*Article*

# Nucleosome Positioning around Transcription Start Site Correlates with Gene Expression Only for Active Chromatin State in Drosophila Interphase Chromosomes

**Victor G. Levitsky [1,2,\*], Tatyana Yu. Zykova [3], Yuri M. Moshkin [1,3] and Igor F. Zhimulev [2,3,\*]**

1 Department of System Biology, Institute of Cytology and Genetics, 630090 Novosibirsk, Russia; moshkin@bionet.nsc.ru

2 Department of Natural Sciences, Novosibirsk State University, 630090 Novosibirsk, Russia

3 Department of the Structure and Function of Chromosomes, Institute of Molecular and Cellular Biology, 630090 Novosibirsk, Russia; vatolina@mcb.nsc.ru

\* Correspondence: levitsky@bionet.nsc.ru (V.G.L.); zhimulev@mcb.nsc.ru (I.F.Z.)

check for
updates

**Abstract:** We analyzed the whole-genome experimental maps of nucleosomes in *Drosophila melanogaster* and classified genes by the expression level in S2 cells (RPKM value, reads per kilobase million) as well as the number of tissues in which a gene was expressed (breadth of expression, BoE). Chromatin in 5′-regions of genes we classified on four states according to the hidden Markov model (4HMM). Only the Aquamarine chromatin state we considered as Active, while the rest three states we defined as Non-Active. Surprisingly, about 20/40% of genes with 5′-regions mapped to Active/Non-Active chromatin possessed the minimal/at least modest RPKM and BoE. We found that regardless of RPKM/BoE the genes of Active chromatin possessed the regular nucleosome arrangement in 5′-regions, while genes of Non-Active chromatin did not show respective specificity. Only for genes of Active chromatin the RPKM/BoE positively correlates with the number of nucleosome sites upstream/around TSS and negatively with that downstream TSS. We propose that for genes of Active chromatin, regardless of RPKM value and BoE the nucleosome arrangement in 5′-regions potentiates transcription, while for genes of Non-Active chromatin, the transcription machinery does not require the substantial support from nucleosome arrangement to influence gene expression.

## 1. Introduction

Nucleosomes are the basic packaging units of chromatin. Their positioning has a critical impact on the regulation of transcription. A wide variety of chromatin proteins can affect nucleosome positions [1,2]. Besides the competitive or cooperative binding of other proteins [3], ATP-dependent chromatin-remodeling complexes (remodelers) [4,5] overrule intrinsic DNA sequence preferences of histone octamers [6]. Thus, these preferences make only a modest contribution to the pattern of nucleosome positioning in vivo (≈20%) [7,8], a stronger impact comes from the pattern of histone variants and modifications and the opportunities of a number of non-histone proteins to interact with DNA [2,9].

Intrinsic DNA sequence preferences due to the thermodynamic costs associated with wrapping stiff DNA around the histone octamer influence on nucleosome positions in vivo [2,9]. A wide variety of chromatin proteins can affect nucleosome positions, most notably ATP-dependent chromatin-remodeling complexes [1,5]. Nucleosome positioning have a critical impact on regulation

of transcription. The combination of DNA context dependent and chromatin protein-influenced positioning fine-tunes the accessibility of gene regulatory regions and their competence for transcription machinery [10] affecting both initiation and elongation of transcription.

Experimental data on nucleosome positioning in genomes come from a number of related techniques [1,10,11]. For example, the digestion of chromatin with micrococcal nuclease (MNase) [11] is used to detect nucleosome positions. It is generally accepted that promoters are nucleosome-depleted; the region of diminished nucleosome occupancy observed at proximal promoter is referred to as nucleosome-depleted region (NDR) [1]. In most genes NDRs are flanked by two positioned nucleosomes (positions Nuc−1 and Nuc+1 relative TSS) [2,12–17]. These nucleosomes only for expressed genes are highly phased relative to their transcription start sites (TSSs) [17]. Analysis of various species supported this concept [8,18–21]. However, besides the expression level, many other factors could influence nucleosome arrangement around TSS. For example, in comparison to wild type, yeast strains carrying mutations in genes with known or suspected roles in nucleosome biology revealed the significant change of nucleosome arrangement [22]. We expect that full genomic separation of genes into groups [23,24] according to inherent chromatin environment influences gene activity and nucleosome arrangement.

Maps of chromatin states were a consequence of recent advances in experimental and bioinformatics approaches for genome analysis. First, we earlier developed an approach to simultaneously map the interband material cytologically and in the genome using transposon insertion tags [25]. This allowed exact localization of insertion sites both on cytological and physical maps. Previously, we showed that polytene chromosomes and interphase chromosomes from dividing cells display identical organization [24,26,27]. Namely, interbands from polytene chromosomes and the corresponding DNA sequences from cell line chromosomes share similar features in terms of localization of open chromatin-type proteins. Second, efforts of the modENCODE project [28] have produced genome-wide profiling data for many proteins that specifically localized to bands or interbands in interphase chromosomes [23,24,27,29].

