Supporting Information

Henikoff et al. 10.1073/pnas.1110731108

SI Experimental Procedures

Modified Solexa Library Protocol. The standard Illumina paired-end library preparation protocol (Illumina # PE-930–1001) is used except as follows: (*i*) All Qiagen clean-up steps and gel purifications are omitted. (*ii*) Phenol/chloroform extraction is used to stop reactions followed by S300 (or S200) spin column cleanup and Speed-vac volume reduction. (*iii*) Ampure XP bead steps following adapter ligation and PCR remove excess adapters and primers, respectively. (*iv*) Reaction volumes are 50 μL, except for the PCR amplification step, where the volume is 10–40 μL, using the same concentrations as recommended by Illumina. (*v*) Low-retention 1.5-mL (siliconized) microfuge tubes are used throughout, minimizing tube transfers to reduce losses. Enzymes are from the Illumina PE-930–1001 kit or are purchased from standard suppliers, such as New England Biolabs. The following protocol is for eight samples and has been used for 20- to 800-ng starting DNA from micrococcal nuclease (MNase) digested chromatin (based on Quant-it Picogreen fluorescence with RNase added).

Step 1.

End repair (total 50 μL):
- 20 μL DNA sample
- 18.5 μL water
- 5 μL T4 ligase buffer with 10 mM ATP
- 2 μL dNTP mix (10 mM each)
- 1.5 μL T4 DNA Pol (5 U/μL)
- 0.5 μL Klenow fragment (5 U/μL)
- 2.5 μL T4 PNK (10 U/μL)

Master mix (add 11.5 μL to 38.5 μL DNA):
- 42 μL T4 ligase buffer with 10 mM ATP
- 16.8 μL dNTP mix (10 mM each)
- 12.6 μL T4 DNA Pol (5 U/μL)
- 4.2 μL Klenow fragment (5 U/μL)
- 21 μL T4 PNK (10 U/μL)

20 °C 30 min at 20 °C (ice bucket and thermometer "water bath")

Add 50 μL phenol/chloroform/isoamyl (25:24:1, OmniPur cat. #6805), vortex, spin, and remove most organic matter from the bottom. When decanting the aqueous layer, touch the tip to the tube wall to drain remaining organic matter.

Apply the aqueous layer to an S-300 spin column (prespun 1 min at 750 × *g*, GE Healthcare).

Spin 2 min at 3,000 rpm.

Reduce volume to 35 μL in Speedvac (no heat). This takes ~15 min.

Step 2.

Add 3′ A overhang (total 50 μL):
- 35 μL end-repaired DNA sample
- 5 μL 10× NEB Buffer 2
- 10 μL 1 mM dATP
- 0.3 μL Klenow exo- (50 U/μL)

Master mix (add 15 μL to 35 μL DNA):
- 42 μL 10× NEB Buffer 2
- 84 μL 1 mM dATP
- 2.5 μL Klenow exo- (50 U/μL New England Biolabs cat. #M0212M)

37°C 30 min.

Extract with 50 μL Phenol/Chloro/isoamyl and S-300 clean-up.

Reduce volume to 18 μL in Speedvac (no heat). This takes ~30 to 45 min.

Step 3.

Ligate adapters (total 50 μL):
- 18 μL 3′A-extended DNA sample
- 2 μL Illumina PE Adapter mix
- 25 μL Enzymatics Rapid DNA ligase buffer 2×
- 5 μL Enzymatics DNA Ligase (600 U/μL)

Add 2 μL Illumina adapters to each DNA sample.

Note: Adjust adapter amount to give an estimated 10:1 adapter:insert molar ratio.

Ligase mix (add 30 μL to 20 μL DNA + adapters):
- 210 μL Rapid DNA Ligase Buffer 2×
- 42 μL Enzymatics DNA Ligase (600 U/μL)
- Incubate 20 °C for 15 min.

Step 4.

Purify on Ampure XP magnetic beads (Agencourt):
- Add 90-μL bead slurry, pipetting up and down 10×, total 5 min.
- Place in magnetic tube holder. After 2 min, aspirate off the solution.
- When still on the magnet, add 200 μL 70% EtOH, 30 s, aspirate off. Repeat.
- Let dry a few minutes (≤5 min), remove from magnet.
- Disperse in 40 μL 0.1 × TE8 (1 × TE8 = 10 mM Tris pH8, 1 mM EDTA); place in magnetic tube holder. Pull off the solution to a fresh tube.

Step 5.

PCR.

Volume = 10–40 μL (aliquot or entire sample ± Speed-vac reduction). Per 10 μL:
- 6.7 μL Ampure purified DNA
- 2 μL 5× Phusion buffer
- 0.8 μL 2.5mM dNTP
- 0.2 μL Illumina Primer 1.0
- 0.2 μL Illumina Primer 2.0
- 0.1 μL Phusion polymerase (Finnzymes)

Master mix (add 3.3 μL to 6.7 μL DNA):
- 16.8 μL 5× Phusion buffer
- 6.8 μL 2.5 mM dNTP
- 1.8 μL Illumina Primer 1.0
- 1.8 μL Illumina Primer 2.0
- 0.9 μL Phusion polymerase

Cycling (SHPCR60 12–18 cycles):
- 98 °C 30 s
- 98 °C 10 s
- 65 °C 30 s
- 60 °C 30 s

Go to cycle 2, repeat 17×
- 60 °C 5 min
- 8 °C forever.

Step 6.

Perform Ampure clean-up of PCR samples, adding 18-μL Ampure beads, transferring to a 1.5-mL tube for the purification, as above, and eluting with 40 μL 0.1 × TE8. Check on a 2% agarose gel with ethidium bromide added.

