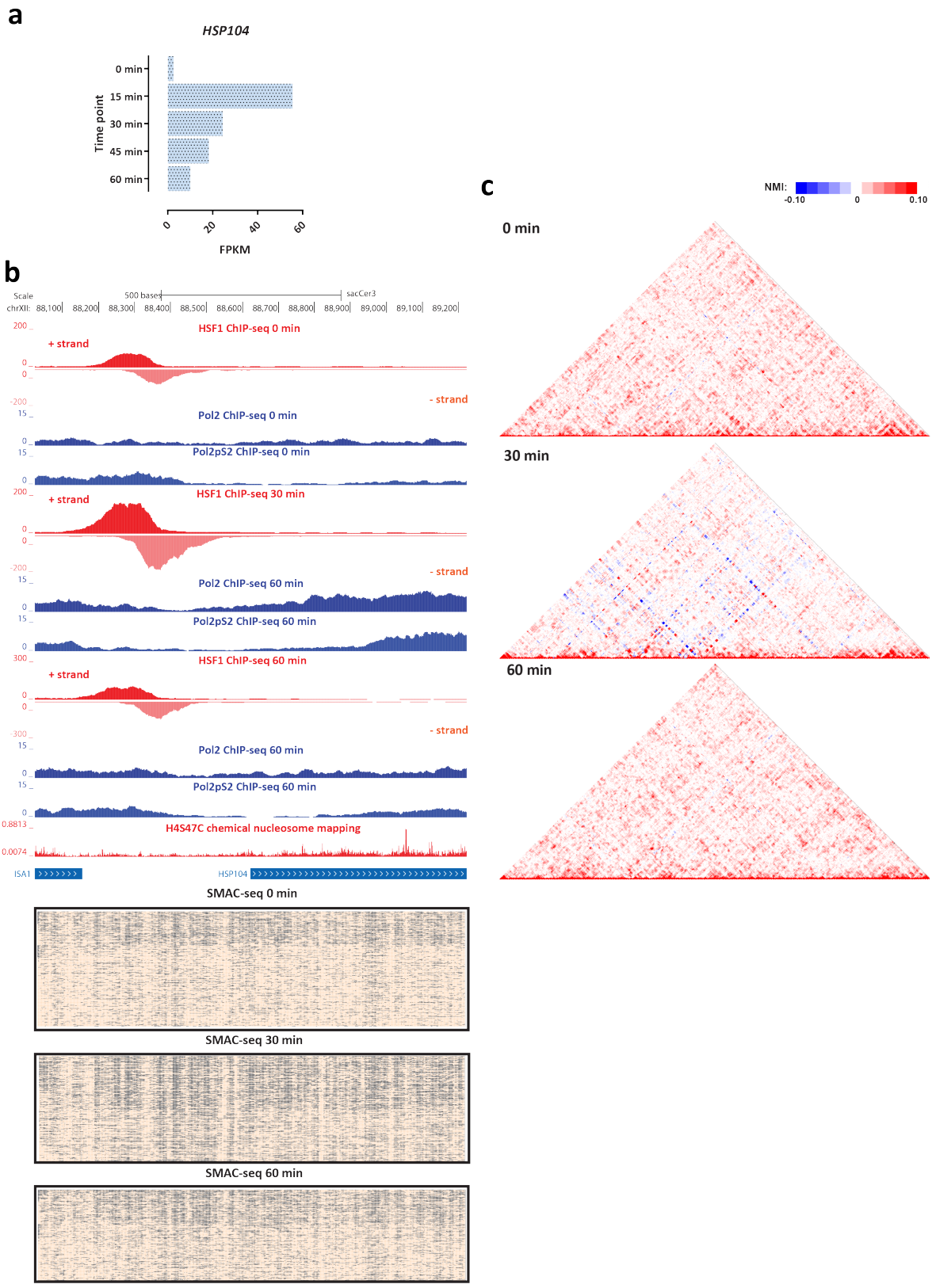
Supplementary Figure 94: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *HSP26* gene.



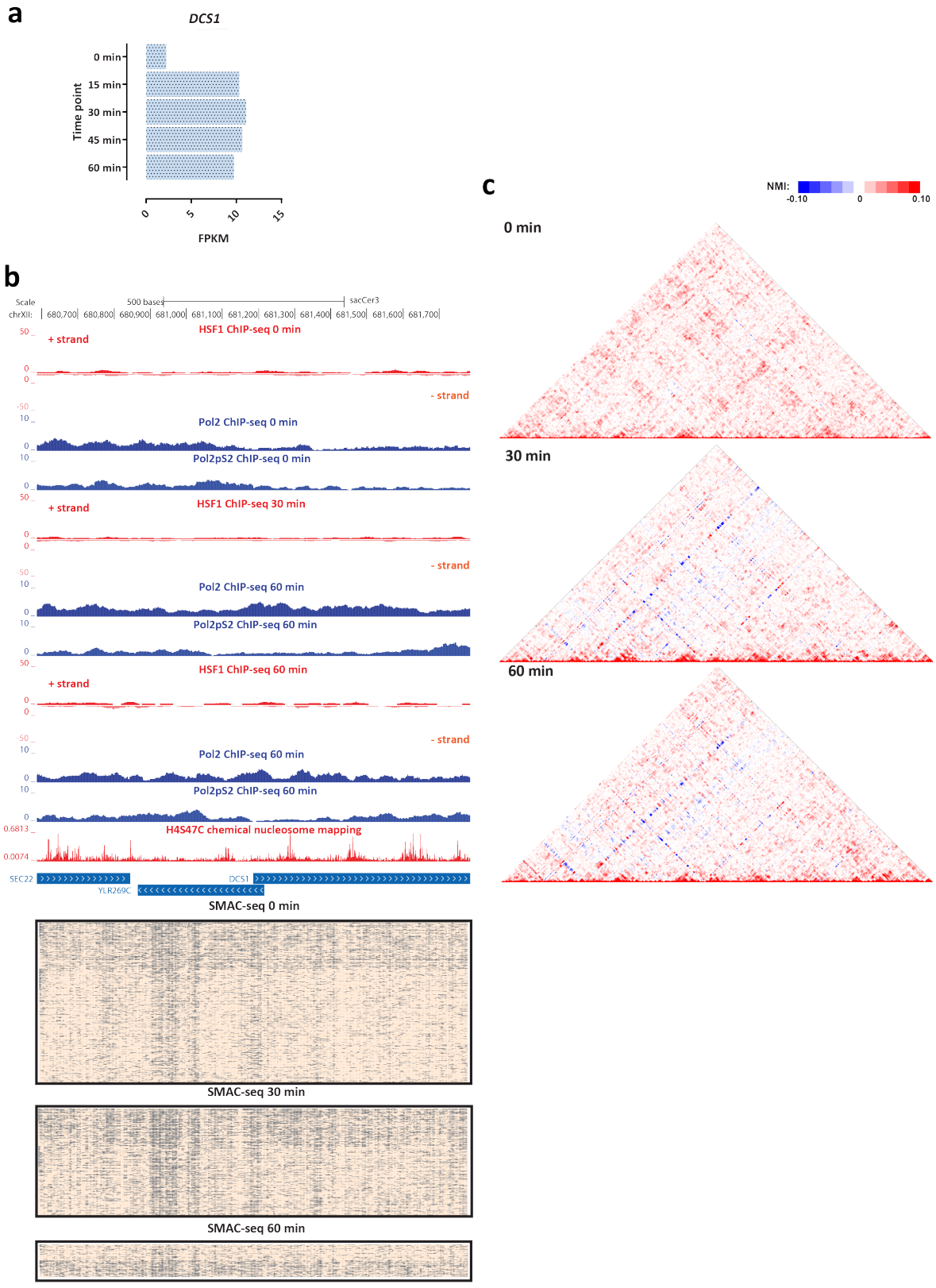
Supplementary Figure 95: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *HSP12* gene.



