

# **Nucleosome fragility is associated with future transcriptional response to developmental cues and stress in *C. elegans***

Tess E. Jeffers<sup>1\*</sup> and Jason D. Lieb<sup>1</sup>

<sup>1</sup> Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

\* Correspondence should be addressed to: Tess Jeffers, [tjeffers@princeton.edu](mailto:tjeffers@princeton.edu)

**Running title** (50 characters or less): Nucleosome Organization and Fragility in *C. elegans*

**Keywords:** Nucleosome Fragility, Nucleosome Resistance, *C. elegans*, Chromatin, MNase

# **ABSTRACT**

Nucleosomes have structural and regulatory functions in all eukaryotic DNA-templated processes. The position of nucleosomes on DNA and the stability of the underlying histone-DNA interactions affect the access of regulatory proteins to DNA. Both stability and position are regulated through DNA sequence, histone post-translational modifications, histone variants, chromatin remodelers, and transcription factors. Here, we explored the functional implications of nucleosome properties on gene expression and development in *C. elegans* embryos. We performed a time-course of micrococcal nuclease (MNase) digestion, and measured the relative sensitivity or resistance of nucleosomes throughout the genome. Fragile nucleosomes were defined by nucleosomal DNA fragments recoverable preferentially in early MNase-digestion time points. We found fragile nucleosomes at locations where we expected to find destabilized nucleosomes, like transcription factor binding sites where nucleosomes compete with DNA-binding factors. Contrary to our expectation, the presence of fragile nucleosomes in gene promoters was anti-correlated with transcriptional activity. Instead, genes with fragile nucleosomes in their promoters tended to be expressed in a context-specific way, operating in neuronal response, the immune system, and stress response. Nucleosome fragility at these promoters was strongly and positively correlated with the AT content of the underlying DNA. There was not a strong correlation between promoter nucleosome fragility and the levels of histone modifications or histone variants. Our data suggest that in *C. elegans* promoters, nucleosome fragility is primarily a DNA-encoded feature that poises genes for future context-specific activation in response to environmental stress and developmental cues.

# INTRODUCTION

The fundamental unit of eukaryotic chromatin is the nucleosome, which consists of 147 bp of DNA wrapped around an octamer of histone proteins (Luger et al. 1997). Nucleosomes have important structural and regulatory functions in organizing the genome and restricting access of regulatory factors to the DNA sequence (Henikoff 2008). As such, the interactions between nucleosomes and DNA strongly influence the regulation of gene expression by determining DNA accessibility for transcription factors and RNA polymerase. In addition to regulated nucleosome assembly and disassembly through the action of histone chaperones and chromatin remodelers, nucleosome stability is influenced by histone modifications, histone variants, DNA features encoded in *cis*, and competition with DNA-binding factors in *trans*. For the purposes of this manuscript, we define “stability” qualitatively, as the propensity of a given nucleosome to remain intact and at that position, rather than to be evicted, disassembled, or translocated to a substantially different position. A complete picture of the mechanisms governing nucleosome stability is fundamental to understanding how gene expression is dynamically regulated.

Nucleosome stability has been studied *in vitro* using sensitivity to enzymatic digestion or salt concentration (Bloom and Anderson 1978; Simon and Felsenfeld 1979; Burton et al. 1978; Li et al. 1993; Jin and Felsenfeld 2007; Wu and Travers 2004; Polach and Widom 1995; 1999; Anderson et al. 2002). Genome-wide adaptations of these methods have been used to identify nucleosome position and stability *in vivo*. More recent studies in yeast, *Drosophila*, plants, and mammals have used varying concentrations of the enzyme micrococcal nuclease (MNase) to identify nucleosomes with differential sensitivity to MNase digestion *in vivo* (Xi et al. 2011; Henikoff et al. 2011; Chereji et al. 2015; Kent et al. 2011; Weiner et al. 2010; Vera et al. 2014; Lombraña et al. 2013; Kubik et al. 2015). Nucleosomes sensitive to low concentrations of MNase have been labeled as “fragile”, and have been associated with transcription factor

binding sites (Vera et al. 2014), active origins of replication (Lombraña et al. 2013), gene promoters (Xi et al. 2011), and genomic sequences with high AT content (Chereji et al. 2015). Thus, both DNA-encoded sequence features and active processes in *trans* influence nucleosome fragility. The relationship between fragility and nucleosome function remains unclear. For example, one study reported fragile nucleosomes at the promoters of repressed stress-response genes during normal growth (Xi et al. 2011), while another found fragile nucleosomes at the promoters of highly transcribed genes in yeast (Kubik et al. 2015).

We performed a timecourse of MNase digestion in *C. elegans* mixed-stage embryos to study the relationship between nucleosome fragility and gene activity in a developing multicellular organism. In our study, fragile nucleosomes were associated with lowly expressed genes and genes expressed in a context-specific fashion. Although we found that competition with *trans* factors promoted nucleosome fragility, our data suggest that the majority of highly fragile nucleosomes in the *C. elegans* embryo are not due to *trans* factors but rather to *cis* features encoded in the DNA sequence. Together, our data indicates that the fragility of nucleosome-DNA interactions may aid in poising genes for induction in response to stress or developmental cues.

## RESULTS

### A digestion timecourse identifies nucleosomes with differential MNase sensitivity

We postulated that functionally distinct nucleosomes in the *C. elegans* could be distinguished by the length of time it took them to be liberated from bulk chromatin by MNase digestion. Previous studies using this approach defined nucleosomes released early in the timecourse as “fragile” and those released later in the timecourse as “resistant” (Xi et al. 2011). To identify nucleosomes of differential sensitivity genome-wide, we isolated mixed-stage embryos from *C. elegans*, treated them with formaldehyde to cross-link the chromatin, isolated nuclei, and

digested the chromatin with MNase (**Figure 1A**). After 2, 4, 8, 15, and 30 minutes of digestion we removed a chromatin aliquot and performed paired-end Illumina sequencing on the mononucleosomal fragments liberated at each time point (**Figure 1B**). Without MNase, chromatin remained intact and undigested. After addition of the enzyme, a stereotypic chromatin ladder rapidly formed, and a small proportion of total chromatin became mononucleosomal. As digestion proceeded, the mononucleosomal fraction increased while polynucleosomal fractions were depleted (**Figure 1C**). We performed two replicate experiments on native chromatin and two replicate experiments on formaldehyde-fixed chromatin samples. Results from the native and fixed chromatin were very similar (**Supplemental Figure 1**). We therefore focused our downstream analysis on fixed chromatin for maximum compatibility with previously-generated datasets.

Mononucleosomal DNA fragments released earliest during the digestion were larger (median size of 2-minute nucleosomal fragments: 180 bp) than fragments released later in the timecourse (median of 30-minute nucleosomes: 155 bp) (**Figure 1D**). Among the digestion timepoints, nucleosome sizes decreased in 10 bp increments, reflecting the MNase digestion preference for WW (AA, AT, TA, or TT) dinucleotides and the 10-11 bp periodicity of the DNA helical turn (McGhee and Felsenfeld 1983; Deniz et al. 2011; Trifonov and Sussman 1980; Ioshikhes et al. 2011) (**Figure 1D**). This is consistent with the model that with increasing lengths of digestion time, MNase will cleave long linkers and any unwrapped ends of nucleosomal DNA. Although the genome-wide occupancy profiles of mononucleosomal fragments were globally similar across the timepoints (**Figure 1B, Supplemental Figure 2**), there were a number of substantial differences in the nucleosome maps among the timepoints (**Supplemental Figure 3, Figure 1E**).

To systematically study nucleosomes of differential sensitivity to MNase, we assigned each

nucleosome both a fragility and a resistance score (**Supplemental Figure 4**). For each timepoint, we first called nucleosome positions and then assigned each nucleosome an occupancy score (see Methods for details). The fragility score for a nucleosome is defined by subtracting the average occupancy score of the intermediate timepoints (4, 8, and 15 min) from the occupancy score of the 2 min timepoint. Conversely, a resistance score is computed by subtracting the average occupancy score of the intermediate timepoints from that of the 30 min timepoint (**Figure 1E**). Thus, fragility and resistance scores were generally reciprocal to each other at a given nucleosome, but not necessarily so. We defined the top 10% of nucleosomes with the highest fragility or resistance scores as “fragile” or “resistant” nucleosomes, respectively (**Figure 1F**).

## ***Trans*-factors increase nucleosome fragility**

We sought to address whether nucleosome fragility was a consequence of competition with DNA binding proteins and other *trans* factors. *Trans*-acting factors disrupt nucleosomes by competing with histones for binding to the DNA sequence (Simpson 1990; Adams and Workman 1995). We first examined regions of the genome where we expected to find nucleosomes destabilized by competition with other DNA-binding factors, for example at transcription factor binding sites (TFBS). We collected a set of 35,062 TFBS bound at any stage of *C. elegans* development, as identified by transcription factor (TF) ChIP-seq from the modENCODE consortium (Araya et al. 2014). TFBS in the *C. elegans* genome on average show strong affinity to histones *in vitro* (Locke et al. 2013). A nucleosome occupancy model based solely on DNA sequence also predicted *C. elegans* TFBS to be nucleosome bound (Kaplan et al. 2009) (**Figure 2A**). *In vivo*, however, these sites show a dip in nucleosome occupancy, consistent with the footprint of TF binding. Moreover, TFBS had high fragility scores on average (**Figure 2B**). These data are in agreement with previous reports from yeast to humans that transcription factors compete with nucleosomes for access to DNA (Barozzi et al.

2014; Wang et al. 2012; Ozonov and van Nimwegen 2013). To further investigate the relationship between TF binding and fragility we broke TFBS into groups depending on the number of TFs bound at a site. Although the majority of TFBS identified in *C. elegans* are bound by a single factor, some sites may be bound by many different TFs (Chen et al. 2014; Araya et al. 2014; Boyle et al. 2014). Fragility scores increased as a function of the number of TFs bound at a single TFBS (**Figure 2C**).

We found that TFBS had high fragility scores despite their intrinsic preference for nucleosome formation *in vitro*. One possible explanation is that transcription factors destabilize nucleosomes at their binding sites, causing the fragility at TFBS. Alternatively, TFBS may contain DNA sequences that disfavor nucleosome formation *in vivo*, thereby increasing nucleosome fragility. To distinguish among these possibilities, we identified a set of TFBS specifically bound at different developmental stages (**Figure 2D**). We hypothesized that if active competition with TFs increases nucleosome fragility, then TFBS bound by TFs only in the embryo should be fragile in embryos, whereas TFBS bound only in the L4 larval stage should not be fragile in embryos (**Figure 2E, top**). Alternatively, if DNA sequence influences nucleosome fragility then the embryo-specific and L4-specific TFBS should be equally fragile in embryos (**Figure 2E, bottom**). Due to their high fragility scores and dynamic nature, we focused our analysis on HOT regions, TFBS where significant enrichments (false discovery rate <5%) in multiple transcription factor binding sites are observed (Araya et al. 2014). We found that embryo-specific HOT regions had high nucleosome fragility and low nucleosome occupancy (**Figure 2F, Supplemental Figure 5**). In contrast, L4-specific HOT regions showed lower nucleosome fragility in embryonic samples and higher nucleosome occupancy than embryo-specific HOT regions (**Figure 2G**). These results support the hypothesis that active competition with transcription factors *in vivo* contributes to nucleosome fragility despite their intrinsically nucleosome favoring properties *in vitro*.

