Published in final edited form as:

Nat Commun. 2013 ; 4: 1464. doi:10.1038/ncomms2469.

A subset of *Drosophila* Myc sites remain associated with mitotic chromosomes co-localized with insulator proteins

Jingping Yang, **Elizabeth Sung**, **Paul G. Donlin-Asp**, and **Victor G Corces**

Department of Biology, Emory University, 1510 Clifton Road NE, Atlanta GA 30322

Abstract

Myc has been characterized as a transcription factor that activates expression of genes involved in pluripotency and cancer, and as a component of the replication complex. Here we find that Myc is present at promoters and enhancers of *D. melanogaster* genes during interphase. Myc co-localizes with Orc2, which is part of the pre-replication complex, during G1. As is the case in mammals, Myc associates preferentially with paused genes, suggesting that it may also be involved in the release of RNAPII from promoter proximal pausing in *Drosophila*. Interestingly, about 40% of Myc sites present in interphase persists during mitosis. None of the Myc mitotic sites correspond to enhancers and only some correspond to promoters. The rest of mitotic Myc sites overlap with binding sites for multiple insulator proteins that are also maintained in mitosis. These results suggest alternative mechanisms to explain the role of Myc in pluripotency and cancer.

Myc has been extensively studied as an oncogene that plays critical roles in cancer initiation and metastasis of may different types of tumors [1, 2]. Myc is a sequence specific DNA binding protein [3] that can bind to both the canonical (CACGTG or CATGTG) and non-canonical CA--TG E box sequences [4]. Although Myc has been found to regulate various cellular processes including cell growth, cell proliferation, and cell differentiation, the mechanisms by which it elicits neoplastic transformation are not well understood.

Myc is a basic-helix-loop-helix leucine zipper (bHLH-Zip) transcription factor that regulates the expression of protein coding genes and microRNAs [5–8]. Genomic searches for Myc target genes have uncovered a role for this protein in the regulation of hundreds of genes involved in cell cycle progression, differentiation, apoptosis, DNA repair, angiogenesis, chromosome instability, and ribosome biogenesis [1, 9, 10]. Results suggest that Myc may regulate expression of these genes, at least in part, by interacting with P-TEFb to release

---

Corresponding author: Victor G. Corces, Department of Biology, Emory University, 1510 Clifton Road NE, Atlanta, GA 30322. Phone number 404 727 4250. Fax number 404 727 2880. vcorces@emory.edu.

Contributions

V.G.C. planned the project; J.Y., P.G.D. and E.S. conducted experiments; J.Y. performed data analysis; J.Y. and V.G.C. wrote the manuscript and discussed the results.

Competing financial interests

The authors declare no competing financial interests.

Accession Codes

Sequence data have been deposited in NCBI's Gene Expression Omnibus (GEO) under accession numbers GSE32584 and GSE39521.

RNA polymerase II (RNAPII) from promoter proximal pausing and bringing about productive elongation [11–13]. Most genome-wide studies of Myc in mammalian cells have focused on its presence at promoter regions [14–16]. However, Myc is also found in non-promoter sequences. For example, Myc is enriched in the first intron of genes, and about 10% of Myc sites are present in intergenic regions (>100 kb from genes) in human B cells [17]. In mouse fibroblasts, 30.4% of Myc sites are intergenic (>1 kb from genes) and 22.4% are intragenic (1 kb downstream of TSS to 3′ end) [18]. These non-promoter Myc sites may function as transcriptional regulatory elements, such as enhancers, but their role has not been studied in detail.

In addition to its role in transcription of genes encoding proteins involved in DNA replication, Myc may also regulate this process directly. Cells overexpressing Myc become polyploidy but do not enter mitosis [19]. In *Drosophila*, Myc is required for endoreplication [20, 21], and it has been suggested that the role of Myc in replication is independent of transcription [22]. In human cells, the Myc protein interacts with the pre-replication complex and it has been shown to be required for recruitment of Mcm proteins at specific loci [22, 23], but whether this is a general phenomenon, and Myc co-localizes genome-wide with the origin recognition complex, has not been investigated. Myc controls a variety of cellular processes required for cell differentiation and is essential for cellular reprogramming to induce pluripotency and stem cell renewal [24–26]. The role of Myc in these cellular processes may be a consequence of its effects on gene expression at the local level but other evidence suggests that Myc can also affect chromatin more globally [27]. An important property of Myc that has been largely ignored when considering potential mechanisms by which this protein can affect gene expression is that it remains bound to DNA during mitosis [28, 29], raising the question of whether some of the functions ascribed to this protein are actually a consequence of its presence in mitotic chromosomes. However, the location and function of Myc in mitotic cells has not been explored.

Here we examine the distribution of Myc during interphase and mitosis in *Drosophila* Kc cells. We find that Myc co-localizes extensively with Orc2 during interphase, supporting a generalized role for Myc in the pre-replication complex. In addition to promoters, Myc is also present at enhancers of *Drosophila* genes. Interestingly, only a specific subset of interphase Myc sites remain in mitotic chromosomes. Mitotic Myc sites include a fraction of promoter regions and aligned insulators, where several insulator proteins co-localize within a 300 bp region. These results suggest that Myc may have an as of yet unappreciated role in the maintenance of chromosome structure and epigenetic information during the cell cycle that may explain some of its effects in tumorigenesis and pluripotency.

Results

Myc is present at the promoters of paused genes

In order to study changes in the distribution of Myc, we performed chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) in *D. melanogaster* Kc cells using an antibody against this protein. To distinguish possible roles of Myc in mitosis versus other stages of the cell cycle, we partially synchronized the cells, labeled them with antibodies to H3S10ph or Lamin Dm0, and separated the mitotic and interphase populations using FACS. Results of ChIP-seq experiments were quantitated and confirmed at three different genomic locations by ChIP-qPCR (Supplementary Figure S1). In order to determine whether the role of Myc in transcription and replication is conserved in *Drosophila*, we first mapped the binding sites for this protein in interphase when the cells are undergoing a process of biomass accumulation and preparing for the next cell cycle. We identified approximately 4000 Myc binding sites across the genome of interphase cells (Supplementary Data S1). Analysis of these data indicates that Myc associates with coding and non-coding genes (Figure 1a, left and right panels, respectively, Supplementary Data S2). Myc binds preferentially to promoter proximal regions (between TSSs and −200 bp). Only 8% of Myc sites fall in exons, including 5′ and 3′ UTRs. In addition to promoter regions, Myc also binds significantly in introns (21% of sites) and intergenic regions (17% of sites) (Figure 1b).