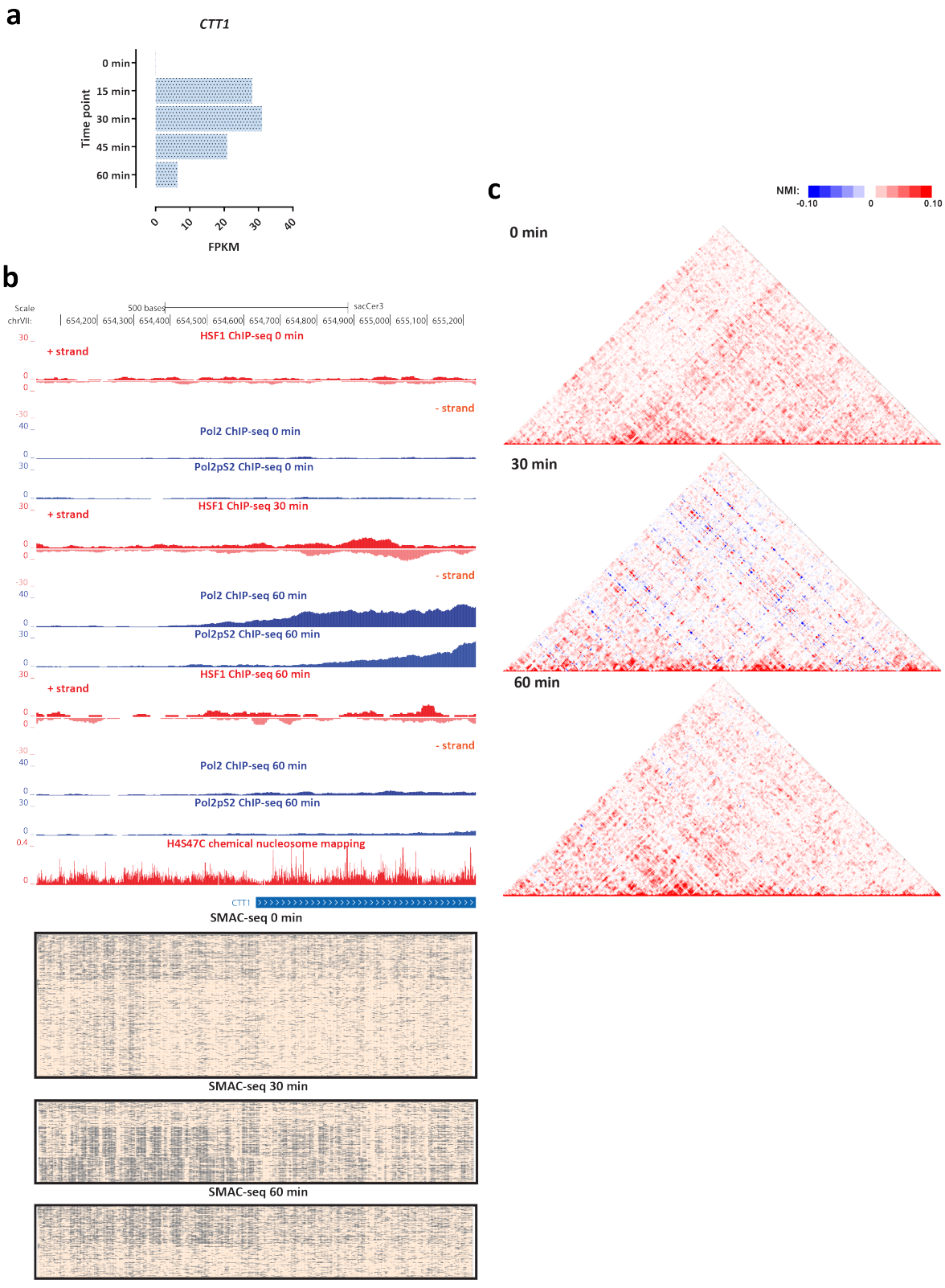
Supplementary Figure 96: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *HSP31* gene.



