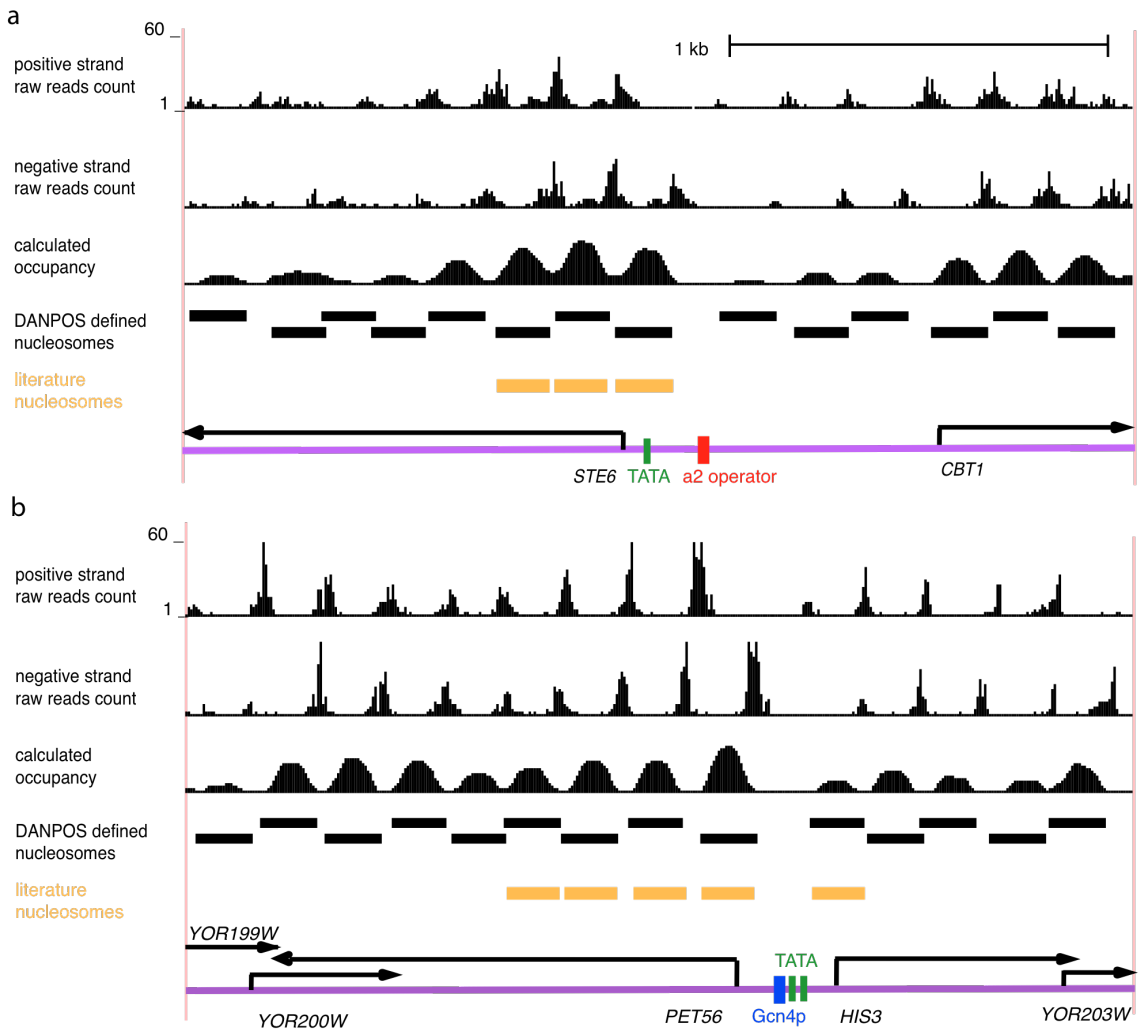
Supplemental Note 1: Comparison between DANPOS and nucleR in defining nucleosome position, occupancy, and fuzziness

