MYC overexpression leads to increased chromatin interactions at superenhancers and MYC binding sites

Yi Xiang See, Kaijing Chen and Melissa J Fullwood

Genome Res. published online February 3, 2022
Access the most recent version at doi:10.1101/gr.276313.121

MYC overexpression leads to increased chromatin interactions at superenhancers and MYC binding sites

Yi Xiang See[1,2,3], Kaijing Chen[1,2,3], Melissa J. Fullwood[1,2,3,4*]

[1]School of Biological Sciences, Nanyang Technological University, Singapore.

[2]Cancer Science Institute of Singapore, National University of Singapore, Singapore.

[3]NUS Centre for Cancer Research, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

[4]Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A*STAR), Singapore.

* Corresponding author (mfullwood@ntu.edu.sg)

Running Title

MYC overexpression increases chromatin interactions

Keywords

MYC, cancer, chromatin interactions, superenhancers

Abstract

The *MYC* oncogene encodes for the MYC protein and is frequently dysregulated across multiple cancer cell types, making it an attractive target for cancer therapy. *MYC* overexpression leads to MYC binding at active enhancers, resulting in a global transcriptional amplification of active genes. Since superenhancers are frequently dysregulated in cancer, we hypothesized that MYC preferentially invades into superenhancers and alters the cancer genome organization. To that end, we performed ChIP-seq, RNA-seq, 4C-seq and SIQHiC (Spike-in Quantitative Hi-C) on the U2OS osteosarcoma cell line with tetracycline-inducible *MYC*. *MYC* overexpression in U2OS cells modulated histone acetylation and increased MYC binding at superenhancers. SIQHiC analysis revealed increased global chromatin contact frequency, particularly at chromatin interactions connecting MYC binding sites at promoters and enhancers. Immunofluorescence staining showed that MYC molecules formed punctate foci at these transcriptionally active domains after *MYC* overexpression. These results demonstrate the accumulation of overexpressed MYC at promoter-enhancer hubs and suggest that MYC invades into

enhancers through spatial proximity. At the same time, the increased protein-protein interactions may strengthen these chromatin interactions to increase chromatin contact frequency. *CTCF* siRNA knockdown in *MYC* overexpressed U2OS cells demonstrated that removal of architectural proteins can disperse MYC and abrogate the increase in chromatin contacts. By elucidating the chromatin landscape of MYC driven cancers, we can potentially target MYC associated chromatin interactions for cancer therapy.

Introduction

Dysregulation of the *MYC* oncogene, which encodes for the MYC transcription factor, is common in cancer. Elevated mRNA and protein expression of *MYC* is seen across most cancer types in The Cancer Genome Atlas (TCGA) dataset (Schaub et al. 2018), making it an attractive target for cancer therapy. However, there are a multitude of causes for *MYC* dysregulation, including focal amplification of the *MYC* gene (Beroukhim et al. 2010), aberrations in the web of oncogenic signalling pathways regulating *MYC* (Kress et al. 2015), and the acquisition of new enhancer regulatory elements in spatial proximity to the *MYC* gene locus through deletions and translocations (Affer et al. 2014). The MYC protein itself is notoriously termed as “undruggable”, because of the unstructured nature of the protein and the lack of enzymatic binding sites or prominent pockets for small molecule inhibitor binding (McKeown and Bradner 2014).

MYC usually binds to the promoters of actively transcribing genes at physiological levels. Upon overexpression, MYC not only increases binding at active promoters, but “invades” into active enhancers as well, resulting in a global transcriptional amplification of active genes (Lin et al. 2012). Canonical MYC binding sites become saturated and MYC occupies non-canonical binding sites with lower affinity, resulting in an upregulation of genes associated with malignant transformation (Walz et al. 2014). Enhancer invasion has also been observed in the MYC family member MYCN (Zeid et al. 2018).

At present, the mechanism underlying MYC enhancer invasion has not been well studied. Enhancers recruit transcription factors to activate linearly distant gene promoters via long-range

chromatin interactions (Carter et al. 2002; Plank and Dean 2014). Recent studies have defined a category of enhancers known as superenhancers: clusters of enhancers with exceptionally high enrichment of transcriptional activators (Hnisz et al. 2013; Loven et al. 2013; Whyte et al. 2013). Superenhancers are associated with a more extensive network of chromatin interactions to multiple distant target genes (Cao et al. 2017), to serve as regulatory hubs to govern processes important to cell identity (Whyte et al. 2013). This led us to question whether MYC enhancer invasion might preferentially occur at superenhancers, and if this might alter the cancer genome architecture.

In this study, we aimed to investigate the immediate changes to the enhancer and chromatin interaction landscapes after *MYC* overexpression, particularly at superenhancers, using Spike-In Quantitative Hi-C (SIQHiC), a modified Hi-C approach that normalizes chromatin interactions by cell count.

Results

MYC overexpression leads to increased MYC binding at superenhancers

We profiled the MYC enhancer binding landscape changes associated with *MYC* overexpression using the U2OS osteosarcoma cell line inserted with a tetracycline-inducible *MYC* cassette, because these cells are not addicted to *MYC* expression and have been used in previous studies to investigate the effects of *MYC* overexpression (Walz et al. 2014; Lorenzin et al. 2016). Doxycycline treatment for 30 hours increased *MYC* gene expression by about 35 times compared to vehicle treatment, together with an increase in MYC protein expression (Supplemental Figure S1A-C). Hereafter, we refer to vehicle-treated and doxycycline-treated U2OS cells as Low *MYC* cells and High *MYC* cells respectively. RNA-seq identified 548 significantly upregulated and 842 significantly downregulated transcripts after *MYC* overexpression (Figure 1A), in line with previous observations by Walz et al. (2014). MYC is known to repress a subset of genes, for example through its association with MIZ1 (Walz et al. 2014). Although downregulated transcripts outnumbered upregulated transcripts (Figure 1B), only 12.7% of the downregulated transcripts were direct targets of MYC, indicating that direct repression was not the main mechanism for downregulation. Most of these

downregulated transcripts were lowly expressed, with reads per kilobase per million (RPKM) values below 1. Since *MYC* overexpression had been shown to increase total RNA content per cell (Loven et al. 2012), transcripts that were upregulated to a smaller extent than the increase in total RNA content would appear to be downregulated (Kress et al. 2015). In contrast to the downregulated transcripts, 24.5% of the upregulated transcript promoters were bound by MYC at endogenous *MYC* expression levels, and increased to 36.9% after *MYC* overexpression (Figure 1B).

We performed a gene set enrichment analysis (GSEA) to determine the pathways dysregulated by *MYC* overexpression. Previously published gene sets of *MYC*-regulated genes were significantly enriched in our dataset (Supplemental Figure S1D). Consistent with previous literature (Eilers and Eisenman 2008), *MYC* overexpression activated pathways involved in cell proliferation, including ribosome biogenesis, translation, mitochondrial biogenesis and splicing (Supplemental Figure S1E). *MYC* overexpression activated the MAPK14 pathway, which is implicated in osteoblast motility (Rodríguez-Carballo et al. 2016) (Supplemental Figure S1E), and downregulated cell adhesion genes (Supplemental Figure S1F), suggesting a shift towards a cell migration phenotype.

We performed H3K27ac ChIP-seq to identify active *cis*-regulatory elements (CREs) in Low *MYC* and High *MYC* cells. H3K27ac peaks within 2.5 kb of transcription start sites were labelled as active promoter peaks, while the remaining peaks were labelled as enhancer peaks. We identified superenhancers using a similar approach to ROSE (rank ordering of superenhancers) (Loven et al. 2013; Whyte et al. 2013). Briefly, enhancer peaks within 12.5 kb of each other were stitched together, and superenhancers were separated from typical enhancers based on their H3K27ac ChIP-seq signal (Methods, Supplemental Figure S2A). In total, we identified around 28,500 stitched enhancers in both Low *MYC* and High *MYC* cells, of which around 7,300 stitched enhancers were lost and gained after *MYC* overexpression (Supplemental Figure S2C). We identified 890 superenhancers in the Low *MYC* cells and 725 superenhancers in the High *MYC* cells (Supplemental Figure S2E). Although many stitched enhancers were gained or lost, almost all High *MYC* superenhancers overlapped with Low *MYC* stitched enhancers (Figure 1C), indicating that superenhancers remain highly acetylated and novel superenhancers were not activated after *MYC* overexpression. 62% of the High *MYC*

superenhancers gained between 1 and 13 H3K27ac peaks after *MYC* overexpression (Figure 1C), with 15% merging multiple Low *MYC* stitched enhancers (Supplemental Figure 2F), resulting in superenhancers with significantly more constituent H3K27ac peaks (Supplemental Figure 2G).