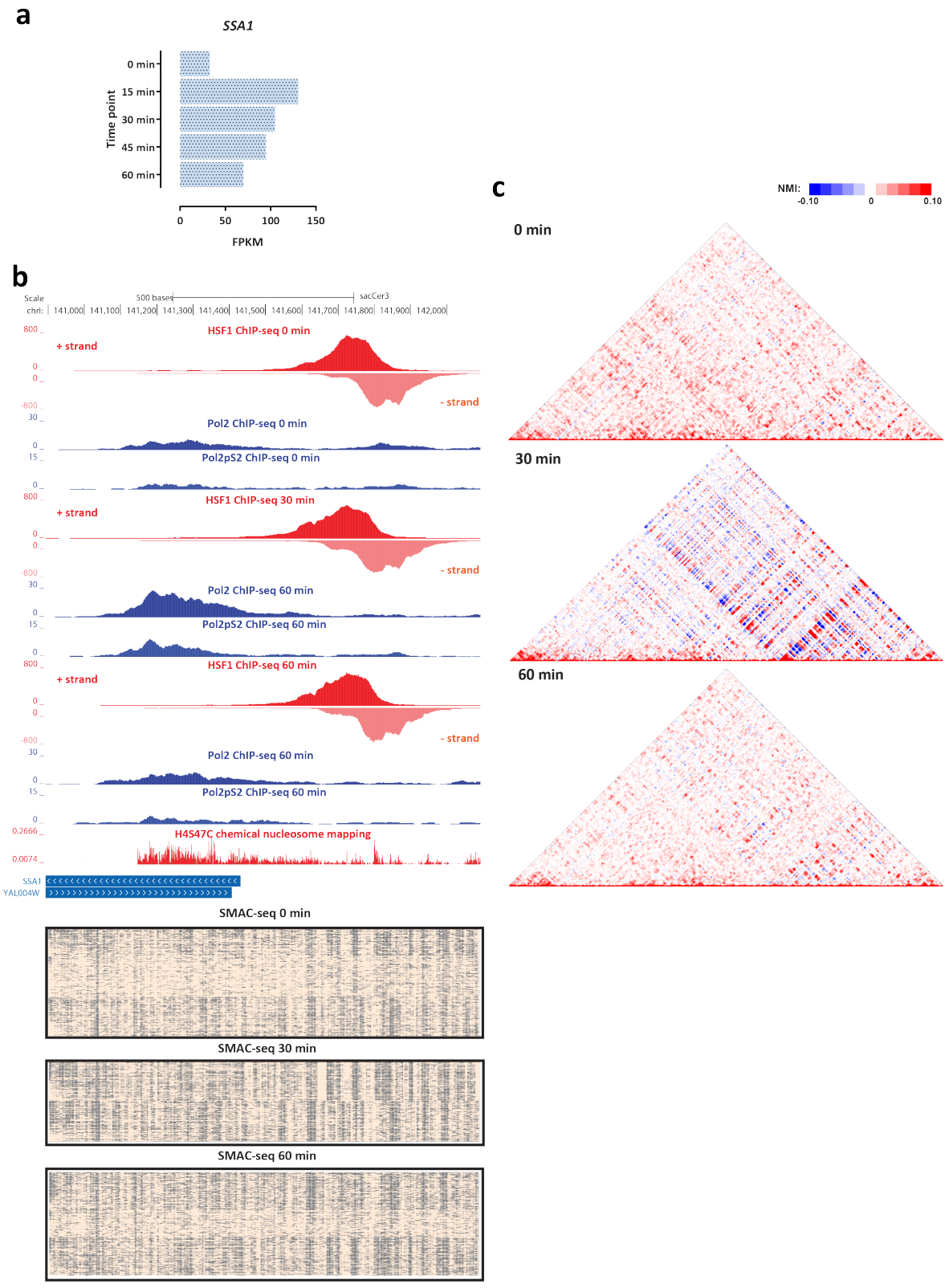
Supplementary Figure 97: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *HSP104* gene.



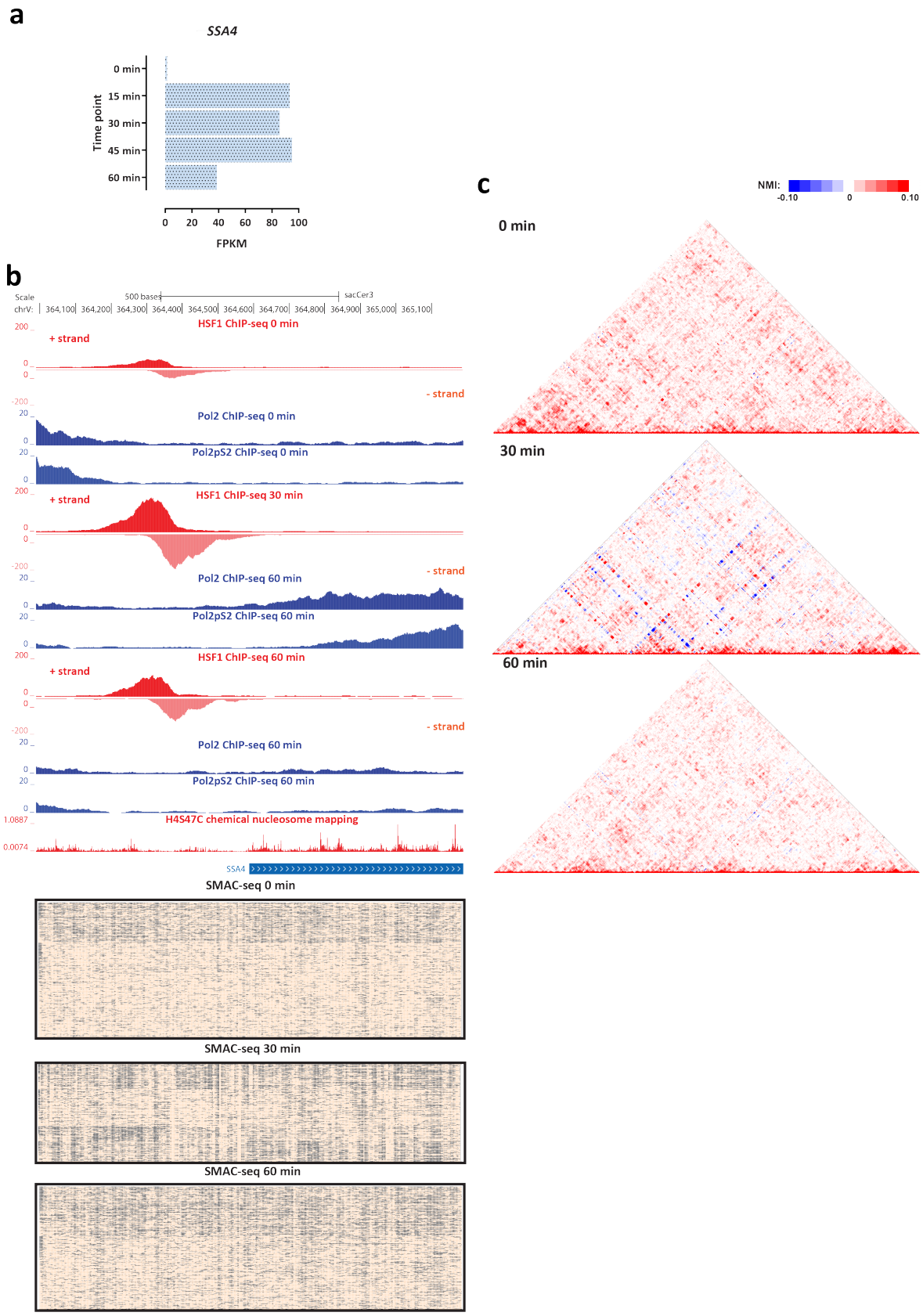
Supplementary Figure 98: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *DSC1* gene.