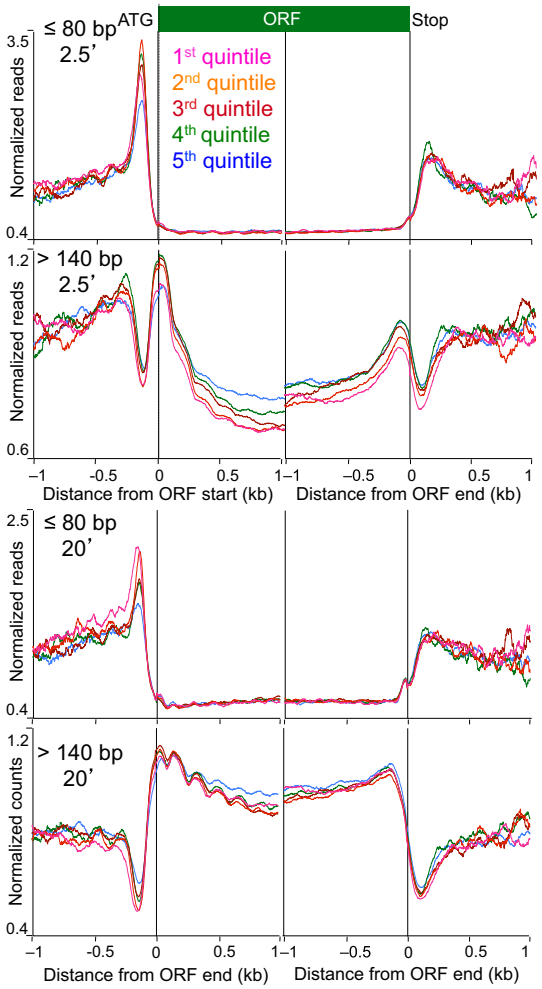

Fig. S1. Genic profiles of subnucleosomal particles and nucleosomes. Genes were aligned at ORF boundaries and subnucleosome-sized (≤80 bp) and nucleosomal (>140 bp) fragments from total nuclei were averaged at 1-bp intervals over a 2-kb region surrounding each ORF end (1). Expression data were obtained from GEO GAM552681 (2) using 4,835 "verified" ORFs. Genes were rank-ordered by expression and averaged within quintiles. Subnucleosomal particles are seen to occupy nucleosome depleted regions both upstream and downstream of ORFs, with low occupancy levels for the most weakly expressed genes with otherwise very little dependence of nucleosome occupancy on expression level. Similarly, there is a lack of distinction between nucleosomal quintiles in occupancy and positioning, which is consistent with results of profiling low-salt soluble active chromatin from *Drosophila* cells, where active genes showed the same average profile regardless of the level of gene expression (3).

1. Weber CM, Henikoff JG, Henikoff S (2010) H2A.Z nucleosomes enriched over active genes are homotypic. *Nat Struct Mol Biol* 17:1500–1507.
2. Tsankov AM, Thompson DA, Socha A, Regev A, Rando OJ (2010) The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol* 8:e1000414.
3. Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K (2009) Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res* 19:460–469.

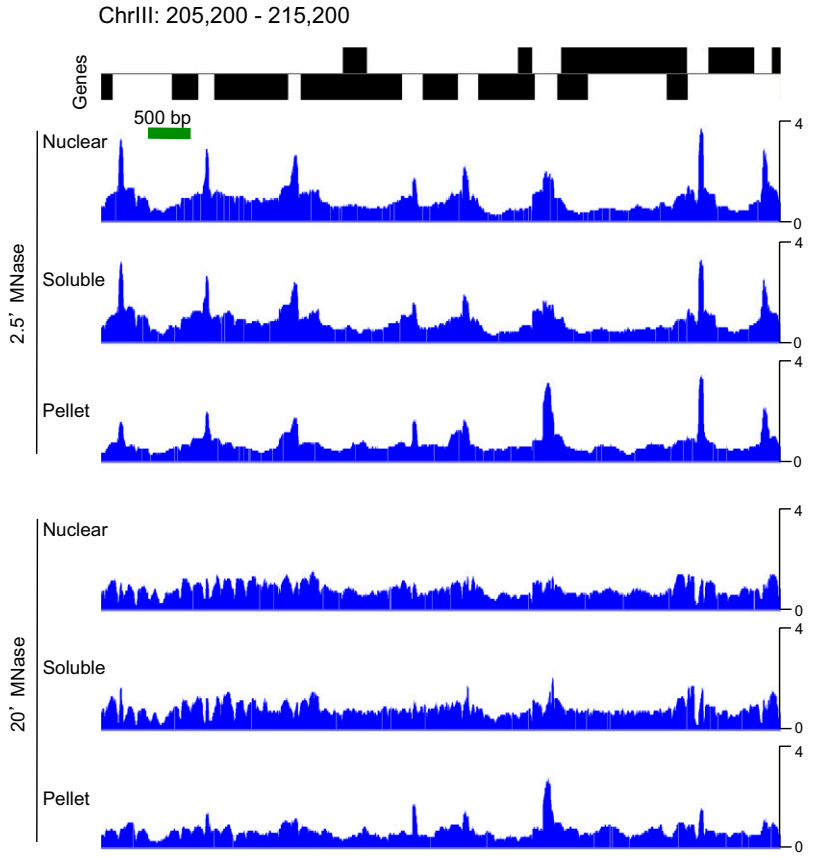

Fig. S2. Typical landscapes for soluble and insoluble chromatin are similar. The same region shown in Fig. 2 for size classes of nuclear DNA is shown here for all reads from nuclear, soluble, and pellet DNA. This comparison indicates that the large majority of the genome is accurately represented in the soluble fraction, although a few quantitative differences are seen. The *y* axis is normalized counts derived from the following total number of mapped paired-end reads: nuclear (~25 million), soluble (~25 million), pellet (~10 million).

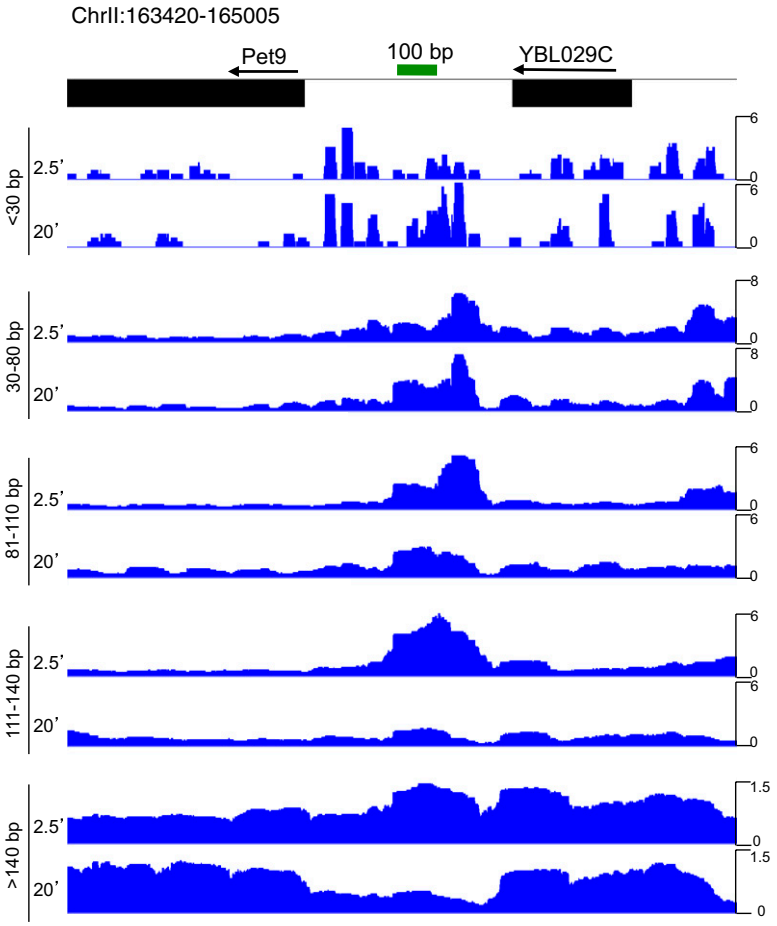

Fig. S3. An intergenic region that shows nucleosome eviction and small particle enrichment. See the legend to Fig. 3*A* for details.