Next, we looked into MYC occupancy at CREs. *MYC* overexpression led to a loss of 1,797 MYC binding sites and a gain of 6,349 MYC binding sites, of which 3,148 binding sites were gained at promoters and 1,665 at stitched enhancers (Figure 1D). The proportion of MYC binding sites at stitched enhancers increased from 17.6% to 21.4%, whereas the proportion of MYC binding sites at promoters remained similar (61.5% to 61.8%) (Supplemental Figure 3A), recapitulating previous observations of enhancer invasion. 21.9% of superenhancers overlap with MYC binding sites, increasing to 46.2% after *MYC* overexpression (Figure 1E). 26% of superenhancers gained MYC binding sites after *MYC* overexpression, compared to only 4% of typical enhancers (Figure 1F, Supplemental Figure S3B-C). H3K27ac ChIP-seq profiles of MYC binding sites show that MYC binds directly adjacent to constituent H3K27ac peaks and not randomly between the constituent peaks within the superenhancers (Supplemental Figure S3D). To test whether the preferential binding of MYC at superenhancers was due to size, we randomly shuffled the genomic intervals of CREs within the same chromosome and overlapped them with MYC binding sites. Significantly more actual superenhancers overlapped with MYC binding sites than 1000 iterations of randomly shuffled genomic loci (Figure 1E), indicating that MYC preferentially binds at superenhancers regardless of their size. Together, these results indicate that overexpressed MYC preferentially accumulates at H3K27ac peaks in superenhancers.

We further compared the MYC binding sites that were pre-existing, lost or gained after *MYC* overexpression. Pre-existing MYC binding sites at promoters and enhancers consistently had the highest H3K27ac and MYC ChIP-seq signals, while gained MYC binding sites had lower signals (Supplemental Figure S3E-G), indicating that open chromatin demarcated by H3K27ac is required for MYC binding. Lost promoter MYC binding sites had high H3K27ac signal comparable to pre-existing and gained MYC binding sites (Supplemental Figure S3G). This suggests that H3K27ac is required but insufficient for MYC binding at promoters, and other factors are involved in the

recruitment of MYC, such as binding site affinity and recruitment by other transcription factors (Lorenzin et al. 2016).

We performed motif analysis at MYC ChIP-seq peaks to ascertain whether there may be different chromatin factors recruiting MYC to promoters and enhancers. Promoter MYC binding sites were more enriched for other transcription factor binding motifs compared to typical enhancer and superenhancer MYC binding sites (Supplemental Figure S4A-B). 35 motifs occurred at more than 50% of promoter MYC binding sites, whereas no single motif occurred at more than 50% of enhancer MYC binding sites (Supplemental Figure S4A-B). Top enriched motifs at promoters included known promoter binding factors such as the KLF/SP family of transcription factors, which contain guanine/cytosine (G/C) rich sequences. Similarly, top enriched motifs at enhancers contained G/C rich sequences, such as MAZ, MIZ1 and EGR2. Chromatin interactions frequently connect CpG rich promoters and enhancers together (Pachano et al. 2021), and these interactions may be mediating the recruitment of MYC from promoter binding sites to enhancers. Gained MYC binding sites were slightly less enriched for transcription factor motifs than pre-existing MYC binding sites (Supplemental Figure S4F), while lost MYC binding sites were poorly enriched for transcription factor motifs, suggesting that MYC is likely recruited to these binding sites by other transcription factors (Supplemental Figure S4G). However, the same transcription factor motifs were enriched for in lost, pre-existing and gained MYC binding sites, indicating that recruitment of supraphysiological MYC to gained binding sites by new interactors is unlikely (Supplemental Figure S4E).

Spike-in Quantitative Hi-C (SIQHiC) reveals increased global chromatin contact frequency per cell after *MYC* overexpression

In traditional Hi-C techniques, cross-sample normalization is based on the assumption that chromatin interactions are largely stable across biological conditions (Lun and Smyth 2015; Stansfield et al. 2018). However, perturbation of factors involved in chromatin organization such as CTCF and cohesin can result in global changes in chromatin contact frequency. We hypothesized that increased global MYC binding at CREs not only increases transcription of associated genes but also stabilizes

chromatin interactions at these genes, resulting in similar global changes in chromatin contact frequency.

Cell-count normalization techniques have been developed for RNA-seq (Lovén et al. 2012) and ChIP-seq (Orlando et al. 2014) to account for global changes in transcription and occupancy respectively. Cell-count normalized transcription analyses had previously revealed a global increase in transcription after *MYC* overexpression (Loven et al. 2012; Nie et al. 2012). These methods involve mixing a control sample from an orthogonal species into the experimental human samples at a fixed cell number ratio. The recently published AQuA-HiChIP protocol (Absolute Quantification of Architecture Hi-ChIP) (Gryder et al. 2019) utilizes the same strategy of spiking in mouse cells into human samples, running on the assumption that contact frequency for untreated mouse cells should be the same across samples.

In our study, we adapted the AQuA-HiChIP protocol (Gryder et al. 2019) to perform Spike-In Quantitative Hi-C (SIQHiC) on Low *MYC* and High *MYC* cells in duplicate, to elucidate the global changes in chromatin contact frequency. In brief, mouse 3T3 cells were cross-linked and mixed into each sample of human cells at a ratio of 1:4, before continuing with the Arima Hi-C kit protocol (Figure 2A, Methods), preserving this cell number ratio throughout all subsequent steps of library preparation. If global chromatin contact frequency is unchanged, we expect to obtain a similar human to mouse chromatin contact ratio (HMR) across all samples. Hence, the relative difference between the HMR of different samples reflects the differences in global chromatin contact frequency and can be used to normalize the samples (Figure 2B).

Ambiguous read pairs mapping to both species consistently made up a small percentage of the total reads per sample (~0.3%) and were discarded (Supplemental Table S1). Using SIQHiC, we observe a 2.1 to 4.4 fold increase in HMR after *MYC* overexpression, suggesting an equivalent increase in global chromatin contacts (Supplemental Table S2). Hi-C chromatin interaction heatmaps of Low *MYC* and High *MYC* cells appear to show a decrease in chromatin interactions after *MYC* overexpression (Figure 2C, left). After SIQHiC normalization, we observed an increase in chromatin

interactions instead (Figure 2D, right). As expected, Low *MYC* duplicates had similar HMR (33.4 and 36.4), reflecting consistent chromatin contact frequencies at physiological *MYC* levels (Supplemental Table S2). Conversely, High *MYC* duplicates had higher and more variable HMR (146.9 and 78.1) after doxycycline induction of *MYC* expression. (Supplemental Table S2).

MYC overexpression weakens TAD insulation and strengthens a set of chromatin loops

Given that *MYC* overexpression increased chromatin contact frequency, we wanted to know whether *MYC* overexpression reshapes the global chromatin landscape. A/B compartment analysis showed few instances of compartment switching between Low *MYC* and High *MYC* cells, with Chromosome 8 shown as an example in Supplemental Figure S5A. However, *MYC* overexpression weakened insulation at topologically associating domain (TAD) boundaries (Supplemental Figure S5B), resulting in a loss of 2,912 TAD boundaries (Supplemental Figure S5C). Lost and common TAD boundaries showed similar loss of insulation, as seen by the reduced amplitude of TAD separation score, while insulation at gained boundaries remained relatively unchanged (Supplemental Figure S5B). This indicates that inactive compartments were not activated by *MYC* overexpression, but chromatin contact frequency increased interactions across TAD boundaries.

Next, we identified 7,266 and 7,910 significant chromatin loops in Low *MYC* and High *MYC* cells respectively, with approximately 60% of these loops common between both conditions (Supplemental Figure S5D). We performed an Aggregate Plot Analysis (APA) to show the aggregate signal of the chromatin loops that were “gained” and “lost” after *MYC* overexpression, or “common” between both conditions. In this analysis, the 105kb x 105kb contact matrices surrounding the midpoints of all loops in each loop set are overlapped such that the center pixel (P) of the APA plot shows the aggregate signal of the midpoints of these loops. The lower left corner of the plot (LL) represents the contact frequency for shorter interactions in the vicinity of the loop set and gives an indication of the local random interaction frequencies. Without SIQHiC normalization, APA plots show decreased peak signal enrichment (P) at chromatin loops after *MYC* overexpression (Supplemental Figure S6A-B). Using SIQHiC to normalize for cell count revealed a 2.2 and 3.3 fold increase in peak signal enrichment (P) at common and gained loop sets instead (Supplemental Figure

S6C-D). Using the ratio of signal at P to the average signal at LL (P2LL) to normalize peak signals against local background interactions, we observed a smaller increase in chromatin contact frequency at common and gained loop sets compared to SIQHiC normalization (1.4 fold and 2.1 fold respectively) (Figure 2D-E), indicating that *MYC* overexpression increased local random interactions but to a lesser extent than loop interactions. Taken together, APA analysis showed that *MYC* overexpression increased global chromatin contact frequency as a whole, but strengthened these chromatin loops in particular.

We zoomed in on the *MYC* gene locus since *MYC* is known to be regulated by superenhancers through chromatin interactions (Shi et al. 2013; Herranz et al. 2014; Zhang et al. 2016; Bahr et al. 2018). We generated virtual 4C plots from our Hi-C data using the *MYC* gene promoter as the viewpoint, to visualize the chromatin interactions at this locus. Without normalization, virtual 4C showed a general decrease in chromatin interaction frequencies (Figure 3A). However, SIQHiC normalization of the Hi-C reads revealed a general increase in virtual 4C signal instead (Figure 3A).