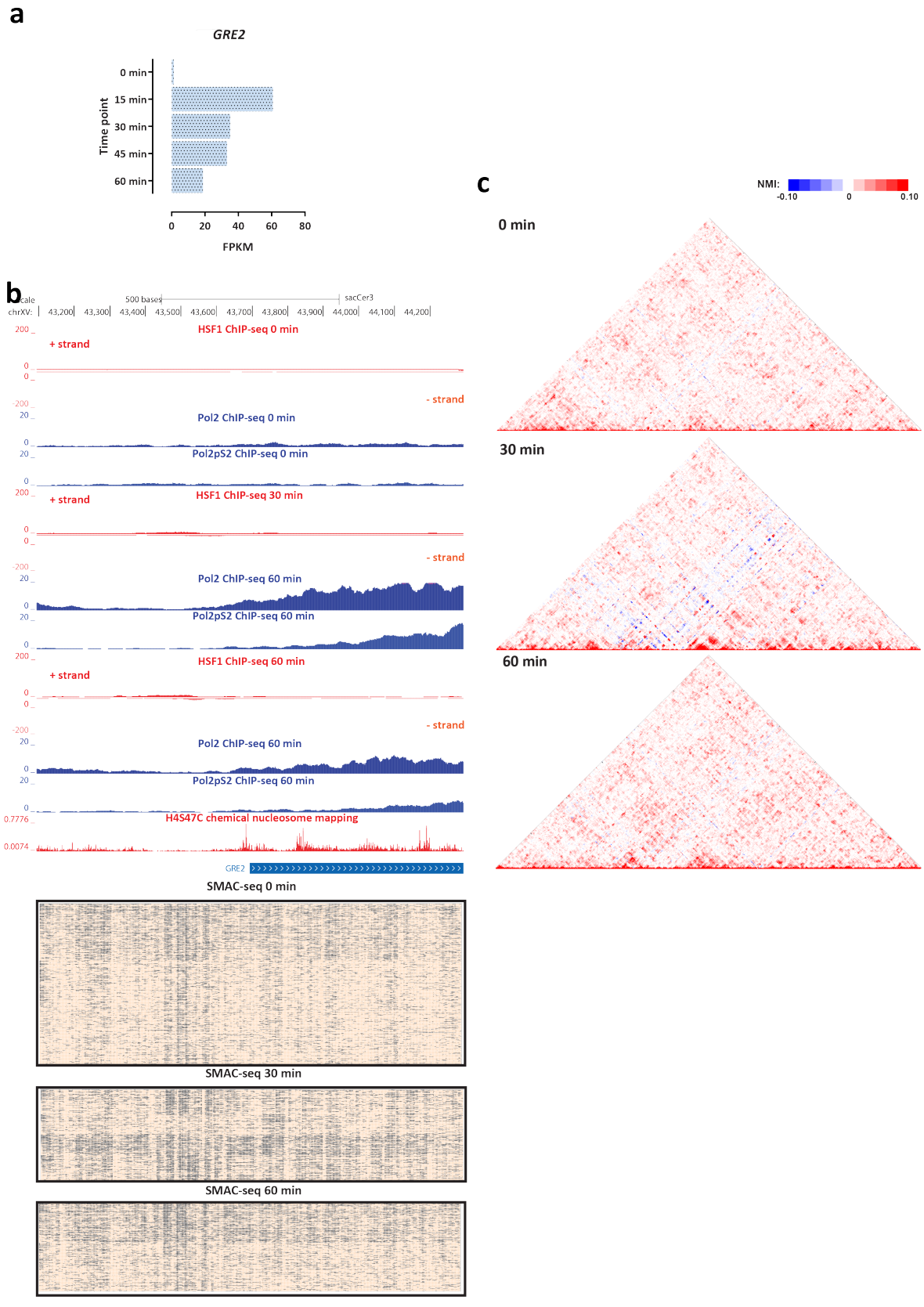
Supplementary Figure 99: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *CTT1* gene.



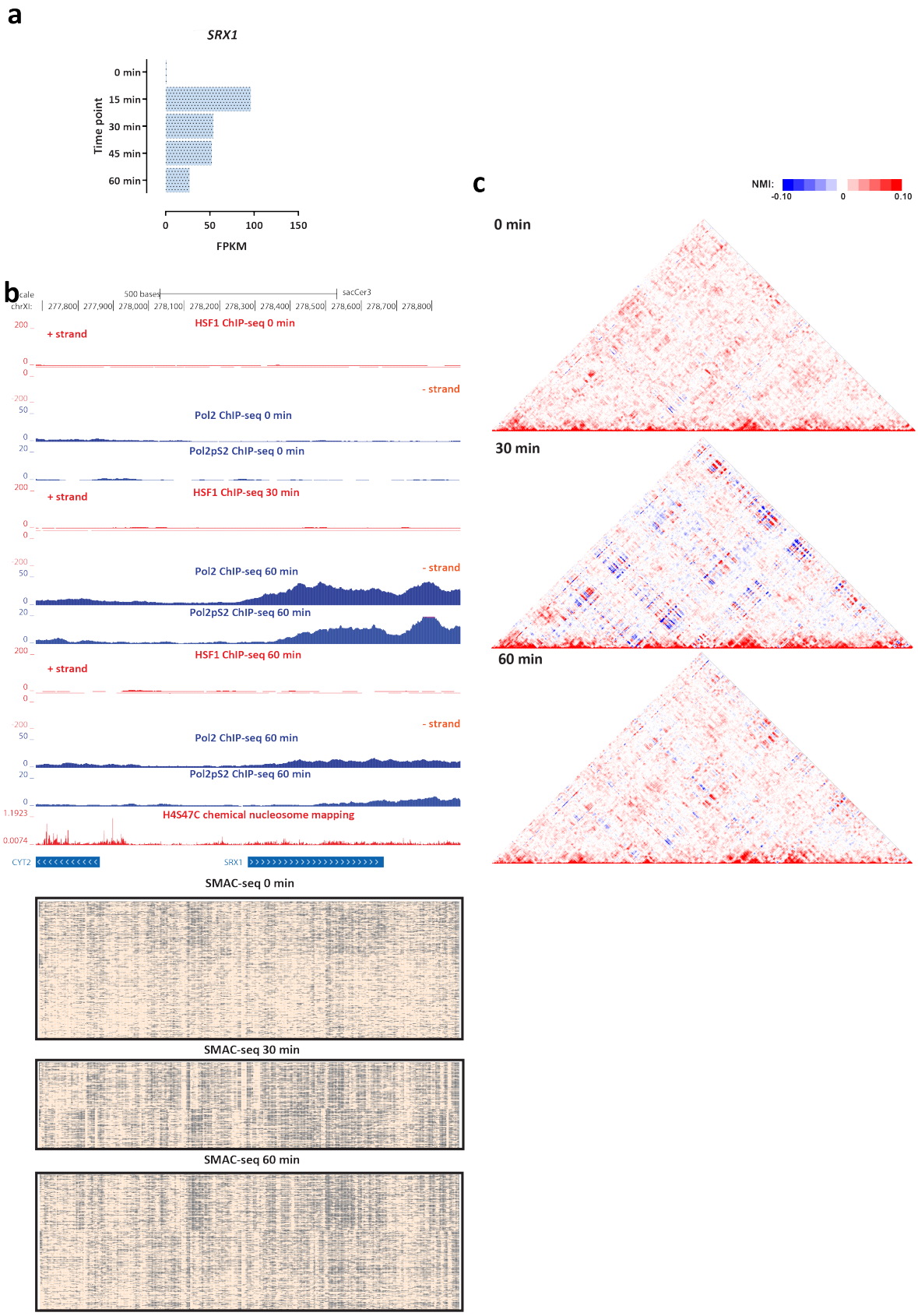
Supplementary Figure 100: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *SSA1* gene.