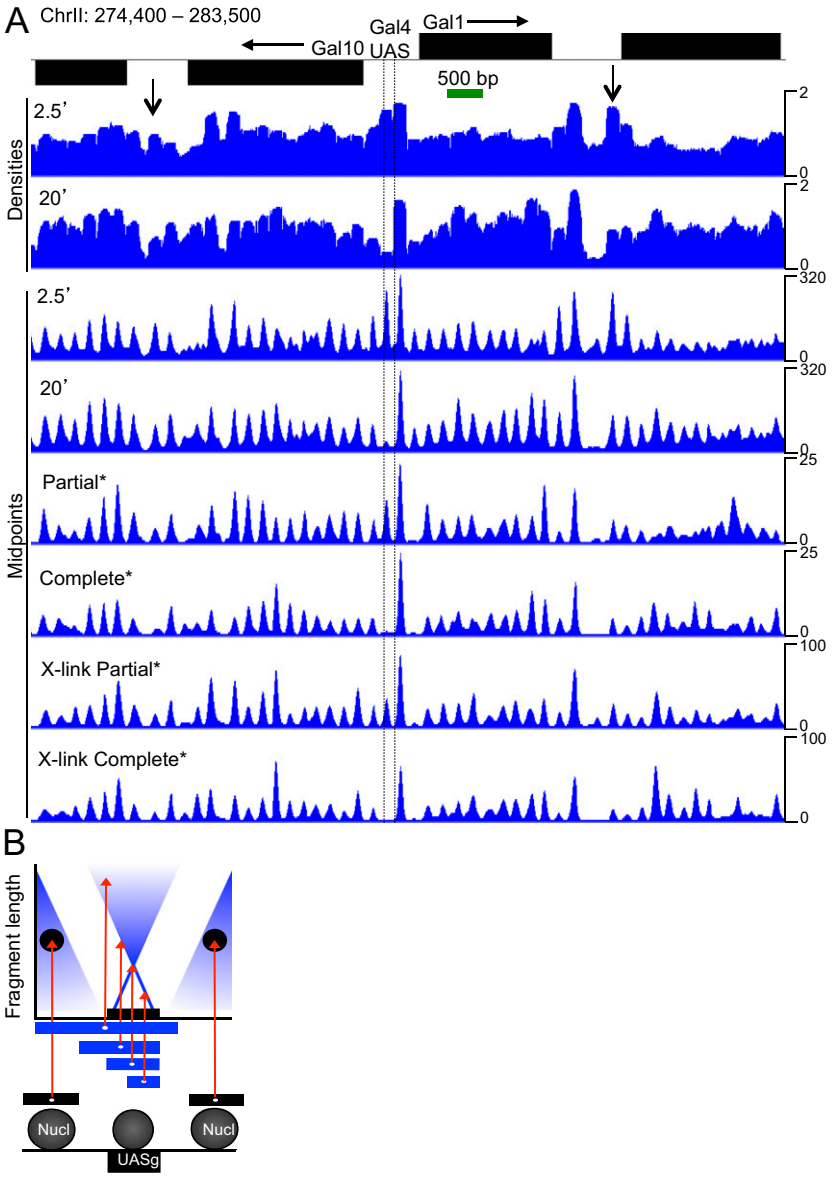

Fig. S4. The Gal4 UAS nucleosome is fragile and precisely positioned. (*A*) MNase protection data described in the present study, shown as either 1-bp resolution occupancy maps or paired-end midpoint profiles, and calculated as midpoint profiles for single-end datasets obtained from the Gene Expression Omnibus (GEO) (marked by asterisks, GSM648301-2 and 648309–10) (1). "Partial" and "Complete" refer to relative digestion levels and "X-link" refers to formaldehyde cross-linking of cells. An extended region around the Gal4 is shown, where vertical arrows point to nearby examples of evident fragility seen in all datasets. A Gaussian kernel density function with bandwidth of 20 and window of 10 was applied to both single-end and paired-end reads. (*B*) Schematic diagram interpreting the V-plots shown in Fig. 4*B*. A fragment that spans an MNase-protected region results in a dot placed in the central sector. The left diagonal results from fragments cleaved precisely on the right side of the protected region, and likewise for the right diagonal, so that their intersection creates a V, which represents the minimal protected region.

1. Xi Y, Yao J, Chen R, Li W, He X (2011) Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. *Genome Res* 21:718–724.

Summary statistics of MNase-protected fragments spanning Abf1 midpoints[1]

Dataset	MNase (min)	Number	Fragment Length	SD	Distance	SD	Nucleosome Positioning
Abf1	2.5'	781290	141.809	83.65	0.575	49.059	Strong
	20'	310097	116.879	69.009	2.255	42.149	
High[2]	2.5'	671231	153.908	88.198	-0.512	52.36	Moderate
	20'	347151	135.834	72.141	1.556	45.601	
Low[3]	2.5'	485743	187.651	86.165	-4.419	59.37	None
	20'	440425	161.009	65.293	-3.255	49.941	
Random[4]	2.5'	808512	196.21	81.958	-7.376	57.5	None
	20'	653803	164.188	65.223	-2.069	48.774	

[1]Fragment lengths and distance of each fragment midpoint from the center of each binding site motif for each dataset are included.
[2]From the 278 highest-scoring Abf1 motifs that were not in the positive set (see Fig. S6).
[3]From the 278 lowest-scoring Abf1 motifs that were not in the positive set (see Fig. S6).
[4]From a sample of randomly chosen genomic positions.

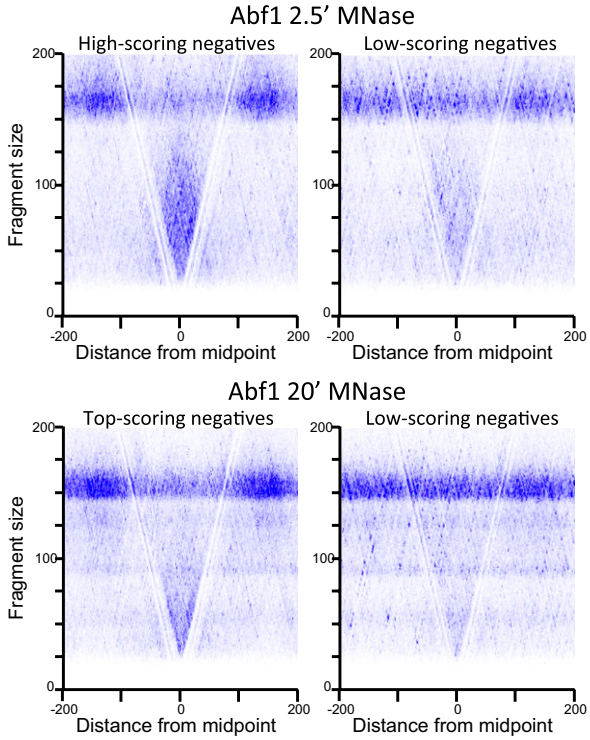

Fig. S5. V-plots based on "unbound" Abf1 consensus motifs. To obtain negative control V-plots, an Abf1 position-specific scoring matrix (1) was used to search the *Saccharomyces cerevisiae* genomic sequence with MAST using the default E < 10 cutoff level of significance (2). Two samples of 278 sites each were chosen from the resulting hit list. Before sampling, we excluded any sites that were present in the set of 278 Abf1 binders used in constructing the Abf1 V-plot shown in Fig. 4*A*. One sample was the 278 highest-scoring negatives in the search and the other was the 278 lowest-scoring negatives. In both datasets, the high-scoring negatives show the distinctive Abf1 pattern, including phasing of neighboring nucleosomes, whereas the low-scoring negatives show only a weak V and slight phasing of nucleosomes. This finding suggests that a large fraction of true Abf1 sites were not detected in the ChIP dataset used for motif identification (3), as previously noted in other studies (4, 5).