Since Hi-C assays the entire genome, it requires very high sequencing depth for sufficient resolution and limits chromatin loop detection within a specific region of interest (Babu and Fullwood 2015; Sati and Cavalli 2017). Hence, we performed 4C-seq at a randomly selected MYC bound superenhancer near the *PROC* gene which gained a chromatin loop to the *CYP27C1* gene after *MYC* overexpression. 4C-seq showed that the gained Hi-C chromatin loop was already present in Low *MYC* cells, but *MYC* overexpression increased the interaction frequency of this loop (Figure 3B). 4C-seq also identified additional chromatin loops at the MYC bound superenhancer locus that were not significant in the Hi-C data (Figure 3B, blue arrows). Taken together, *MYC* overexpression strengthens a set of gained chromatin loops.

SIQHiC normalization shows reduced chromatin contact frequency after *CTCF* siRNA knockdown

To validate our SIQHiC normalization approach, we transfected High *MYC* cells with either small interfering RNA (siRNA) targeting *CTCF* (siCTCF) or non-targeting siRNA (siControl) and

performed SIQHiC in duplicate. siCTCF knockdown reduced *CTCF* expression by 80% compared to siControl (Supplemental Figure S7A). SIQHiC normalization showed that global chromatin contact frequency in siCTCF cells was reduced by 55% compared to siControl cells (Supplemental Tables S3 and S4). 74.3% (4,633 / 6,236) of TAD boundaries from siControl cells remained after *CTCF* siRNA knockdown (Supplemental Figure S7B), but 47.1% (5,624 / 11,929) of chromatin loops were lost (Supplemental Figure S7C). TAD boundaries lost after *CTCF* siRNA knockdown showed a decrease in insulation, but insulation at the common and gained TAD boundaries remained mostly unchanged (Supplemental Figure S7D). These results are similar to previous reports of insulation preservation at selected TAD boundaries when CTCF is incompletely depleted (Zuin et al. 2014; Khoury et al. 2020). Without normalization, siCTCF appeared to selectively decrease chromatin contact frequency at siControl-unique chromatin loops and increase chromatin contact frequency at a small set of siCTCF-unique loops, while contact frequency at common loops remained unchanged (Supplemental Figure S7C,E-F), similar to results from Zuin et al. (2014) and Kubo et al. (2021). After applying SIQHiC normalization, we observed a global decrease in chromatin contact frequency at all loop sets instead (Supplemental Figure S7G-H), indicating that the siCTCF unique loops were not being induced after CTCF depletion but rather reduced to a lesser extent compared to siControl unique and common loops. This demonstrates the necessity for cell-count normalization when perturbing chromatin architectural factors.

Chromatin loops tend to connect MYC binding sites at superenhancers

We overlapped all Hi-C loop anchors with MYC binding sites and found that the proportion of all chromatin loops with MYC binding sites increased from 21.8% to 32.5% after *MYC* overexpression (Figure 4A). Gained loops were particularly enriched in MYC binding, with 47.6% of these loops occupied by MYC in High *MYC* cells, which was considerably higher than the proportion of randomly shuffled chromatin loop coordinates overlapping with MYC binding sites (Figure 4A). Notably, more gained loops were occupied by MYC at physiological MYC levels (34.3%) compared to common (16.8%) and lost loops (14.3%) (Figure 4A). Additionally, 7.6% of the gained loops had MYC binding at both anchors at physiological MYC levels, and increased to 15.7% after *MYC*

overexpression (Figure 4B). APA analysis revealed a 1.54 fold increase in chromatin contact frequency at loops connecting MYC binding sites, compared to a 1.32 fold increase at loops with no MYC binding (Supplemental Figure S8A). Gene Ontology analysis showed that MYC bound genes at lost chromatin loops were associated with negative regulation of transcription (Supplemental Figure S8B). MYC bound genes at gained chromatin loops were associated with previously shown *MYC* dysregulated pathways such as transcription, translation and cell-cell adhesion (Supplemental Figure S1E-F, S8C). These results suggest that chromatin interactions connecting MYC bound genomic loci are preferentially strengthened after *MYC* overexpression.

Next, we wanted to know whether chromatin interactions are preferentially altered at the enhancers invaded by MYC. We annotated the Hi-C loops according to their overlap with promoters and stitched enhancers. *MYC* overexpression increased chromatin looping between promoters (11% to 20%), between stitched enhancers (13% to 18%) and between promoters and stitched enhancers (18% to 22%) (Figure 4C). In particular, the proportion of CRE loops within the lost loop set was similar to the common loop set, whereas the proportion of CRE loops within the gained loop set was markedly higher (Supplemental Figure S9A), showing that chromatin interactions are preferentially strengthened between CREs. Consistent with previous research (Cao et al. 2017), a markedly higher proportion of superenhancers (42.0 %) was associated with Hi-C loops compared to typical enhancers (12.7%) and promoters (15.0%), and *MYC* overexpression further increased the proportion of superenhancers with loops to 66.2% (Figure 4D). MYC bound CREs were more associated with chromatin loops compared to non MYC bound CREs, particularly after *MYC* overexpression (Supplemental Figure S9B). Seeing as superenhancers are larger than typical enhancers, we compared the constituent H3K27ac peaks within typical and superenhancers. Similarly, more MYC bound superenhancer constituent H3K27ac peaks overlapped with Hi-C loops (27.3% - 39.9%) compared to MYC bound typical enhancer constituents (13.2% - 21.2%) and promoter loci (15.4% - 28.0%) (Supplemental Figure S9B-C).

Since MYC binds at both promoters and enhancers, we looked at promoter-enhancer chromatin loops (P-E loops) in detail. We overlapped P-E loops with MYC binding sites and

categorized the loops as being bound proximal to the promoter loop anchor or bound at both loop anchors (Figure 5A). *MYC* overexpression increased the proportion of P-E loops with only proximal MYC binding sites from 16% to 26%, while P-E loops bound at both anchors increased from 3% to 16% (Supplemental Figure 9D). Although lost P-E loops gained MYC binding sites at the promoter proximal loop anchor after *MYC* overexpression (22% to 36%), few of these loci gained MYC at the distal anchor as well (3% to 8%) (Figure 5B). In contrast, gained P-E loops acquired more distal MYC binding sites (9% to 23%) (Figure 5B). Taken together, our results suggest that *MYC* overexpression preferentially strengthens chromatin interactions between MYC binding sites at promoters and enhancers.

MYC overexpression increased chromatin contact frequency regardless of transcription activity

Previous studies have linked transcription activity with the formation of chromatin interactions, possibly through cohesin binding and positioning (Busslinger et al. 2017; Isoda et al. 2017; Heinz et al. 2018) and *cis*-regulatory element mobility (Gu et al. 2018; Nagashima et al. 2019). Since *MYC* overexpression leads to a global increase in transcription, we wanted to know whether the observed increase in chromatin interactions are a result of this increase in transcriptional activity. Firstly, we compared significantly up- and downregulated transcripts ($P<0.05$, abs(beta) >1) with non-regulated transcripts ($P>0.95$, transcripts per million>1). Hi-C loops are not preferentially gained or lost at *MYC* regulated transcripts compared to *MYC* non-regulated transcripts (Supplemental Figure S10A). However, more MYC bound transcripts gained Hi-C loops (Supplemental Figure S10B) compared to transcripts without MYC binding (Supplemental Figure S10C), indicating that MYC binding contributed to chromatin looping regardless of differential transcript expression.

When transcripts were stratified into tertiles based on transcript expression, top tertile transcript promoters overlapped better with Hi-C loops compared to the bottom tertile (25.7% and 16.1% respectively), demonstrating that transcriptional activity positively correlates with chromatin looping (Supplemental Figure S11A). However, we found that MYC bound transcripts were more associated with gained Hi-C loops compared to transcripts without MYC, regardless of expression levels (Figure 5C-D). Transcripts with pre-existing or gained MYC binding after *MYC* overexpression

were more associated with gained Hi-C loops regardless of expression levels (Supplemental Figure S11B-C), indicating that the direct binding of MYC increases chromatin contact frequency regardless of the level of transcriptional activity.

In contrast, among transcripts that lost MYC binding after *MYC* overexpression, bottom and middle tertile transcripts more associated with lost Hi-C loops and less associated with gained Hi-C loops, whereas top tertile transcripts were associated with gained Hi-C loops (Supplemental Figure S11D), demonstrating the contribution of transcriptional activity to chromatin contact frequency when MYC is not present. Together, these results show that transcriptional activity correlates with chromatin interactions but the direct binding of MYC also increases chromatin contact frequency through other mechanisms.