We simulated 23,774,388 reads in total for 54,668 nucleosomes with different occupancy, fuzziness, and MNase digestion conditions. Clonal reads that were exceptionally preferred in PCR amplification were also randomly added to some nucleosomes (Supplemental Fig. 2). By the default parameters, DANPOS defined all the 54,668 nucleosome positions, whereas nucleR reported an excessive number of positions with 195,577 in total. This result is reasonable because nucleR was specifically designed for detecting multiple potential nucleosome positions for each given nucleosome unit; whereas DANPOS was tailored to infer the most probable position of a given nucleosome unit. We therefore only kept one position closest to each true dyad for nucleR. We then asked to what degree DANPOS and nucleR could reproduce the true nucleosome position, occupancy, and fuzziness from the simulated data. Based on 54,668 nucleosome positions in the simulated data, nucleR defined about 85% of nucleosomes within 20bp distance from the true dyads, whereas DANPOS achieves a much higher percentage of 99% with the same distance cutoff (Supplemental Fig. 3a). Although nucleR considers both occupancy and fuzziness, it only provides one positioning score for either lower occupancy and/or higher fuzziness. In contrast, DANPOS can generate separate scores, with the occupancy value being reads coverage and fuzziness score being the standard deviation of read positions relative to dyad in each peak (Jiang and Pugh 2009). Correlation analysis indicated that the occupancy value and fuzziness score defined by DANPOS agrees very well with the simulated data (Supplemental Fig. 3 b, d). In contrast, the positioning score of nucleR is a reasonable indicator of occupancy (R= 0.81) but not fuzziness (R=-0.22), (Supplemental Fig. 3c and 3e), probably because different methods were used for the estimation of nucleosome fuzziness.

Supplementary table 1: single nucleotide differential signal improves the detection of fuzziness changes.

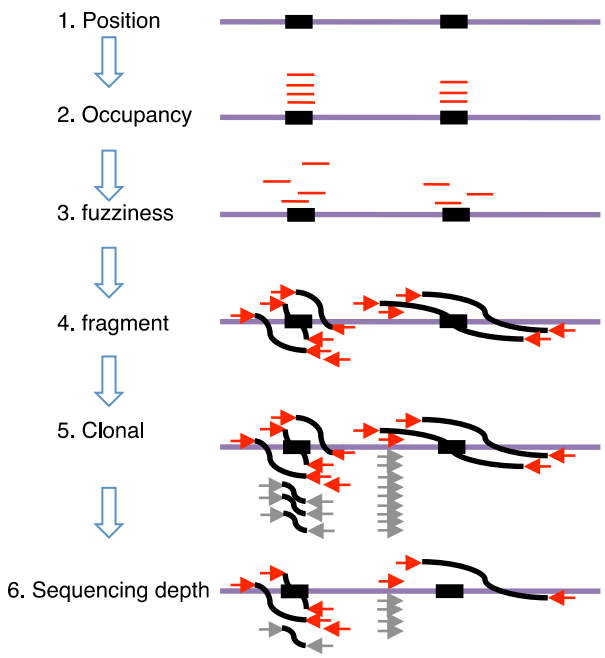
nucleosome changes	Function term	Gene Count	Gene Percent	Q value
fuzziness $P<1x10^{-7}$ and single nucleotide differential P value $>1x10^{-7}$ (75 genes)	fatty acid beta-oxidation	3	4.00	$2.70x10^{-1}$
	vacuolar protein catabolic process	6	8.00	$2.70x10^{-1}$
	lipid oxidation	3	4.00	$2.80x10^{-1}$
	fatty acid oxidation	3	4.00	$2.80x10^{-1}$
	fatty acid catabolic process	3	4.00	$2.80x10^{-1}$
	generation of precursor metabolites and energy	9	12.00	$3.00x10^{-1}$
	oxidative phosphorylation	5	6.67	$3.30x10^{-1}$
	cellular response to heat	7	9.33	$3.40x10^{-1}$
	respiratory electron transport chain	4	5.33	$3.50x10^{-1}$
	coenzyme metabolic process	6	8.00	$3.90x10^{-1}$
fuzziness $P<1x10^{-7}$ (123genes)	energy derivation by oxidation of organic compounds	13	10.57	$3.90x10^{-3}$
	generation of precursor metabolites and energy	16	13.01	$4.70x10^{-3}$
	hexose metabolic process	10	8.13	$1.70x10^{-2}$
	glucose metabolic process	9	7.32	$1.80x10^{-2}$
	monosaccharide metabolic process	10	8.13	$2.30x10^{-2}$
	oxidation reduction	15	12.20	$7.40x10^{-2}$
	monosaccharide catabolic process	6	4.88	$8.30x10^{-2}$
	response to abiotic stimulus	14	11.38	$8.30x10^{-2}$
	vacuolar protein catabolic process	8	6.50	$9.00x10^{-2}$
	oxidative phosphorylation	6	4.88	$9.70x10^{-2}$
fuzziness $P<1x10^{-7}$ and single nucleotide differential P value $<1x10^{-7}$ (76 genes)	energy derivation by oxidation of organic compounds	11	14.47	$8.50x10^{-4}$
	generation of precursor metabolites and energy	13	17.11	$1.50x10^{-3}$
	glucose metabolic process	7	9.21	$3.00x10^{-2}$
	acetyl-CoA metabolic process	5	6.58	$3.40x10^{-2}$
	hexose metabolic process	7	9.21	$6.90x10^{-2}$
	oxidation reduction	11	14.47	$8.60x10^{-2}$
	monosaccharide metabolic process	7	9.21	$8.60x10^{-2}$
	tricarboxylic acid cycle	4	5.26	$1.00x10^{-1}$
	acetyl-CoA catabolic process	4	5.26	$1.00x10^{-1}$
	cofactor metabolic process	8	10.53	$1.10x10^{-1}$