Myc has been extensively characterized as a transcription factor in mammals. In *Drosophila* 53% of Myc sites are located in the promoter regions of genes and 4% in the 5′UTR (Figure 1b). These sites of Myc associate with genes involved in rRNA synthesis, cell cycle, and development (Supplementary Data S2). To understand how Myc regulates adjacent genes, we examined their expression status. Generally Myc associates with genes that also have RNAPII at the TSS, suggesting they have undergone transcription initiation. However, the difference in RNAPII levels between the promoter region and the gene body is larger for genes with Myc than for genes without Myc (Figure 1a). These observations suggest that Myc-associated genes have high pausing indexes. In mammalian cells, Myc plays a role in the release of RNAPII from promoter-proximal pausing [13]. We therefore examined whether *Drosophila* Myc is also associated with paused genes by determining the pausing index, which is a measure of the difference in RNAPII levels between the promoter and gene body. The results indicate that Myc preferentially associates with paused genes in *Drosophila*. More Myc-associated genes show high pausing index than the average genome level. Around 41% of Myc associated genes show pausing indexes higher than 1, while only around 16% of all genes show a pausing index larger than 1 (K-S test, $p<2x10^{-16}$) (Figure 1c). Although Myc-associated genes show high pausing indexes, they are also highly expressed. About 50% of Myc-associated genes belong to the group with the highest transcription levels (Figure 1d), suggesting that RNAPII in Myc-associated paused genes is quickly released into productive elongation.

The role of Myc at non-promoter regions

As previously observed in mammalian cells [17, 18], *Drosophila* Myc also binds to non-promoter regions of coding genes (Figure 1b). These intergenic or intronic Myc sites are not located in promoter regions of non-coding genes. Out of 1144 intergenic or intronic Myc sites only 31 are found adjacent to non-coding RNA genes. Genome-wide studies of Myc distribution have been carried out in different types of mammalian cells but no detailed analysis of the function of these non-promoter sites is available. We therefore examined the distribution of histone modifications and RNAPII enrichment at these non-promoter Myc sites. Results indicate that a subset of these Myc sites show chromatin signatures characteristic of enhancers (Figure 2a and 2b). Myc sites at promoter regions are enriched in H3K4me3 and RNAPII. In contrast, a subset of Myc sites at non-promoter regions show

H3K4me1 while H3K4me3 and RNAPII are absent (Figure 2b). Enrichment in H3K4me1 without H3K4me3 is characteristic of enhancers [30]. These Myc sites also contain H3K27ac, indicating they correspond to active enhancers (Figure 2b). We have previously identified all enhancers in *Drosophila* Kc cells [31] and we used this information here to compare enhancers with and without Myc. The results suggest that most active enhancers containing H3K27ac also have Myc, while inactive enhancers lacking H3K27ac are depleted of Myc (Figure 2c). In fact, if we consider the subset of Myc-containing enhancers located in introns in order to be able to assign enhancers to their target genes, Myc-containing enhancers are present in the most actively transcribed genes (Figure 2d). Thus, a subset of non-promoter Myc sites appears to be present at active enhancers that regulate highly transcribed genes. However, not all the non-promoter Myc sites show enhancer-like chromatin features. A second subset of non-promoter Myc sites (denoted with a question mark in Figure 2b) is depleted of H3K4me1, H3K4me3, and H3K27ac and, therefore, it does not correspond to enhancers or promoters.

Myc associates with Orc2 genome-wide in *D. melanogaster*

In mammalian cells, the Myc protein interacts with the pre-replication complex and it has been found at specific DNA replication origins with Origin recognition complex (Orc) proteins [22]. To test whether this is also the case in *Drosophila*, and some of the sites with unknown function in Figure 2B correspond to replication origins, we compared the binding profiles of Myc and the Orc2 component of the complex. The signals for the two proteins show significant overlap in specific regions of the genome (Figure 3a). We then examined genome-wide correlations between the two proteins using heatmaps to visualize the information. The results indicate that Orc2 is present at most Myc sites in the genome and vice versa (Figure 3b). Thus, Myc co-localizes with Orc2 genome-wide in *Drosophila* cells at both promoter and non-promoter regions.

A subset of Myc sites is bound to chromosomes during mitosis

Myc has been shown to be present in mitotic chromosomes [28, 29] but its specific distribution in chromatin during mitosis has never been analyzed. The presence of Myc in mitotic chromosomes may be critical for its role in transcription. To gain further insights into mechanisms by which Myc affects gene expression, we mapped Myc binding sites in mitotic cells and compared the distribution of this protein between interphase and mitosis. The results indicate that all Myc mitotic sites are also occupied by this protein during interphase, but not all Myc sites in interphase are retained in mitosis (Figure 4a). Myc sites in the genome can therefore be classified as interphase-specific (Class I) or common to mitosis and interphase (Class II). We will refer to this second group as mitotic sites, although they are also present in interphase (Figure 4a). In interphase, the average enrichment of Myc at Class II sites is only about half of the average enrichment at Class I sites, but the enrichment is significantly higher than background (Figure 4b, top panel). In mitosis, there is no significant enrichment for Myc at Class I sites (Figure 4b, bottom panel). The mechanism by which Myc persists at only a subset of interphase sites may depend on the specific recognition sequence present at each class of sites. We therefore examined potential differences in the consensus motif at Class I and Class II sites. E boxes can be found in about 70% of Myc sites in either class, but the two classes show different preferences for specific sequences. Myc sites in Class I preferentially contain the canonical E box (CATGTG/CACGTG). In contrast, Myc sites in Class II sites are depleted of the canonical E box and, instead, show enrichment for the non-canonical E box (Figure 4c). The preference in utilization of each E box type is significant (chi square $p<0.0001$) and may represent the underlying mechanism to select Myc sites that will be maintained during the cell cycle.

Two classes of Myc sites with different roles in transcription

To gain insights into possible functional differences between interphase-specific and mitotic Myc sites, we first performed GO analysis for genes associated with each class. Class I Myc sites are enriched at genes involved in ribosome biogenesis, which is important for biomass accumulation in G1. This is consistent with reports for Myc-regulated genes in interphase cells in mammals [9]. However, genes associated with Class II Myc sites are not enriched for this category (Figure 5a). In addition to their presence at different target genes, the two groups of Myc sites may also affect gene expression by different mechanisms. In interphase cells, Myc-associated genes generally show higher pausing index than the average gene in the genome. We then parsed genes into two groups based on their association with Class I or Class II Myc sites. The results suggest that genes associated with Class I Myc sites (interphase-specific) still show a high pausing index, whereas genes associated with Class II Myc sites (those also present in mitosis) have slightly but statistically significant lower pausing indexes (K-S test, $p=2\text{x}10^{-6}$) (Figure 5b). Class I Myc sites may then be involved in the release of paused RNAPII for productive elongation in interphase cells, whereas Class II Myc sites may play a different regulatory function in transcription that is dependent on the presence of Myc protein in mitotic chromosomes.