1. Chen CY, et al. (2008) Discovering gapped binding sites of yeast transcription factors. *Proc Natl Acad Sci USA* 105:2527–2532.
2. Bailey TL, Gribskov M (1998) Combining evidence using p-values: Application to sequence homology searches. *Bioinformatics* 14:48–54.
3. Harbison CT, et al. (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* 431:99–104.
4. Ganapathi M, et al. (2011) Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. *Nucleic Acids Res* 39: 2032–2044.
5. Goh WS, Orlov Y, Li J, Clarke ND (2010) Blurring of high-resolution data shows that the effect of intrinsic nucleosome occupancy on transcription factor binding is mostly regional, not local. *PLOS Comput Biol* 6:e1000649.

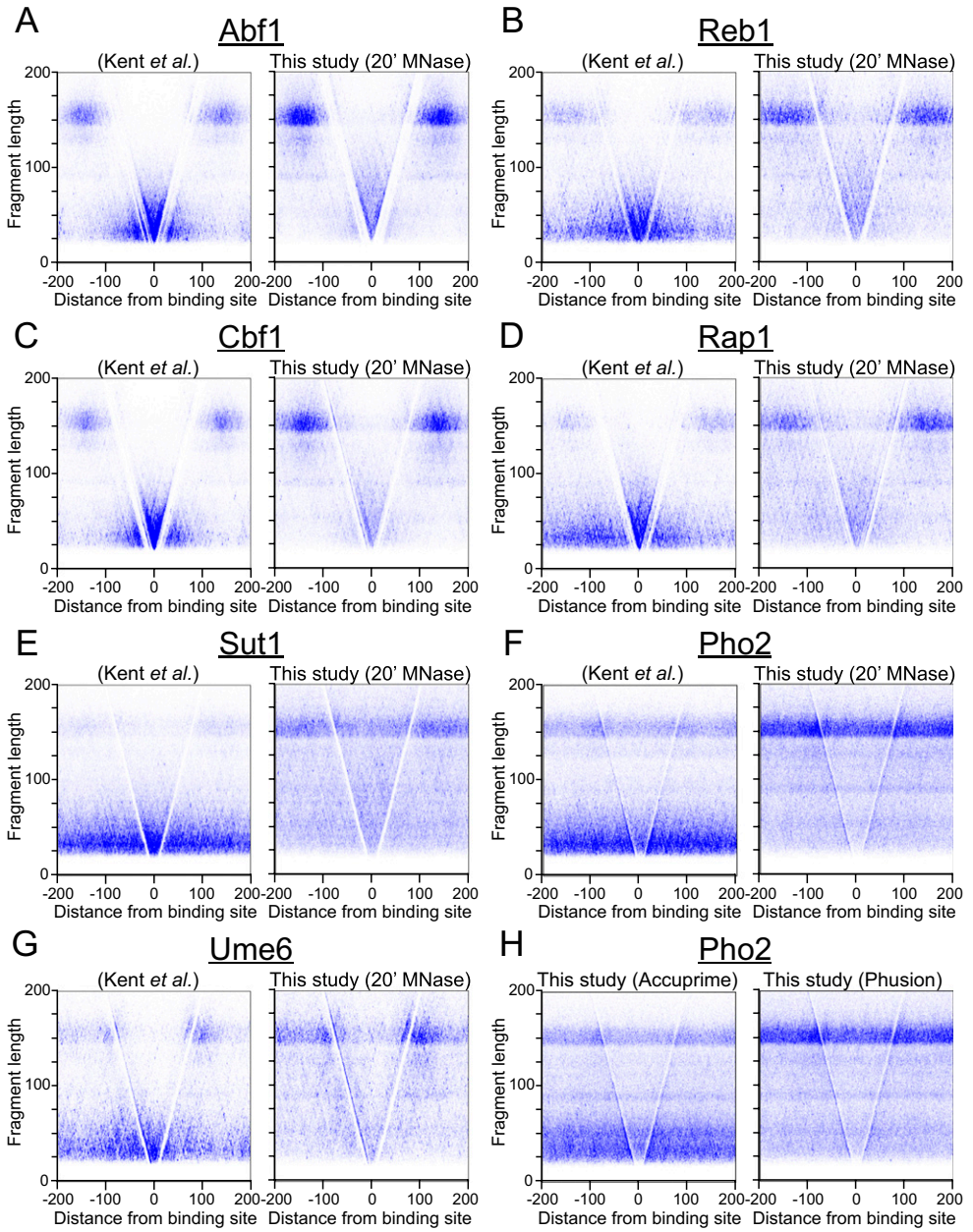

Fig. S6. V-plot displays using sequence datasets with different fragment size distributions. (*A–G*) Paired-end sequencing data from MNase-digested yeast nuclei described by Kent et al. (1) were obtained from the GEO Short Read Archive (SRA020615.3) and aligned to the yeast genome. V-plots for this dataset are shown together with the 20 min V-plots for the transcription factors highlighted in the text. In this study, 10 μg of DNA derived from MNase-treated chromatin DNA obtained from pooled cultures was used to construct standard Illumina libraries. Note that similar features are evident in this dataset as in the current study. Equal numbers of reads were used to construct each V-plot pair. (*H*) Efficiency of PCR amplification affects the size distribution of sequenced fragments. An adapter-linked sample derived from the 20-min MNase digestion sample was split into two equal aliquots after removal of adapters with Ampure beads. One aliquot was amplified with Accuprime DNA Polymerase (Invitrogen) and the other half with Phusion DNA Polymerase (New England Biolabs) to generate sequencing libraries. The size distribution of the Accuprime library closely matched that of the Kent et al. (1) dataset, as is evident from a comparison of the Pho2 V-plots with those in *F*.

1. Kent NA, Adams S, Moorhouse A, Paszkiewicz K (2011) Chromatin particle spectrum analysis: A method for comparative chromatin structure analysis using paired-end mode next-generation DNA sequencing. *Nucleic Acids Res* 39:e26.