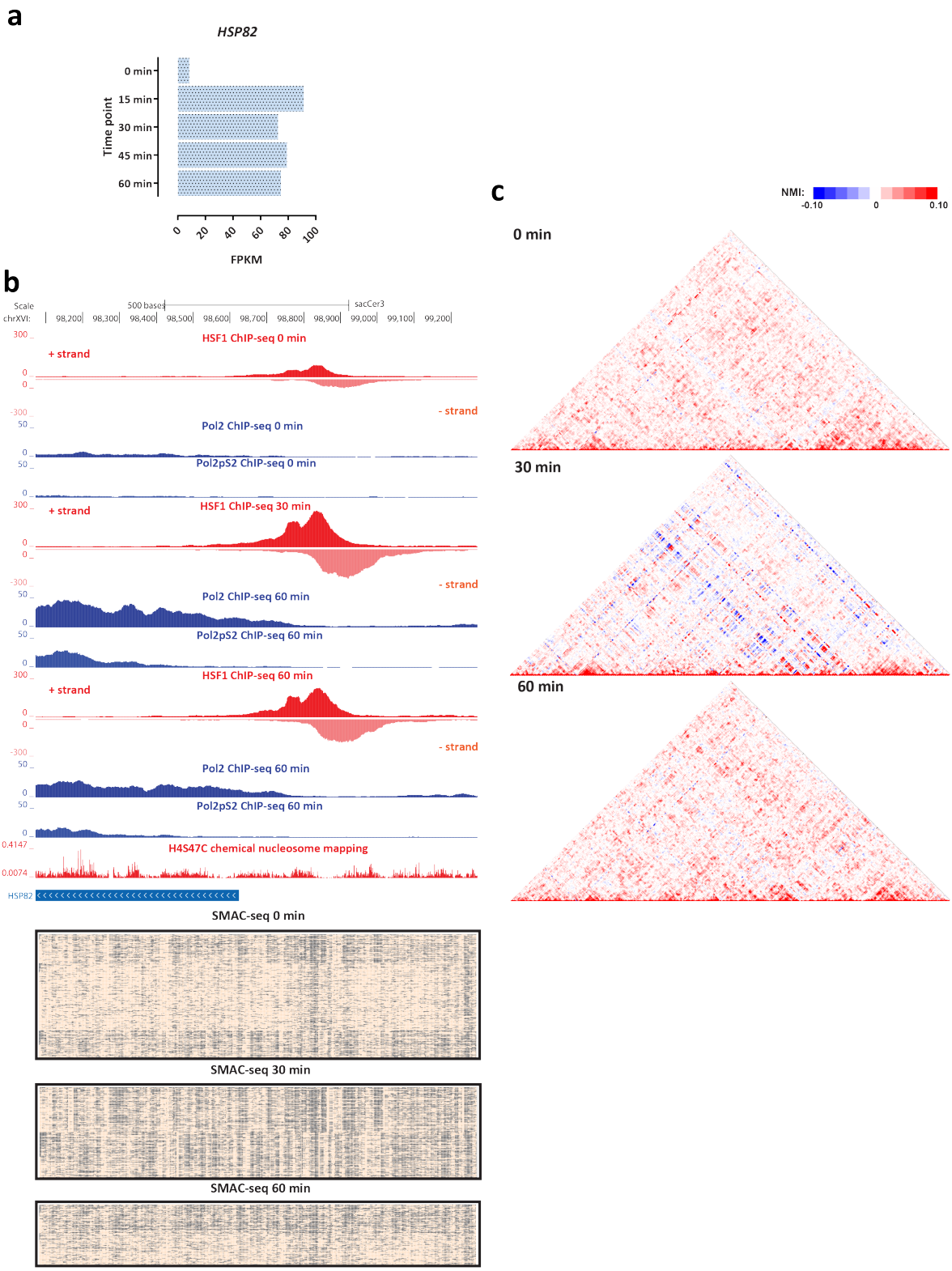
Supplementary Figure 101: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *SSA4* gene.