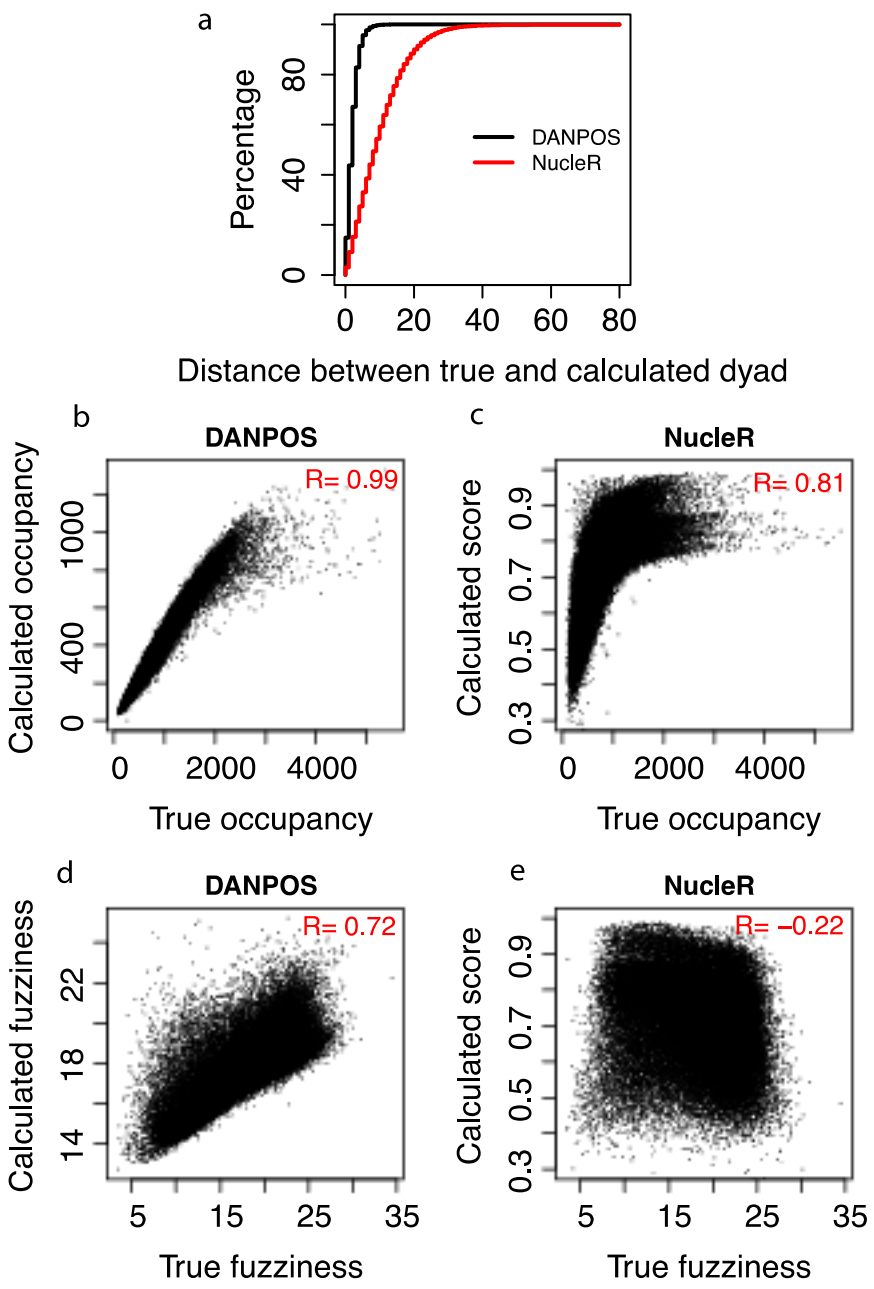
Supplemental Figure 1: DANPOS reproduces published nucleosome maps. Count of raw reads and calculated nucleosome occupancy was plotted as black area. Black bars represent nucleosome positions defined by DANPOS based on real nucleosome data of yeast grown under various conditions (Kaplan et al. 2008). Yellow bars represent nucleosome positions defined in previous literatures. Genes, TATA box, Matalpha2 operator, and Gcn4p binding sites were also labeled on the DNA (purple line).