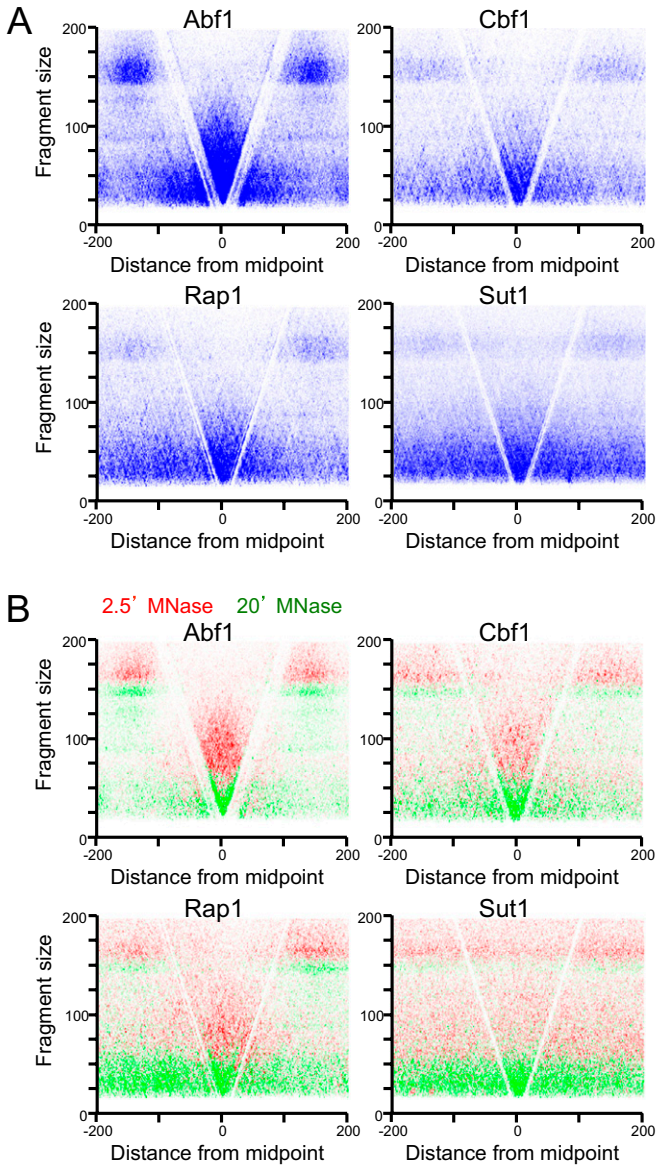

Fig. S7. Transcription factor midpoint vs. length maps for solubilized chromatin. (*A*) Composite of 2.5- and 20-min time points. (*B*) Separated time points (2.5 min red, 20 min green). See the legend to Fig. 5 for details.

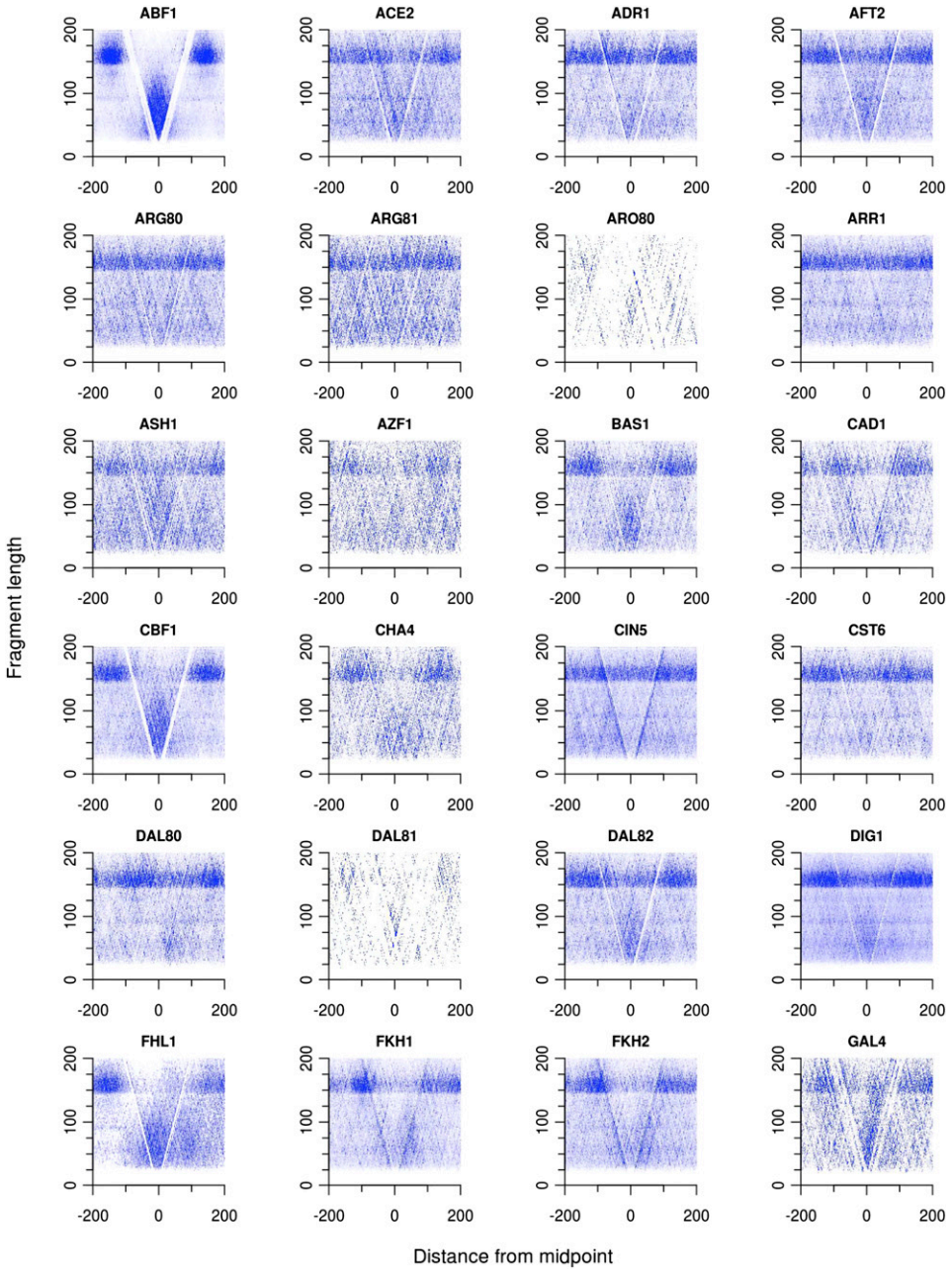

Fig. S8. (Continued)