These advances in cytology and bioinformatics allowed partitioning the genome into four chromatin types referred to as Aquamarine, Lazurite, Malachite, and Ruby [23,30]. We have earlier shown that genes referred to Aquamarine and Ruby types had notably different levels and patterns of expression. Namely, the median expression level of genes whose 5'-ends were located in Aquamarine chromatin type was substantially higher than that for Ruby type. Similarly, the mean number of tissues for "interband" Aquamarine genes was substantially higher than that for "band" Ruby genes. Hence, the Aquamarine chromatin was interpreted as open, i.e., it harbored regulatory regions of active genes. On the contrary, Ruby chromatin contained mostly tissue-specific genes, intergenic spacers, and introns. Two remaining types respected to gene bodies (Lazurite) and intermediate domains between other three types of chromatin (Malachite) [23,24,30]. All of this permits to carry out an analysis of localization and activity of genes as well as chromatin states and interphase polytene chromosome structures. In the current study we further classified four states [23,30] into "Active" (Aquamarine) and all the rest "Non-Active" (i.e., at most moderately active Lazurite/Malachite and completely non-active Ruby).

The canonical nucleosome arrangement Nuc−1/NDR/Nuc+1 is typical for housekeeping genes; stress-responsive genes have noncanonical nucleosome arrangement due to delocalized, i.e., uniformly distributed nucleosomes [9]. Partial/complete digestion of Drosophila chromatin with high/low concentrations of MNase revealed whole-genome maps of MNase sensitive/resistant types of nucleosomes, respectively [21]. We performed analysis of these maps and classified all genes according to expression level and the chromatin state of promoter DNA and found relatively large fractions of "silent" genes and expressed genes referred to Active and Non-Active chromatin environment, respectively. It was found that the regular/fuzzy nucleosome arrangement in promoters referred to Active/Non-Active chromatin states is conserved, irrespective to either the expression level or the number of expressed tissues.

## 2. Results

### 2.1. Gene Classification by Chromatin States and Expression Measures

We proceeded from the earlier analysis of one cell line [21] to the analysis of multiple tissues. We analyzed 13,574 protein-coding genes from FlyBase with RNA-seq profile for 29 tissues [31] and cell types [32], and unique chromosome position of TSS (see Section 4).

We classified all genes according to RPKM value (reads per kilobase million) in S2 cells [30] on three classes:

- Silent (RPKM = 0, 6972 genes);
- Modest (1 ≤ RPKM ≤ 25, 3718);
- High (RPKM > 25, 3584).

We computed the breadth of expression (BoE) as the number of tissues in which a gene was expressed. We considered genes satisfied the criterion RPKM > 3 to compute BoE. We classified genes by BoE on three classes:

- Tissue-specific (BoE ≤ 6, 5534 genes);
- Modest (6 > BoE ≤ 24, 4278);
- Constitutive (BoE >24, 3942).

To incorporate the chromatin landscape in the analysis we considered the whole-genome 4-state fragmentation model 4HMM [23,30]. We computed the number of genes that were masked by certain state at a specific distance from TSS. For any state a number of genes remained approximately constant within the region (−300; +200) around TSS (Figure S1A). According to 4HMM the Aquamarine chromatin below will be referenced to as "Active", and the remaining three states as "Non-Active". Hence, below we consider 6565/6209 genes with this region entirely encompassed by the Active domains ("Active" genes) or any other domain ("Non-Active" genes).

The distribution of Active genes into classes of RPKM value and BoE demonstrated that the majority of genes had at least modest BoE or RPKM (81.5%, Figure 1A, RPKM > 0 or BoE > 6). On the contrary, the majority of Non-Active genes were silent or tissue-specific (58.3%, Figure 1B). Hence, about 20% of Active genes were either silent or tissue-specific, and about 40% of Non-Active genes revealed either positive RPKM value or modest/high BoE.
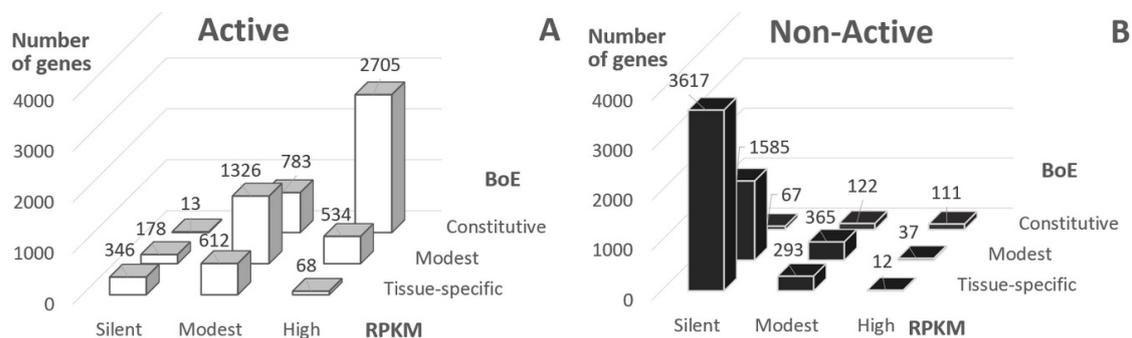


**Figure 1.** Classification of genes according to expression level (RPKM, reads per kilobase million) and breadth of expression (BoE). BoE was computed for 29 tissues; RPKM values respected to S2 cells; panels (**A**,**B**) show "Active" and "Non-Active" genes, these classes respect to Aquamarine and Lazurite/Malachite/Ruby chromatin states according to 4HMM [23,30] (see Section 4).