## Nucleosome fragility increases throughout heat-shock genes upon induction

The preceding analysis found a correlation between transcription factor binding and nucleosome fragility. We next sought to test the relationship between fragile nucleosomes and *trans* factors more explicitly. Moderate transcription levels have been shown to cause displacement and repositioning of nucleosomes in gene bodies (Reinberg and Sims 2006). At extremely highly transcribed genes, such as heat shock responsive genes after induction, it has been proposed that RNA Pol II molecules occupy the entire gene body (Merz et al. 2008; Cole et al. 2014; Schwabish and Struhl 2004; Kristjuhan and Svejstrup 2004). We hypothesized that nucleosome fragility should increase at gene bodies after inducing high levels of transcription, as a result of nucleosome competition with transcribing RNA polymerase II (Pol II). To test whether we could induce nucleosome fragility, we designed a heat shock experiment in conjunction with an MNase-seq timecourse (**Figure 3A**).

Heat shock in *C. elegans* activates HSF-1 and HSF-2, two homologues of the mammalian HSF1 transcription factor, which bind heat shock elements (HSE) in the promoters of heat shock-responsive genes to upregulate their expression (Åkerfelt et al. 2010). Upon heat shock, the heat-shock-response genes undergo rapid chromatin remodeling and colocalize with the nuclear pore complex (Rohner et al. 2013). Using RNA-seq, we identified 14 genes that are rapidly upregulated after a brief (20 minute) heat shock at 34 °C (**Figure 3B, Supplemental Figure 6**). We then analyzed how fragility scores changed at those genes after heat shock (**Supplemental Figure 7**). Though nucleosome occupancy remained largely unchanged, we found nucleosome fragility dramatically increased both 5' and 3' of heat-shock genes, as well as in the gene body itself (**Figure 3C, 3D**). Notably, promoter and +1 nucleosome fragility increased on average genome-wide, although gene-body fragility was specific to the set of heat



shock-induced genes (**Figure 3E**). This result suggests that the rapid induction of these genes increases nucleosome competition with Pol II. This mode of transcription has been suggested to remove the entire histone octamer, rather than FACT-mediated H2A-H2B recycling (Kulaeva et al. 2010; Kireeva et al. 2002). Alternatively, previous studies of the Hsp70 locus in *Drosophila* have shown that heat shock induces rapid transcription-independent loss of gene-body nucleosomes (Petesch and Lis 2008). Future experiments with transcription inhibitors may clarify the exact mechanism by which gene body nucleosomes become fragile after heat shock. Our results demonstrate that nucleosome fragility can be modulated by *trans*-acting factors like transcription factors and RNA polymerase II, and is not solely dependent on DNA sequence.

# **Nucleosome fragility near genes is anti-correlated with expression**

We found high fragility scores at the types of genomic locations where destabilized nucleosomes had been previously reported, namely transcription factor binding sites and the gene bodies of newly induced genes (Urnov and Wolffe 2001). We investigated the genome-wide distribution of fragile nucleosomes (nucleosomes with the highest 10% of fragility scores; **Figure 1F**) in detail. Fragile nucleosomes were enriched 5' and 3' of genes, specifically at the promoter -2, -1, and +1 nucleosomes, and at the terminal nucleosome (TN) and TN+1 nucleosomes (**Figure 4, Supplemental Figures 1, 8**). Resistant nucleosomes (nucleosomes with the highest 10% of resistance scores (**Figure 1F**) were enriched in gene bodies (**Supplemental Figures 1, 8**).

With increasing distance from the promoter, nucleosomes become less sharply positioned (Mavrich et al. 2008; Yuan et al. 2005). To investigate whether the higher nucleosome resistance scores at highly expressed genes were a consequence of more consistent nucleosome positioning, we asked whether fragility or resistance scores were correlated with

the standard deviation of the nucleosome center, or “fuzziness”. Neither nucleosome fragility nor nucleosome resistance scores were correlated with the fuzziness of the nucleosome at the intermediate timepoint (fragility vs. fuzziness:  $R = 0.03$ , resistance vs. fuzziness:  $R = -0.06$ ), suggesting that the susceptibility or resistance to MNase digestion is not a direct consequence of more or less well-positioned nucleosomes (**Supplemental Figure 9**).

Given the positive relationship between fragility and transcription factor binding (**Figure 2C**), and given that induction resulted in increased nucleosome fragility throughout heat-shock genes (**Figure 3C, 3D**), we expected that nucleosome fragility would be enriched in the promoters or gene bodies of highly expressed genes in the embryo, which exhibit high levels of TF and Pol II binding as measured by ChIP-seq (**Supplemental Figure 10**) (Ho et al. 2014). To our surprise, gene expression levels on average were anti-correlated with nucleosome fragility at both promoters and gene bodies (**Figures 4A, 4B, 4F, Supplemental Figure 12**.  $R = -0.17$ ) and positively correlated with nucleosome resistance (**Figures 4A, 4B, 4G, Supplemental Figure 12**.  $R = 0.11$ ). For example, in contrast to our earlier observation at the heat shock genes, we failed to find a correlation between ongoing expression and nucleosome fragility at the gene bodies of the most highly transcribed genes (compare **Figure 3D** to **Supplemental Figure 11**). Perhaps only newly-induced genes display gene-body fragility, or extremely high levels of transcription are required to induce fragility in gene bodies.

Although overall nucleosome fragility scores were high 5' and 3' of all genes, including at the majority of TFBS (**Figure 2**), fragile nucleosomes occurred preferentially at the promoters of lowly-expressed genes (**Figures 4B, 4F, Supplemental Figure 10**). No single transcription factor profiled in the embryo significantly overlapped the distribution of fragile nucleosomes (**Figure 1F, Supplemental Figure 10**). Though previous reports have suggested that the histone variant H2A.Z may act to promote nucleosome instability (Jin and Felsenfeld 2007; Jin

et al. 2009; Xi et al. 2011), we did not observe a significant overlap between previously-identified H2A.Z-containing nucleosomes (Ho et al. 2014) and fragile nucleosomes (**Figures 4D, Supplemental Figure 13**). These data collectively suggest that two separate mechanisms account for nucleosome fragility depending on the genomic context. In places where nucleosomes are directly in competition with transcription factors (**Figure 2**) or in the bodies of exceptionally highly expressed or newly induced genes (**Figure 3**) fragility arises through competition with transcription factors or other DNA-associated proteins. By contrast, fragility 5' and 3' of genes at locations with few TF binding events appears to be determined by another mechanism, which we explored next.

# **Nucleosome fragility is correlated to *cis*-encoded DNA features**

We hypothesized that *cis* features may be responsible for the fragility of nucleosomes at the promoters of lowly-expressed genes. We examined the DNA sequences occupied by fragile and resistant nucleosomes, and compared these to sequences occupied by all nucleosomes in the genome (**Figure 5**). Compared to the set of all nucleosomes, DNA sequences occupied by fragile nucleosomes had lower GC sequence content on average, a feature favoring nucleosome formation (**Figure 5A**) (Deniz et al. 2011). We then asked whether these sequences were likely to form nucleosomes based on a previously reported *in vitro* reconstitution assay (Locke et al. 2013). We found that DNA sequences occupied by fragile nucleosomes in the embryo were generally less occupied *in vitro* (**Figure 5C**). Finally, we observed that sequences occupied by fragile nucleosomes were less conserved across nematodes than DNA sequences occupied by the set of all nucleosomes (**Figure 5E**).

Poly(dA:dT) tracts disrupt nucleosome formation and tend to increase transcription of downstream genes (Raveh-Sadka et al. 2012), while TATA box motifs in yeast are associated

with bendable promoters sensitive to chromatin remodelers (Tirosh et al. 2007; Albert et al. 2007). In *C. elegans*, the number of T-block motifs (3 to 5 consecutive thymine nucleotides, often spaced at 10 bp periodicity) have been positively correlated with expression: genes with more than 5 T-blocks have fivefold higher expression than genes with fewer than 4 T-blocks (Grishkevich et al. 2011), presumably through a reduction in promoter nucleosome occupancy. T-blocks were not enriched at fragile or resistant nucleosomes, whereas TATA box motifs were enriched at fragile nucleosomes (**Supplemental Figure 14**).

In addition to DNA-encoded *cis* features, promoter fragility at lowly-transcribed genes may be influenced by epigenetic features associated with these nucleosomes. Through comparison with previously generated datasets, we asked whether any histone post-translational modifications, histone variants, or chromatin states were positively associated with nucleosome fragility or resistance (Ho et al. 2014; Ooi et al. 2010). Only “low signal” chromatin states and chromatin extracted with 80 mM salt (another method proposed to identify unstable nucleosomes) were associated with fragile nucleosomes (**Supplemental Figures 15, 16, 17**) (Ooi et al. 2010; Ho et al. 2014). Though longer linkers were weakly correlated with increased fragility levels (**Supplemental Figure 18**), the GC content of the nucleosome was the strongest predictor of overall nucleosome fragility score (**Supplemental Figure 19**). Taken together, our data indicate that fragile nucleosomes in gene promoters are correlated with high AT content and TATA box motifs. It seems that most promoters are fragile at least in part due to high AT content (see residual fragility in **Figure 2G**), but that this *cis* effect of DNA sequence becomes apparent only at sites where the observation is not confounded by TF binding.

## **Fragile nucleosomes are associated with genes expressed in future or context-specific situations**

To infer potential functional implications of nucleosome fragility in the developing embryo, we

next asked which genes were significantly associated with fragile nucleosomes. Fragile nucleosomes were enriched at the -1 nucleosome and resistant nucleosomes were enriched at the +1 nucleosome (**Figure 4**). We identified two sets of genes that contain fragile and resistant nucleosomes, respectively, +/- 500 bp from their transcript start site. We then identified enriched GO terms in each of the gene sets (Supek et al. 2011; Al-Shahrour et al. 2004). Genes with fragile nucleosomes were enriched for GO terms related to neuronal response, immune response, and stress response genes (“sensory perception of chemical stimulus”, “defense response”, “pharynx development”, “immune system process”) (**Figure 6A**). In contrast, genes with resistant nucleosomes were enriched for embryogenesis and cell cycle related terms (“mitotic cell cycle”, “RNA processing”, “regulation of developmental process”, “organic substance transport”) (**Figure 6B**). A complementary analysis confirmed the association between high fragility scores in the promoter and genes that function in context-specific processes. Unbiased k-means clustering of promoters based on their fragility scores identified a cluster containing lowly-expressed genes with high fragility scores that are expressed in response to perception of chemical stimulus, cognition, and neurological processes (**Supplemental Figure 20**). This second class of fragile nucleosomes in the embryo is unlikely to be fragile due to the action of *trans* factors because of their anti-correlation with expression and transcription factor binding. Rather, fragile nucleosomes were associated with lowly-transcribed genes that are expressed in a future context-specific fashion during stress response or development.