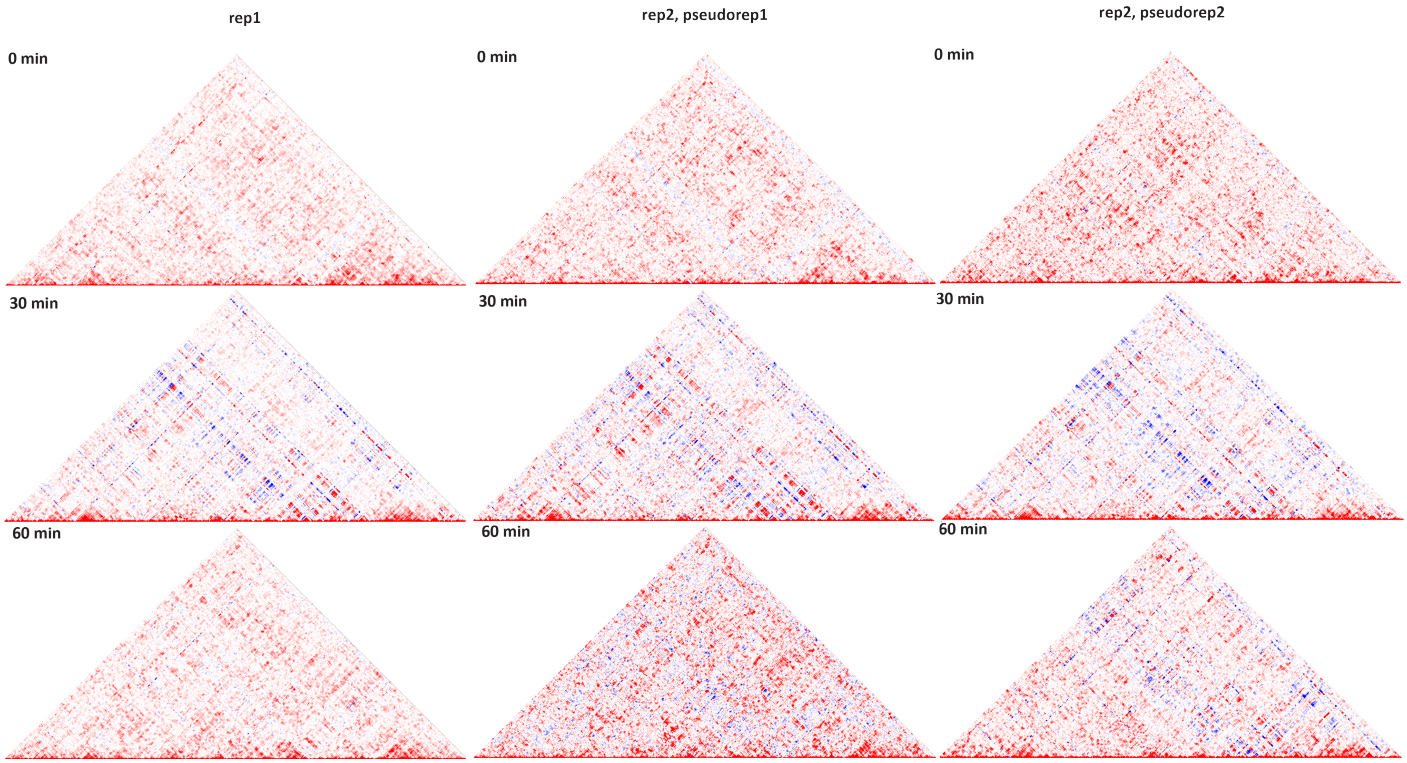
Supplementary Figure 102: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *GRE2* gene.



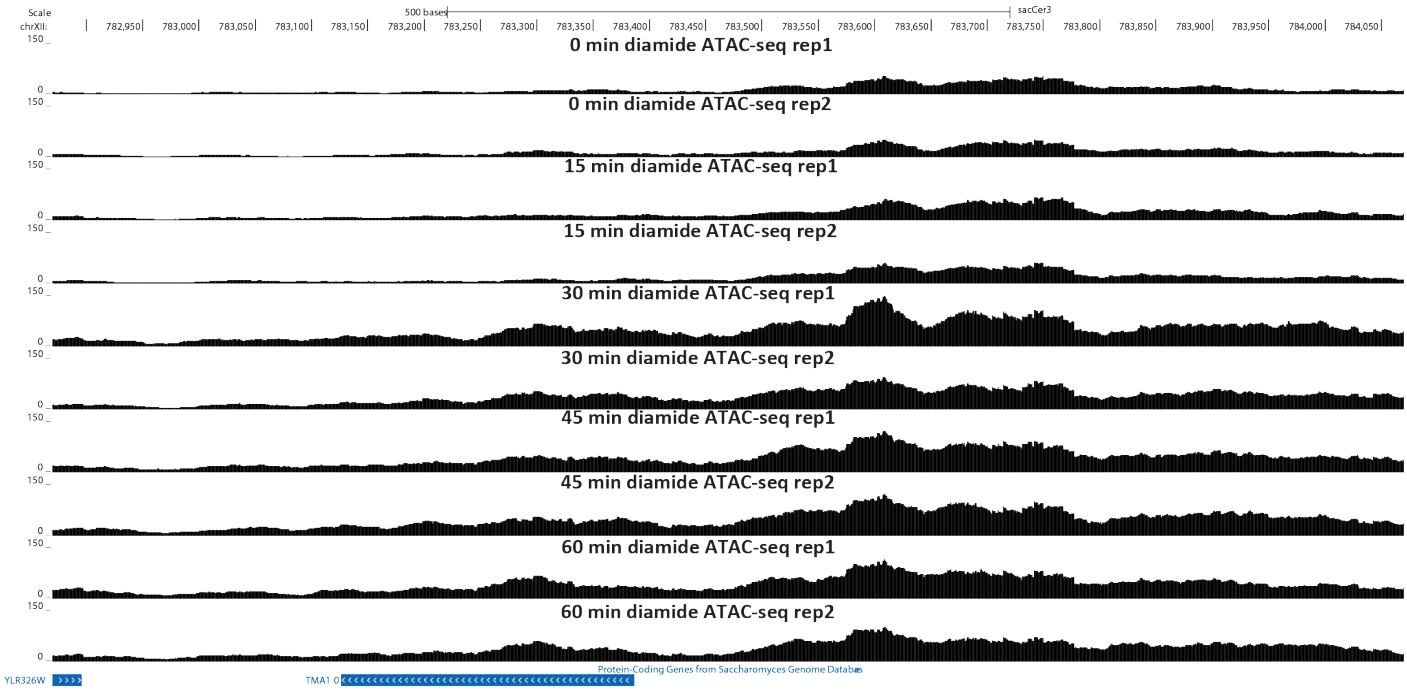
Supplementary Figure 103: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *SRX1* gene.