### 2.2. Irrespective to Expression Measures, Nucleosome Arrangement Is Conservative for Active Genes

We appealed to experimental data on the mapping of resistant/sensitive nucleosomes (see Section 4, [21]) to explain the diverse classification of genes by the chromatin state and expression

measures (Figure 1). We computed the number of nucleosome sites mapped for various RPKM/BoE classes in the region (−300; +200) relative to TSS for Active and Non-Active genes.

Figure 2A demonstrates the profiles of the average number of resistant and sensitive nucleosome sites for Active genes. The strong nucleosome positioning site downstream TSS (the nucleosome center position +135, Nuc+1) is clearly visible for all datasets. The slight blurring of the peak we found only for "silent" genes. According to the previous observation [21], profiles for resistant and sensitive nucleosome sites show the notable difference in the region upstream TSS. The depletion of resistant nucleosome sites just upstream TSS are clear for all datasets besides the "silent". In the upstream region, peaks for centers of sensitive nucleosomes are located in positions −35 (Nuc−1) and −185 (Nuc−2) for genes with the modest/high RPKM. The majority of silent Active genes were tissue-specific (346, Figure 1A). These genes revealed the arrangement similar to that of genes with the modest/high RPKM.

Figure 2B shows profiles of the average number of resistant/sensitive nucleosome sites for Non-Active genes. These profiles are substantially divergent from those of Active genes (Figure 2A). Namely, we did not find

- the strong nucleosome formation sites downstream TSS;
- the depletion of resistant nucleosome sites upstream TSS;
- the regular ladder of sensitive nucleosomes upstream TSS.

Note that peaks at positions Nuc−2, Nuc−1, and Nuc+1 are completely absent in genes with modest/high RPKM or in "modest BoE"/constitutive genes.

Thus, for Non-Active genes respecting modest/high RPKM or "modest BoE"/constitutive classes (Figure 2B), we did not detect the arrangement revealed above for Active genes (Figure 2A). The comparison of profiles for Active/Non-Active genes (Figure 2) revealed that irrespective to RPKM/BoE Active genes possessed the common pattern of nucleosome arrangement that was absent in Non-Active genes.

### 2.3. Correlations between Expression Measures and Nucleosome Positioning

In this section, we study dependencies between the number of mapped nucleosome sites around TSS and expression measures RPKM/BoE. This analysis was performed separately for Active and Non-Active genes. We applied Kendall's rank correlation coefficients (CC) (see Section 4) to reveal significant dependencies. Figure 3 shows profiles of CC between the RPKM value and the number of resistant/sensitive sites for three classes of BoE. We found several differences between profiles of CC in Active and Non-Active genes. Thus, for Active genes we found:

- strongly significant negative CC for resistant nucleosomal sites around TSS [−120; +80] for tissue-specific and modest; moderate negative CC for sensitive nucleosomal sites for tissue-specific [−30; +50] and modest [+10; +70];
- strongly significant positive CC downstream TSS [+70; +200] for tissue-specific and modest;
- moderately significant positive CC for sensitive nucleosomal sites upstream TSS in tissue-specific [−210; −180]) and modest [−160; −150].

In Non-Active genes the 1st and 3rd trends are not observed; the 2nd trend is even opposite for tissue-specific and is absent for the rest genes.

Constitutive Active and Non-Active genes did not show significant correlation (Figure 3). However, the pattern of Active genes downstream TSS is similar to that of genes with lower BoE ([+130; +160] in Figure 3). As we expected, the tendency to a positive correlation is completely absent in this region in constitutive Non-Active genes.