To confirm the association between fragile nucleosomes and future context-specific expression with an independent method, we used the publicly available modENCODE transcriptome sequencing data from seven different life stages to define a set of “developmentally regulated” genes and a set of “stably expressed” genes (**Figure 6C**) (Pérez-Lluch et al. 2015; Gerstein et al. 2014; Spencer et al. 2011). We hypothesized that if promoter nucleosome fragility is related

to context-specific expression as our GO analysis suggested, then we should find higher fragility signals near developmentally regulated genes. When we plotted the average fragility scores around these genes, we indeed saw higher nucleosome occupancy and fragility signals at developmentally regulated genes as compared to the set of stably expressed genes (**Figure 6D**). While both sets of genes have fragile promoters, our data indicate that fragility is enriched at genes that tend to be expressed specifically during development, stress, or environmental stimulus response. Together, we suggest that these sequences may reflect a specialized promoter architecture that is primarily determined by high AT content, which acts to allow future disruption of nucleosome stability, and thereby the rapid induction of gene expression in a context-specific fashion (**Figure 7**).

## DISCUSSION

We performed an MNase digestion timecourse, a simple modification to the traditional MNase digestion assay, in *C. elegans* embryos. Our experiment measured which individual nucleosomes were most quickly released from their polynucleosome context after exposure to MNase. Sensitivity to MNase digestion, and thereby fragility or resistance as defined in this study, could be determined by a number of factors. These include (1) a DNA sequence in the linker region that is preferentially cut by MNase; (2) longer linker regions; (3) nucleosome instability caused by low DNA-histone affinity; or (4) instability caused by competition with transcription factors. As such, nucleosome fragility or resistance is likely to be a reasonable proxy for the underlying stability of the nucleosome. Two lines of evidence suggest that nucleosome fragility reflects nucleosome instability. First, we found that nucleosomes can be made fragile by competition with transcription factors and RNA polymerase II. Second, we observed that nucleosomes can be made unstable *in cis* by being wrapped around nucleosome-disfavoring DNA sequences with high AT content and TATA-box motifs. All of these factors have been shown to cause nucleosome instability in previous studies (Widom 2002; Ozonov

and van Nimwegen 2013).

We performed our experiments in nuclei derived from whole embryos, which reflect a mixture of cell types and creates challenges for data interpretation. However, our core conclusions stand regardless. First, the DNA sequence underlying the data is the same across all cell types and therefore our conclusions regarding the *cis* contribution to nucleosome fragility are derived from and shared among all cells in the embryo. Second, previous precisely-timed developmental RNA-seq experiments show that certain genes are stably expressed in every cell type of the developing embryo, while other gene classes are not expressed at all during the time that the embryos used in our sample were collected, but instead are poised for expression later in development (Spencer et al. 2011; Gerstein et al. 2014). Therefore, for specific genes, we can say definitively that they were “on” or “off” in our sample, and make general conclusions accordingly. Third, the RNA-seq data we used is also derived from mixed embryos, and therefore quantitatively matches our fragility and resistance data. The same applies to the modENCODE chromatin data; the embryos used in this study were staged specifically to match the embryos used in the modENCODE studies. While future studies of pure populations of cells may reveal additional relationships between nucleosome properties and gene expression, the conclusions we make here with whole embryos are valid and are compatible with previously generated TF and histone ChIP-seq datasets.

### *MNase resistant nucleosomes are correlated with expression in the embryo*

We found a class of nucleosomes that required relatively long durations of MNase digestion to be removed from chromatin. Traditional expectations might be that unstable nucleosomes would be found in the body of transcribed genes, and stable nucleosomes in silent, heterochromatic genomic regions. However, recent reports illustrate that nucleosomes in the gene body of transcribed genes are consistently well-positioned and highly occupied due to a number of

factors, including the activity of RNA Polymerase II and the histone chaperone FACT (FACilitates Chromatin Transcription) (Bai and Morozov 2010; Jiang and Pugh 2009). Indeed, we found stable nucleosomes enriched in the gene body of actively transcribed housekeeping genes. During transcription, Pol II disrupts an H2A-H2B dimer that remains bound by FACT and is rapidly reassembled in the wake of Pol II (Formosa 2012). Histone modifications such as H3K36me3, which we found positively correlated with stable nucleosomes, are also thought to contribute to nucleosome stability and maintenance of transcription fidelity (Lieb and Clarke 2005; Lickwar et al. 2009). Together, our measurements are in agreement with an emerging picture of highly regulated nucleosome stability throughout the genome, which is likely critical for regulation of DNA templated events like transcription, splicing, and DNA replication (Bintu et al. 2011; Kwak et al. 2013; Tilgner et al. 2009; Chen et al. 2010; Eaton et al. 2010).

### *MNase sensitive fragile nucleosomes are 5' enriched and anti-correlated with expression*

Differential MNase digestion and salt fractionation have been previously used to probe nucleosome-DNA stability. Results from yeast (Xi et al. 2011; Weiner et al. 2010; Kubik et al. 2015), plants (Vera et al. 2014), mouse (Lombrana et al. 2013; Deng et al. 2015), worm, (Ooi et al. 2010) and fly (Chereji et al. 2015; Henikoff et al. 2009) have identified highly labile nucleosomes in 5' and 3' "nucleosome free" regions. In yeast, Xi et al. observed that fragile nucleosomes were associated with H2A.Z containing promoter nucleosomes, believed to be involved in stress response (Zhang et al. 2005; Li et al. 2005). In vertebrates, nucleosomes containing both H3.3 and H2A.Z histone variants are unstable (Jin and Felsenfeld 2007; Jin et al. 2009). We did not observe a correlation between H2A.Z incorporation and nucleosome fragility in *C. elegans* as measured by our assay. Rather, H2A.Z distribution is strongly biased towards active genes (Whittle et al. 2008; Liu et al. 2011). This distinction could be due to a divergence in H2A.Z properties between yeast and *C. elegans* (Zlatanova and Thakar 2008). It



is also possible that the distribution of H2A.Z-containing nucleosomes used for this study (measured in Ho et al.) was comprised of the particularly stable, homotypic type of H2A.Z nucleosomes (Ishibashi et al. 2009).

Overall, we observed an anti-correlation between promoter nucleosome fragility and gene expression. Previous reports disagree about the relationship between nucleosome fragility and expression. Studies in yeast, *C. elegans*, *Drosophila*, and Maize have used salt profiling and different MNase concentrations to identify a positive correlation between promoter fragility and expression (Kubik et al. 2015; Ooi et al. 2010; Henikoff et al. 2009; Vera et al. 2014). However, Xi et al used the same approach and observed the opposite effect. One possible explanation for the discrepancy is that very light MNase digestion or salt profiling techniques may recover fragments from nucleosome-depleted loci, like the nucleosome-free regions in the promoters of active genes. If those regions are excluded, for example by slightly longer digestion times, higher enzyme concentrations, or by filtering out sub-nucleosomal fragments, truly unstable nucleosomes may become more apparent at the promoters of stress and context-responsive genes. We also point out that the relation between fragility and other methods used to measure nucleosome properties, such as profiling of salt-sensitive nucleosomes in *Drosophila* and *C. elegans*, remains unclear (Ooi et al. 2010; Henikoff et al. 2009). Both differential MNase digestion experiments and salt profiling experiments endeavor to measure some aspect of nucleosome stability but a rigorous study is needed to compare results from the two methods.

A previous study (Xi et al. 2011) identified nucleosome fragility at nearly one-third of all promoters of protein-coding genes, and enriched at the promoters of genes involved in stress response. When we assessed the types of functional annotations that were enriched at promoters with fragile nucleosomes in *C. elegans* embryos, we identified GO terms related to context-specific expression: sensory perception of chemical stimulus, defense response,

immune system process. Based on our findings in conjunction of those with Xi et al., we propose that nucleosome fragility may serve to poise genes for rapid activation in response to developmental or external stimuli. This is consistent with previous work investigating the transcriptional activation of mammalian primary response genes, where unstable nucleosomes are used to achieve rapid induction independent of chromatin remodeling complexes (Ramirez-Carrozzi et al. 2009).

We found high fragility scores at the -2, -1, and +1 nucleosomes of developmentally regulated genes in comparison to stably expressed housekeeping genes. Previous work found that the promoters of developmentally regulated genes lack the histone post-translational modifications associated with active genes, like H3K4me3 (Pérez-Lluch et al. 2015). Similarly, we were unable to find an association between fragile nucleosomes and any histone post-translational modifications examined by the modENCODE group. Given the increased fragility of these nucleosomes, it is possible that (1) these nucleosomes at developmentally regulated genes were lost from standard chromatin preparation protocols and are thus underrepresented in the histone ChIP, or (2) developmentally regulated genes use promoter nucleosome fragility as a mechanism for gene regulation.

Our results are reminiscent of previous reports from yeast, which propose that promoter structures can generally be classified as containing depleted proximal nucleosomes (DPN) or occupied proximal nucleosomes (OPN) (Tirosh and Barkai 2008). In yeast, DPN genes have low transcriptional plasticity (defined as the capacity to modulate transcription levels upon changing conditions), well positioned nucleosomes, and are enriched for TF binding sites and H2A.Z. In contrast, OPN genes have high transcriptional plasticity, higher evolutionary divergence, higher nucleosome turnover, and were sensitive to chromatin regulation. The yeast DPN genes may correspond to the set of stably-expressed genes we defined in *C. elegans*,

which that have depleted proximal nucleosomes (**Figure 6D top**). The yeast OPN genes may correspond to the set of developmentally regulated genes we defined in *C. elegans*, which have high promoter fragility and highly occupied proximal nucleosomes (**Figure 6D, bottom**). To our knowledge, OPN and DPN-type promoters have not been described or defined in *C. elegans*. Our results are consistent with a model in which nucleosome instability is encoded at the promoters of DPN-type genes, potentiating the high transcriptional plasticity observed at these sites. The presence of these promoter structures in yeast, human, and now *C. elegans* suggests a well-conserved strategy that uses nucleosome architecture to regulate the dynamics of gene expression.

## **MATERIALS AND METHODS**

### **Worm strains and growth in liquid culture**

Wild type *N2* worms were obtained from the *Caenorhabditis* Genome Center and maintained at 20° C in liquid culture as previously described (Whittle et al. 2008; Ercan et al. 2011). Mixed-stage embryos were isolated from gravid adults by bleach hypochlorite treatment and fixed with 2% formaldehyde for 30 minutes at room temperature.

### **MNase Digestion Time Course**

MNase digestion was performed as previously described (Ercan et al. 2011), with slight alterations. Micrococcal nuclease (Worthington LS004798) was resuspended in water at 50 U/uL and frozen in individual aliquots at -80°C. To control for variability in enzyme activity, individual aliquots were removed from the freezer and thawed on ice, and never reused. Mixed-stage embryos were incubated with chitinase (Sigma Cat #C6137), washed, and dounced in dounce buffer (0.35 M sucrose, 15 mM HEPES-KOH pH 7.5, 0.5 mM EGTA, 5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.5% TritonX-100, 0.25% NP-40) to extract nuclei. Nuclei

were pelleted and washed with MNase digestion buffer (110 mM NaCl, 40 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 50 mM HEPES-KOH pH 7.5). MNase was added, and at each timepoint (0, 2, 4, 8, 15, or 30 minutes after enzyme addition) a fraction of the reaction was removed, quenched with EDTA, and stored on ice. Samples were treated with Proteinase K for 2 hours at 55°C, then incubated overnight at 65°C to reverse crosslinks. DNA was isolated from RNA and proteins using phenol:chloroform extraction and RNase A treatment for 1 hour at 37°C. Mononucleosome-sized fragments (100 to 200 bp) were extracted from a 2% agarose gel and purified using a Qiagen gel extraction kit.