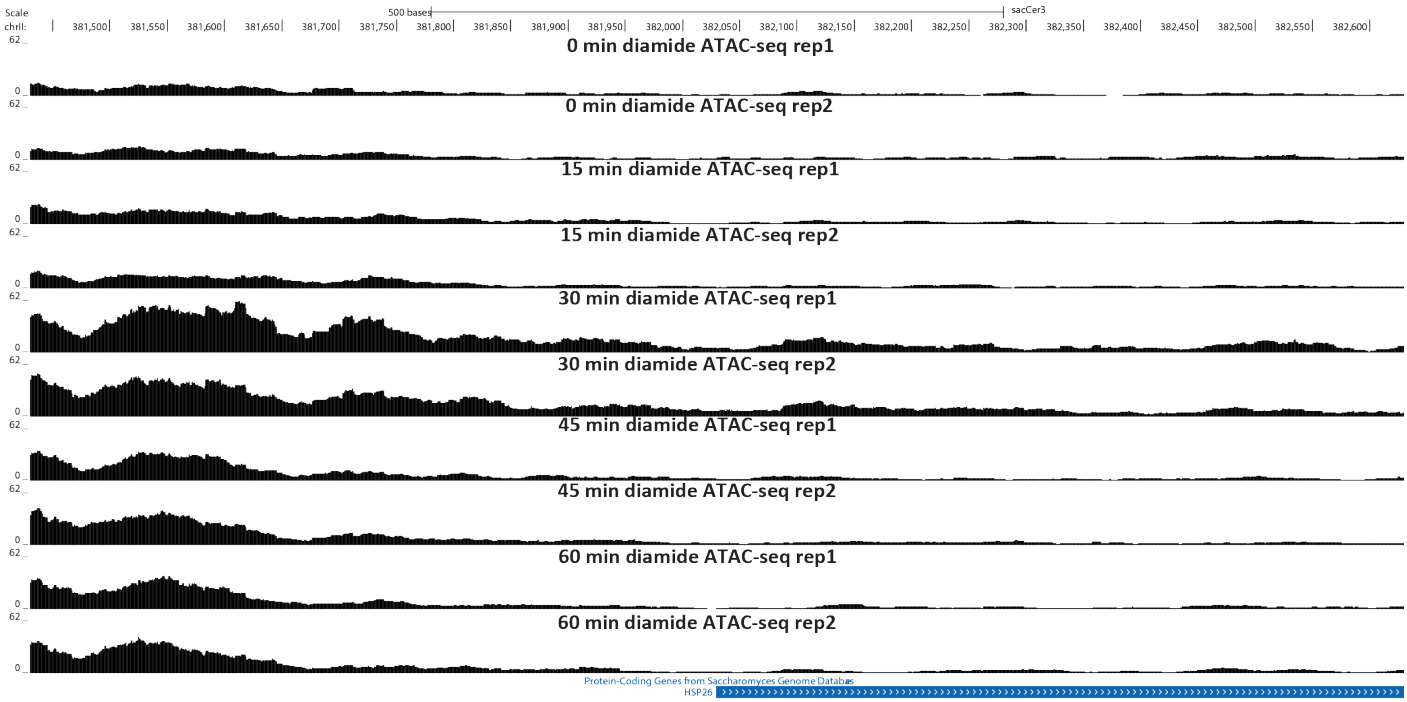
Supplementary Figure 104: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response. Shown are changes in RNA Polymerase and HSF1 occupancy (measured by ChIP-seq), SMAC-seq profiles (1-bp resolution, 10-bp aggregate scores) and NMI profiles in the vicinity of the *HSP82* gene.



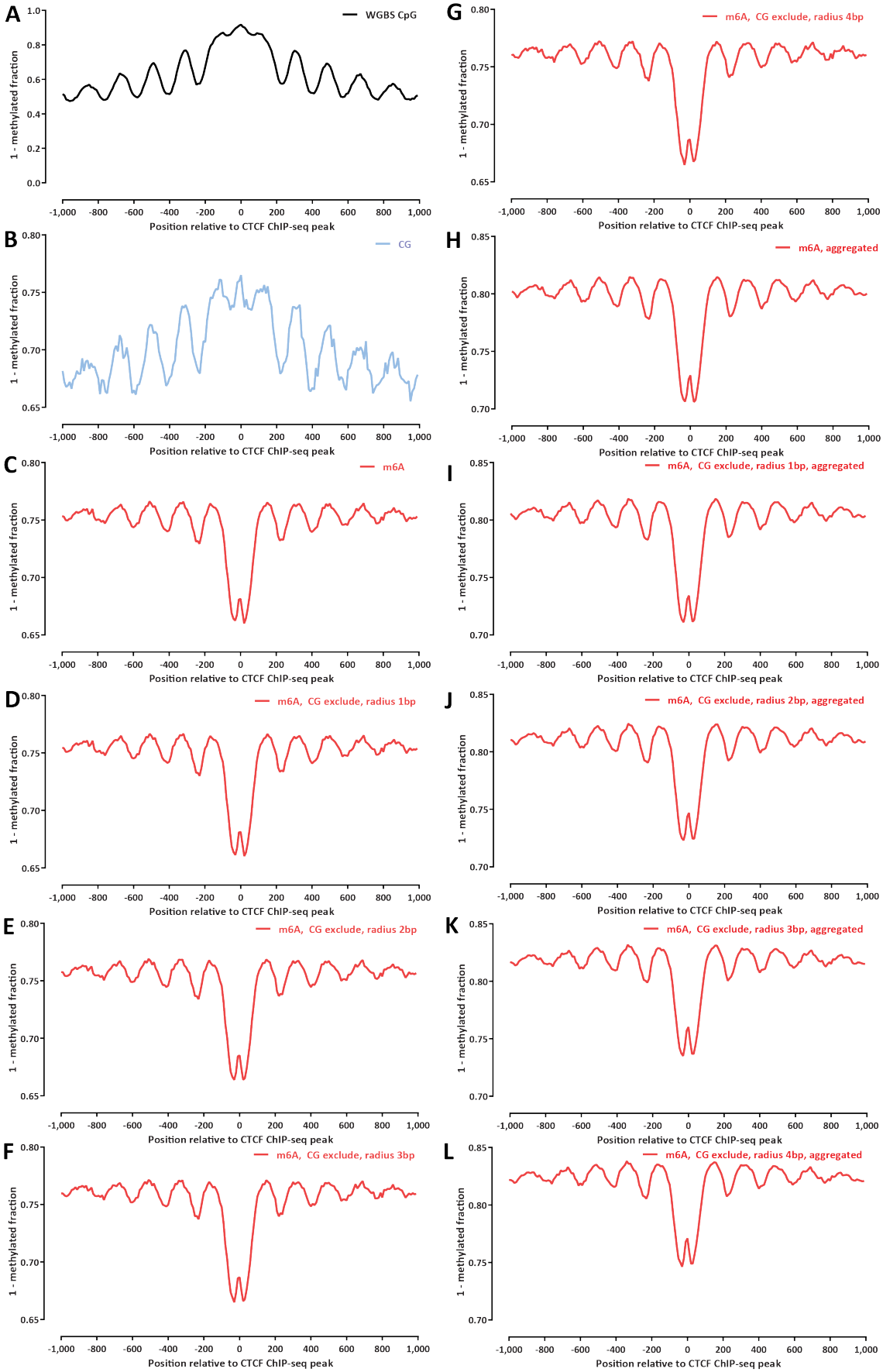
Supplementary Figure 105: Coordinated changes in chromatin accessibility and nucleosomal occupancy during the yeast stress response around the *HSP82* gene. Shown are NMI profiles in the vicinity of the *HSP82* gene for the first replicate of the diamide time course as well as for pseudoreplicates (generated by randomly splitting reads in two halves) from the second replicate of the diamide time course.