Figure S2 shows profiles of CC between BoE and the number of resistant/sensitive nucleosomal sites for three classes of RPKM value. Though correlations for the BoE, in general, are less significant compared to those for the RPKM value (Figure 3), the pattern for Active genes still is clearer than that for Non-Active genes.
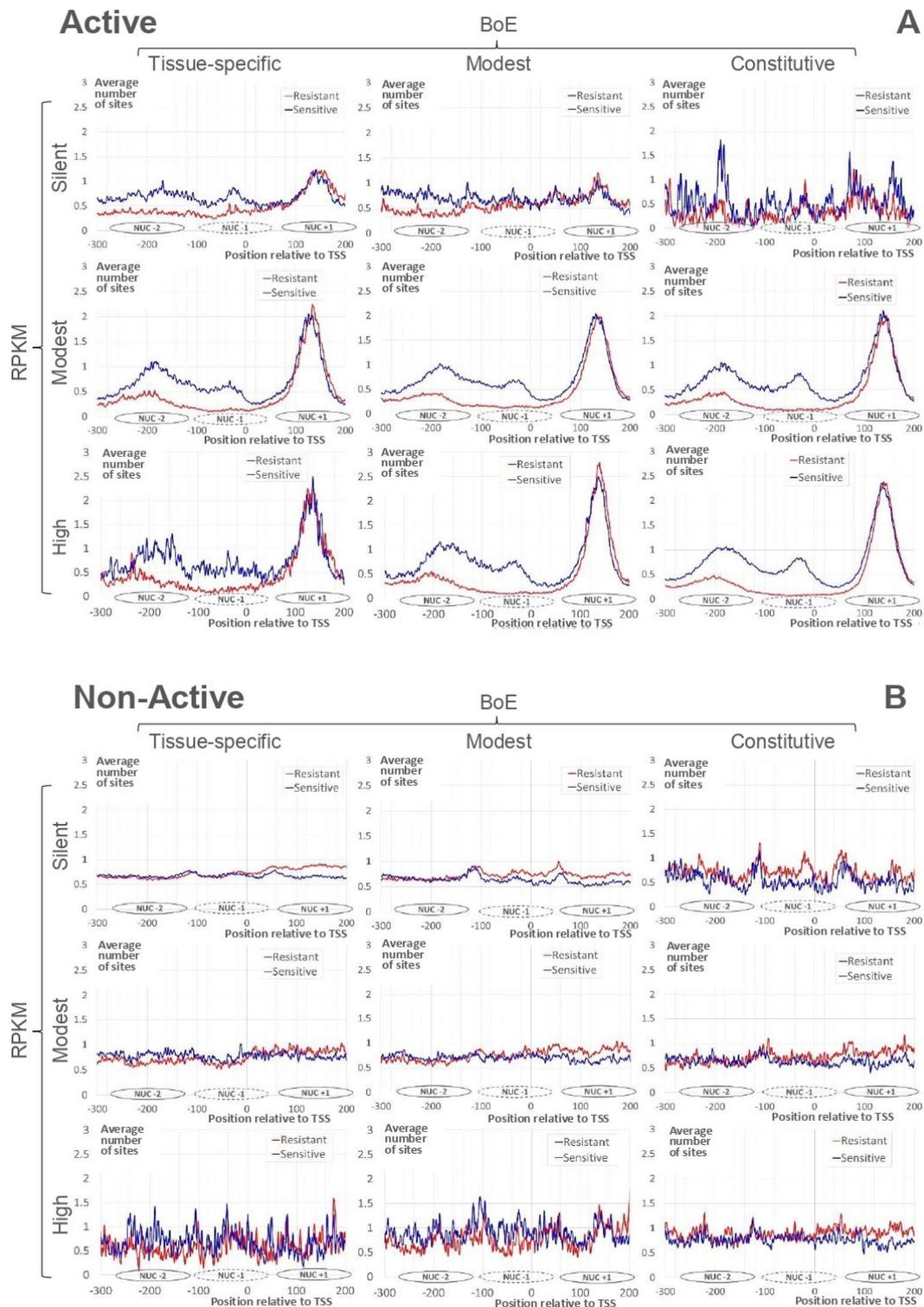
**Figure 2.** Average numbers of mapped nucleosome sites for "Active" (**A**) and "Non-Active" (**B**) genes (see Section 4). Red/blue lines mark resistant/sensitive nucleosomes [21]. X axes denote the position relative to TSS. Y axes show the moving mean over a window of 3 nt for average number of nucleosome sites. The headers of columns/rows mean the classification of genes by BoE/RPKM. Ovals near X axes denote approximate positions of nucleosomes upstream (Nuc−2, Nuc−1) and downstream (Nuc+1) TSS.
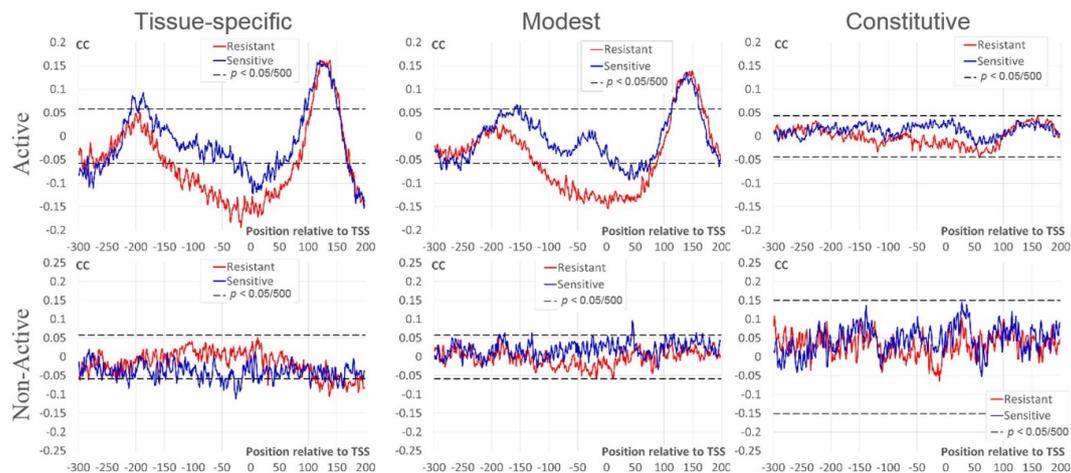
**Figure 3.** Correlation coefficients (CC) between the expression level (RPKM) and the number of mapped nucleosomes. The top/bottom rows denote datasets of Active/Non-Active genes (see Section 4). The left, middle, and right columns mark intervals of the expression breadth (BoE). X axes show the position relative to TSS, Y axes denote the moving mean of CC over a window of three nt. Red/blue lines denote resistant/sensitive nucleosomes [21]. Dashed black lines mark the threshold of Bonferroni corrected *p*-value for CC ($p < 0.05/500$).