## **Heat shock**

Mixed-stage embryos were isolated as described and split into two pools. One pool was incubated at 34°C for 20 minutes with intermittent brief mixing, while the other pool nutated at room temperature. After 20 minutes, an aliquot from each pool was saved for RNA-seq, while the remaining embryos were fixed for 30 minutes in 2% formaldehyde at room temperature.

## **RNA isolation**

Embryos were dropped into TRIzol (Life Technologies) and flash frozen in liquid nitrogen after incubation for 20 minutes at room temperature or 34°C heat shock. Embryos were homogenized by thawing at 37°C and refreezing in liquid nitrogen 3x. Total RNA was isolated using a TRIzol/chloroform extraction followed by RNeasy Mini (Qiagen) preparation with On Column DNaseI Digestion (Qiagen).

## **Illumina Library Preparation**

Individual libraries were prepared with unique barcodes for each timepoint from the timecourse. MNaseTC libraries were prepared from 100 ng of gel-extracted DNA using the Illumina TruSeq DNA library preparation kit v2 (FC-121-2001) according to manufacturer instructions. Ampure

beads were used for the final purification in lieu of gel purification. RNA-seq libraries were prepared from 2 ug of total RNA using the Illumina TruSeq RNA library preparation kit v2 (RS-122-2001) according to manufacturer instructions. Individual samples were barcoded with unique 6bp index sequences contained within the sequencing adapters. Individual libraries were then pooled at equimolar ratios for paired-end multiplex sequencing.

## **Illumina Sequencing and Post-Processing**

Paired-end sequencing was performed by the Princeton University Sequencing Core Facility according to Illumina protocols. Paired end reads were mapped to the UCSC Oct. 2010 (WS220/ce10) genome release using Bowtie (v1.1.2) with stringent multimapping parameters:

```
bowtie -q -X 2000 --fr -p 1 -S -n 2 -e 70 -l 28 --pairtries 100 --maxbts 125 -k 1 -m 1 --un /Unmapped_Reads.fastq --phred33-quals /ce10 -1 /read1.fastq -2 /read2.fastq
```

## **Nucleosome analysis**

Reads with insert sizes between 100 and 250 bp were kept for downstream analysis.

Replicates were first processed individually, then pooled after confirming a high degree of correlation between replicates. Nucleosome analysis was performed as described previously (Kaplan et al. 2010; Gossett and Lieb 2012).

*Coverage:* Nucleosome coverage was calculated by extending the filtered mapped reads to their fragment length and measuring the sum of reads covering each bp. To normalize for variation between samples, nucleosome coverage was scaled by 1/(mean coverage), yielding a mean nucleosome coverage of 1.0.

*Dyads:* Dyads are approximated as the center of a paired-end fragment. The number of dyads at each base pair was scaled by 1/(mean dyad density), then Gaussian smoothed with a standard deviation of 20 bp.

*Nucleosome calls:* Nucleosome positions were identified from dyad density maps using a previously reported greedy algorithm (Albert et al. 2007; Gossett and Lieb 2012). Using the local maxima of the dyad density as the nucleosome center  $p$ , the size of the nucleosome (the nucleosome-protected region) was determined by measuring the average length of all reads that covered the nucleosome center. The standard deviation of the nucleosome center (the nucleosome “fuzziness”) was calculated for each called nucleosome as the standard deviation of dyads around the mean. Nucleosome occupancy was defined as the number of dyads that fell within 50 bp of the nucleosome center.

*Boundary nucleosomes:* Using these called nucleosome positions 5′ and 3′ boundary nucleosomes were identified for the 20,578 RefSeq annotated genes. 5′ +1 nucleosomes were identified as the first nucleosome call with a dyad coordinate downstream of the 1st coding exon. Similarly, the 3′ boundary nucleosome was identified as the first nucleosome call with a dyad coordinate upstream of the TTS. Because *C. elegans* utilizes trans-splicing, the 5′ end of mature polyadenylated mRNAs, and the RefSeq annotations used in this analysis, do not reflect the exact base pair position of transcription initiation. Although recent studies have used novel methods to identify the true transcription initiation sites (Chen et al. 2013; Kruesi et al. 2013; Saito et al. 2013), the TSS annotations are only known for a subset of expressed genes in a small number of stages (Chen et al. only tested embryos, only 31.7% of genes had a TSS in at least 1 of 3 stages tested in Kruesi et al., Saito et al., only tested embryo and adult). For completeness, we chose to instead investigate the full set of known genes using their first coding exon as an alignment point.

## **Nucleosome Fragility and Resistance scores**

The pooled “intermediate” nucleosome profile was generated by pooling the reads from the 4, 8, and 15 minute time points from each replicate. The pooled reads were used to generate

average nucleosome positions, fuzziness scores, and occupancies. To identify regions of the genome that were liberated earlier or later than average, we subtracted the occupancy of the pooled sample from either the 2 minute (2m – pool = Fragility score) or the 30 minute samples (30m – pool = Resistance score). To highlight regions significantly enriched with this signal, we considered the 10% of nucleosomes with the highest fragility or resistance scores as Fragile or Resistant Nucleosomes.

### Gene ontology analysis

Gene lists were uploaded to the FatiGO web server ([babelomics.bioinfo.cipf.es](http://babelomics.bioinfo.cipf.es)) and compared against the background set of all *C. elegans* genes (Al-Shahrour et al. 2004). P-values were calculated using the Fisher's exact test, and corrected for multiple testing using the FDR procedure of Benjamini and Hochberg (Benjamini and Hochberg 1995). Corrected p-values and GO terms were then input in to REVIGO to reduce and visualize significantly enriched GO clusters (Supek et al. 2011).

### Stable and developmentally regulated genes

Pre-normalized transcriptome sequencing data was downloaded from: <https://www.encodeproject.org/comparative/transcriptome/> (Gerstein et al. 2014; Spencer et al. 2011). For each gene, we calculated the coefficient of variation (CV):  $c_v = \frac{\sigma}{\mu}$ . We took the 1000 genes with the highest CVs as the set of developmentally regulated genes, and the set of 1000 genes with the lowest CVs as the set of stably expressed genes.

### RNA-seq analysis

Unstranded mRNA libraries were prepared from total RNA for RNA-seq using the Illumina TruSeq RNA Library Preparation Kit v2 (RS-122-2001). RNA-seq reads were mapped to the *C.*

*C. elegans* WS220 Gene Annotation Model using Tophat2 (v0.7) (Trapnell et al. 2012). The resulting alignment files were quantified using HT-Seq (v0.4.1) and the RefSeq gene annotations for WS220 (Anders et al. 2015). Total read counts per gene were normalized for differential expression using DESeq2 (v1.0.19) in R (v3.0.1) (Love et al. 2014).

## **Additional datasets**

A brief description of the additional publicly available datasets used in this study and their accession numbers can be found in **Supplemental Table S1**.

## **DATA ACCESS**

The sequencing data from this study have been submitted to the NCBI Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) under accession number SRP072274. Additionally, processed and raw data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE79567.

## **ACKNOWLEDGEMENTS**

We thank Coleen Murphy and members of the Lieb laboratory for feedback and comments on the manuscript. This project was supported by NIH NHGRI grant U01 HG004270. TEJ was supported by NHGRI predoctoral training grant T32HG003284-10S1. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

## **AUTHOR CONTRIBUTIONS**

Experiments were designed by TEJ and JDL. Experiments and bioinformatics analyses were performed by TEJ with guidance and feedback from JDL. Manuscript was prepared by TEJ, and edited by JDL.



## DISCLOSURE DECLARATION

The authors declare no conflicts of interest.

## FIGURE LEGENDS

### Figure 1. An MNase digestion time course on *C. elegans* embryos

(A) Mixed-stage embryos were collected from gravid hermaphrodites by bleach treatment. Dissociated nuclei from mixed-stage embryos were incubated with MNase for 2, 4, 8, 15, or 30 minutes. (B) Paired-end reads from each timepoint were mapped to the *C. elegans* genome, normalized, and Gaussian smoothed for display. High signals represent regions of the genome protected from MNase digestion. Region plotted: chr IV position 12,074,951 to 12,084,347. (C) Representative image of an N2 embryo MNase digestion timecourse after gel electrophoresis. For each timepoint, mononucleosome-sized fragments were excised from the gel (white box) and used for paired-end Illumina DNA sequencing. Size markers (M) are indicated. (D) Mononucleosome fragments are shorter with increasing MNase digestion time, in 10 bp increments. (E) Calculation of fragility and resistance scores. Fragility: for each nucleosome, the average occupancy of the intermediate timepoints is subtracted from the 2 minute timepoint. Resistance: for each nucleosome, the average occupancy of the intermediate timepoints is subtracted from the 30 minute timepoint. Intermediate timepoints are 4, 8, and 15 minutes. Region plotted: chr IV position 12,076,980 to 12,078,364. (F) Distribution of fragility and resistance scores at all nucleosomes. The top 10% of each class (shaded in green and orange, respectively) were considered “Fragile” or “Resistant”.

### Figure 2. Competition with transcription factors influences nucleosome fragility

(A) Average reconstituted nucleosome occupancy scores (Locke et al. 2013) and computational nucleosome occupancy model scores (Kaplan et al. 2009) at 35,062 regions bound at any stage

by any number of transcription factors. **(B)** Average fragility, resistance, and intermediate nucleosome occupancy scores are plotted around the same set of intervals from **(A)**. **(C)** Nucleosome fragility scores are higher at sites bound by more transcription factors. Boxplot of average fragility or resistance scores at groups of sites bound by different numbers of transcription factors. N = number of regions in each category. **(D)** Cartoon characterization of how embryo-specific and L4-specific HOT regions were identified. Embryo HOT region: binding sites bound only in the embryo. L4 HOT region: binding sites bound only in the L4 stage. **(E)** Model to distinguish whether *trans* (**top**) or *cis* (**bottom**) effects result in nucleosome fragility at a given nucleosome in the embryo. Hypothetical fragility scores are represented. **(F)** Fragility, resistance, and nucleosome occupancy scores measured in the embryo at 119 regions found to be specifically HOT in the embryo. **(G)** Fragility, resistance, and intermediate nucleosome occupancy scores at 88 regions found to be specifically HOT in the L4 stage worm.

### **Figure 3. Heat shock increases nucleosome fragility at the promoter and gene body of upregulated genes**

**(A)** Experimental overview. Mixed-stage embryos were either incubated at room temperature (RT) or heat shocked at 34°C (HS) for 20 minutes. Subsequently, embryos were fixed and used for an MNase-seq timecourse or stored in TRIzol and used for RNA-seq. **(B)** mRNA-seq of differentially expressed genes after a 20 minute HS at 34°C. Significantly differentially expressed genes ( $p_{adj} < 0.1$ ) shown in red. **(C)** Increased fragility scores after heat shock in the coding region of F33E5.4 and F33E5.5, two divergently transcribed hsp-70 orthologues. Region plotted: chr II position 11,749,925 to 11,770,394. **(D)** Nucleosome fragility and nucleosome occupancy at 11 significantly differentially expressed genes with and without heat shock. **(E)** Nucleosome fragility and nucleosome resistance at all 20,768 coding genes with and without heat shock.