Chromatin contacts can be disrupted by removing CTCF

Since chromatin interactions were connecting MYC binding sites together at these regions of high transcriptional activity, we wanted to know whether the overexpressed MYC were being spatially constrained at these regions or merely diffused throughout the nucleus. Immunofluorescence staining showed that MYC formed discrete punctate spots within the nucleus after *MYC* overexpression, overlapping regions with high RNA polymerase II signal (Supplemental Figure S12). These punctate spots were lost after *CTCF* siRNA knockdown, indicating that these foci can be disrupted by removing chromatin architectural factors (Supplemental Figure S12).

To confirm this, we looked at the siCTCF-mediated changes in chromatin interaction frequencies at the common and gained loops after *MYC* overexpression. SIQHiC-normalized APA analysis showed similar reductions in chromatin contact frequency at common and gained loops, which was not evident in the non-normalized APA plots (Figure 6A-D). SIQHiC normalized virtual 4C plots of interactions at the *MYC* promoter illustrated the decrease in chromatin contact frequency (Figure 6E), recapitulating previous observations in CTCF-depleted SEM cells (Hyle et al. 2019). Although CTCF depletion reduced *MYC* expression in SEM cells (Hyle et al. 2019), *MYC* was not downregulated in siCTCF cells because of exogenous *MYC* expression from the tetracycline-inducible

system (Supplemental Figure S7A). While maintaining high *MYC* expression, we show that reduction of chromatin architectural proteins such as CTCF can reduce chromatin interactions at MYC foci and disperse MYC. Taken together, we propose a model where MYC accumulates at canonical promoter binding sites, binds to non-canonical binding sites at spatially adjacent superenhancers, and forms protein-protein interactions with other transcription factors to stabilize chromatin interactions at these domains (Figure 6F).

Discussion

MYC overexpression has long been described as a key driving force for oncogenesis in multiple cell types, resulting in tumours that become highly dependent on elevated *MYC* expression (Gabay et al. 2014; Bradner et al. 2017), but it is still unclear how overexpressed MYC shifts from regulating its canonical target genes to activate oncogenes. Recent studies have shown that overexpressed MYC invades into distal enhancer regulatory elements to differentially regulate gene expression (Lin et al. 2012; Nie et al. 2012; Sabo et al. 2014). Since cancer cells frequently acquire novel, cancer-specific superenhancer signatures that are associated with cancer initiation and maintenance (Ooi et al. 2016; Tsang et al. 2019; Raisner et al. 2020), we wondered whether MYC binds preferentially to superenhancers to perform its oncogenic functions. In this paper, we use the U2OS cell line with a tetracycline inducible MYC system, which has been widely used for the study of transcriptional regulation by MYC (Walz et al. 2014; Lorenzin et al. 2016; de Pretis et al. 2017; Nie et al. 2020) and is a reasonable model for establishing a parsimonious understanding of MYC enhancer invasion. We show that overexpressed MYC indeed invades into superenhancers in particular and leads to changes in the enhancer landscape and chromatin interactome. In particular, superenhancers gained more constituent H3K27ac peaks and increased enrichment of MYC binding after *MYC* overexpression.

Although there was a loss and gain of more than 7000 stitched enhancers after *MYC* overexpression, we observed no *de novo* superenhancer formation at H3K27ac unmarked loci. This is consistent with results from Poli et al. (2018) using TERT-immortalized human mammary epithelial

(IMEC) cells, where cancer-associated *de novo* enhancers were not activated in *MYC* overexpressed IMEC cells but marked by H3K27ac only after deriving secondary mammospheres. Earlier studies have shown that MYC does not bind to closed chromatin (Guccione et al. 2006; Lin et al. 2012; Soufi et al. 2012), suggesting that MYC invades into active superenhancers but is unable to activate silent oncogenic superenhancers by itself.

Since superenhancers are associated with abundant long-range chromatin interactions and high interaction frequencies (Schmitt et al. 2016; Beagrie et al. 2017; Cao et al. 2017), we wanted to know whether MYC superenhancer invasion alters the three-dimensional chromatin interactome. Using SIQHiC - a modified Hi-C protocol normalizing for cell count between samples, we observed increased global chromatin contact frequency after *MYC* overexpression. Chromatin interactions at MYC binding sites were preferentially strengthened and connected MYC binding sites together, notably between promoters and enhancers. Kieffer-Kwon et al. (2017) previously showed that *MYC* upregulation during B cell activation is accompanied by a two-fold increase in chromatin loops, particularly the formation of B cell associated chromatin loops. In this work, *MYC* overexpression alone strengthened pre-existing loops but did not increase the total number of chromatin loops, suggesting that other chromatin factors are required to reconfigure the chromatin interaction landscape in cancer as well as in B cell activation.

Previous studies have shown that *CTCF* siRNA knockdown does not deplete CTCF protein completely and leads to limited disruption of TAD boundaries (Zuin et al. 2014; Khoury et al. 2020), and this was recapitulated in our study. In contrast, degron-mediated CTCF protein degradation resulted in wide scale weakening of TAD insulation (Nora et al. 2017; Kubo et al. 2021). Although *CTCF* siRNA knockdown had limited effects on TAD structure, SIQHiC-normalization revealed a global decrease in chromatin contact frequency at all chromatin loop sets, demonstrating the importance of cell count normalization in analysing the perturbation of chromatin architectural factors.

Previous research has shown that active transcription may play a role in maintaining the three-dimensional genome organization (Busslinger et al. 2017; Isoda et al. 2017; Gu et al. 2018;

Heinz et al. 2018; Nagashima et al. 2019). Here, we show that transcriptional activity does correlate with chromatin interactions. Although *MYC* overexpression increases global transcription of active genes, we observed increased chromatin contact frequency regardless of transcription activity, indicating that direct binding of MYC contributes to chromatin interactions via other mechanisms.

MYC overexpression is a common event in oncogenesis across multiple cancers, but cannot alone induce malignant transformation. Our results suggest that other cancer initiating mutations are required to reconfigure the epigenetic and chromatin interaction landscape, with MYC playing a role in maintaining these epigenetic alterations by stabilizing these transient changes in chromatin interactions. These cancers would in turn become reliant on MYC and other participants of these chromatin interactions. If MYC expression is inhibited, we suspect that the reduction in protein-protein interactions will reduce these aberrant chromatin interactions.

Alternatively, these chromatin interactions can be perturbed by inhibiting other architectural factors involved. In this work, *CTCF* knockdown reduced chromatin contact frequency and dispersed MYC, supporting this line of investigation. A recent study has identified curaxins as a class of anti-cancer drugs that can disrupt chromatin looping by intercalating with DNA and interfering with CTCF binding (Kantidze et al. 2019), without causing DNA damage (Gasparian Alexander et al. 2011; Kantidze et al. 2019; Lu et al. 2021). Curaxins also exert anti-cancer activity through trapping of the histone chaperone FACT complex on chromatin (Gasparian Alexander et al. 2011). Crucially, both chromatin looping disruption and FACT complex trapping led to reduced expression of *MYC* family genes (Carter et al. 2015; Kantidze et al. 2019; Wang et al. 2020). Although complete degradation of CTCF is lethal in normal cells (Moore et al. 2012; Kemp et al. 2014), curaxins were well tolerated in xenograft models across multiple cancer cell types (Dermawan et al. 2016; Kim et al. 2016; Barone et al. 2017) and recently completed phase I clinical trials (Sarantopoulos et al. 2020), making it a promising drug against *MYC*-addicted cancers.

Taken together, our manuscript has demonstrated that *MYC* overexpression leads to MYC invasion into superenhancers and SIQHiC uncovered an increase in chromatin interactions at these

regions. We have also shown the utility of SIQHiC in investigating chromatin architectural proteins such as CTCF. Our results lay the groundwork for further research to elucidate the role of MYC in maintaining cancer-specific chromatin interactions in established cancers, and ways to therapeutically target these chromatin interactions. We anticipate that SIQHiC will continue to be useful for uncovering changes in chromatin contact frequency when perturbing chromatin architectural factors and using drugs such as curaxins.

Methods

Cell lines

U2OS osteosarcoma cells (HTB-96, ATCC) with a doxycycline-inducible *MYC* system were kindly provided by Elmar Wolf's lab from Universität Würzburg, Germany. Cell lines were authenticated by STR profiling and mycoplasma testing was performed using the MycoAlert™ PLUS Mycoplasma Detection Kit (Lonza). U2OS cells were grown in DMEM supplemented with 10% tetracycline-free foetal bovine serum (Clontech), 100U/ml penicillin and 100µg/ml streptomycin. *MYC* expression was induced in U2OS cells by the addition of 1µg/ml doxycycline (Clontech) for 30h. *CTCF* siRNA knockdown was performed using ON-TARGETplus SMARTpool Human CTCF (10664) siRNA (Dharmacon, L-020165-00-0005). Cells were transfected with 5nM *CTCF* siRNA using Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to manufacturer's instructions, with medium replacement after 24 hours, and cells were harvested after 48 hours.