Mitotic Myc sites are present at promoters but not enhancers

To further explore functional differences between the two classes of Myc sites, we clustered all the sites with histone modifications characteristic of enhancers and promoters using *k*-means clustering. The results reveal 5 clusters of Myc sites (Figure 5c). Class I Myc sites are present in three different clusters whereas class II Myc sites associate with two clusters. Class I sites are present at enhancers (Cluster I), promoters (Cluster III) and a cluster lacking either characteristic (Cluster II). Class II sites are present at promoters (Cluster IV) and a cluster of unknown function (Cluster V). Therefore, Myc is present at enhancers only during interphase and persists during mitosis at only a specific subset of all promoters occupied during interphase (Figure 5c). Interestingly, Myc-associated promoters in both interphase and mitosis appear to cluster in two groups with high or low levels of H3K4me3. In *Drosophila*, enhancers defined as sequences enriched in H3K27ac and H3K4me1 but lacking H3K4me3, are typically found within intronic regions [32]. Consistent with the clustering results, 55% of Class I non-promoter Myc sites are in introns while 70% of Class II non-promoter Myc sites fall in intergenic regions ($p<0.0001$) (Figure 5d). These results suggest that Class II Myc sites do not function as enhancers and they may play a different role in the genome independent of transcription. The possibility of a different role for Class II sites is supported by the observation that these sites are further apart from each other compared to Class I sites (Figure 5e).

Myc sites of unknown function associate with insulators

A subset of Myc sites in both Class I and Class II are not present at either enhancers or promoters. A third type of regulatory sequences found in eukaryotic cells is represented by insulators. To test whether the Myc sites of unknown function are present at insulators, we parsed ChIP-seq datasets of *Drosophila* insulator proteins BEAF-32, dCTCF, Su(Hw), GAF and CP190 with the clusters shown in Figure 5C. The results show a dramatic difference between Class I and Class II sites (Figure 6a). Class I sites associate preferentially with GAF, both at enhancers and promoters where this protein has been shown to be present [33], as well as in Cluster II containing sites not present at these two types of regulatory sequences. A subset of interphase-specific Class I promoter sites present in Cluster III, those containing high levels of H3K4me3 and presumably actively transcribed, contain BEAF-32 instead of GAF. Class II sites, on the other hand, associate with insulator proteins other than GAF (Figure 6a). In particular, all Class II sites, including those in Cluster V, contain all four insulator proteins tested, Su(Hw), BEAF-32, dCTCF and CP190 (Figure 6a and 6b).

Myc mitotic sites associate with mitotic insulator sites

The results presented above suggest a strong association between Class II Myc sites and sites of specific insulator proteins from interphase cells. Since Class II sites persist during mitosis, we wondered whether insulator proteins also remain at these sites during mitosis. To test this possibility, we mapped the binding of insulator proteins in mitosis and compared the distribution of Class II Myc sites with datasets of insulator protein localization in mitotic chromosomes. The results indicate that Myc overlaps extensively with insulator proteins during mitosis (Figure 6c). All mitotic Myc sites contain dCTCF, DREF and CP190, and a subset also contains BEAF-32. A fourth insulator protein, Su(Hw), is not present in chromosomes during mitosis (Supplementary Figure S2). Interestingly, a subset of the sites where Myc and Orc2 co-localize during interphase are sites where Myc persists during mitosis (Figure 6c).

Mitotic Myc sites localize at borders of topological domains

The role of Myc during mitosis may be local i.e. to mark a subset of promoters or origins of replication for rapid resumption of transcription or assembly of the pre-replication complex at the beginning of G1. Alternatively, Myc may play a more global role in chromatin organization. Recent work suggests that eukaryotic chromosomes during interphase are organized into topological association domains, characterized by high frequency of interactions, and separated by domain borders [34–38]. These borders are enriched in insulator proteins, which may contribute to the formation of boundaries that separate topological association domains. It is possible that some of this organization persist during mitosis, and that insulator proteins contribute to the maintenance of chromosome architecture during the cell cycle. The fact that Myc persists at the same genomic sites as insulator proteins in mitotic chromosomes suggest that it may also be present at domain borders. To test this hypothesis, we compared the distribution of Class I and Class II Myc sites with respect to domain borders previously defined in *Drosophila* embryonic nuclei [38]. The results suggest that this is indeed the case (Figure 6d). Class II Myc sites that remain on chromosomes during mitosis are significantly enriched at domain borders, whereas interphase-specific

Class II sites are significantly enriched inside domains. These results could be interpreted to suggest that a specific subset of Myc sites may remain bound to chromosomes during mitosis to organize the higher order structure of chromatin. Alternatively, the presence of Myc at domain borders may be a consequence, rather than a cause, of chromosome organization.

Discussion

Myc is a bHLH-Zip sequence-specific DNA binding protein that plays a crucial role in the regulation of critical cellular processes such as cell growth, cell division and cell differentiation. Importantly, an increase of Myc levels in the cell leads to oncogenic transformation [1]. The Myc protein is present in most or all proliferating cells of normal tissues, and its expression depends on the existence of mitogenic signals. In addition, expression of Myc is sufficient to induce cell division of most normal cells. Since Myc interacts with DNA in a sequence-specific manner, its role in cell proliferation has been explained based on its ability to control the expression of specific genes by activation or repression of transcription. Interestingly, recent results suggest that the role of Myc in gene expression is not to turn on the transcription of specific genes but rather to amplify the transcriptional output of genes that are already being expressed [39, 40]. The effects of Myc in transcription and replication have been rationalized on the basis of its involvement in the control of promoter-proximal pausing of RNAPII and its effects on chromatin structure at the level of histone covalent modifications. Myc can induce H4 acetylation [23, 41], which correlates with an increase of H4K20me2 and a transient increase of H4K20me1 [41]. H4K20me1 can function at the crossroad of genome integrity, cell cycle, and transcription [42], and H4K20me2 is recognized by Orc1, which is a component of the Orc complex mediating pre- DNA replication licensing. The bromo adjacent homology (BAH) domain of Orc1 specifically recognizes H4K20me2, a property common to BAH domains present within diverse metazoan Orc1 proteins [43]. The sole enzyme that catalyzes H4K20me1 is Setd8 (also known as PR-Set7 or KMT5a), which is an essential mediator of Myc-induced epidermal differentiation. Deletion of Setd8 in Myc-overexpressing skin cells blocks proliferation and differentiation [44].

Although the ability of Myc to act as a sequence-specific transcription factor and elicit changes in the 10 nm chromatin fiber may account for many of its effects on cell function, the finding of Myc in the proteome of mitotic chromosomes [28] represents an interesting puzzle. It is possible that Myc persistence on chromatin during mitosis has no relevance to its role in nuclear biology. On the other hand, several aspects of the distribution of Myc in mitotic chromosomes offer tantalizing explanations for some of its effects on transcription and replication. By comparing Myc binding sites in cells at interphase and mitosis we find two distinct groups of Myc sites. Class I sites only harbor Myc during interphase but become devoid of this protein during mitosis. These sites are adjacent to genes involved ribosome biogenesis, which have been reported to be cell type and species independent Myc targets [45]. In contrast, Class II Myc sites that persist during mitosis associate with genes that play roles in cell cycle or cell differentiation. In mammalian cells, this includes genes important for maintaining pluripotency and reprogramming [45]. It is possible that the presence of Myc at these genes during mitosis serves to preserve epigenetic memory of their expression necessary for the maintenance of cell identity.