## 3. Discussion

The development of whole-genome maps of chromatin states [23,33–35] and the availability of whole transcriptome profiling data has [31,32,36] created a rich resource for bioinformatics analyses. Earlier, Filion et al. [33] suggested that the chromatin states act as guides that helped to target DNA binding factors to specific regions of the genome, even though the cognate binding motifs were broadly distributed. In the current study, we supposed that this statement refers to histones too. Particularly, the chromatin state may also guide the gene expression through the nucleosome positioning. Hence, for various chromatin states we analyzed the association between the expression measures and the number of mapped nucleosome sites in specific position relative to TSS. The gene expression we measured with the expression level in S2 cells (RPKM value) and the breadth of expression (BoE, the number of expressed tissues among 29 tissue types annotated in FlyBase database [32]). We analyzed the distribution of domains of chromatin states relative TSS for the 4-state [23,30] and 5-state [33] maps of chromatin states. We selected the criterion of entire overlap of the region (−300; +200) relative TSS with domains of certain chromatin state and partitioned all genes (Figure S1) on "Active" and "Non-Active" classes. These datasets we further classified by RPKM value and BoE (Figure 1 and Supplementary Figure S3). Although, the Active genes generally had higher RPKM and BoE than Non-Active genes,

- about 20% of Active genes were either silent or tissue-specific;
- about 40% of Non-Active genes had at least modest RPKM or BoE.

Then, along the length of promoter, we computed the profiles of the average numbers of resistant/sensitive nucleosome sites for Active and Non-Active genes and categorized them by RPKM value and BoE. Notably, these subsets revealed clear diversity in the nucleosome arrangement around TSS (Figure 2 and Supplementary Figure S4). We concluded that the regular nucleosome positioning is the specific feature of the Active chromatin state.

Next, we tested the significance of correlations between the number of resistant/sensitive nucleosomal sites [21] and RPKM/BoE for Active and Non-Active genes in various promoter positions. We performed this analysis separately for every class of RPKM/BoE (Figure 1). We confirmed that Active and Non-Active genes had substantially different dependencies between the number of nucleosomal sites and the expression level. Whereas Active genes revealed the conservative pattern of significant dependencies in various positions of promoter, Non-Active genes did not have this pattern (Figure 3 and

Supplementary Figure S5). The similar trend we found for the breadth of expression (Supplementary Figures S2 and S6), i.e., the conservative pattern of nucleosome positioning around TSS, is more prominent for Active genes than for Non-Active genes.

Our conclusions are in agreement with a number of recent reports:

- It was estimated [37] that almost half of the human genome contained regular arrays of nucleosomes that enriched in active chromatin domains. Since the presence of transcription factor (TF) binding sites (BS) strongly supports gene expression, the analysis examined the positioning of nucleosomes around these BS using ChIP-seq data for 35 different TFs. As a result, TF occupancy was closely related to the strength of nucleosome positioning.
- The genome-wide location of nucleosomes during zebrafish embryogenesis indicated that well-positioned nucleosome arrays appeared on thousands of promoters during genome activation [38]. It is important to remember that both models [23,30,33] of chromatin states were based on whole genome profiles of a number of non-histone proteins, i.e., the pattern of BS of these proteins should be different in Active in Non-Active genes.
- The recent analysis of chromatin folding for various epigenetic states of chromatin [39] supported the importance of chromatin state annotation for analysis of genome-wide nucleosome organization. In particular, the classification of genomic domains in Drosophila cells into transcriptionally active, inactive, or Polycomb-repressed states revealed distinct chromatin organizations for each state.

Earlier [21], genes ordered by their expression levels showed marked dichotomy, with highly expressed genes exhibiting much more pronounced nucleosome positioning around the TSS. Based on our analysis, we propose two distinct mechanisms associated with mapping of gene promoter to domains of Active or Non-Active chromatin states. Regardless of the expression level or breadth of expression, the nucleosome positioning in the promoter:

- for Active state potentiates transcription, i.e., a more regular nucleosome arrangement in promoter points to higher expression level and higher number of expressed tissues;
- for Non-Active state does not influence gene expression, i.e., the expression level/breadth and the nucleosome positioning are independent.

## 4. Materials and Methods

### 4.1. Whole Genome Data Analysis

We took the nucleosome mapping data [21] (GEO AC GSE69336) as the whole-genome profiles for chromosome arms 2L, 2R, 3L, 3R, and X. These data included the results for MNase-sensitive/resistant nucleosome positioning sites (low/high concentration of MNase) for S2 cells.

We used the dataset of RNA-seq data from the modENCODE database [28], FB2013_03 release of FlyBase [40]. This table represents expression levels (RPKM values) for 29 tissues (FBlc0000206) [31] and S2 cells (FBlc0000260) [32]. We considered 13,574 protein-coding genes from the FlyBase database [41] with distinct TSS and present RNA-seq data (Supplementary Table S1). We applied the threshold RPKM = 3 for all tissues to compute the breadth of expression (BoE), the number of tissues with RPKM values exceeded the threshold.