**Figure 4. Fragility is enriched 5' and 3' of genes, and on average is anti-correlated with gene expression**

(A) Log<sub>2</sub> DESeq-normalized number of read counts measured by mRNA-seq at 20,785 genes, ordered by their relative expression. (B) Heatmap of fragility scores (green) at. Genes were aligned by the center of the first nucleosome downstream from the transcript start site, known as the +1 or 5' Boundary Nucleosome (yellow line). (C) Same as in (B), except resistance scores are plotted in red. (D) Same as (B), except for HTZ-1 input-normalized ChIP-seq signals (Ho et al Nature 2014). (E) Same as (B), except the average GC content (as a percentage of 100%) in 5 bp windows is plotted. (F) Fragility and (G) resistance scores around the 5' and 3' boundary nucleosomes averaged over expression quintiles (highest expressed 20% in dark orange or green, lowest expressed 20% in lightest orange or green). Quintile 1: 0 to 4.5 normalized counts. Quintile 2: 4.5 to 65. Quintile 3: 65 to 619. Quintile 4: 619 to 2209. Quintile 5: > 2209.

**Figure 5. Fragile nucleosomes are found at AT-rich, nucleosome disfavoring, and poorly conserved promoters**

(A) Histogram of average GC content at fragile (green) or all nucleosomes (grey). (B) Same as A for resistant (orange) or all nucleosomes (grey). (C) Histogram of *in vitro* nucleosome occupancy scores at fragile (green) or all nucleosomes (grey). (D) Histogram of *in vitro* nucleosome occupancy scores at resistant (orange) or all nucleosomes (grey). (E) Histogram of PhastCons 7-way conservation score at fragile nucleosomes (green) or all nucleosomes (grey). (F) Same as E for resistant nucleosomes (orange) or all nucleosomes (grey).

**Figure 6. Fragile nucleosomes are enriched at genes that will be expressed in the future and in specific contexts**

(A) Top 10 Gene Ontology biological process functional annotation terms associated with genes with fragile nucleosomes. (B) Top 10 gene ontology biological process functional

annotation terms associated with at genes with resistant nucleosomes. **(C)** Bar plot representation of expression levels and coefficient of variation (CV) for R07H5.2 and F26F2.10. R07H5.2 has a low CV, and is an example of a stably-expressed gene; F26F2.10 has a high CV, and is an example of a developmentally regulated gene. **(D)** Average plot of fragility and nucleosome occupancy scores at 1000 stably-expressed genes (top), or developmentally regulated genes (bottom) as determined by their coefficient of variation across 7 different life stages: early embryo, late embryo, larval stages L1, L2, L3, L4, and young adult.

## Figure 7

We propose a model whereby nucleosome fragility is determined by two distinct mechanisms, one that operates in *cis* at all genes, and one that operates in *trans* at a subset of genes. **Left:** Competition in *trans* with transcription factors and polymerase machinery destabilizes nucleosomes at the promoters of actively transcribed genes that tend to be stably expressed.. **Right:** Condition-specific and developmentally regulated genes contain promoters with high levels of nucleosome fragility, determined primarily in *cis* by high AT content. Green line: high AT content is sequence-encoded at all promoters, but is highest at condition-specific genes. Orange cylinders: resistant nucleosomes found in the gene body of highly and stably expressed genes. Green cylinders: fragile nucleosomes compete (single arrow) with transcription factors and RNA Pol II at stably expressed genes. Fragile nucleosomes at condition-specific genes “treadmill” on the DNA (three arrows) due to destabilizing DNA elements like TATA-box motifs and high AT content.

# REFERENCES

- Adams CC, Workman JL. 1995. Binding of Disparate Transcriptional Activators to Nucleosomal DNA Is Inherently Cooperative. *Molecular and Cellular Biology* **15**: 1405–1421.
- Al-Shahrour F, Diaz-Uriarte R, Dopazo J. 2004. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* **20**: 578–580.
- Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF. 2007. Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**: 572–576.
- Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.
- Anderson JD, Thåström A, Widom J. 2002. Spontaneous Access of Proteins to Buried Nucleosomal DNA Target Sites Occurs via a Mechanism That Is Distinct from Nucleosome Translocation. *Molecular and Cellular Biology* **22**: 7147–7157.
- Araya CL, Kawli T, Kundaje A, Jiang L, Wu B, Vafeados D, Terrell R, Weissdepp P, Gevirtzman L, Mace D, et al. 2014. Regulatory analysis of the *C. elegans* genome with spatiotemporal resolution. *Nature* **512**: 400–405.
- Åkerfelt M, Morimoto RI, Sistonen L. 2010. Heat shock factors: integrators of cell stress, development and lifespan. *Nat Rev Mol Cell Biol* **11**: 545–555.
- Bai L, Morozov AV. 2010. Gene regulation by nucleosome positioning. *TRENDS in Genetics* **26**: 476–483.
- Barozzi I, Simonatto M, Bonifacio S, Yang L, Rohs R, Ghisletti S, Natoli G. 2014. Coregulation of Transcription Factor Binding and Nucleosome Occupancy through DNA Features of Mammalian Enhancers. *Molecular Cell* 1–14.
- Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* **57**: 289–300.
- Bintu L, Kopaczynska M, Hodges C, Lubkowska L, Kashlev M, Bustamante C. 2011. The elongation rate of RNA polymerase determines the fate of transcribed nucleosomes. *Nature Structural & Molecular Biology* **18**: 1394–1399.
- Bloom KS, Anderson JN. 1978. Fractionation of Hen Oviduct Chromatin into Transcriptionally Active and Inactive Regions after Selective Micrococcal Nuclease Digestion. *Cell* **15**: 141–150.
- Boyle AP, Araya CL, Brdlik C, Cayting P, Cheng C, Cheng Y, Gardner K, Hillier LW, Janette J, Jiang L, et al. 2014. Comparative analysis of regulatory information and circuits across distant species. *Nature* **512**: 453–456.
- Burton DR, Butler MJ, Hyde JE, Phillips D, Skidmore CJ, Walker IO. 1978. The interaction of core histones with DNA: equilibrium binding studies. *Nucleic Acids Research* **5**: 3643–3664.

- 1 Chen RAJ, Down TA, Stempor P, Chen QB, Egelhofer TA, Hillier LW, Jeffers TE, Ahringer J.  
2 2013. The landscape of RNA polymerase II transcription initiation in *C. elegans* reveals  
3 promoter and enhancer architectures. *Genome Research*.
- 4 Chen RAJ, Stempor P, Down TA, Zeiser E, Feuer SK, Ahringer J. 2014. Extreme HOT regions  
5 are CpG-dense promoters in *C. elegans* and humans. *Genome Research* **24**: 1138–1146.
- 6 Chen W, Luo L, Zhang L. 2010. The organization of nucleosomes around splice sites. *Nucleic  
7 Acids Research* **38**: 2788–2798.
- 8 Chereji RV, Kan T-W, Grudniewska MK, Romashchenko AV, Berezhikov E, Zhimulev IF, Guryev  
9 V, Morozov AV, Moshkin YM. 2015. Genome-wide profiling of nucleosome sensitivity and  
10 chromatin accessibility in *Drosophila melanogaster*. *Nucleic Acids Research* **44**: 1036–  
11 1051.
- 12 Cole HA, Ocampo J, Iben JR, Chereji RV, Clark DJ. 2014. Heavy transcription of yeast genes  
13 correlates with differential loss of histone H2B relative to H4 and queued RNA polymerases.  
14 *Nucleic Acids Research* **42**: 12512–12522.
- 15 Deng T, Zhu ZI, Zhang S, Postnikov Y, Huang D, Horsch M, Furusawa T, Beckers J, Rozman J,  
16 Klingenspor M, et al. 2015. Functional compensation among HMGN variants modulates the  
17 DNase I hypersensitive sites at enhancers. *Genome Research* **25**: 1295–1308.
- 18 Deniz Ö, Flores O, Battistini F, Pérez A, Soler-López M, Orozco M. 2011. Physical properties of  
19 naked DNA influence nucleosome positioning and correlate with transcription start and  
20 termination sites in yeast. *BMC Genomics* **12**: 489.
- 21 Eaton ML, Galani K, Kang S, Bell SP, MacAlpine DM. 2010. Conserved nucleosome positioning  
22 defines replication origins. *Genes & Development* **24**: 748–753.
- 23 Ercan S, Lubling Y, Segal E, Lieb JD. 2011. High nucleosome occupancy is encoded at X-linked  
24 gene promoters in *C. elegans*. *Genome Research* **21**: 237–244.
- 25 Formosa T. 2012. Biochimica et Biophysica Acta. *BBA - Gene Regulatory Mechanisms* **1819**:  
26 247–255.
- 27 Gerstein MB, Rozowsky J, Yan K-K, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C,  
28 Li JJ, et al. 2014. Comparative analysis of the transcriptome across distant species. *Nature*  
29 **512**: 445–448.
- 30 Gossett AJ, Lieb JD. 2012. In Vivo Effects of Histone H3 Depletion on Nucleosome Occupancy  
31 and Position in *Saccharomyces cerevisiae* ed. O.J. Rando. **8**: e1002771.  
32 <http://dx.plos.org/10.1371/journal.pgen.1002771.s016>.
- 33 Grishkevich V, Hashimshony T, Yanai I. 2011. Core promoter T-blocks correlate with gene  
34 expression levels in *C. elegans*. *Genome Research* **21**: 707–717.
- 35 Henikoff JG, Belsky JA, Krassovsky K, MacAlpine DM, Henikoff S. 2011. Epigenome  
36 characterization at single base-pair resolution. *Proceedings of the National Academy of  
37 Sciences* **108**: 18318–18323.