The striking overlap of Myc and insulator sites during mitosis points to a more complex role for this protein in mitotic chromatin. Insulators have been shown to mediate long-range intra- and inter-chromosomal interactions [46]. Although the role of some of these interactions may be to regulate enhancer-promoter contacts, the finding of insulators at the boundaries of topological chromosome domains points to a larger and more complex function of these proteins in higher-order chromatin organization. The presence of Myc together with insulator proteins at these sites in mitotic chromosomes may explain some Myc-dependent phenotypes, including its effects on genome integrity. Mouse iPS cell lines induced with Myc show a significantly higher frequency of translocation than those induced without c-Myc [47]. This result is Myc dependent, as deletion of Myc box II reduces the translocation frequency [48]. Myc overexpression also induces telomeric aggregation in the interphase nucleus [49]. These effects of Myc on genomic integrity suggest that Myc may play a role in chromosome higher order structure that may depend on its presence at insulator sites during interphase and/or mitosis. This conclusion is further supported by the observation that mitotic Myc sites are enriched at the borders of topological chromosome domains, which are also enriched in insulator proteins. Domain boundaries are more accessible to the insertion of transposable elements and allow higher expression of transgenes, suggesting that they represent regions of the genome with more open higher-order chromatin [35]. Together, these observations agree with a model by which insulators organize the chromatin in the interphase nucleus by mediating interactions that create chromosomal domains. Transition from interphase to mitosis involves a condensation of the chromatin that nevertheless maintains this organization via the persistence of insulator proteins at domain boundaries. The boundary regions contain more open chromatin that may become accessible to components of the transcription and replication apparatus earlier at the end of M phase. The maintenance of Myc at these domain boundaries may ensure that adjacent genes are transcribed early at the M/G1 transition and a subset of replication origins assemble pre-replication complexes by recruiting Orc2 and, perhaps, determining replication timing. Additional experiments will be necessary to test this speculative but plausible model.

Methods

Cell culture and flow cytometry

Drosophila Kc167 cells were grown at 25°C in CCM3 media (Hyclone) to a density of $2X10^6$. For cell synchronization, the culture was treated with hydroxyurea (1 mg/ml in ethanol to a final concentration of 15 ng/ml) for 16 hr, incubated for 8 hr with nocodazole (5 mg/ml in DMSO, to a final concentration of 2 ng/ml) and harvested. For flow cytometry, cells were fixed for 10 min in 1% formaldehyde, blocked in suspension for 30 min in blocking buffer, incubated overnight with rabbit α-H3S10ph at 1:5000 or mouse α-Lamin Dm0 at 1:500, washed 3 x 15 min in blocking buffer, and then incubated with secondary antibody Alexa Fluor 488 α-rabbit at 1:5000. After a 30 min incubation in blocking buffer plus propidium iodide (0.1 mg/ml), samples were passed several times through a 25-gauge syringe to reduce clumping and sorted on a FACSAria II cell sorter. Enrichment of the mitotic and interphase cell populations was carried out by visualization of the mitotic marker H3S10ph by immunofluorescence microscopy, showing 97–99% purity (Supplementary Figure S2).

ChIP-seq analysis

ChIP was performed with ~4 × 10^7 cells. Cells were cross-linked with 1% formaldehyde for 10 min at room temperature. Nuclear lysates were sonicated to generate 200– 1000 bp DNA fragments. ChIP was then performed with 6 μL of *Drosophila* α-Myc antibody (Santa Cruz Biotechnology, sc-28208), α-CP190, α-CTCF or α-BEAF-32 antibodies [50, 51]. Libraries were prepared with the Illumina TruSeq DNA Sample Preparation Kit. Fragments in the 200–300 bp ranged were selected and sequenced in an Illumina HiSeq sequencer at the HudsonAlpha Institute for Biotechnology.

Bioinformatics analyses

Sequences were aligned to *Drosophila* dm3 using Bowtie. The output map files were converted to bed format for each chromosome arm using the VancouverShort package (http://vancouvershortr.sourceforge.net/). Peaks were called using CCAT3.0 [52] with the enrichment parameter set to 15. Myc-associated genes were defined as genes with Myc binding sites between −200 bp and the TSS or in the 5′UTR region.

In addition to the *Drosophila* Myc data obtained in this study, we used several datasets obtained from public sources. Orc2 ChIP-seq (modENCODE_2755), RNAPII ChIP-chip (modENCODE_328) and RNA expression in Kc cells (modENCODE_3305) were obtained from modENCODE. ChIP-seq data sets for H3K4me1, H3K4me3 and H3K27ac are from GSE36374. ChIP-seq data sets for insulator proteins are from GSE30740, GSE32584 and GSE39664. To build heatmaps, values for each ChIP-seq dataset were extracted for the 2000 bp region around the summit of peaks using custom R scripts (available upon request) and heatmap graphs were created using TreeView. Clusters in Figures 5c and 6a were created by *k*-means clustering using Cluster3.0 based on the mean values of the 300 bp around Myc sites for the samples listed.