Protein extraction and Western blot

Cells were lysed in RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific) at 4°C for 30 minutes, with vortexing every 10 minutes. Cell lysate was centrifuged to remove cell debris. Protein concentration was measured using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). 50 µg of protein was loaded into 12% SDS-PAGE gel for electrophoresis and transferred onto PVDF membrane. The membrane was blocked with 5% non-fat milk in Tris Buffered Saline with 0.1% Tween 20 (TBST) for 1 hour at room temperature, followed by incubation with primary antibodies at 4°C overnight. The following primary antibodies were used: anti-MYC (abcam, ab32072, diluted

1:1000) and anti-vinculin (Sigma-Aldrich, V9131, diluted 1:200). The membrane was washed thrice with TBST, followed by incubation with either mouse (Cell Signalling Technologies, 7076, diluted 1:5000) or rabbit (Cell Signalling Technologies, 7074, diluted 1:5000) HRP-conjugated secondary antibodies at room temperature for 1 hour. After washing thrice with TBST, bands were imaged using Clarity Western ECL Substrate (Bio-Rad) on the ImageQuant™ LAS 500 (GE healthcare).

Immunofluorescence staining

Cells were grown on a coverslip, fixed in 4% formaldehyde for 15 minutes and permeabilized with 0.2% Triton X-100 in Tris Buffered Saline (TBST). Cells were then blocked with 1% Bovine Serum Albumin (BSA) in 0.2% TBST for 3 hours at room temperature and incubated with primary antibodies against MYC (abcam, ab32072) or RNA polymerase II (BioLegend, 664906) at room temperature for 2 hours. Coverslips were washed thrice with 0.2% TBST, incubated with anti-mouse secondary antibodies conjugated with Alexa Fluor 555 (Life Technologies, A21422) or anti-rabbit secondary antibodies conjugated with Alexa Fluor 488 (Life Technologies, A11034) at room temperature for 1 hour, followed by incubation with DAPI (Sigma-Aldrich) for 10 minutes. Coverslips were mounted onto slides and imaged on the Nikon A1R confocal microscope at 100X magnification.

Reverse transcription and quantitative polymerase chain reaction (RT-qPCR)

RNA was extracted from U2OS cells with or without *MYC* induction in triplicate, using the RNeasy Mini Kit (QIAGEN). RNA was reverse transcribed into complementary DNA using qScript cDNA Synthesis Kit (Quantabio). Quantitative PCR was performed on the QuantStudio 5 (Thermo Fisher Scientific) using the GoTaq qPCR Mastermix (Promega), with *TBP* as the housekeeping control. Primer sequences are listed in Supplemental Table S5.

Circular chromosome conformation capture (4C-seq) experimental procedure

4C-seq was performed in duplicate on U2OS cells with or without *MYC* induction as previously described (Cao et al. 2017). 40 million cells were cross-linked in PBS supplemented with 1% formaldehyde at room temperature for 10 minutes, followed by quenching with 0.25M glycine for 5

minutes. Cells were washed thrice with PBS supplemented with 1mM phenylmethylsulfonyl fluoride (Sigma-Aldrich). Nuclei were isolated by lysing cells in 4C Lysis Buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 5mM EDTA, 0.5% Igepal CA-630) at 4°C for 10 minutes, followed by homogenization using a dounce homogenizer (Wheaton) (50 strokes using Pestle B).

Nuclei were permeabilized at 37°C with the addition of 0.3% SDS for 1 hour, followed by 2% Triton-X 100 for 1 hour. Primary enzyme digestion with HindIII-HF (NEB) was done at 37°C for 18h, before proximity ligation in dilute conditions with T4 DNA Ligase (Thermo Fisher Scientific) at 16°C overnight. Crosslinking was reversed with the addition of 55µg/ml Proteinase K (Ambion) at 65°C for 4 hours and 37°C overnight, and treated with RNase A at 37°C for 1 hour. DNA was purified using phenol/chloroform extraction followed by ethanol precipitation to yield the 3C library. Secondary enzyme digestion with DpnII (NEB) was done on the 3C library at 37°C overnight, followed by proximity ligation and de-crosslinking as described above. 4C libraries were generated for each viewpoint through nested inverse-PCR using Phusion DNA Polymerase (Thermo Fisher Scientific), with primers listed in Supplemental Table S5. DNA fragments between 200 and 1000bp were isolated from the 4C libraries by gel excision after running the 4C libraries on a 4-20% gradient TBE gel (Thermo Fisher Scientific). The gel slices were shredded and macerated in Tris EDTA buffer at 37°C overnight and DNA was precipitated using ethanol. The 4C libraries were multiplexed and single-end 1 x 150bp sequencing was performed on the Illumina MiSeq.

RNA-seq experimental procedure

Total RNA was extracted from U2OS cells with or without *MYC* induction in duplicate, using the RNeasy Mini Kit (QIAGEN). Ribosomal RNA depletion and library preparation was performed using the TruSeq Stranded Total RNA LT Kit (Illumina) according to the manufacturer's protocol. Paired-end 2 x 100bp sequencing was performed on the Illumina HiSeq 2500.

ChIP-seq experimental procedure

H3K27ac and MYC ChIP-seq were performed in duplicate on U2OS cells with or without *MYC* induction. Cells were fixed in 1% formaldehyde and sonicated using the truChIP Chromatin Shearing

Kit (Covaris) on the ME220 Focused-Ultrasonicator (Covaris) according to the manufacturer's protocol. H3K27ac and MYC bound DNA were immunoprecipitated using anti-H3K27ac (abcam, ab4729, Lot: GR150367-1) and anti-MYC (Santa Cruz, sc-764, Lot: H1712), with anti-IgG (Santa Cruz, sc-2027, Lot: H2615) as the negative control. 3.5µg of each antibody was rotated with 15µl of Dynabeads Protein G magnetic beads (Invitrogen) at 4°C overnight and washed thrice with Beads Wash Buffer (0.1% Triton X-100 in PBS) to remove unbound antibodies. Antibody-bound beads were combined with sonicated chromatin from 5 million cells and rotated at 4°C overnight. Beads were then washed thrice with Shearing Buffer D3 (Covaris), once with High Salt Washing Buffer (50mM HEPES pH 7.5, 350mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS), once with Lithium Chloride Wash Buffer (10mM Tris pH 8.0, 250mM LiCl, 1mM EDTA, 0.5% NP-40, 0.5% Sodium Deoxycholate) and once with Tris-EDTA buffer. Chromatin was eluted from the magnetic beads in 100µl elution buffer (50mM Tris pH 8.0, 10mM EDTA, 1% SDS) supplemented with 2µl of 0.5mg/ml RNase A (QIAGEN) at 55°C for 2 hours. Chromatin was then decrosslinked with the addition of 2µl of 20mg/ml Proteinase K (Ambion) at 55°C for 4 hours and 37°C overnight. For total input, chromatin from 500,000 cells were treated with RNase A and decrosslinked similarly. DNA was purified using the MinElute PCR Purification Kit (QIAGEN). ChIP-seq libraries were prepared using the ThruPLEX DNA-seq Kit (Rubicon) and sequenced paired-end 2 x 100bp on the HiSeq 2500 (Illumina).

SIQHiC experimental procedure

SIQHiC was performed in duplicate on U2OS cells with or without *MYC* induction, with the addition of untreated mouse 3T3 cells. Human U2OS cells and mouse 3T3 cells were counted using the Countess II automated cell counter (Thermo Fisher Scientific) and fixed with 2% formaldehyde using the Arima Hi-C Kit (Arima). 1 million fixed mouse 3T3 cells were added to each sample of 4 million fixed human U2OS cells before subsequent steps of restriction enzyme digest, biotin end filling and ligation using the Arima-HiC kit (Arima Genomics) according to the manufacturer's protocol. Libraries were prepared using the KAPA Hyper-Prep Kit (KAPA), according to the Arima-HiC kit protocol. SIQHiC libraries were sequenced paired-end 2 x 150bp on the HiSeq 4000 (Illumina).

Genome assembly

4C-seq, ChIP-seq, RNA-seq and SIQHiC libraries were all mapped to the hg19 human genome assembly, in line with previous published datasets on MYC enhancer invasion that were analyzed using this assembly. We do not expect significant differences when the libraries are mapped to the newer GRCh38 genome assembly because a previous analysis of TCGA data had shown significant concordance between hg19 and GRCh38 versions (Gao et al. 2019).

4C-seq data processing

4C-seq libraries were mapped to the hg19 genome using BWA-MEM (Li 2013) version 0.7.5a-r405 using default settings and significant 4C interactions were identified using the r3Cseq pipeline ($q < 0.05$) (Thongjuea et al. 2013) on the CSI NGS portal (An et al. 2020) (https://csibioinfo.nus.edu.sg/).