- 1 Henikoff S. 2008. Nucleosome destabilization in the epigenetic regulation of gene expression.  
2 *Nat Rev Genet* **9**: 15–26.
- 3 Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. 2009. Genome-wide profiling of salt  
4 fractions maps physical properties of chromatin. *Genome Research* **19**: 460–469.
- 5 Ho JWK, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn K-A, Minoda A, Tolstorukov MY,  
6 Appert A, et al. 2014. Comparative analysis of metazoan chromatin organization. *Nature*  
7 **512**: 449–452.
- 8 Ioshikhes I, Hosid S, Pugh BF. 2011. Variety of genomic DNA patterns for nucleosome  
9 positioning. *Genome Research* **21**: 1863–1871.
- 10 Ishibashi T, Dryhurst D, Rose KL, Shabanowitz J, Hunt DF, Ausió J. 2009. Acetylation of  
11 Vertebrate H2A.Z and Its Effect on the Structure of the Nucleosome. *Biochemistry* **48**:  
12 5007–5017.
- 13 Jiang C, Pugh BF. 2009. Nucleosome positioning and gene regulation: advances through  
14 genomics. *Nat Rev Genet* **10**: 161–172.
- 15 Jin C, Felsenfeld G. 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z.  
16 *Genes & Development* **21**: 1519–1529.
- 17 Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-  
18 containing nucleosomes mark “nucleosome-free regions” of active promoters and other  
19 regulatory regions. *Nature Publishing Group* **41**: 941–945.
- 20 Kaplan N, Hughes TR, Lieb JD, Widom J, Segal E. 2010. Contribution of histone sequence  
21 preferences to nucleosome organization: proposed definitions and methodology. *Genome*  
22 *biology* **11**: 1–12.
- 23 Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes  
24 TR, Lieb JD, Widom J, et al. 2009. The DNA-encoded nucleosome organization of a  
25 eukaryotic genome. *Nature* **458**: 362–366.
- 26 Kent NA, Adams S, Moorhouse A, Paszkiewicz K. 2011. Chromatin particle spectrum analysis:  
27 a method for comparative chromatin structure analysis using paired-end mode next-  
28 generation DNA sequencing. *Nucleic Acids Research* **39**: e26–e26.
- 29 Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. 2002.  
30 Nucleosome Remodeling Induced by RNA Polymerase II: Loss of the H2A/H2B Dimer  
31 during Transcription. *Molecular Cell* **9**: 541–552.
- 32 Kristjuhan A, Svejstrup JQ. 2004. untitled. **23**: 4243–4252.
- 33 Kruesi WS, Core LJ, Waters CT, Lis JT, Meyer BJ. 2013. Condensin controls recruitment of  
34 RNA polymerase II to achieve nematode X-chromosome dosage compensation. *eLife* **2**:  
35 e00808–e00808.
- 36 Kubik S, Bruzzone MJ, Jacquet P, Falcone J-L, Rougemont J, Shore D. 2015. Nucleosome  
37 Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast.



1        *Molecular Cell* **60**: 422–434.

2        Kulaeva OI, Hsieh F-K, Studitsky VM. 2010. RNA polymerase complexes cooperate to relieve  
3        the nucleosomal barrier and evict histones. *Proceedings of the National Academy of*  
4        *Sciences* **107**: 11325–11330.

5        Kwak H, Fuda NJ, Core LJ, Lis JT. 2013. Precise Maps of RNA Polymerase Reveal How  
6        Promoters Direct Initiation and Pausing. *Science* **339**: 950–953.

7        Li B, Pattenden SG, Lee D, Gutierrez J, Chen J, Seidel C, Gerton J, Workman JL. 2005.  
8        Preferential occupancy of histone variant H2AZ at inactive promoters influences local  
9        histone modifications and chromatin remodeling. *Proceedings of the National Academy of*  
10        *Sciences* **102**: 18385–18390.

11        Li W, Nagaraja S, Delcuve GP, Hendzel MJ, Davie JR. 1993. Effects of histone acetylation,  
12        ubiquitination and variants on nucleosome stability. *Biochem J* **296**: 737–744.

13        Lickwar CR, Rao B, Shabalin AA, Nobel AB, Strahl BD, Lieb JD. 2009. The Set2/Rpd3S  
14        Pathway Suppresses Cryptic Transcription without Regard to Gene Length or Transcription  
15        Frequency ed. L. Rusche. *PLoS ONE* **4**: e4886.

16        Lieb JD, Clarke ND. 2005. Control of Transcription through Intragenic Patterns of Nucleosome  
17        Composition. *Cell* **123**: 1187–1190.

18        Liu T, Rechtsteiner A, Egelhofer TA, Vielle A, Latorre I, Cheung MS, Ercan S, Ikegami K,  
19        Jensen M, Kolasinska-Zwierz P, et al. 2011. Broad chromosomal domains of histone  
20        modification patterns in *C. elegans*. *Genome Research* **21**: 227–236.

21        Locke G, Haberman D, Johnson SM, Morozov AV. 2013. Global remodeling of nucleosome  
22        positions in *C. elegans*. *BMC Genomics* **14**: 1–1.

23        Lombraña R, Almeida R, Revuelta I, Madeira S, Herranz G, Saiz N, Bastolla U, Gómez M.  
24        2013. High-resolution analysis of DNA synthesis start sites and nucleosome architecture at  
25        efficient mammalian replication origins. *The EMBO Journal* **32**: 2631–2644.

26        Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for  
27        RNA-seq data with DESeq2. *Genome biology* **15**: 550.

28        Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the  
29        nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260.

30        Mavrich TN, Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh  
31        BF. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout  
32        the yeast genome. *Genome Research* **18**: 1073–1083.

33        McGhee JD, Felsenfeld G. 1983. Another Potential Artifact in the Study of Nucleosome Phasing  
34        by Chromatin Digestion with Micrococcal Nuclease. *Cell* **32**: 1205–1215.

35        Merz K, Hondele M, Goetze H, Gmelch K, Stoeckl U, Griesenbeck J. 2008. Actively transcribed  
36        rRNA genes in *S. cerevisiae* are organized in a specialized chromatin associated with the  
37        high-mobility group protein Hmo1 and are largely devoid of histone molecules. *Genes &*