The pausing index of genes was calculated using ChIP-chip data sets of RNAPII in Kc cells obtained from modENCODE. RNAPII at TSSs (P_{TSS}) was calculated as the mean enrichment of RNAPII at the 200 bp region around each TSS. RNAPII in the gene body (P_{body}) was calculated as the mean enrichment of RNAPII from +200 bp to the end of the gene. The pausing index is defined as the different between the P_{TSS} and P_{body}. Motif analysis of Myc binding sites was performed using Myc peak summits extend 50 bp on either side. The resulting 100 bp sequence for each peak was used to search for E boxes using a custom Perl script available upon request. Gene ontology analysis for Myc associated genes was performed with DAVID (http://david.abcc.ncifcrf.gov). Flybase IDs were used to determine statistically enriched biological process categories on the basis of a background list of all annotated genes in the *Drosophila* genome. In order to calculate the different expression groups and the differences in expression levels between these groups, we separated all the annotated genes into five groups with the same number of genes in each group. The expression score was obtained from ModEncode and was created based on the normalized tilling array data from 25 cell lines and 30 developmental stages. Here we use the expression score for Kc cells. The range of the expression score (log10(score) for each group is as follows: Group 1 <=1.65, Group 2 1.65–2.13, Group 3 2.13–2.44, Group 4 2.44–2.82, and Group 5 2.82–4.53.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank members of the lab for helpful discussions and suggestions during this study. We also thank The Genomic Services Lab at the HudsonAlpha Institute for Biotechnology for their help in performing Illumina sequencing of ChIP-seq samples. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM035463. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Dang, Chi V. MYC on the Path to Cancer. Cell. 2012; 149:22–35. [PubMed: 22464321]
2. Wolfer A, Ramaswamy S. MYC and Metastasis. Cancer Research. 2011; 71:2034–2037. [PubMed: 21406394]
3. Blackwell T, Kretzner L, Blackwood E, Eisenman R, Weintraub H. Sequence-specific DNA binding by the c-Myc protein. Science. 1990; 250:1149–1151. [PubMed: 2251503]
4. Blackwell TK, et al. Binding of myc proteins to canonical and noncanonical DNA sequences. Molecular and Cellular Biology. 1993; 13:5216–5224. [PubMed: 8395000]
5. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature. 2005; 435:839–843. [PubMed: 15944709]
6. Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB. MicroRNA regulation of a cancer network: Consequences of the feedback loops involving miR-17- 92, E2F, and Myc. Proceedings of the National Academy of Sciences. 2008; 105:19678–19683.
7. Lovén J, et al. MYCN-regulated microRNAs repress estrogen receptor-α (ESR1) expression and neuronal differentiation in human neuroblastoma. Proceedings of the National Academy of Sciences. 2010; 107:1553–1558.
8. Chang TC, et al. Widespread microRNA repression by Myc contributes to tumorigenesis. Nat Genet. 2008; 40:43–50. [PubMed: 18066065]
9. van Riggelen J, Yetil A, Felsher DW. MYC as a regulator of ribosome biogenesis and protein synthesis. Nat Rev Cancer. 2010; 10:301–309. [PubMed: 20332779]
10. Meyer N, Penn LZ. Reflecting on 25 years with MYC. Nat Rev Cancer. 2008; 8:976–990. [PubMed: 19029958]
11. Gargano B, Amente S, Majello B, Lania L. P-TEFb is a Crucial Co-Factor for Myc Transactivation. Cell Cycle. 2007; 6:2031–2037. [PubMed: 17700062]
12. Kanazawa S, Soucek L, Evan G, Okamoto T, Peterlin BM. c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis. Oncogene. 2003; 22:5707–5711. [PubMed: 12944920]
13. Rahl PB, et al. c-Myc Regulates Transcriptional Pause Release. Cell. 2010; 141:432–445. [PubMed: 20434984]
14. Guccione E, et al. Myc-binding-site recognition in the human genome is determined by chromatin context. Nat Cell Biol. 2006; 8:764–770. [PubMed: 16767079]
15. Kidder BL, Yang J, Palmer S. Stat3 and c-Myc Genome-Wide Promoter Occupancy in Embryonic Stem Cells. PLoS ONE. 2008; 3:e3932. [PubMed: 19079543]

16. Li Z, et al. A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. Proceedings of the National Academy of Sciences. 2003; 100:8164–8169.

17. Zeller KI, et al. Global mapping of c-Myc binding sites and target gene networks in human B cells. Proceedings of the National Academy of Sciences. 2006; 103:17834–17839.

18. Perna D, et al. Genome-wide mapping of Myc binding and gene regulation in serum-stimulated fibroblasts. Oncogene. 2012; 31:1695–1709. [PubMed: 21860422]

19. Li Q, Dang CV. c-Myc Overexpression Uncouples DNA Replication from Mitosis. Molecular and Cellular Biology. 1999; 19:5339–5351. [PubMed: 10409725]

20. Pierce SB, et al. dMyc is required for larval growth and endoreplication in Drosophila. Development. 2004; 131:2317–2327. [PubMed: 15128666]

21. Maines JZ, Stevens LM, Tong X, Stein D. Drosophila dMyc is required for ovary cell growth and endoreplication. Development. 2004; 131:775–786. [PubMed: 14724122]

22. Dominguez-Sola D, et al. Non-transcriptional control of DNA replication by c-Myc. Nature. 2007; 448:445–451. [PubMed: 17597761]

23. Swarnalatha M, Singh AK, Kumar V. The epigenetic control of E-box and Myc-dependent chromatin modifications regulate the licensing of lamin B2 origin during cell cycle. Nucleic Acids Research. 2012

24. Smith KN, Singh AM, Dalton S. Myc Represses Primitive Endoderm Differentiation in Pluripotent Stem Cells. Cell Stem Cell. 2010; 7:343–354. [PubMed: 20804970]

25. Moumen M, et al. The Proto-Oncogene Myc Is Essential for Mammary Stem Cell Function. STEM CELLS. 2012; 30:1246–1254. [PubMed: 22438054]

26. Varlakhanova NV, et al. myc maintains embryonic stem cell pluripotency and self-renewal. Differentiation. 2010; 80:9–19. [PubMed: 20537458]

27. Varlakhanova NV, Knoepfler PS. Acting locally and globally: Myc's ever-expanding roles on chromatin. Cancer Res. 2009; 69:7487–7490. [PubMed: 19773445]

28. Ohta S, et al. The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. Cell. 2010; 142:810–821. [PubMed: 20813266]

29. O'Donovan KJ, Diedler J, Couture GC, Fak JJ, Darnell RB. The Onconeural Antigen cdr2 Is a Novel APC/C Target that Acts in Mitosis to Regulate C-Myc Target Genes in Mammalian Tumor Cells. PLoS ONE. 2010; 5:e10045. [PubMed: 20383333]

30. Heintzman ND, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007; 39:311–318. [PubMed: 17277777]

31. Kellner WA, Ramos E, Van Bortle K, Takenaka N, Corces VG. Genome-wide phosphoacetylation of histone H3 at Drosophila enhancers and promoters. Genome Research. 2012; 22:1081–1088. [PubMed: 22508764]

32. Kharchenko PV, et al. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature. 2011; 471:480–485. [PubMed: 21179089]

33. Negre N, et al. A cis-regulatory map of the Drosophila genome. Nature. 2011; 471:527–531. [PubMed: 21430782]

34. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485:376–380. [PubMed: 22495300]

35. Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol Cell. 2012; 48:471–484. [PubMed: 23041285]