a) Literature nucleosomes on *STE6* promoter through conventional methods. The binding of Matalpha2 to its operator in the promoter of *STE6* leads to the placement of an array of nucleosome in the transcription direction (Shimizu et al. 1991; Cooper et al. 1994). The most proximal nucleosome in the array was mapped immediately upstream of TSS at the location from approximately -24 to -164bp.

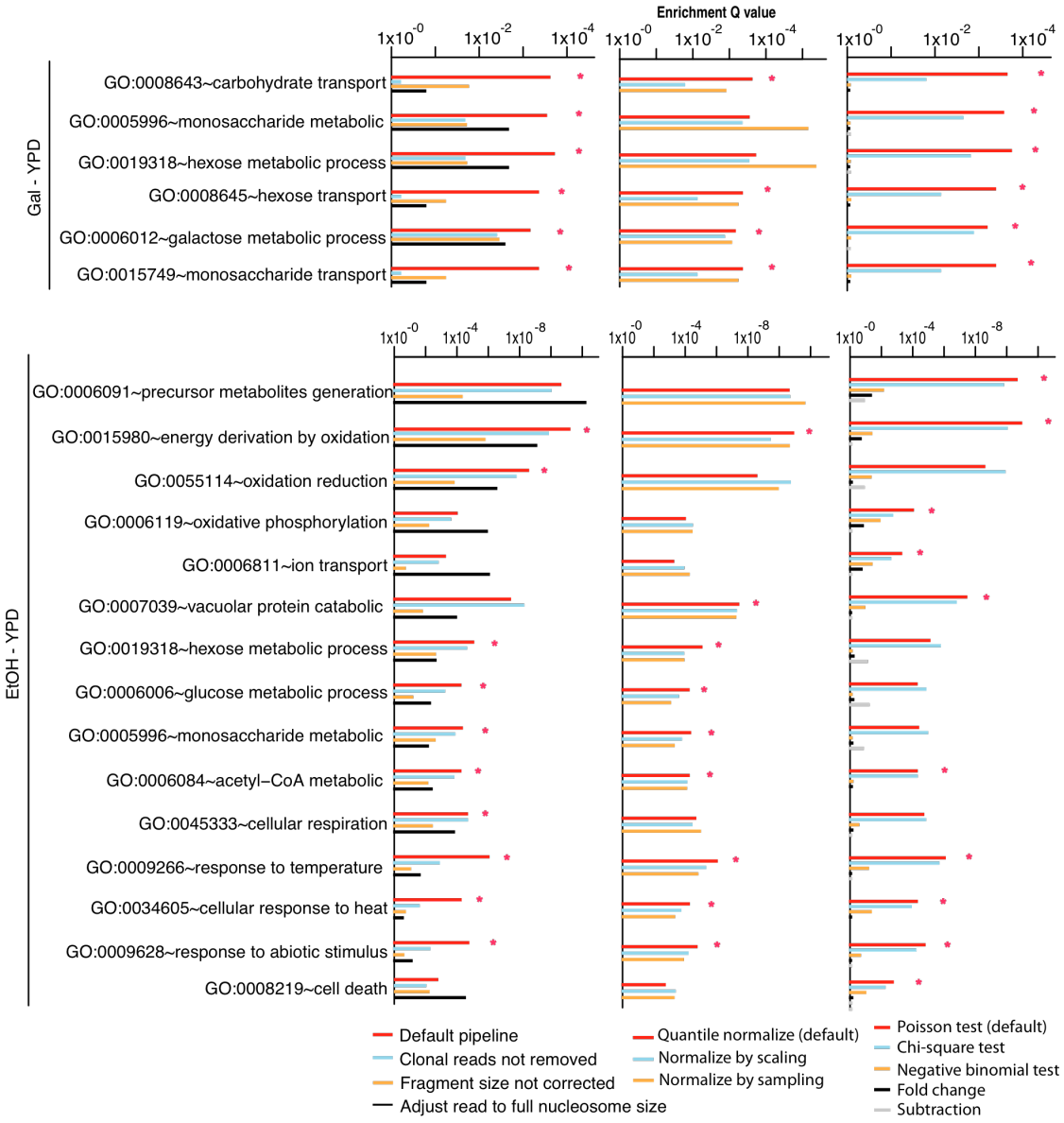
b) Literature nucleosomes on *HIS3* promoter through tiling array hybridization. An array of well positioned nucleosomes were shown to surround the Gcn4p binding site in nucleosome free region (NFR) (Yuan et al. 2005; Kim et al. 2006).



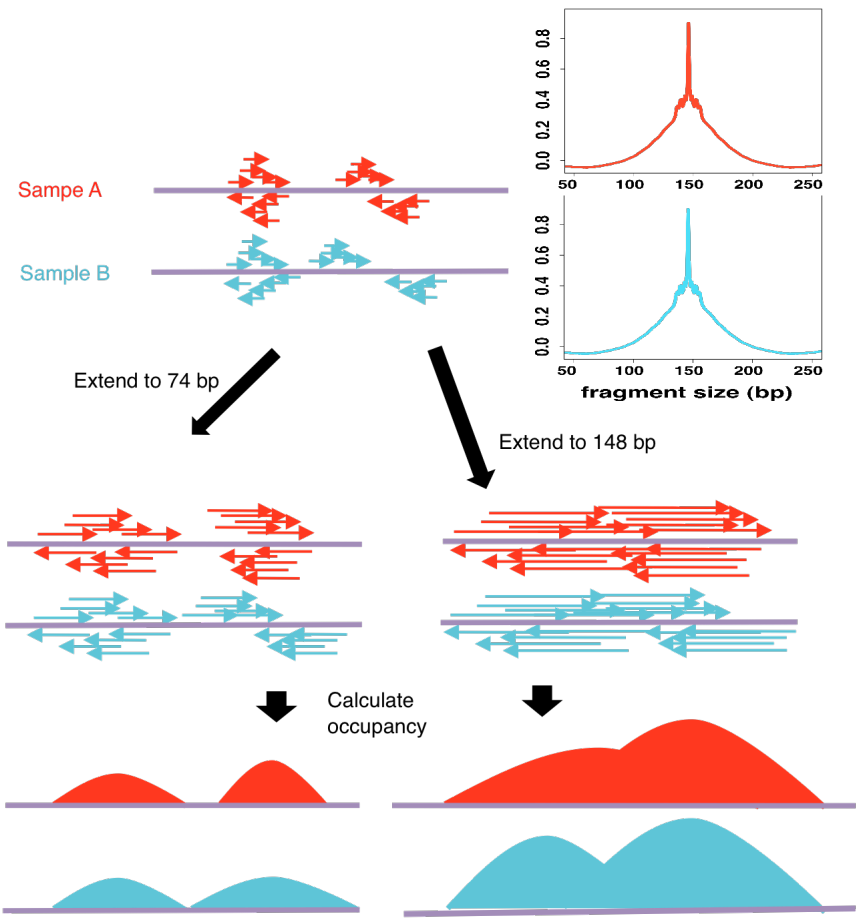
Supplemental Figure 2: Flowchart of the MNase sequencing reads simulation. The 1) positions (black bar), 2) occupancy (red lines count), and 3) fuzziness levels (red lines positions) of each nucleosome unit were determined by a nucleosome reference map derived from a pool of real nucleosome data. 4) A pair of sequencing reads (red arrows) is assigned to each nucleosome (red line), allowing variable fragment size (black curve). Single-end reads were also assigned to a small portion of nucleosomes to accommodate the fact that sometimes only one end of the paired-end reads can be mapped. 5) Clonal reads (gray arrows) that were exceptionally preferred in the PCR amplification were also randomly added to some nucleosomes. 6) Finally, reads were randomly sampled to a specific sequencing depth from the reads generated at step 5).