We applied the 4-state chromatin map (4HMM) [23,30]. We considered the mapping of 5′-regulatory region of genes to domains of Aquamarine (or Active) state as the marker of an active gene. In particular, genes possessing region (−300; +200) around TSS entirely mapped within domains of Active and any other states (Lazurite, Malachite, and Ruby) we referred to as Active and Non-Active, respectively. Additionally, we considered the 5-state chromatin map [33] (the Yellow/Red and Green/Blue/Black states defined Active and Non-Active genes, respectively).

*4.2. Statistical Analysis*

We computed Kendall's rank correlation coefficients (CC) to estimate the significance of correlations between the expression measure (BoE or RPKM) and the number of nucleosomes mapped in a specific position. To estimate the significance of Kendall's rank CC we applied the normal approximation [42] with the mean zero and the variance *D*:

$D = \frac{2 \times (2 \times N + 5)}{9 \times N \times (N-1)}$, here *N* denotes the sample size.

## 5. Conclusions

We took the whole-genome map of chromatin states and referred "Active" and "Non-Active" genes according to the overlap of their 5'-regulatory regions with domains of Active and Non-Active chromatin states. Our analysis confirmed that classification of all genes according chromatin states allowed deducing the better explanation for the influence of nucleosome positioning on gene transcription activity. We found that irrespective to the expression level or the number of expressed tissues:

- Active genes demonstrated clear pattern of nucleosome arrangement around TSS; the regular nucleosome positioning correlates with the expression level/breadth;
- Non-Active genes did not show such clear specificity in nucleosome positioning; the correlation between the regular nucleosome positioning and expression level or breadth either is absent, not conservative, or shows even the opposite significant trend compared to that of Active genes.

The analysis of 5-state model of Filion et al. [33] supported these conclusions (Supplementary Figures S1B and S3–S6).

**Supplementary Materials:** Supplementary materials can be found at http://www.mdpi.com/1422-0067/21/23/9282/s1.

**Author Contributions:** Conceptualization, V.G.L., Y.M.M. and I.F.Z.; Methodology, Software, Validation, Formal analysis, V.G.L., Data curation, V.G.L. and T.Y.Z.; Writing—original draft preparation, V.G.L., Writing—review and editing, V.G.L., T.Y.Z. and I.F.Z. All authors have read and agreed to the published version of the manuscript.

## Abbreviations

| | |
|---|---|
| 4HMM | Four state hidden Markov model |
| BoE | Breadth of expression |
| MNase | Micrococcal nuclease |
| Nuc | Nucleosome |
| NDR | Nucleosome-depleted region |
| RPKM | Reads per kilobase million |
| TSS | Transcription start site |

## References

1. Hughes, A.L.; Rando, O.J. Mechanisms Underlying Nucleosome Positioning In Vivo. *Annu. Rev. Biophys.* **2014**, *43*, 41–63. [CrossRef] [PubMed]
2. Jiang, C.; Pugh, B.F. Nucleosome positioning and gene regulation: Advances through genomics. *Nat. Rev. Genet.* **2009**, *10*, 161–172. [CrossRef] [PubMed]
3. Morozov, A.V.; Fortney, K.; Gaykalova, D.A.; Studitsky, V.M.; Widom, J.; Siggia, E.D. Using DNA mechanics to predict in vitro nucleosome positions and formation energies. *Nucleic Acids Res.* **2009**, *37*, 4707–4722. [CrossRef] [PubMed]

4.  Hughes, A.L.; Jin, Y.; Rando, O.J.; Struhl, K. A functional evolutionary approach to identify determinants of nucleosome positioning: A unifying model for establishing the genome-wide pattern. *Mol. Cell* **2012**, *48*, 5–15. [CrossRef] [PubMed]

5.  Struhl, K.; Segal, E. Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* **2013**, *20*, 267–273. [CrossRef]

6.  Moshkin, Y.M.; Chalkley, G.E.; Kan, T.W.; Reddy, B.A.; Ozgur, Z.; Van Ijcken, W.F.J.; Dekkers, D.H.W.; Demmers, J.A.; Travers, A.A.; Verrijzer, C.P. Remodelers Organize Cellular Chromatin by Counteracting Intrinsic Histone-DNA Sequence Preferences in a Class-Specific Manner. *Mol. Cell. Biol.* **2011**, *32*, 675–688. [CrossRef]

7.  Zhang, Y.; Moqtaderi, Z.; Rattner, B.P.; Euskirchen, G.; Snyder, M.; Kadonaga, J.T.; Liu, X.S.; Struhl, K. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat. Struct. Mol. Biol.* **2009**, *16*, 847–852. [CrossRef]

8.  Valouev, A.; Johnson, S.M.; Boyd, S.D.; Smith, S.E.; Fire, A.Z.; Sidow, A. Determinants of nucleosome organization in primary human cells. *Nat. Cell Biol.* **2011**, *474*, 516–520. [CrossRef]

9.  Radman-Livaja, M.; Rando, O.J. Nucleosome positioning: How is it established, and why does it matter? *Dev. Biol.* **2010**, *339*, 258–266. [CrossRef]