36. Lieberman-Aiden E, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009; 326:289–293. [PubMed: 19815776]

37. Nora EP, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012; 485:381–385. [PubMed: 22495304]

38. Sexton T, et al. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell. 2012; 148:458–472. [PubMed: 22265598]

39. Lin CY, et al. Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell. 2012; 151:56–67. [PubMed: 23021215]

40. Nie Z, et al. c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells. Cell. 2012; 151:68–79. [PubMed: 23021216]
41. Frye M, Fisher AG, Watt FM. Epidermal Stem Cells Are Defined by Global Histone Modifications that Are Altered by Myc-Induced Differentiation. PLoS ONE. 2007; 2:e763. [PubMed: 17712411]
42. Beck DB, Oda H, Shen SS, Reinberg D. PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. Genes & Development. 2012; 26:325–337. [PubMed: 22345514]
43. Kuo AJ, et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. Nature. 2012; 484:115–119. [PubMed: 22398447]
44. Driskell I, et al. The histone methyltransferase Setd8 acts in concert with c-Myc and is required to maintain skin. EMBO J. 2012; 31:616–629. [PubMed: 22117221]
45. Ji H, et al. Cell-Type Independent MYC Target Genes Reveal a Primordial Signature Involved in Biomass Accumulation. PLoS ONE. 2011; 6:e26057. [PubMed: 22039435]
46. Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell. 2009; 137:1194–1211. [PubMed: 19563753]
47. Chen Q, et al. Recurrent trisomy and Robertsonian translocation of chromosome 14 in murine iPS cell lines. Chromosome Research. 2011; 19:857–868. [PubMed: 22009222]
48. Guffei A, et al. c-Myc-dependent formation of Robertsonian translocation chromosomes in mouse cells. Neoplasia (New York, NY). 2007; 9:578–588.
49. Louis SF, et al. c-Myc induces chromosomal rearrangements through telomere and chromosome remodeling in the interphase nucleus. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:9613–9618. [PubMed: 15983382]
50. Wood AM, et al. Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. Mol Cell. 2011; 44:29–38. [PubMed: 21981916]
51. Van Bortle K, et al. Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. Genome Res. 2012; 22:2176–2187. [PubMed: 22722341]
52. Xu H, et al. A signal-noise model for significance analysis of ChIP-seq with negative control. Bioinformatics. 2010; 26:1199–1204. [PubMed: 20371496]

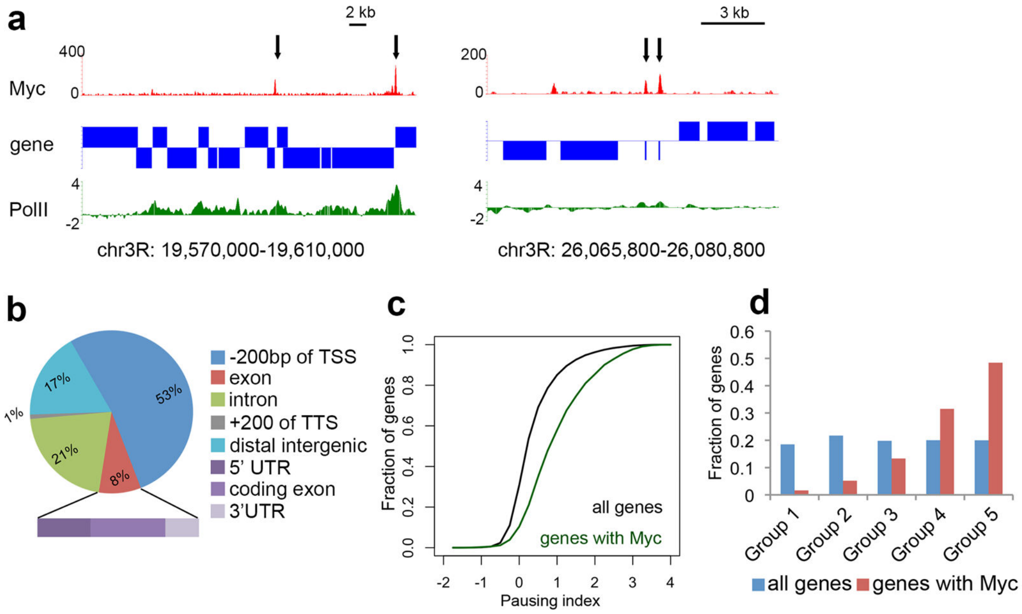

Figure 1.

Characteristics of Myc-associated genes. (a) Examples of Myc-associated genes. The signal for Myc is represented as the number of raw reads from ChIP-seq and the signal for RNAPII is represented by the log2 enrichment from ChIP-chip. On the gene track, the genes above the line are transcribed from the plus strand and the genes below the line are transcribed from the minus strand. The two arrows on the left point to two protein coding genes associated with Myc that also show high accumulation of RNAPII at the TSS. The two arrows on the right point to two non protein-coding microRNA genes associated with Myc. (b) Genome wide distribution of Myc binding sites with respect to various gene landmarks. Distal intergenic region means regions that are at least 200 bp away from genes. (c) Cumulative curve of pausing index for all coding genes in the genome (black) or coding genes associated with Myc (green). The two distributions are significantly different (K-S test, $p<2x\ 10^{-16}$). (d) Distribution of expression levels of Myc target genes. All genes in the genome were sorted according to their expression score and binned into five groups (Group 1 with the lowest expression and Group 5 with the highest expression). Myc target genes were assigned to one of the groups if their expression scores fall into the range of expression levels for that group.

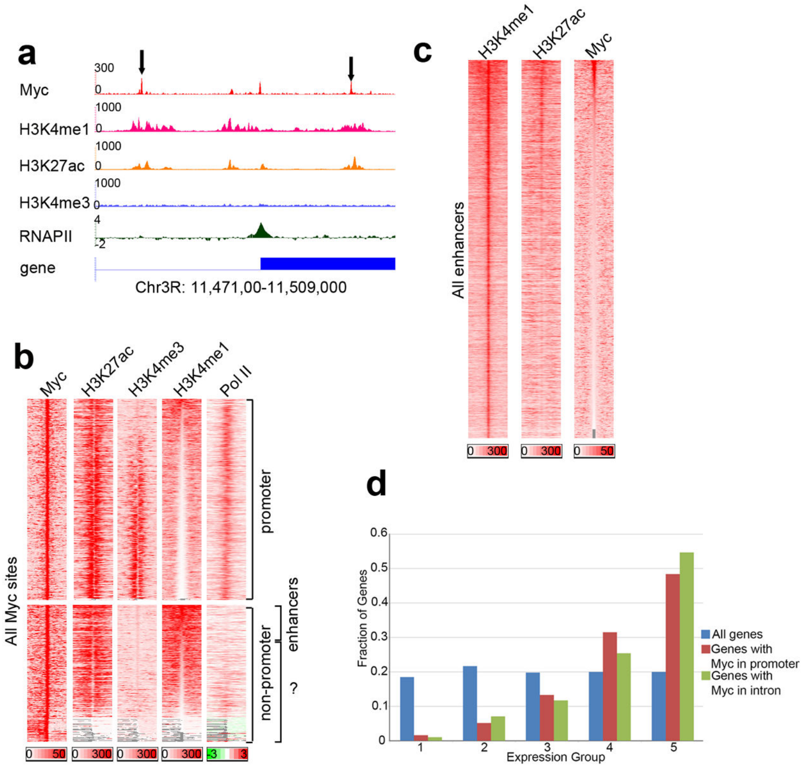

Figure 2.