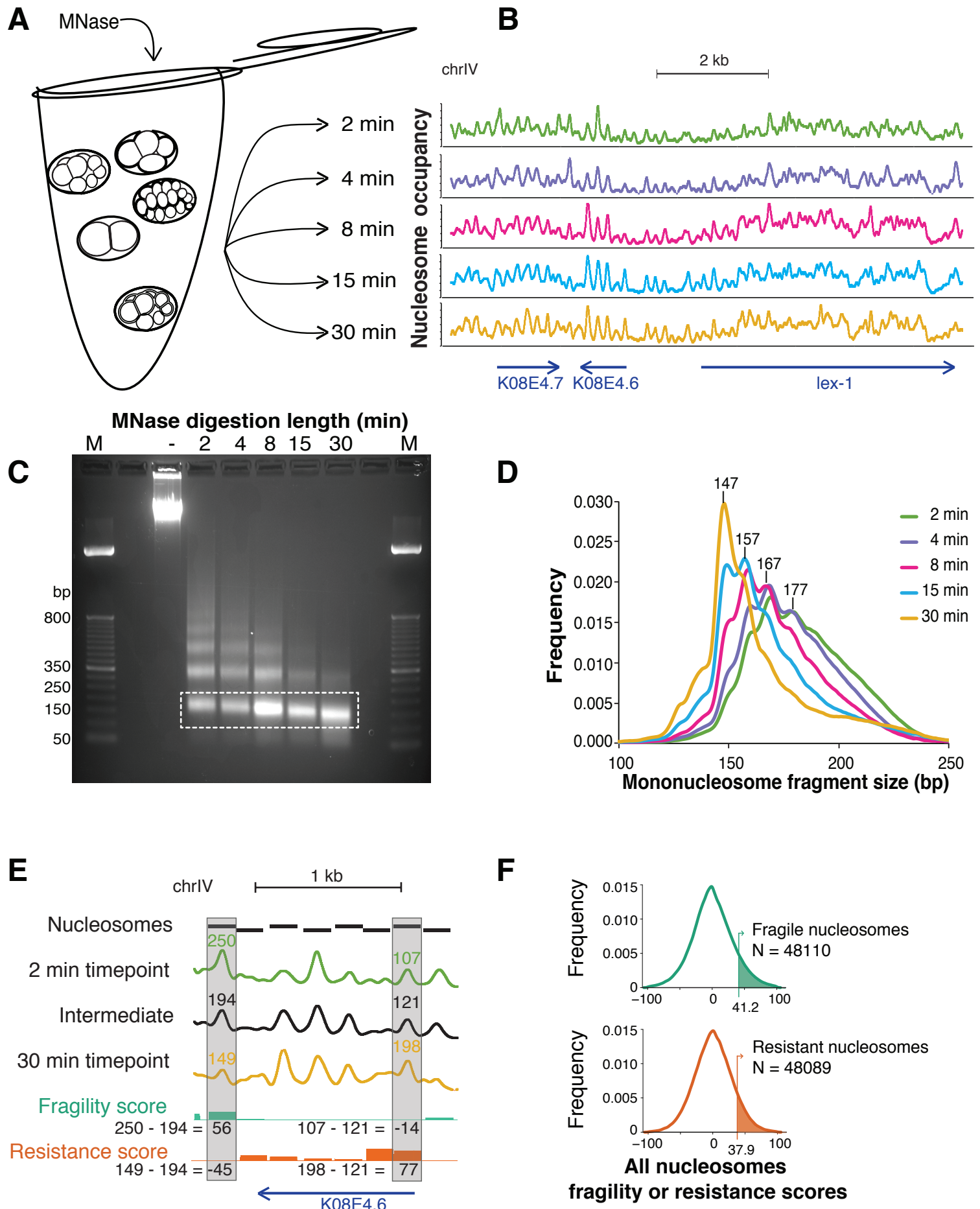


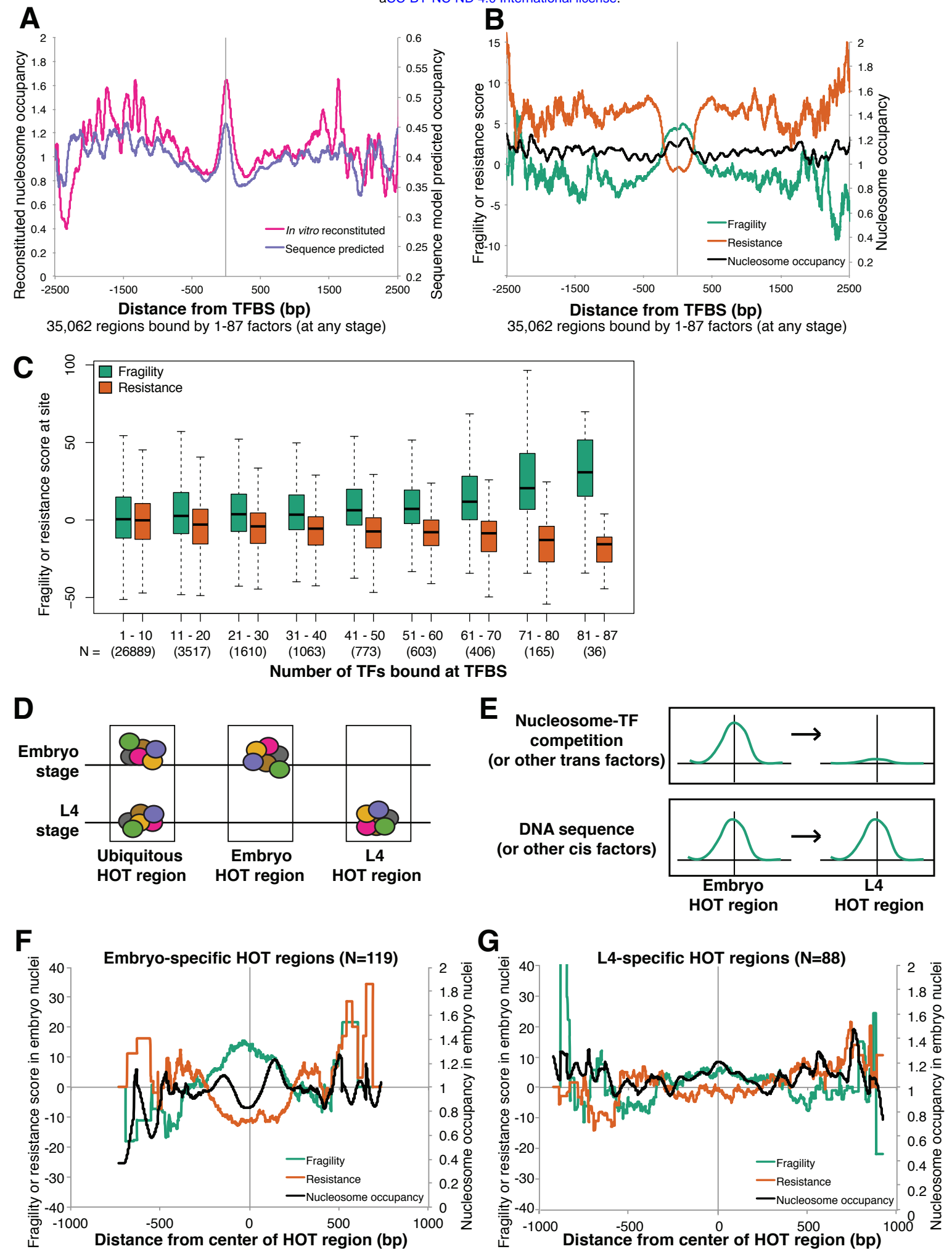
- 1        *Development* **22**: 1190–1204.
- 2        Ooi SL, Henikoff JG, Henikoff S. 2010. A native chromatin purification system for epigenomic  
3        profiling in *Caenorhabditis elegans*. *Nucleic Acids Research* **38**: e26–e26.
- 4        Ozonov EA, van Nimwegen E. 2013. Nucleosome Free Regions in Yeast Promoters Result from  
5        Competitive Binding of Transcription Factors That Interact with Chromatin Modifiers ed. I.  
6        Ioshikhes. *PLoS Comput Biol* **9**: e1003181.
- 7        Petesch SJ, Lis JT. 2008. Rapid, Transcription-Independent Loss of Nucleosomes over a Large  
8        Chromatin Domain at Hsp70 Loci. *Cell* **134**: 74–84.
- 9        Pérez-Lluch S, Blanco E, Tilgner H, Curado J, Ruiz-Romero M, Corominas M, Guigó R. 2015.  
10        Absence of canonical marks of active chromatin in developmentally regulated genes.  
11        *Nature Publishing Group* **47**: 1158–1167.
- 12        Polach KJ, Widom J. 1995. Mechanism of Protein Access to Specific DNA Sequences in  
13        Chromatin: A Dynamic Equilibrium Model for Gene Regulation. *Journal of Molecular Biology*  
14        **254**: 130–149.
- 15        Polach KJ, Widom J. 1999. Restriction Enzymes as Probes of Nucleosome Stability and  
16        Dynamics. *Methods in Enzymology* **304**: 278–298.
- 17        Ramirez-Carrozzi VR, Braas D, Bhatt DM, Cheng CS, Hong C, Doty KR, Black JC, Hoffmann A,  
18        Carey M, Smale ST. 2009. A Unifying Model for the Selective Regulation of Inducible  
19        Transcription by CpG Islands and Nucleosome Remodeling. *Cell* **138**: 114–128.
- 20        Raveh-Sadka T, Levo M, Shabi U, Shany B, Keren L, Lotan-Pompan M, Zeevi D, Sharon E,  
21        Weinberger A, Segal E. 2012. Manipulating nucleosome disfavoring sequences allows fine-  
22        tune regulation of gene expression in yeast. *Nat Genet* **44**: 743–750.
- 23        Reinberg D, Sims RJ. 2006. de FACTo Nucleosome Dynamics. *Journal of Biological Chemistry*  
24        **281**: 23297–23301.
- 25        Rohner S, Kalck V, Wang X, Ikegami K, Lieb JD, Gasser SM, Meister P. 2013. Promoter- and  
26        RNA polymerase II-dependent hsp-16 gene association with nuclear pores in  
27        *Caenorhabditis elegans*. *The Journal of Cell Biology* **200**: 589–604.
- 28        Saito TL, Hashimoto SI, Gu SG, Morton JJ, Stadler M, Blumenthal T, Fire A, Morishita S. 2013.  
29        The transcription start site landscape of *C. elegans*. *Genome Research* **23**: 1348–1361.
- 30        Schwabish MA, Struhl K. 2004. Evidence for Eviction and Rapid Deposition of Histones upon  
31        Transcriptional Elongation by RNA Polymerase II. *Molecular and Cellular Biology* **24**:  
32        10111–10117.
- 33        Simon RH, Felsenfeld G. 1979. A new procedure for purifying histone pairs H2A + H2B and H3  
34        + H4 from chromatin using hydroxylapatite. *Nucleic Acids Research* **6**: 689–696.
- 35        Simpson RT. 1990. Nucleosome positioning can affect the function of a cis-acting DNA element  
36        in vivo. *Nature* **343**: 387–389.

- 1 Spencer WC, Zeller G, Watson JD, Henz SR, Watkins KL, McWhirter RD, Petersen S,  
2 Sreedharan VT, Widmer C, Jo J, et al. 2011. A spatial and temporal map of *C. elegans*  
3 gene expression. *Genome Research* **21**: 325–341.
- 4 Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO Summarizes and Visualizes Long Lists  
5 of Gene Ontology Terms ed. C. Gibas. *PLoS ONE* **6**: e21800.
- 6 Tilgner H, Nikolaou C, Althammer S, Sammeth M, Beato M, Valcárcel J, Guigó R. 2009.  
7 Nucleosome positioning as a determinant of exon recognition. *Nature Structural &  
8 Molecular Biology* **16**: 996–1001.
- 9 Tirosh I, Barkai N. 2008. Two strategies for gene regulation by promoter nucleosomes. *Genome  
10 Research* **18**: 1084–1091.
- 11 Tirosh I, Berman J, Barkai N. 2007. The pattern and evolution of yeast promoter bendability.  
12 *TRENDS in Genetics* **23**: 318–321.
- 13 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL,  
14 Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq  
15 experiments with TopHat and Cufflinks. *Nature Protocols* **7**: 562–578.
- 16 Trifonov EN, Sussman JL. 1980. The pitch of chromatin DNA is reflected in its nucleotide  
17 sequence. *Proceedings of the National Academy of Sciences* **77**: 3816–3820.
- 18 Urnov FD, Wolffe AP. 2001. Chromatin remodeling and transcriptional activation: the cast (in  
19 order of appearance). *Oncogene* **20**: 2991–3006.
- 20 Vera DL, Madzima TF, Labonne JD, Alam MP, Hoffman GG, Girmurugan SB, Zhang J,  
21 McGinnis KM, Dennis JH, Bass HW. 2014. Differential Nuclease Sensitivity Profiling of  
22 Chromatin Reveals Biochemical Footprints Coupled to Gene Expression and Functional  
23 DNA Elements in Maize. *The Plant Cell* **26**: 3883–3893.
- 24 Wang J, Zhuang J, Iyer S, Lin X, Whitfield TW, Greven MC, Pierce BG, Dong X, Kundaje A,  
25 Cheng Y, et al. 2012. Sequence features and chromatin structure around the genomic  
26 regions bound by 119 human transcription factors. *Genome Research* **22**: 1798–1812.
- 27 Weiner A, Hughes A, Yassour M, Rando OJ, Friedman N. 2010. High-resolution nucleosome  
28 mapping reveals transcription-dependent promoter packaging. *Genome Research* **20**: 90–  
29 100.
- 30 Whittle CM, McClinic KN, Ercan S, Zhang X, Green RD, Kelly WG, Lieb JD. 2008. The Genomic  
31 Distribution and Function of Histone Variant HTZ-1 during *C. elegans* Embryogenesis ed. A.  
32 Akhtar. *PLoS Genetics* **4**: e1000187.
- 33 Widom J. 2002. Role of DNA sequence in nucleosome stability and dynamics. *Quart Rev  
34 Biophys* **34**.
- 35 Wu C, Travers A. 2004. A “one-pot” assay for the accessibility of DNA in a nucleosome core  
36 particle. *Nucleic Acids Research* **32**: e122–e122.
- 37 Xi Y, Yao J, Chen R, Li W, He X. 2011. Nucleosome fragility reveals novel functional states of

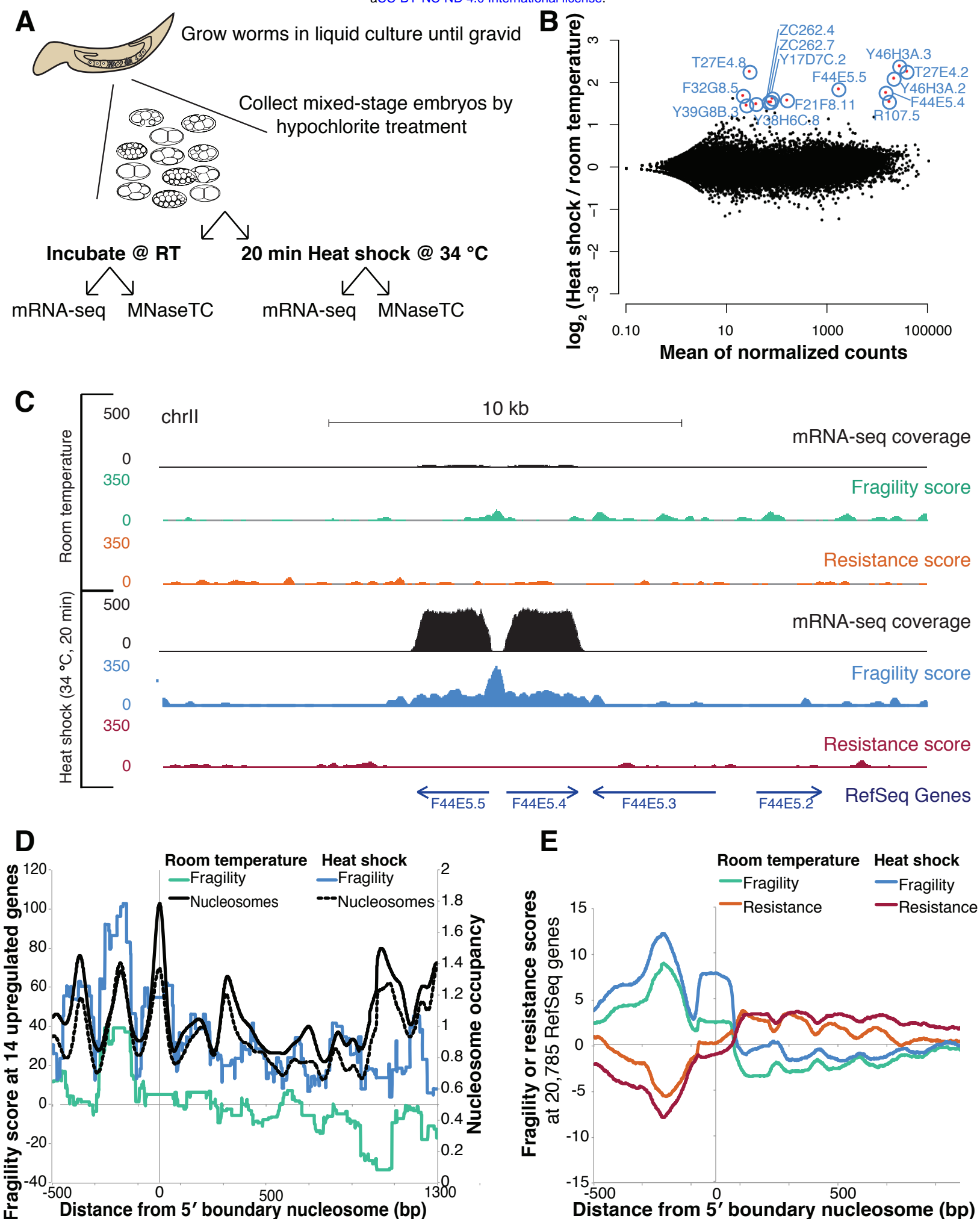
- 1 chromatin and poises genes for activation. *Genome Research* **21**: 718–724.
- 2 Yuan G-C, Liu Y-J, Dion MF, Slack MD, Wu LF, Altschuler SJ, Rando OJ. 2005. Genome-scale
- 3 Identifications of Nucleosome Positions in *S. cerevisiae*. *Science* **309**: 626–630.
- 4 Zhang H, Roberts DN, Cairns BR. 2005. Genome-Wide Dynamics of Htz1, a Histone H2A
- 5 Variant that Poises Repressed/Basal Promoters for Activation through Histone Loss. *Cell*
- 6 **123**: 219–231.
- 7 Zlatanova J, Thakar A. 2008. H2A.Z: View from the Top. *Structure* **16**: 166–179.
- 8