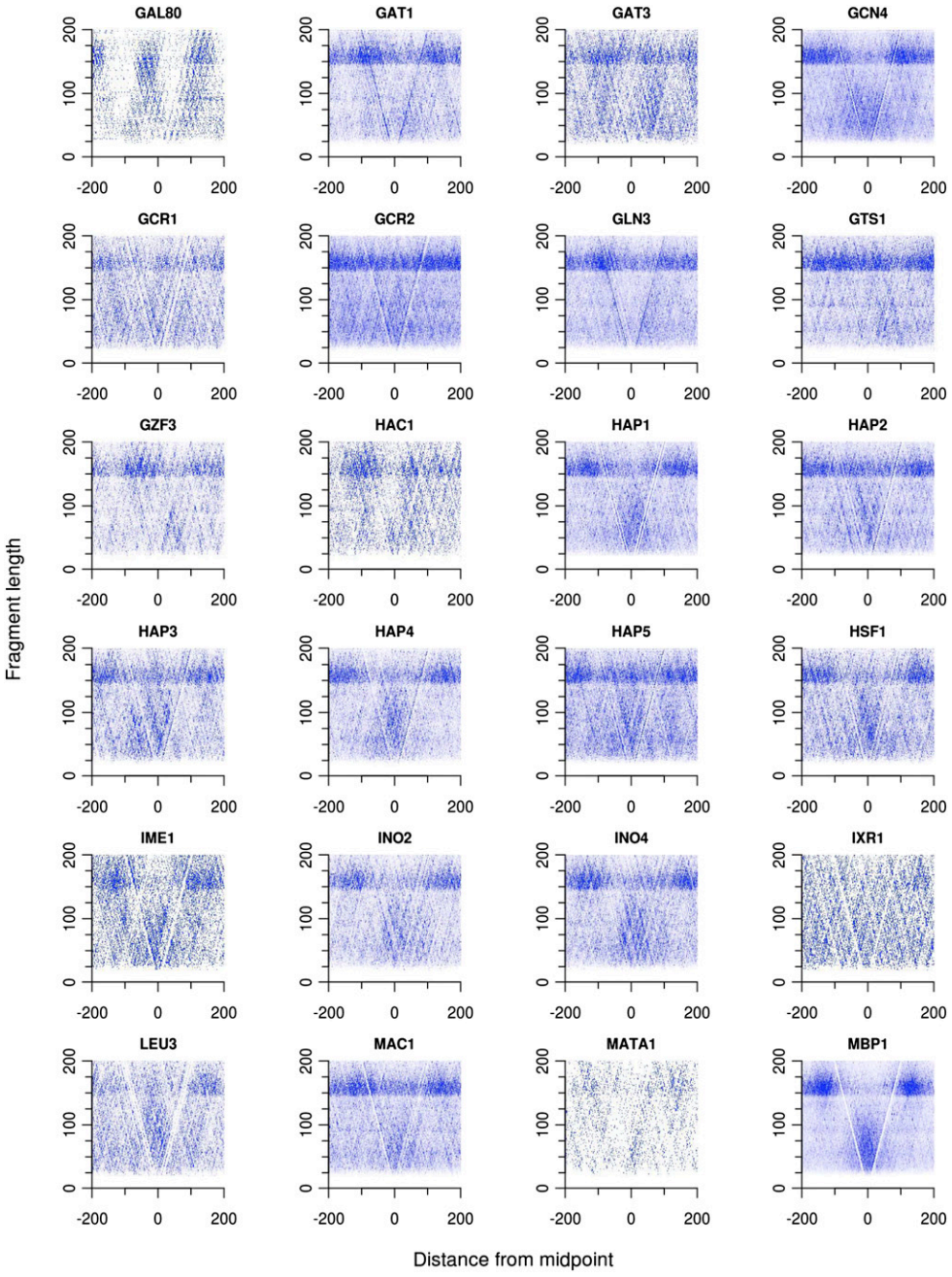

Fig. S8. (Continued)

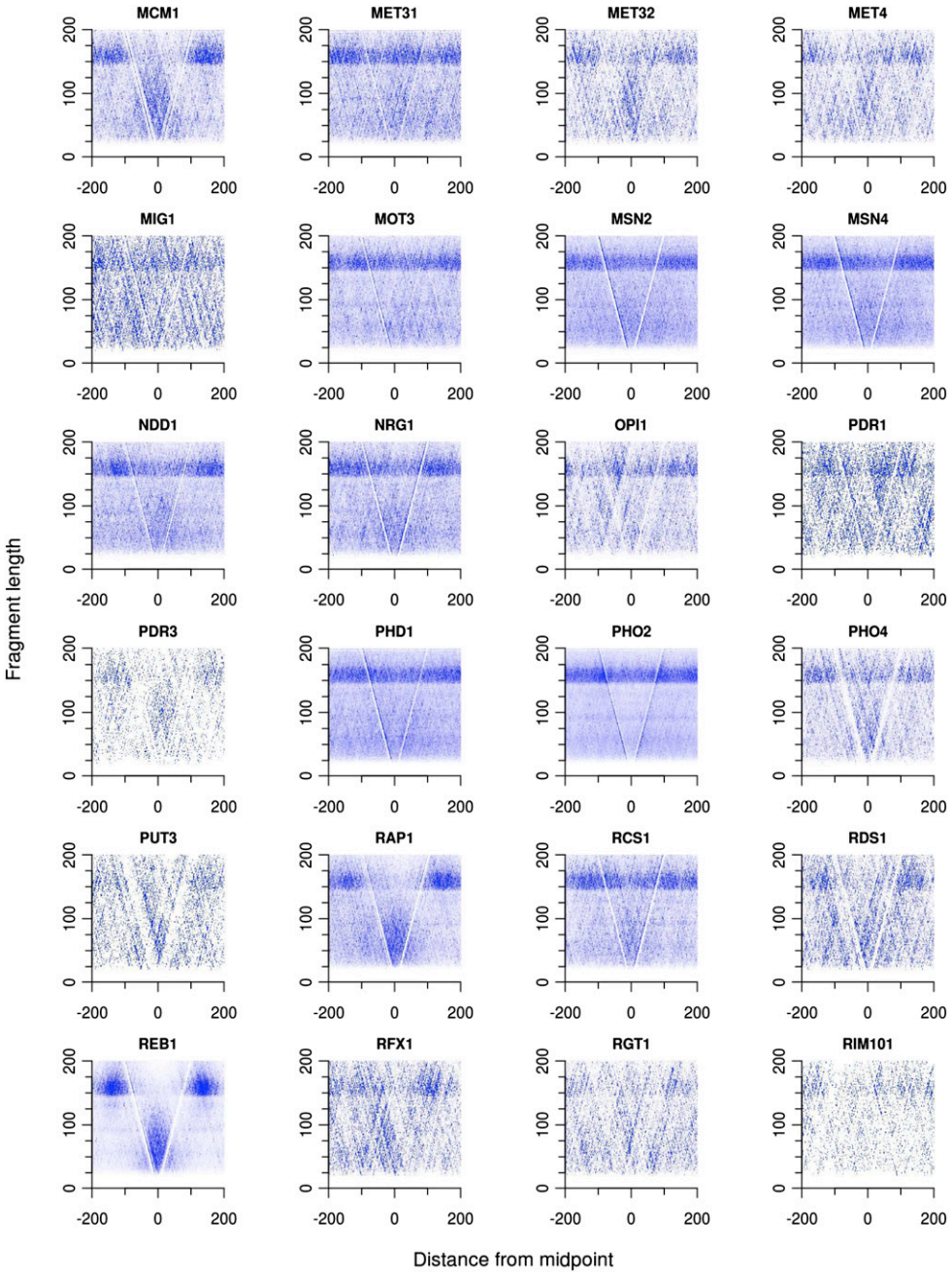

Fig. S8. (Continued)