A subset of Myc sites at non-promoter regions have characteristics of enhancers. (a) Examples of non-promoter Myc sites in the genome. The signals on the tracks of Myc, H3K4me1, H3K27ac and H3K4me3, are represented by the raw reads from ChIP-seq. The signal for RNAPII is represented by the log2 value from ChIP-chip. The arrows represent two Myc sites that display enhancer chromatin signatures (presence of H3K4me1/H3K27ac and absence of H3K4me3). (b) Heatmaps showing the chromatin features at promoter and non-promoter Myc sites. Each panel represents 2 kb upstream and downstream of the Myc sites. The sites are ordered by signal of H3K4me1. (c) Heatmaps showing chromatin features at all identified enhancers in Kc cells. The sites are ordered by signal of Myc. (d) Fraction of genes containing Myc at the promoter or enhancer expressed at different levels compared to all genes in the genome. Group 1 genes are expressed at low levels and Group 5 genes at the highest levels (see Methods).

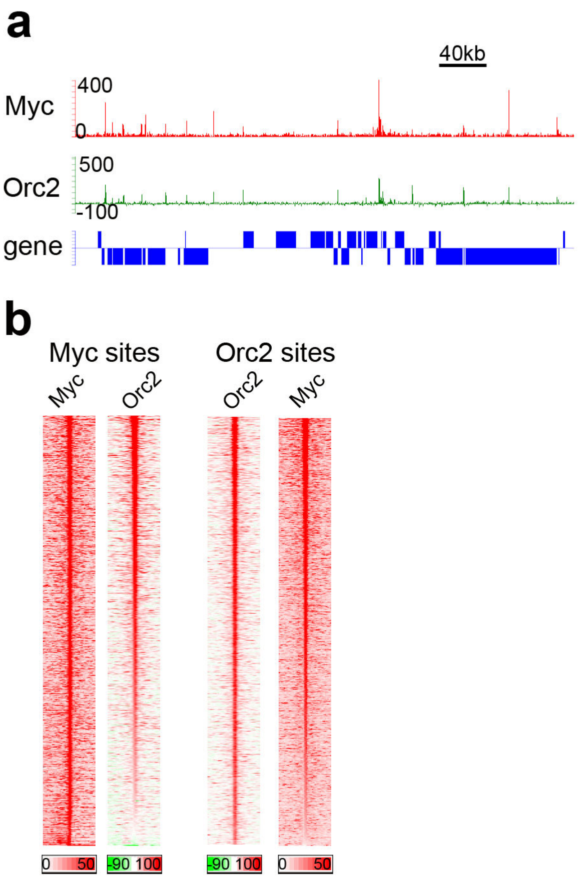

Figure 3.

Myc associates with Orc2 genome wide. (a) Snapshot of a region of the *Drosophila* genome showing the distribution of Orc2 sites compared with Myc. The signals represent the number of raw reads from ChIP-seq data sets across the region. (b) Heatmaps of Myc and Orc2 signal at all Myc and Orc2 binding sites. Each panel represents 2 kb upstream and downstream of the anchor sites. The two panels on the left are the signals at all Myc binding sites in Kc cells discovered in this study ordered by Orc2 signal intensity. The two panels on the right are the signals at all Orc2 binding sites obtained from modENCODE ordered by Myc signal.

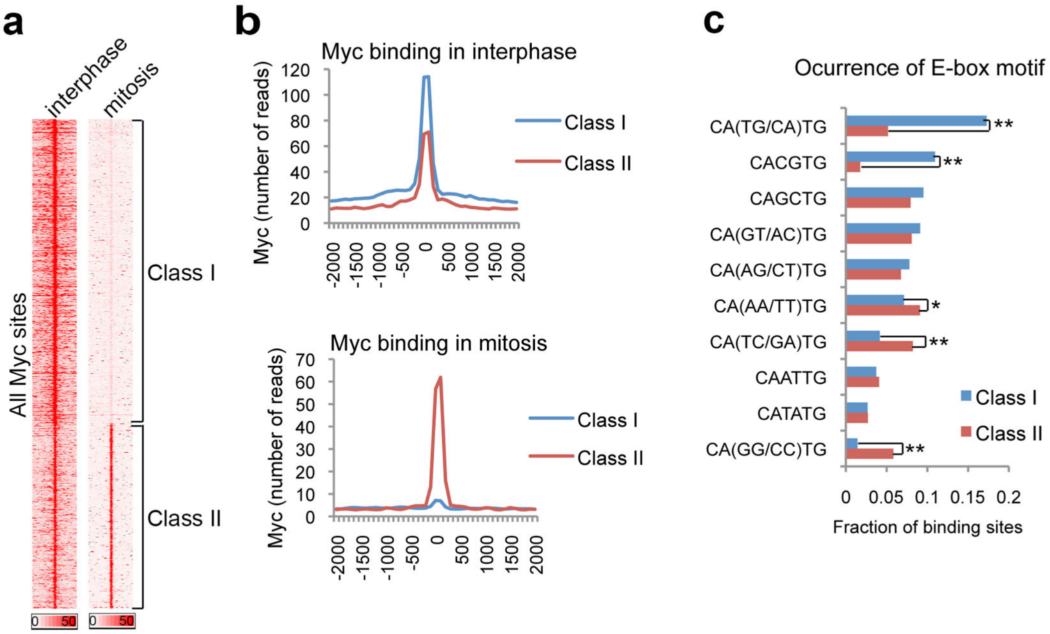

Figure 4.

Properties of Myc sites in interphase and mitotic chromosomes. (a) Heatmap showing signals of Myc in interphase or mitosis at all the Myc site in the genome. The information indicates the existence of two groups of Myc sites in the genome, one is interphase-specific (Class I) and the second one is common to interphase and mitosis (Class II). (b) Binding intensity at Myc sites during interphase or mitosis plotted from the information displayed in panel A. The X axis represents distance from Myc sites and ‘0’ is the summit of Myc sites. Negative values indicate upstream and positive values indicate downstream of the Myc sites. (c) Usage of different binding motifs by the Myc protein in interphase (Class I) or mitosis (Class II). The significance of this difference was tested by the chi square test for each motif (** p<0.01, * p<0.05).