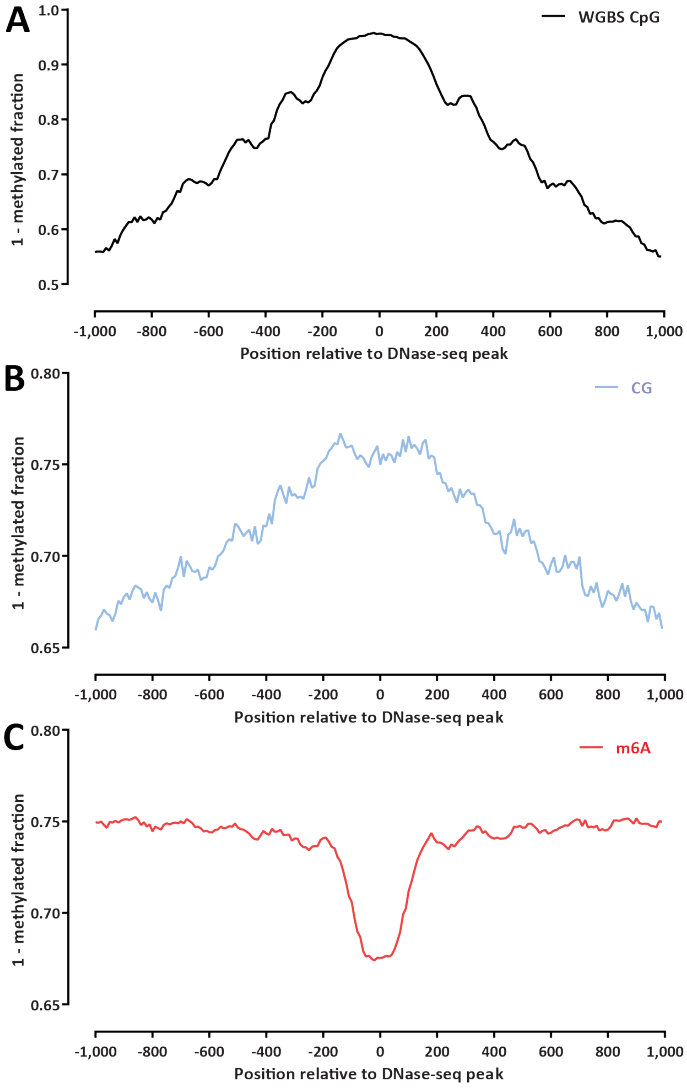
Supplementary Figure 106: Changes in chromatin accessibility during the yeast stress response around the *TMA10* gene as measured by ATAC-seq.



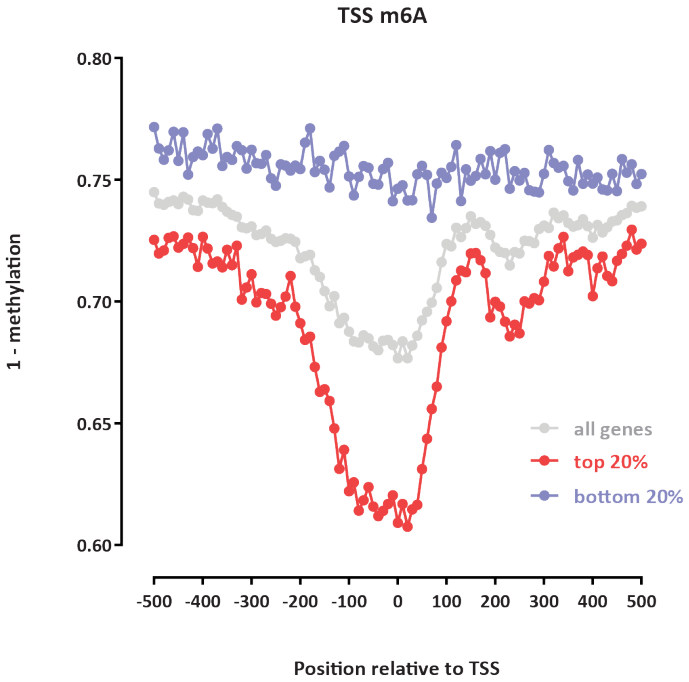
Supplementary Figure 107: Changes in chromatin accessibility during the yeast stress response around the *HSP26* gene as measured by ATAC-seq.


A
WGBS CpG
B
CG
C
m6A
D
m6A, CG exclude, radius 1bp
E
m6A, CG exclude, radius 2bp
F
m6A, CG exclude, radius 3bp
G
m6A, CG exclude, radius 4bp
H
m6A, aggregated
I
m6A, CG exclude, radius 1bp, aggregated
J
m6A, CG exclude, radius 2bp, aggregated
K
m6A, CG exclude, radius 3bp, aggregated
L
m6A, CG exclude, radius 4bp, aggregated
1 - methylated fraction
Position relative to CTCF ChIP-seq peak
-1,000 -800 -600 -400 -200 0 200 400 600 800 1,000

Supplementary Figure 108 *(preceding page)*: Detection of chromatin accessibility features around CTCF ChIP-seq peak summits in human GM12878 cells using m^{6}A-SMAC-seq. (a) Whole-genome bisulfite sequencing data (obtained from the ENCODE Consortium) (b) CpG methylation (from m^{6}A-SMAC-seq experiment generated by this study) (c) m^{6}A methylation (from m^{6}A-SMAC-seq experiment generated by this study), average profile; (d) m^{6}A methylation (excluding A positions within 1 bp of a CpG dinucleotide), average profile; (e) m^{6}A methylation (excluding A positions within 2 bp of a CpG dinucleotide), average profile; (f) m^{6}A methylation (excluding A positions within 3 bp of a CpG dinucleotide), average profile; (g) m^{6}A methylation (excluding A positions within 4 bp of a CpG dinucleotide), average profile; (h) m^{6}A methylation (from m^{6}A-SMAC-seq experiment generated by this study), aggregate profile; (i) m^{6}A methylation (excluding A positions within 1 bp of a CpG dinucleotide), aggregate profile; (j) m^{6}A methylation (excluding A positions within 2 bp of a CpG dinucleotide), aggregate profile; (k) m^{6}A methylation (excluding A positions within 3 bp of a CpG dinucleotide), aggregate profile; (l) m^{6}A methylation (excluding A positions within 4 bp of a CpG dinucleotide), aggregate profile.



Supplementary Figure 109: Average profiles of endogenous DNA methylation and m^{6}A-SMAC-seq around DNase hypersensitive sites in human GM12878 cells. (a) Whole-genome bisulfite sequencing data (obtained from the ENCODE Consortium); (b) CpG methylation (from m^{6}A-SMAC-seq experiment generated by this study); (c) m^{6}A methylation (from m^{6}A-SMAC-seq experiment generated by this study).



Supplementary Figure 110: Measurement of chromatin accessibility around transcription start sites using m^{6}A-SMAC-seq human GM12878 cells. Shown are the average m^{6}A-SMAC-seq around all protein coding genes, as well as the top 20% and the bottom 20% of genes in GM12878 cells (as determined using RNA-seq measurements provided by the ENCODE Consortium).