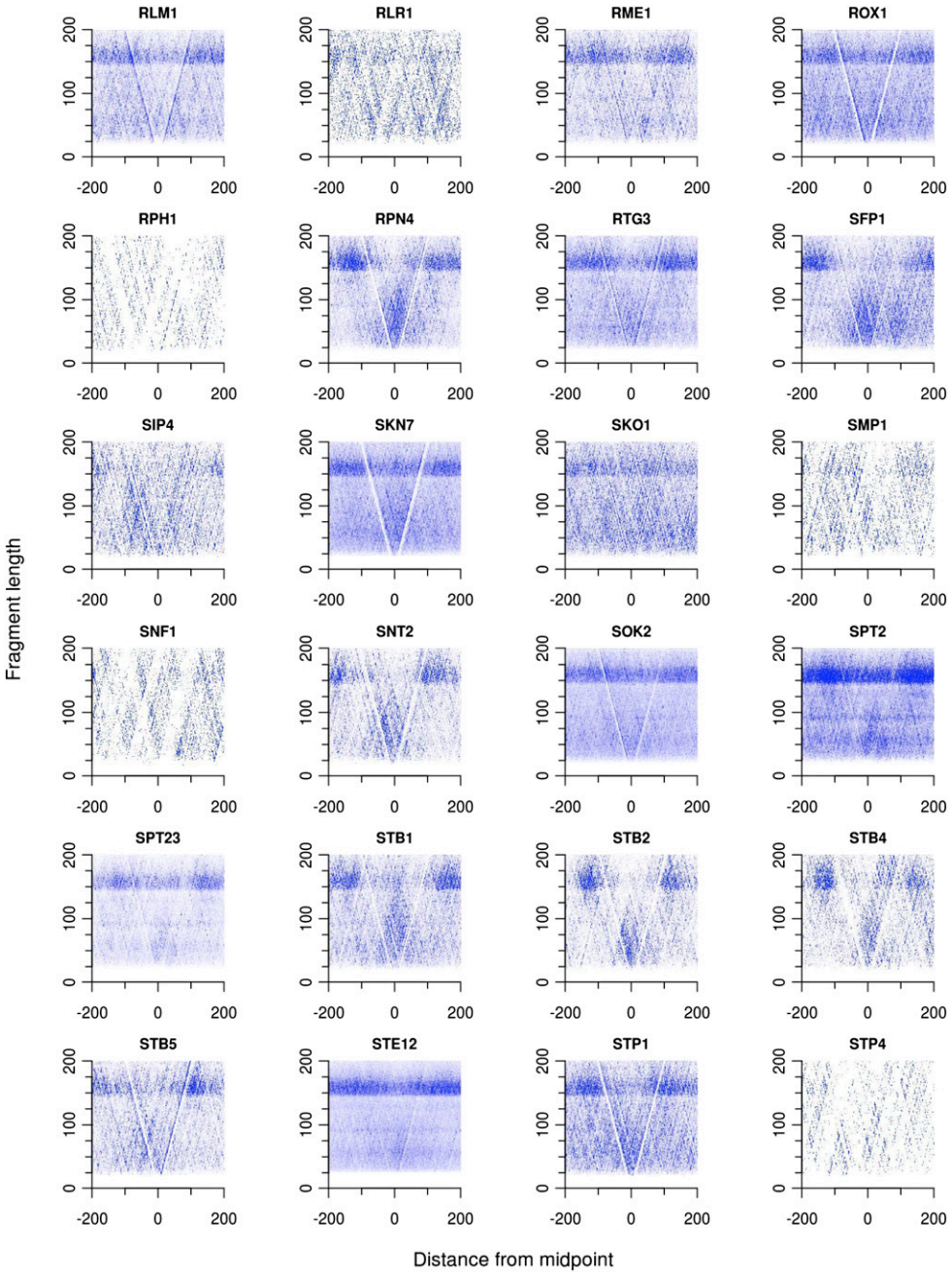

Fig. S8. (Continued)