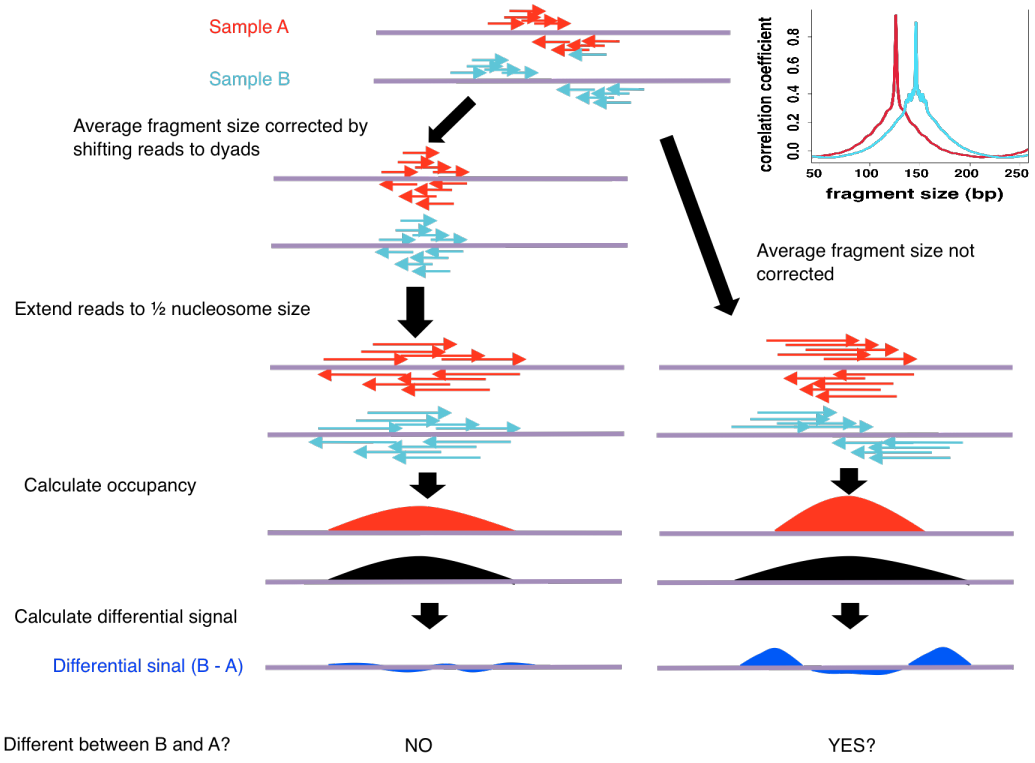
Supplemental Figure 3: Comparison between DANPOS and nucleR based on simulated nucleosome data indicates that DANPOS performs better in defining nucleosome position (a), occupancy (b,c), and fuzziness (d,e). Spearman correlation coefficient was labeled in each scatter plot (b,c,d,e). For each nucleosome unit, true occupancy and fuzziness were determined by total number of simulated reads within the nucleosome unit (147 bp) and standard deviation of simulated read positions relative to the nucleosome center, respectively.