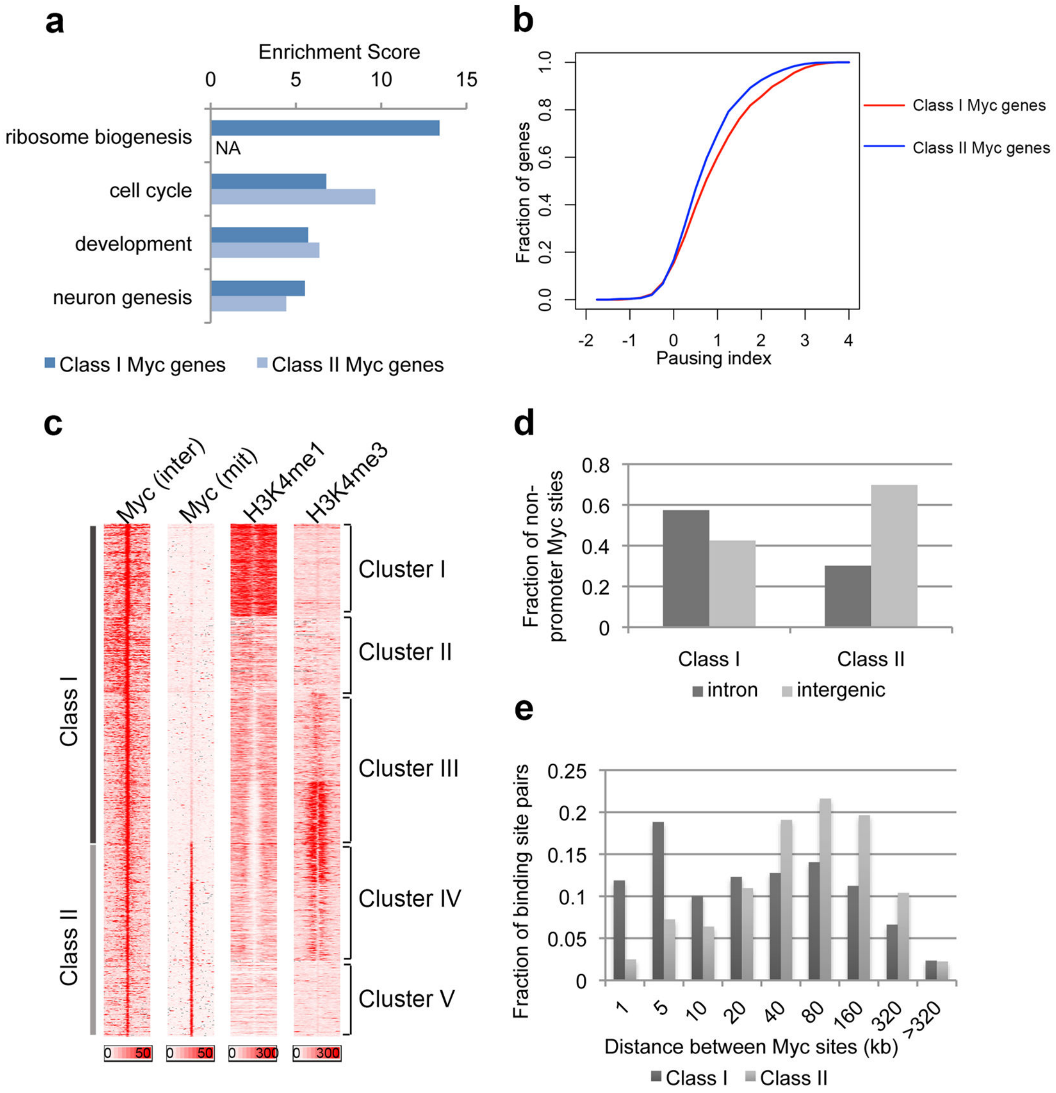

Figure 5.

Myc sites occupied during interphase and mitosis have different characteristics. (a) Gene ontology of protein coding genes associated with the two classes of Myc sites. (b) Cumulative curve of pausing index for protein coding genes associated with Class I (red) or Class II (blue) Myc sites. The two distributions are significantly different (K-S test, $p<2x\ 10^{-6}$). (c) Heatmaps showing chromatin features of Class I and Class II Myc sites. Each panel represents 2 kb upstream or downstream of the Myc sites. Clusters were created using Cluster 3.0 based on the signal value for the listed features at the Myc sites. (d) Distribution of Class I and Class II non-promoter Myc sites in introns or intergenic regions. (e) Distance between Myc site pairs for Class I and Class II Myc sites.

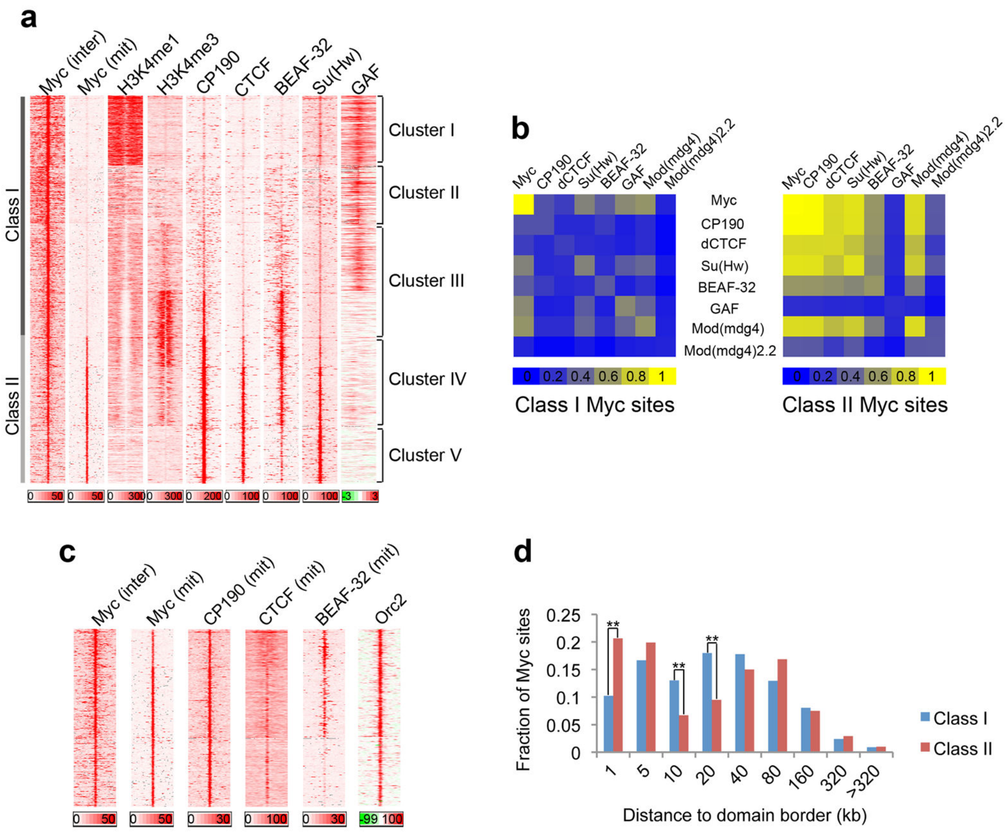

Figure 6.

Myc sites present in mitotic chromosomes associate with insulator proteins. (a) Heatmaps showing chromatin features at Class I and Class II Myc sites. Each panel represents 2 kb upstream or downstream of the Myc sites. Signals are represented by the number of raw reads from ChIP-seq or enrichment log2 value from ChIP-chip for the listed proteins or histone modifications. (b) Heatmaps showing the overlap of the listed protein at the two classes of Myc sites. (c) Insulator proteins and Myc persist on chromatin during mitosis at Class II Myc sites. (d) Distribution of Class I and Class II Myc sites with respect to topological domain boundaries. The significance was examined for each distance category by the chi square test (** p<0.01, * p<0.05).