10. Barozzi, I.; Simonatto, M.; Bonifacio, S.; Yang, L.; Rohs, R.; Ghisletti, S.; Natoli, G. Coregulation of Transcription Factor Binding and Nucleosome Occupancy through DNA Features of Mammalian Enhancers. *Mol. Cell* **2014**, *54*, 844–857. [CrossRef]

11. Tsompana, M.; Buck, M.J. Chromatin accessibility: A window into the genome. *Epigenet. Chromatin* **2014**, *7*, 1–16. [CrossRef] [PubMed]

12. Barski, A.; Cuddapah, S.; Cui, K.; Roh, T.Y.; Schones, D.E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. High-resolution profiling of histone methylations in the human genome. *Cell* **2007**, *129*, 823–837. [CrossRef] [PubMed]

13. Yuan, G.-C.; Liu, Y.-J.; Dion, M.F.; Slack, M.D.; Wu, L.F.; Altschuler, S.J.; Rando, O.J. Genome-Scale Identification of Nucleosome Positions in S. cerevisiae. *Science* **2005**, *309*, 626–630. [CrossRef] [PubMed]

14. Mavrich, T.N.; Ioshikhes, I.P.; Venters, B.J.; Jiang, C.; Tomsho, L.P.; Qi, J.; Schuster, S.C.; Albert, I.; Pugh, B.F. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **2008**, *18*, 1073–1083. [CrossRef]

15. Mavrich, T.N.; Jiang, C.; Ioshikhes, I.P.; Li, X.; Venters, B.J.; Zanton, S.J.; Tomsho, L.P.; Qi, J.; Glaser, R.L.; Schuster, S.C.; et al. Nucleosome organization in the Drosophila genome. *Nat. Cell Biol.* **2008**, *453*, 358–362. [CrossRef]

16. Lee, W.; Tillo, D.; Bray, N.; Morse, R.H.; Davis, R.W.; Hughes, T.R.; Nislow, C. A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* **2007**, *39*, 1235–1244. [CrossRef]

17. Schones, D.E.; Cui, K.; Cuddapah, S.; Roh, T.-Y.; Barski, A.; Wang, Z.; Wei, G.; Zhao, K. Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell* **2008**, *132*, 887–898. [CrossRef]

18. Zhang, T.; Zhang, W.; Jiang, J. Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiol.* **2015**, *168*, 1406–1416. [CrossRef]

19. Li, G.; Liu, S.; Wang, J.; He, J.; Huang, H.; Zhang, Y.; Xu, L. ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis. *Plant J.* **2014**, *78*, 706–714. [CrossRef]

20. Liu, M.-J.; Seddon, A.E.; Tsai, Z.T.-Y.; Major, I.T.; Floer, M.; Howe, G.A.; Shiu, S.-H. Determinants of nucleosome positioning and their influence on plant gene expression. *Genome Res.* **2015**, *25*, 1182–1195. [CrossRef]

21. Chereji, R.V.; Kan, T.-W.; Grudniewska, M.; Romashchenko, A.V.; Berezikov, E.; Zhimulev, I.F.; Guryev, V.; Morozov, A.V.; Moshkin, Y.M. Genome-wide profiling of nucleosome sensitivity and chromatin accessibility in Drosophila melanogaster. *Nucleic Acids Res.* **2015**, *44*, 1036–1051. [CrossRef] [PubMed]

22. Van Bakel, H.; Tsui, K.; Gebbia, M.; Mnaimneh, S.; Hughes, T.R.; Nislow, C. A Compendium of Nucleosome and Transcript Profiles Reveals Determinants of Chromatin Architecture and Transcription. *PLoS Genet.* **2013**, *9*, e1003479. [CrossRef] [PubMed]

23. Zhimulev, I.F.; Zykova, T.Y.; Goncharov, F.P.; Khoroshko, V.A.; Demakova, O.V.; Semeshin, V.F.; Pokholkova, G.V.; Boldyreva, L.V.; Demidova, D.S.; Babenko, V.N.; et al. Genetic Organization of Interphase Chromosome Bands and Interbands in Drosophila melanogaster. *PLoS ONE* **2014**, *9*, e101631. [CrossRef] [PubMed]

24. Zykova, T.Y.; Levitsky, V.G.; Belyaeva, E.S.; Zhimulev, I.F. Polytene Chromosomes—A Portrait of Functional Organization of the Drosophila Genome. *Curr. Genom.* **2018**, *19*, 179–191. [CrossRef]

25. Semeshin, V.F.; Demakov, S.A.; Alonso, M.P.; Belyaeva, E.S.; Bonner, J.J.; Zhimulev, I.F. Electron microscopical analysis of Drosophila polytene chromosomes. *Chromosoma* **1989**, *97*, 396–412. [CrossRef]

26. Vatolina, T.Y.; Boldyreva, L.V.; Demakova, O.V.; Demakov, S.A.; Kokoza, E.B.; Semeshin, V.F.; Babenko, V.N.; Goncharov, F.P.; Belyaeva, E.S.; Zhimulev, I.F. Correction: Identical Functional Organization of Nonpolytene and Polytene Chromosomes in Drosophila melanogaster. *PLoS ONE* **2011**, *6*, 25960. [CrossRef]