# Figure 1





# Figure 3





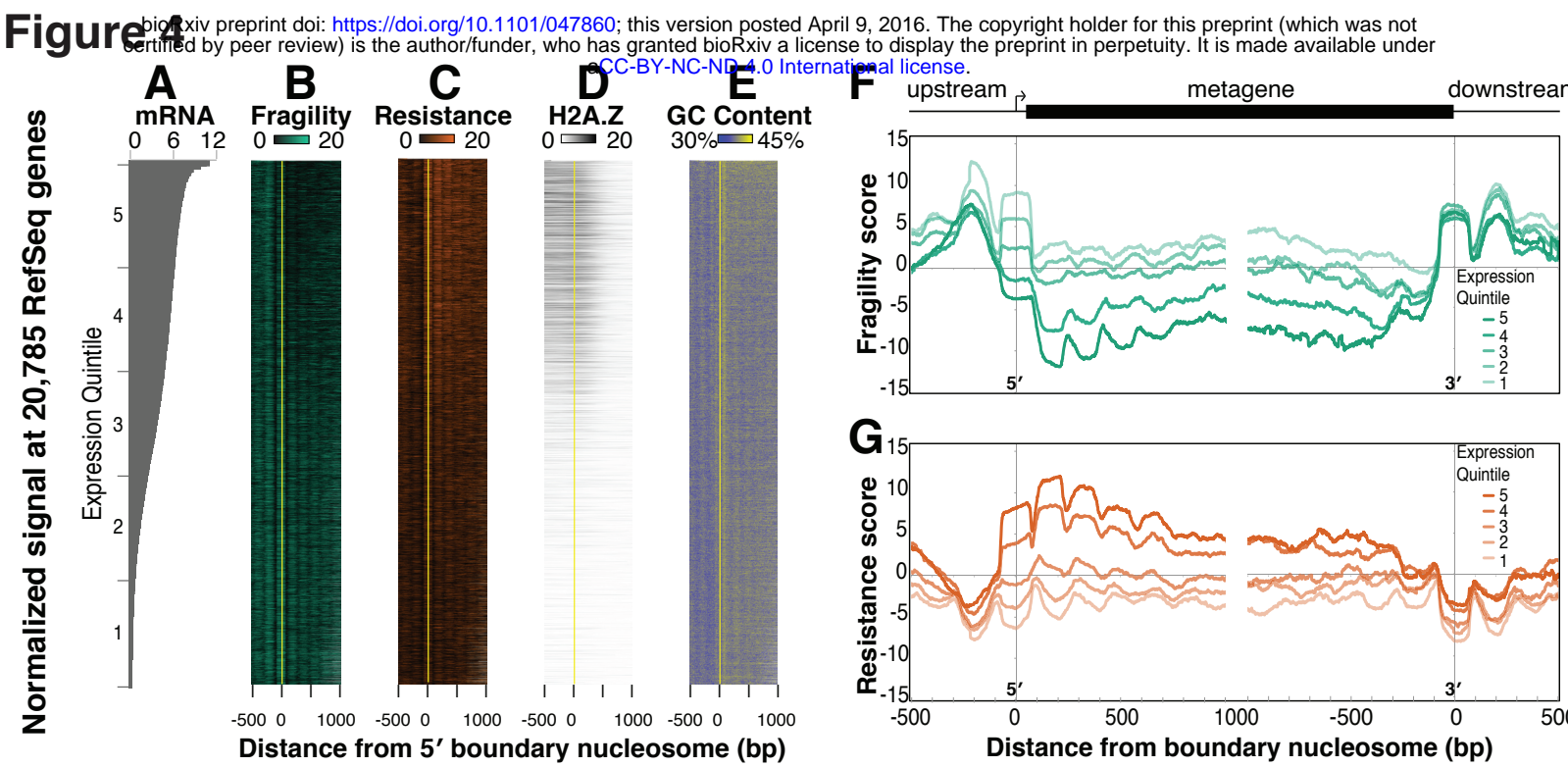
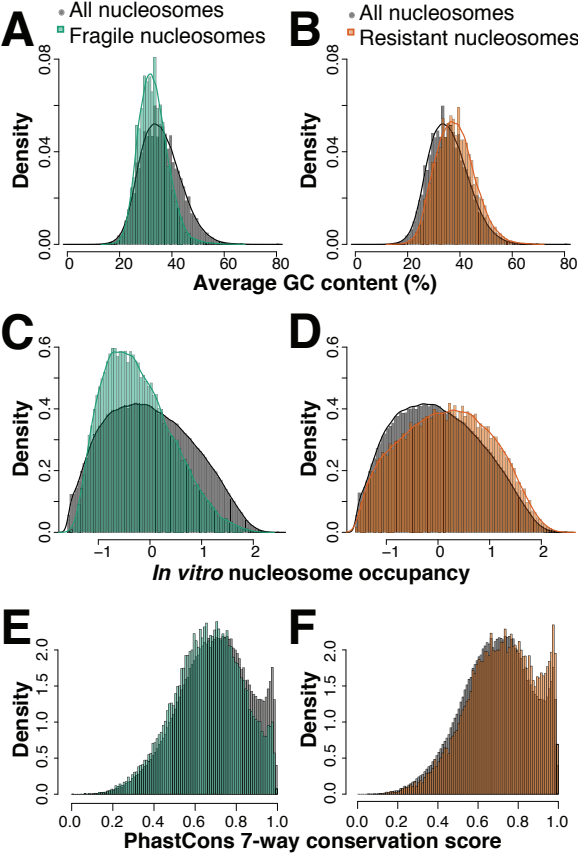
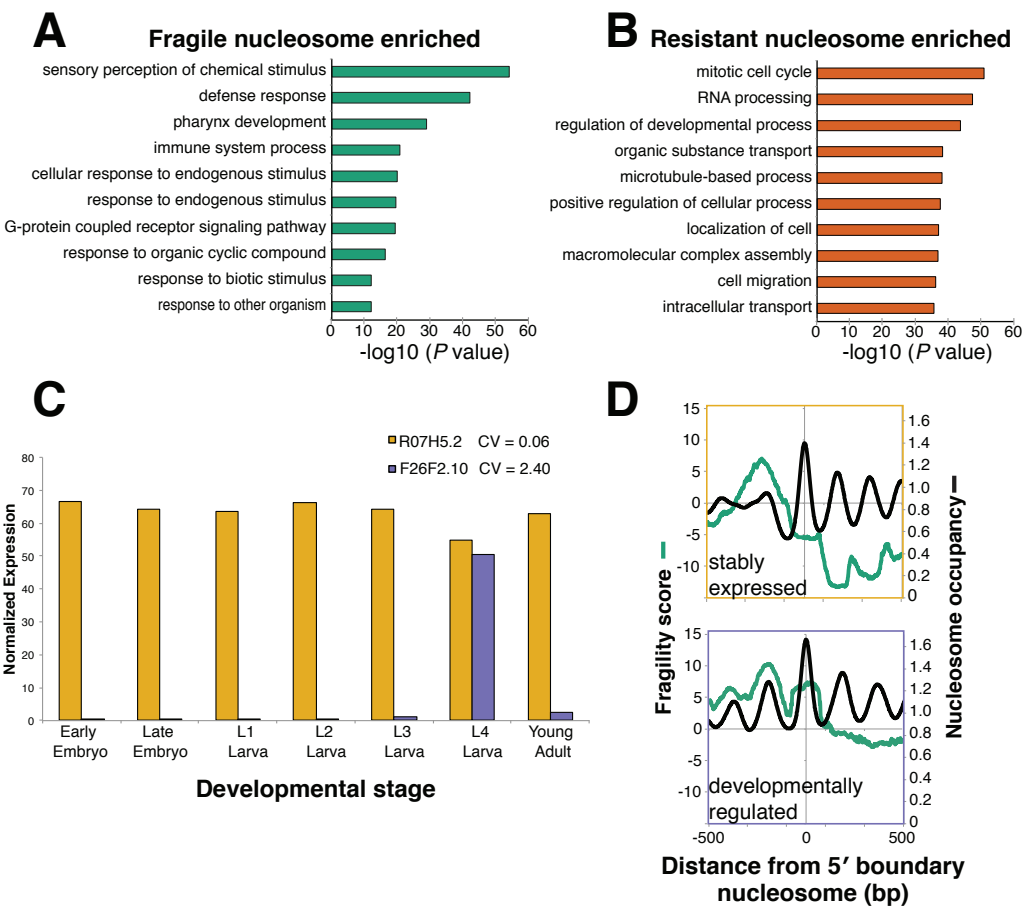


Figure 5



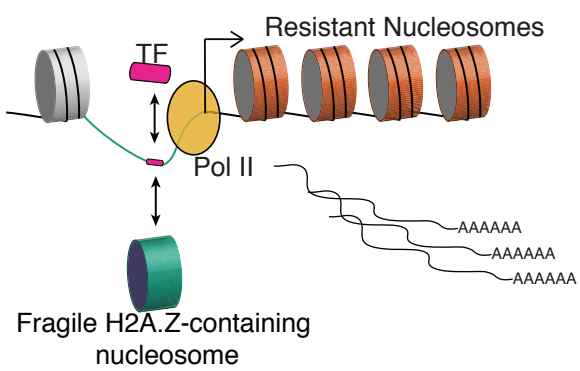


# Figure 6



**Stably expressed genes**

Promoter nucleosomes made unstable due to competition with TFs, Pol II, or other *trans*-acting factors like chromatin remodelers



**Condition-specific, developmentally regulated, stress response genes**

Promoter nucleosomes made unstable due to sequence features encoded in *cis*