ChIP-seq data processing

ChIP-seq libraries were mapped to the hg19 genome using BWA-MEM (Li 2013) version 0.7.5a-r405 using default settings. PCR duplicates were removed using SAMtools (Li et al. 2009) version 1.7 and reads falling within the ENCODE consensus blacklisted regions (Encode Project Consortium 2012) were removed using BEDTools (Quinlan and Hall 2010) version 2.26.0. ChIP-seq signals were visualized using deepTools (Ramírez et al. 2014) version 3.2.1 (bamCompare --normalizeUsing RPKM --operation subtract -bs 1). BAM files of biological replicates were merged before calling ChIP-seq peaks using MACS2 (Zhang et al. 2008) version 2.1.2. H3K27ac ChIP-seq peaks within 2kb of transcription start sites were labelled as active promoter peaks while the rest were labelled as enhancer peaks.

Superenhancers were called as described previously (Cao et al. 2017). Briefly, H3K27ac enhancer peaks within 12.5kb of each other were stitched together and ranked based on H3K27ac enrichment. The point where a line with slope 1 is tangential to the ranked H3K27ac enrichment curve was chosen as a cutoff to separate superenhancers from typical enhancers.

To identify differential H3K27ac peaks, BAM files from all replicates and conditions were merged and a common list of H3K27ac peaks was called using MACS2. Read counts for the common list of

H3K27ac peaks were obtained using Rsubread (Liao et al. 2019). Differential H3K27ac peaks were identified using DESeq2 (Love et al. 2014) version 1.24.0 (padj<0.01 and padj<0.01 respectively).

Motif analysis was performed to find the frequency of HOCOMOCO Human v11 Core motifs (Kulakovskiy et al. 2017) occurring at MYC binding sites using FIMO (Grant et al. 2011) with default parameters.

RNA-seq data processing

Total RNA-seq libraries were mapped to the hg19 genome and read counts for UCSC RefSeq transcripts were obtained using kallisto v0.44.0 (Bray et al. 2016). Differentially expressed transcripts were identified using sleuth v0.30.0 (Pimentel et al. 2017) (fdr<0.05, |beta|>1) in R (R Core Team 2021). Gene set enrichment analysis was performed using GSEA v4.1.0 (Subramanian et al. 2005). Gene ontology (GO) functional annotation was performed using DAVID (Huang da et al. 2009).

SIQHiC data processing

SIQHiC libraries were analyzed using the Juicer (v1.5) pipeline (Durand et al. 2016) with some modifications to obtain HIC matrix files. Since SIQHiC libraries include human and mouse DNA, paired reads were separated and mapped as single reads using BWA-MEM (Li 2013) to an artificial reference genome combining hg19 human and mm10 mouse genome sequences. Human and mouse chromosomes were appended with “H” and “M” respectively, to avoid chromosome name duplication. After mapping, reads were paired together again and split into 2 files. Paired reads both mapping to human chromosomes were placed together as Human-Human paired reads (H-H), while paired reads both mapping to mouse chromosomes were placed together as Mouse-Mouse paired reads (M-M). Ambiguous read pairs that separately mapped to different species are discarded. H-H and M-M reads were processed using Juicer separately, filtering out duplicates, intra-fragment reads and reads with MAPQ < 30 to generate HIC matrix files for each species. H-H reads of biological replicates were also merged and processed in the same way to generate HIC matrix files for downstream analyses.

The ratio between human and mouse contacts (HMR) for each sample was calculated. Since the mouse 3T3 cells in each sample come from the same population and were spiked into the human cells

at the same ratio of 1:4, we expect to obtain a similar HMR ratio across all samples. Hence, the relative difference between the HMR of different samples (SIQHiC ratio) reflects the changes in global chromatin contact frequency. The SIQHiC ratio was calculated such that samples with lower HMR are normalized against samples with the highest HMR.

Hi-C contact matrices from merged biological replicates were normalized using the SIQHiC ratio by adding a custom normalization vector to the original contact matrix HIC file using the Juicer 'addNorm' subroutine, where the magnitude of each bin was the SIQHiC ratio raised to the power of (-0.5). In this way, the Hi-C matrices are cell-count normalized to the sample with the highest HMR to prevent scale-up extrapolation errors.

The SIQHiC normalized matrix for each chromosome was extracted from the HIC file using the Juicer 'dump' subroutine, combined together, and finally reassembled into a SIQHiC normalized HIC matrix file using the Juicer 'pre' subroutine. Since the SIQHiC normalization vector for the High *MYC* Hi-C matrix is 1, SIQHiC normalized matrix is equivalent to the non-normalized matrix. Hence, only the Low *MYC* Hi-C matrix was SIQHiC normalized. Original and SIQHiC normalized HIC matrices were balanced using the Knight-Ruiz algorithm for subsequent analyses.

Aggregate peak analyses (APA) was performed using the Juicer 'apa' subroutine using both non-normalized and SIQHiC normalized contact matrices using the parameters "-k KR –u –n 30". APA heatmaps were generated using pheatmap v1.0.12 (Kolde 2019) in R (R Core Team 2021).

Compartments were identified using the Juicer 'eigenvector' and 'pearsons' subroutines at 100kb resolution. Since eigenvalues signs are arbitrary, we correlated eigenvectors with the number of gene promoters within each 100kb bin, such that gene rich regions were assigned positive eigenvalues.

Topologically associating domain (TAD) boundaries were identified using the HiCExplorer (Ramírez et al. 2018) 'hicFindTADs' subroutine using the parameter "--minDepth 50000 --maxDepth 500000 --thresholdComparisons 0.001 --delta 0.01 --correctForMultipleTesting fdr". Lost, common and gained TAD boundaries were identified by intersecting Low *MYC* and High *MYC* TAD boundaries. TAD

separation scores of lost, common and gained TAD boundaries were visualized using deepTools (Ramírez et al. 2014) version 3.2.1.

Loops were called using the Juicer 'hiccups' subroutine with parameter "-k KR -m 1024 --ignore_sparsity". Lost, common and gained loops were identified using pgltools (Greenwald et al. 2017) by merging and intersecting Low *MYC* loops with High *MYC* loops using a 10kb intersecting window at each loop anchor. Loop subsets are listed in Supplemental Table S5. Hi-C loops were annotated according to their overlap with MYC binding sites, promoters, typical enhancers and superenhancers using pgltools.

Virtual 4C tracks were extracted from the Hi-C matrices by using the Juicer 'dump' subroutine, dumping a 5kb window against its entire chromosome using a 5kb bin size.

Data Access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE164777.

Disclosure Declaration

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

M.J.F declares two patents on methodologies related to ChIA-PET. No other conflicts of interest are declared.

Acknowledgements

This research is supported by the RNA Biology Center at the Cancer Science Institute of Singapore, NUS, as part of funding under the Singapore Ministry of Education Academic Research Fund Tier 3 grant awarded to Daniel Tenen (MOE2014-T3-1-006). This research is supported by the NRF Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiative and a Singapore Ministry of Education Academic Research Fund Tier 2 grant awarded to M.J.F. (MOET2EP30120-0009).

Author contributions: Y.X.S. and M.J.F. contributed to the conception and design of the study. Y.X.S. conducted the experiments and developed the SIQHiC protocol. Y.X.S. and K.C. performed the bioinformatics analyses. Y.X.S. and M.J.F. reviewed the data and wrote the manuscript. All authors reviewed and approved the manuscript.

References

Affer M, Chesi M, Chen WD, Keats JJ, Demchenko YN, Tamizhmani K, Garbitt VM, Riggs DL, Brents LA, Roschke AV et al. 2014. Promiscuous MYC locus rearrangements hijack enhancers but mostly super-enhancers to dysregulate MYC expression in multiple myeloma. *Leukemia* **28**: 1725-1735.

An O, Tan KT, Li Y, Li J, Wu CS, Zhang B, Chen L, Yang H. 2020. CSI NGS Portal: An Online Platform for Automated NGS Data Analysis and Sharing. *Int J Mol Sci* **21**.

Babu D, Fullwood MJ. 2015. 3D genome organization in health and disease: emerging opportunities in cancer translational medicine. *Nucleus* **6**: 382-393.

Bahr C, von Paleske L, Uslu VV, Remeseiro S, Takayama N, Ng SW, Murison A, Langenfeld K, Petretich M, Scognamiglio R et al. 2018. A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. *Nature* **553**: 515-520.

Barone TA, Burkhart CA, Safina A, Haderski G, Gurova KV, Purmal AA, Gudkov AV, Plunkett RJ. 2017. Anticancer drug candidate CBL0137, which inhibits histone chaperone FACT, is efficacious in preclinical orthotopic models of temozolomide-responsive and -resistant glioblastoma. *Neuro Oncol* **19**: 186-196.

Beagrie RA, Scialdone A, Schueler M, Kraemer DC, Chotalia M, Xie SQ, Barbieri M, de Santiago I, Lavitas LM, Branco MR et al. 2017. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**: 519-524.

Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M et al. 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* **463**: 899-905.

Bradner JE, Hnisz D, Young RA. 2017. Transcriptional Addiction in Cancer. *Cell* **168**: 629-643.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**: 525-527.

Busslinger GA, Stocsits RR, van der Lelij P, Axelsson E, Tedeschi A, Galjart N, Peters JM. 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* **544**: 503-507.