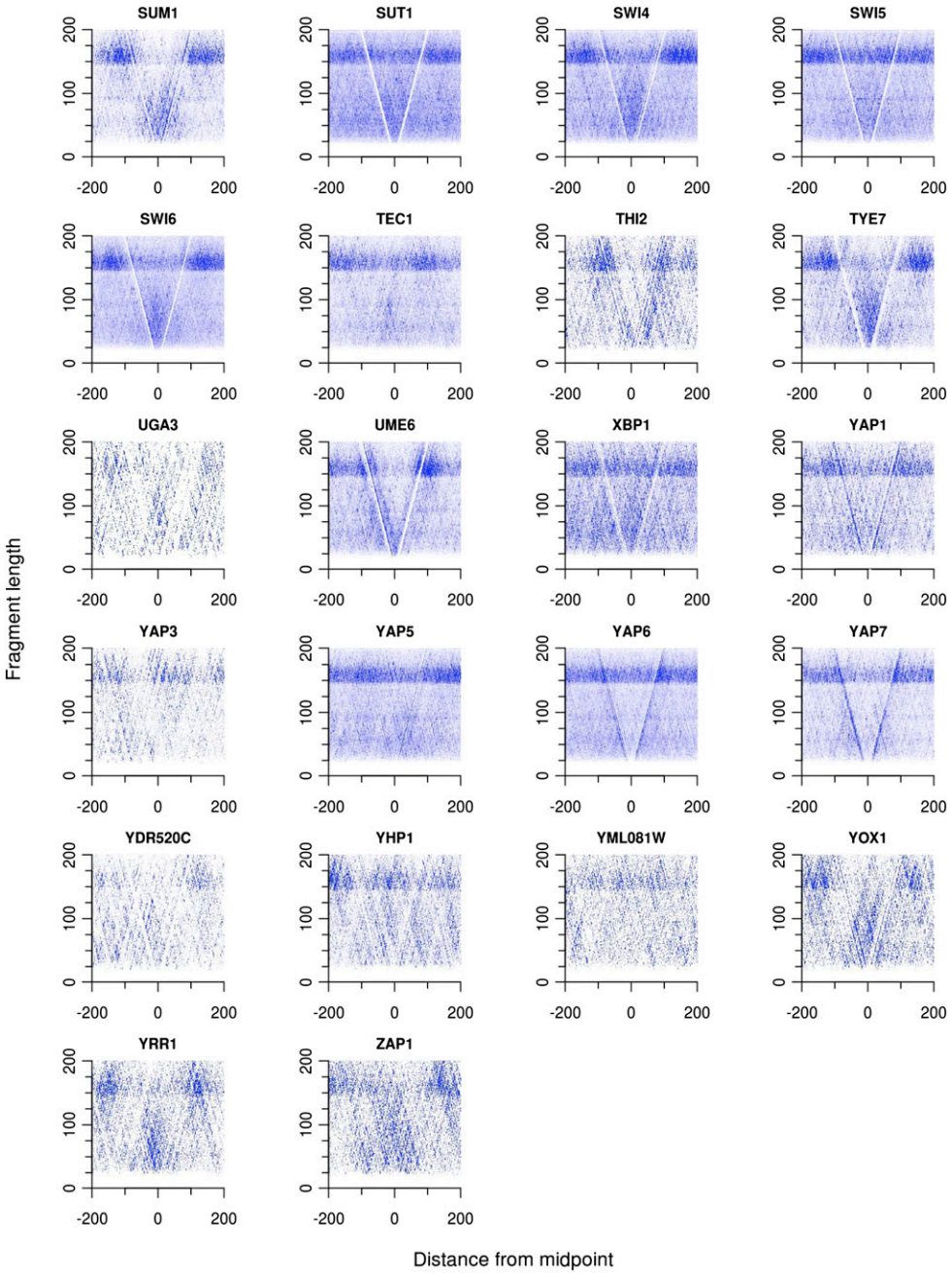

Fig. S8. Comprehensive V-plots for 118 transcription factors. See Figs. 4 and 5 for selected examples drawn from this full set. We attribute the lack of Vs in many of the V-plots to technical limitations, insofar as the binding site identifications were obtained by using a greedy algorithm together with conservation information to find consensus sequences within low-resolution ChIP data (1). The sites that give clear V patterns are most likely those for which the algorithm succeeded, and those without a V might have been more challenging. These 118 transcription factors are the only ones for which the algorithm succeeded in finding a consensus sequence, and so it is likely that many other consensus sequences are too weak for accurate predictions to be made. For example, the AZF1 consensus is TTTTTCTT, and might well have been spurious, identifying T-rich regions, which are common in the yeast genome near promoters, but which might indicate specific AZF1 binding.

1. MacIsaac KD, et al. (2006) An improved map of conserved regulatory sites for Saccharomyces cerevisiae. *BMC Bioinformatics* 7:113.

Table S1. Summary statistics of MNase-protected fragments spanning transcription factor binding-site midpoints

Transcription factor	MNase (min)	Number	Fragment length	SD	Distance	SD	Nucleosome positioning
Abf1	2.5′	781290	141.809	83.65	0.575	49.059	Strong
	20′	310097	116.879	69.009	2.255	42.149	
Aft2	2.5′	379596	182.881	86.713	1.861	58.18	Weak
	20′	261834	152.232	66.797	2.071	49.232	
Cbf1	2.5′	417751	166.836	88.075	−2.994	55.286	Moderate to strong
	20′	231094	135.176	70.354	−2.777	47.554	
Cin5	2.5′	362352	183.676	85.242	0.108	60.151	Very weak
	20′	229066	156.048	62.582	0.305	48.76	
Dig1	2.5′	928089	178.718	86.532	−1.048	58.878	Weak
	20′	609343	151.377	65.191	−0.16	49.295	
Fhl1	2.5′	346483	156.699	86.926	4.244	53.67	Strong, asymmetric
	20′	141788	121.371	68.036	−0.031	42.706	
Fkh1	2.5′	330019	173.677	87.16	−7.089	60.773	Moderate, asymmetric
	20′	168828	151.376	65.103	−6.927	50.849	
Gcn4	2.5′	444519	171.324	87.613	1.875	59.034	Moderate, asymmetric
	20′	254169	143.948	67.881	1.865	49.82	
Mbp1	2.5′	578283	161.574	90.481	0.003	57.855	Strong
	20′	14001	127.595	72.714	−1.296	46.2	
Mcm1	2.5′	135952	169.655	87.654	−3.398	54.037	Moderate to strong
	20′	81551	133.426	66.991	−1.352	42.783	
Mot3	2.5′	335304	183.53	86.379	−1.818	60.213	None
	20′	235505	152.425	65.463	−2.16	49.523	
Msn2	2.5′	603766	180.469	87.062	−5.565	58.843	Very weak
	20′	422441	147.265	66.377	−5.144	48.549	
Msn4	2.5′	559720	179.545	86.746	−7.347	58.977	Very weak
	20′	387265	148.356	66.253	−7.379	48.852	
Nrg1	2.5′	335427	180.739	87.949	−0.412	59.839	Weak
	20′	223324	149.183	67.41	−0.805	49.722	
Phd1	2.5′	932670	183.366	86.387	−3.871	59.592	Very weak
	20′	635875	150.334	65.445	−3.568	48.754	
Pho2	2.5′	1211144	186.86	86.34	−0.943	60.292	Very weak
	20′	840056	157.879	63.469	−0.707	49.587	
Rap1	2.5′	517151	161.193	87.911	−2.472	55.124	Moderate to strong
	20′	232901	129.079	69.223	−0.286	44.756	
Reb1	2.5′	657354	143.099	86.685	−4.457	53.365	Strong
	20′	260869	124.073	71.246	−6.973	46.741	
Rpn4	2.5′	152676	159.734	87.828	4.121	57.493	Moderate, asymmetric
	20′	80728	136.475	70.409	2.752	50.668	
Skn7	2.5′	582884	183.481	87.556	1.192	57.751	Weak
	20′	410176	146.218	67.242	0.959	46.701	
Sok2	2.5′	703896	183.232	86.856	−3.472	59.386	Very weak
	20′	477289	150.013	66.28	−3.376	48.755	
Ste12	2.5′	860405	180.615	86.482	−1.213	59.269	Very weak
	20′	569296	152.301	64.778	−1.175	49.423	
Sut1	2.5′	674817	185.516	87.774	−0.399	58.513	Very weak
	20′	484782	146.494	67.151	−0.141	46.936	
Swi4	2.5′	553035	177.865	88.19	0.042	58.017	Weak
	20′	353482	143.57	67.549	−0.493	47.49	
Swi6	2.5′	688475	175.909	89.016	−1.812	58.302	Moderate
	20′	429072	143.678	67.963	−2.077	47.448	
Tye7	2.5′	150165	166.8	88.059	−1.867	55.292	Moderate
	20′	84823	137.741	69.326	−1.433	48.204	
Ume6	2.5′	287827	179.665	90.632	−2.318	64.331	Moderate, asymmetric
	20′	192830	148.801	69.749	−2.114	54.289	
Yap5	2.5′	360987	176.617	84.94	−3.07	59.207	None
	20′	221289	150.787	64.228	−2.521	49.373	
Yap6	2.5′	434494	183.852	85.33	−0.123	60.723	Very weak
	20′	276176	156.78	63.435	−0.59	49.825	

Fragment lengths and distance of each fragment midpoint from the center of each binding site motif for each dataset are included.