Supplemental Figure 4: The DANPOS default algorithm improves the functional interpretation of Gal-YPD (top) and EtOH-YPD (bottom) dynamic nucleosomes. Each column shows the enrichment of biological processes in genes associated with dynamic nucleosomes identified by different preliminary data processing methods (**left**), data normalization methods (**middle**), or statistical tests for differential signals (**right**); Red stars show the biological processes that are most enriched in dynamic nucleosomes detected by the default pipeline in DANPOS.

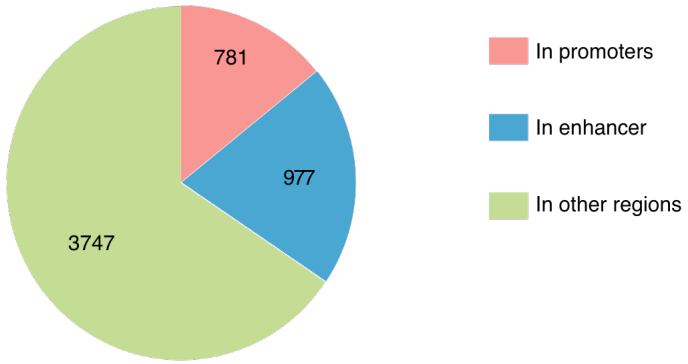


Supplemental Figure 5: Correcting for average fragment size between samples to the same level compensates for technical variation in MNase digestion. The purple line represents the DNA sequence; red and sky-blue arrows show sequencing reads in samples A and B, respectively; red and sky-blue areas represent nucleosome occupancy in the two samples, respectively. Blue areas show differential occupancy signals between samples A and B at single-nucleotide resolution. Distributions of the average fragment size in samples A and B are plotted at the top right.



Supplemental Figure 6: Adjusting the read length to half of the nucleosome size rather than the full nucleosome size improves the signal-to-noise ratio. The purple line represents the DNA sequence; red and sky-blue arrows show sequencing reads in samples A and B, respectively; red and sky-blue areas indicate nucleosome occupancy calculated from sequencing reads in samples A and B, respectively. Distributions of the average fragment sizes in samples A and B are plotted at the top right.

a



b

Dynamic nucleosomes	Term ID	Term Name	Binom FDR
In Promoters	GO:0001503	ossification	3.52×10^{-6}
	GO:0048701	embryonic cranial skeleton morphogenesis	6.23×10^{-4}
	GO:0009880	embryonic pattern specification	1.49×10^{-4}
In Enhancers	GO:0035019	somatic stem cell maintenance	3.65×10^{-6}
	GO:0048863	stem cell differentiation	2.36×10^{-5}
	GO:0009880	embryonic pattern specification	8.40×10^{-4}
In other regions	GO:0060231	mesenchymal to epithelial transition	3.12×10^{-2}
	GO:0072077	renal vesicle morphogenesis	2.18×10^{-2}
	GO:0002076	osteoblast development	2.97×10^{-2}

Supplemental Figure 7: DNAPOS defines functional dynamic nucleosomes during mouse embryonic stem cell differentiation. (a) Genomic distribution of the dynamic nucleosomes. **(b)** Enrichment of function terms for dynamic nucleosomes in each genomic category. The mouse promoter and enhancer regions were derived from a map of the cis-regulatory sequences in the mouse genome (Shen et al. 2012). GREAT (Mclean et al. 2010) was used to analyze the functional significance of dynamic nucleosomes in each group.