Cao F, Fang Y, Tan HK, Goh Y, Choy JYH, Koh BTH, Hao Tan J, Bertin N, Ramadass A, Hunter E et al. 2017. Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. *Sci Rep* **7**.

Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P. 2002. Long-range chromatin regulatory interactions in vivo. *Nat Genet* **32**: 623-626.

Carter DR, Murray J, Cheung BB, Gamble L, Koach J, Tsang J, Sutton S, Kalla H, Syed S, Gifford AJ et al. 2015. Therapeutic targeting of the MYC signal by inhibition of histone chaperone FACT in neuroblastoma. *Science Translational Medicine* **7**: 312ra176-312ra176.

de Pretis S, Kress TR, Morelli MJ, Sabò A, Locarno C, Verrecchia A, Doni M, Campaner S, Amati B, Pelizzola M. 2017. Integrative analysis of RNA polymerase II and transcriptional dynamics upon MYC activation. *Genome research* **27**: 1658-1664.

Dermawan JKT, Hitomi M, Silver DJ, Wu Q, Sandlesh P, Sloan AE, Purmal AA, Gurova KV, Rich JN, Lathia JD et al. 2016. Pharmacological Targeting of the Histone Chaperone Complex FACT Preferentially Eliminates Glioblastoma Stem Cells and Prolongs Survival in Preclinical Models. *Cancer research* **76**: 2432.

Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. 2016. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems* **3**: 95-98.

Eilers M, Eisenman RN. 2008. Myc's broad reach. *Genes Dev* **22**: 2755-2766.

Encode Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.

Gabay M, Li Y, Felsher DW. 2014. MYC Activation Is a Hallmark of Cancer Initiation and Maintenance. *Cold Spring Harbor Perspectives in Medicine* **4**.

Gao GF, Parker JS, Reynolds SM, Silva TC, Wang L-B, Zhou W, Akbani R, Bailey M, Balu S, Berman BP et al. 2019. Before and After: Comparison of Legacy and Harmonized TCGA Genomic Data Commons' Data. *Cell Systems* **9**: 24-34.e10.

Gasparian Alexander V, Burkhart Catherine A, Purmal Andrei A, Brodsky L, Pal M, Saranadasa M, Bosykh Dmitry A, Commane M, Guryanova Olga A, Pal S et al. 2011. Curaxins: Anticancer Compounds That Simultaneously Suppress NF-κB and Activate p53 by Targeting FACT. *Science Translational Medicine* **3**: 95ra74-95ra74.

Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics (Oxford, England)* **27**: 1017-1018.

Greenwald WW, Li H, Smith EN, Benaglio P, Nariai N, Frazer KA. 2017. Pgltools: a genomic arithmetic tool suite for manipulation of Hi-C peak and other chromatin interaction data. *BMC Bioinformatics* **18**: 207.

Gryder BE, Pomella S, Sayers C, Wu XS, Song Y, Chiarella AM, Bagchi S, Chou H-C, Sinniah RS, Walton A et al. 2019. Histone hyperacetylation disrupts core gene regulatory architecture in rhabdomyosarcoma. *Nature Genetics* **51**: 1714-1722.

Gu B, Swigut T, Spencley A, Bauer MR, Chung M, Meyer T, Wysocka J. 2018. Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* **359**: 1050-1055.

Guccione E, Martinato F, Finocchiaro G, Luzi L, Tizzoni L, Dall' Olio V, Zardo G, Nervi C, Bernard L, Amati B. 2006. Myc-binding-site recognition in the human genome is determined by chromatin context. *Nature cell biology* **8**: 764-770.

Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L et al. 2018. Transcription Elongation Can Affect Genome 3D Structure. *Cell* **174**: 1522-1536.e1522.

Herranz D, Ambesi-Impiombato A, Palomero T, Schnell SA, Belver L, Wendorff AA, Xu L, Castillo-Martin M, Llobet-Navás D, Cordon-Cardo C et al. 2014. A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nature Medicine* **20**: 1130-1137.

Hnisz D, Abraham Brian J, Lee Tong I, Lau A, Saint-André V, Sigova Alla A, Hoke Heather A, Young Richard A. 2013. Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **155**: 934-947.

Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44-57.

Hyle J, Zhang Y, Wright S, Xu B, Shao Y, Easton J, Tian L, Feng R, Xu P, Li C. 2019. Acute depletion of CTCF directly affects MYC regulation through loss of enhancer-promoter looping. *Nucleic acids research* **47**: 6699-6713.

Isoda T, Moore AJ, He Z, Chandra V, Aida M, Denholtz M, Piet van Hamburg J, Fisch KM, Chang AN, Fahl SP et al. 2017. Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* **171**: 103-119.e118.

Kantidze OL, Luzhin AV, Nizovtseva EV, Safina A, Valieva ME, Golov AK, Velichko AK, Lyubitelev AV, Feofanov AV, Gurova KV et al. 2019. The anti-cancer drugs curaxins target spatial genome organization. *Nature Communications* **10**: 1441.

Kemp Christopher J, Moore James M, Moser R, Bernard B, Teater M, Smith Leslie E, Rabaia Natalia A, Gurley Kay E, Guinney J, Busch Stephanie E et al. 2014. CTCF Haploinsufficiency Destabilizes DNA Methylation and Predisposes to Cancer. *Cell Reports* **7**: 1020-1029.

Khoury A, Achinger-Kawecka J, Bert SA, Smith GC, French HJ, Luu P-L, Peters TJ, Du Q, Parry AJ, Valdes-Mora F et al. 2020. Constitutively bound CTCF sites maintain 3D chromatin architecture and long-range epigenetically regulated domains. *Nature Communications* **11**: 54.

Kieffer-Kwon K-R, Nimura K, Rao SSP, Xu J, Jung S, Pekowska A, Dose M, Stevens E, Mathe E, Dong P et al. 2017. Myc Regulates Chromatin Decompaction and Nuclear Architecture during B Cell Activation. *Molecular Cell* **67**: 566-578.e510.

Kim M, Neznanov N, Wilfong CD, Fleyshman DI, Purmal AA, Haderski G, Stanhope-Baker P, Burkhart CA, Gurova KV, Gudkov AV et al. 2016. Preclinical Validation of a Single-Treatment Infusion Modality That Can Eradicate Extremity Melanomas. *Cancer research* **76**: 6620.

Kolde R. 2019. pheatmap: Pretty Heatmaps.

Kress TR, Sabo A, Amati B. 2015. MYC: connecting selective transcriptional control to global RNA production. *Nat Rev Cancer* **15**: 593-607.

Kubo N, Ishii H, Xiong X, Bianco S, Meitinger F, Hu R, Hocker JD, Conte M, Gorkin D, Yu M et al. 2021. Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nature Structural & Molecular Biology* **28**: 152-161.

Kulakovskiy IV, Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, Medvedeva YA, Magana-Mora A, Bajic VB, Papatsenko DA et al. 2017. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Research* **46**: D252-D259.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *eprint arXiv:13033997*: arXiv:1303.3997.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**: 2078-2079.

Liao Y, Smyth GK, Shi W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* **47**: e47-e47.

Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA. 2012. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**: 56-67.

Lorenzin F, Benary U, Baluapuri A, Walz S, Jung LA, von Eyss B, Kisker C, Wolf J, Eilers M, Wolf E. 2016. Different promoter affinities account for specificity in MYC-dependent gene regulation. *eLife* **5**: e15161.

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

Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA. 2013. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**: 320-334.

Loven J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, Levens DL, Lee TI, Young RA. 2012. Revisiting global gene expression analysis. *Cell* **151**: 476-482.

Lovén J, Orlando David A, Sigova Alla A, Lin Charles Y, Rahl Peter B, Burge Christopher B, Levens David L, Lee Tong I, Young Richard A. 2012. Revisiting Global Gene Expression Analysis. *Cell* **151**: 476-482.

Lu K, Liu C, Liu Y, Luo A, Chen J, Lei Z, Kong J, Xiao X, Zhang S, Wang Y-Z et al. 2021. Curaxin-Induced DNA Topology Alterations Trigger the Distinct Binding Response of CTCF and FACT at the Single-Molecule Level. *Biochemistry* **60**: 494-499.

Lun ATL, Smyth GK. 2015. diffHic: a Bioconductor package to detect differential genomic interactions in Hi-C data. *BMC Bioinformatics* **16**: 258.

McKeown MR, Bradner JE. 2014. Therapeutic strategies to inhibit MYC. *Cold Spring Harb Perspect Med* **4**.

Moore JM, Rabaia NA, Smith LE, Fagerlie S, Gurley K, Loukinov D, Disteche CM, Collins SJ, Kemp CJ, Lobanenkov VV et al. 2012. Loss of Maternal CTCF Is Associated with Peri-Implantation Lethality of Ctcf Null Embryos. *PLOS ONE* **7**: e34915.

Nagashima R, Hibino K, Ashwin SS, Babokhov M, Fujishiro S, Imai R, Nozaki T, Tamura S, Tani T, Kimura H et al. 2019. Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *J Cell Biol* **218**: 1511-1530.