27. Zhimulev, I.F.; Belyaeva, E.S.; Vatolina, T.Y.; Demakov, S.A. Banding patterns in Drosophila melanogaster polytene chromosomes correlate with DNA-binding protein occupancy. *BioEssays* **2012**, *34*, 498–508. [CrossRef]

28. Celniker, S.E.; modENCODE Consortium; Dillon, L.A.L.; Gerstein, M.B.; Gunsalus, K.C.; Henikoff, S.; Karpen, G.H.; Kellis, M.; Lai, E.C.; Lieb, J.D.; et al. Unlocking the secrets of the genome. *Nat. Cell Biol.* **2009**, *459*, 927–930. [CrossRef]

29. Demakov, S.A.; Zykova, T.Y.; Babenko, V.; Semeshin, V.F.; Belyaeva, E.S.; Zhimulev, I.F. Protein composition of interband regions in polytene and cell line chromosomes of Drosophila melanogaster. *BMC Genom.* **2011**, *12*, 566. [CrossRef]

30. Boldyreva, L.V.; Goncharov, F.P.; Demakova, O.V.; Zykova, T.Y.; Levitsky, V.G.; Kolesnikov, N.N.; Pindyurin, A.V.; Semeshin, V.F.; Zhimulev, I.F. Protein and Genetic Composition of Four Chromatin Types in Drosophila melanogaster Cell Lines. *Curr. Genom.* **2017**, *18*, 214–226. [CrossRef]

31. Brown, J.B.; Boley, N.; Eisman, R.C.; May, G.E.; Stoiber, M.H.; Duff, M.O.; Booth, B.W.; Wen, J.; Park, S.; Suzuki, A.M.; et al. Diversity and dynamics of the Drosophila transcriptome. *Nat. Cell Biol.* **2014**, *512*, 393–399. [CrossRef] [PubMed]

32. Cherbas, L.; Willingham, A.; Zhang, D.; Yang, L.; Zou, Y.; Eads, B.D.; Carlson, J.W.; Landolin, J.M.; Kapranov, P.; Dumais, J.; et al. The transcriptional diversity of 25 Drosophila cell lines. *Genome Res.* **2011**, *21*, 301–314. [CrossRef] [PubMed]

33. Filion, G.J.; Van Bemmel, J.G.; Braunschweig, U.; Talhout, W.; Kind, J.; Ward, L.D.; Brugman, W.; De Castro, I.J.; Kerkhoven, R.M.; Bussemaker, H.J.; et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **2010**, *143*, 212–224. [CrossRef] [PubMed]

34. Kharchenko, P.V.; Alekseyenko, A.A.; Schwartz, Y.B.; Minoda, A.; Riddle, N.C.; Ernst, J.; Sabo, P.J.; Larschan, E.; Gorchakov, A.A.; Gu, T.; et al. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nat. Cell Biol.* **2011**, *471*, 480–485. [CrossRef]

35. Milon, B.; Sun, Y.; Chang, W.; Creasy, T.; Mahurkar, A.; Shetty, A.; Nurminsky, D.; Nurminskaya, M. Map of open and closed chromatin domains in Drosophila genome. *BMC Genom.* **2014**, *15*, 988. [CrossRef] [PubMed]

36. Graveley, B.R.; Brooks, A.N.; Carlson, J.W.; Duff, M.O.; Landolin, J.M.; Yang, L.; Artieri, C.G.; Van Baren, M.J.; Boley, N.; Booth, B.W.; et al. The developmental transcriptome of Drosophila melanogaster. *Nat. Cell Biol.* **2011**, *471*, 473–479. [CrossRef]

37. Gaffney, D.J.; McVicker, G.; Pai, A.A.; Fondufe-Mittendorf, Y.N.; Lewellen, N.; Michelini, K.; Widom, J.; Gilad, Y.; Pritchard, J.K. Controls of Nucleosome Positioning in the Human Genome. *PLoS Genet.* **2012**, *8*, e1003036. [CrossRef]

38. Zhang, Y.; Vastenhouw, N.L.; Feng, J.; Fu, K.; Wang, C.; Ge, Y.; Pauli, A.; Van Hummelen, P.; Schier, A.F.; Liu, X.S. Canonical nucleosome organization at promoters forms during genome activation. *Genome Res.* **2013**, *24*, 260–266. [CrossRef]

39. Boettiger, A.N.; Bintu, B.; Moffitt, J.R.; Wang, S.; Beliveau, B.J.; Fudenberg, G.; Imakaev, M.V.; Mirny, L.A.; Wu, C.-T.; Zhuang, X. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nat. Cell Biol.* **2016**, *529*, 418–422. [CrossRef]

40. The Flybase Database RNA-seq Data. Available online: http://fb2014_03.flybase.org/static_pages/downloads/FB2014_03/genes/gene_rpkm_report_fb_2014_03.tsv.gz (accessed on 4 December 2020).

41. The Genome of D. Melanogaster, Release 5.50 of FlyBase Database. Available online: ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/dmel_r5.50_FB2013_02 (accessed on 4 December 2020).

42. Kendall, M.G. *Rank Correlation Methods*, 4th ed.; Charles Griffin & Co.: London, UK, 1970; ISBN 978-0852641996.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.