Nie Z, Guo C, Das SK, Chow CC, Batchelor E, Simons SSJ, Levens D. 2020. Dissecting transcriptional amplification by MYC. *eLife* **9**: e52483.

Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, Wang R, Green DR, Tessarollo L, Casellas R et al. 2012. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* **151**: 68-79.

Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**: 930-944.e922.

Ooi WF, Xing M, Xu C, Yao X, Ramlee MK, Lim MC, Cao F, Lim K, Babu D, Poon L-F et al. 2016. Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. *Nature Communications* **7**: 12983.

Orlando David A, Chen Mei W, Brown Victoria E, Solanki S, Choi Yoon J, Olson Eric R, Fritz Christian C, Bradner James E, Guenther Matthew G. 2014. Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. *Cell Reports* **9**: 1163-1170.

Pachano T, Sánchez-Gaya V, Ealo T, Mariner-Faulí M, Bleckwehl T, Asenjo HG, Respuela P, Cruz-Molina S, Muñoz-San Martín M, Haro E et al. 2021. Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nature Genetics* **53**: 1036-1049.

Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods* **14**: 687-690.

Plank Jennifer L, Dean A. 2014. Enhancer Function: Mechanistic and Genome-Wide Insights Come Together. *Molecular Cell* **55**: 5-14.

Poli V, Fagnocchi L, Fasciani A, Cherubini A, Mazzoleni S, Ferrillo S, Miluzio A, Gaudioso G, Vaira V, Turdo A et al. 2018. MYC-driven epigenetic reprogramming favors the onset of tumorigenesis by inducing a stem cell-like state. *Nature Communications* **9**: 1024.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* **26**: 841-842.

R Core Team. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Raisner R, Bainer R, Haverty PM, Benedetti KL, Gascoigne KE. 2020. Super-enhancer acquisition drives oncogene expression in triple negative breast cancer. *PLOS ONE* **15**: e0235343.

Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. 2018. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nature Communications* **9**: 189.

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research* **42**: W187-W191.

Rodríguez-Carballo E, Gámez B, Ventura F. 2016. p38 MAPK Signaling in Osteoblast Differentiation. *Frontiers in Cell and Developmental Biology* **4**.

Sabo A, Kress TR, Pelizzola M, de Pretis S, Gorski MM, Tesi A, Morelli MJ, Bora P, Doni M, Verrecchia A et al. 2014. Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* **511**: 488-492.

Sarantopoulos J, Mahalingam D, Sharma N, Iyer RV, Ma WW, Ahluwalia MS, Johnson S, Purmal A, Shpigotskaya P, Hards A et al. 2020. Results of a completed phase I trial of CBL0137 administered intravenously (IV) to patients (Pts) with advanced solid tumors. *Journal of Clinical Oncology* **38**: 3583-3583.

Sati S, Cavalli G. 2017. Chromosome conformation capture technologies and their impact in understanding genome function. *Chromosoma* **126**: 33-44.

Schaub FX, Dhankani V, Berger AC, Trivedi M, Richardson AB, Shaw R, Zhao W, Zhang X, Ventura A, Liu Y et al. 2018. Pan-cancer Alterations of the MYC Oncogene and Its Proximal Network across the Cancer Genome Atlas. *Cell Systems* **6**: 282-300.e282.

Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL et al. 2016. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep* **17**: 2042-2059.

Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP, Shen C, Roe J-S, Minder JL, Mercan F, Wang E, Eckersley-Maslin MA et al. 2013. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes & Development* **27**: 2648-2662.

Soufi A, Donahue G, Zaret KS. 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**: 994-1004.

Stansfield JC, Cresswell KG, Vladimirov VI, Dozmorov MG. 2018. HiCcompare: an R-package for joint normalization and comparison of HI-C datasets. *BMC Bioinformatics* **19**: 279.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**: 15545.

Thongjuea S, Stadhouders R, Grosveld FG, Soler E, Lenhard B. 2013. r3Cseq: an R/Bioconductor package for the discovery of long-range genomic interactions from chromosome conformation capture and next-generation sequencing data. *Nucleic Acids Research* **41**: e132-e132.

Tsang FH, Law CT, Tang TC, Cheng CL, Chin DW, Tam WV, Wei L, Wong CC, Ng IO, Wong CM. 2019. Aberrant Super-Enhancer Landscape in Human Hepatocellular Carcinoma. *Hepatology (Baltimore, Md)* **69**: 2502-2517.

Walz S, Lorenzin F, Morton J, Wiese KE, von Eyss B, Herold S, Rycak L, Dumay-Odelot H, Karim S, Bartkuhn M et al. 2014. Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. *Nature* **511**: 483-487.

Wang J, Sui Y, Li Q, Zhao Y, Dong X, Yang J, Liang Z, Han Y, Tang Y, Ma J. 2020. Effective inhibition of MYC-amplified group 3 medulloblastoma by FACT-targeted curaxin drug CBL0137. *Cell Death & Disease* **11**: 1029.

Whyte Warren A, Orlando David A, Hnisz D, Abraham Brian J, Lin Charles Y, Kagey Michael H, Rahl Peter B, Lee Tong I, Young Richard A. 2013. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **153**: 307-319.

Zeid R, Lawlor MA, Poon E, Reyes JM, Fulciniti M, Lopez MA, Scott TG, Nabet B, Erb MA, Winter GE et al. 2018. Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nature Genetics* **50**: 515-523.

Zhang X, Choi PS, Francis JM, Imielinski M, Watanabe H, Cherniack AD, Meyerson M. 2016. Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. *Nature Genetics* **48**: 176-182.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**: R137.

Zuin J, Dixon JR, van der Reijden MIJA, Ye Z, Kolovos P, Brouwer RWW, van de Corput MPC, van de Werken HJG, Knoch TA, van Ijcken WFJ et al. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 996-1001.

Figure Legends

Figure 1 | *MYC* overexpression leads to differential transcript expression, differential H3K27ac signal and increased MYC binding at superenhancers.

(**A**) Significantly upregulated (red) and downregulated (blue) transcripts after *MYC* overexpression (|beta|>1, FDR<0.05, Wald test). Top 10 significantly regulated transcripts are labelled. (**B**) Number of differentially expressed transcripts (DETs) bound by MYC in Low *MYC* and High *MYC* cells. (**C**) Volcano plot showing H3K27ac peaks with significant changes in H3K27ac signal. ($|\log_2 \text{foldchange}|>1$, FDR<0.01, Wald Test). (**D**) Bar graph and pie chart showing gain and loss of H3K27ac peaks at superenhancers after *MYC* overexpression. (**E**) Bar graph showing gain and loss of MYC binding sites at stitched enhancers and promoters after *MYC* overexpression. (**F**) Proportion of typical enhancers, superenhancers and promoters being bound by MYC compared to randomly shuffled coordinates. Proportion of MYC bound CREs are shown as orange dots. Box plots show 1000 iterations of MYC occupancy at random genomic loci of the same size and on the same chromosome as the actual CREs. (**G**) Pie charts showing gain and loss of MYC ChIP-seq peaks at typical enhancers (left) and superenhancers (right) after *MYC* overexpression.

Figure 2 | SIQHiC (Spike-In Quantitative Hi-C) normalization reveals increased chromatin contact frequency per cell after *MYC* overexpression.

(**A**) Brief overview of the SIQHiC workflow. (**B**) Cartoon illustrating Hi-C contacts before (left) and after (right) SIQHiC normalization. SIQHiC normalization scaled down the Low *MYC* contacts such that the number of mouse contacts in both conditions were the same, thereby revealing an increase in human chromatin contacts. (**C**) Hi-C matrix heatmaps of a region on Chromosome 1. Left panel: No normalization. Right panel: SIQHiC normalization. (**D**) Aggregate Peak Analysis (APA) plots at 5kb resolution showing the aggregate signal of “Lost”, “Common” and “Gained” chromatin loop sets in Low *MYC* and High *MYC* cells identified using the non-normalized Hi-C contact matrices. P: Peak signal at the centre pixel. LL: Average signal of the 3 x 3 square at the lower left corner of the APA plot, representing local background. P2LL: Ratio of P to LL. APA colour scales were normalized by

the LL signal. Loop sets were filtered to remove short loops near the diagonal (shown above each APA plot; numerator: number of filtered loops; denominator: total number of loops). (**E**) Fold change of P2LL ratio between High *MYC* and Low *MYC* cells.

Figure 3 | *MYC* overexpression increases chromatin contact frequency and chromatin loops at superenhancers.

(**A**) Genome browser view of High *MYC* and Low *MYC* cells at the *MYC* locus Tracks show Hi-C chromatin loops, virtual 4C signal of interactions at the *MYC* promoter (red and blue tracks respectively), H3K27ac and MYC ChIP-seq signal. The difference between the virtual 4C signals of High *MYC* and Low *MYC* cells are shown with and without SIQHiC normalization (green tracks). Superenhancers (red) and typical enhancers (black) are shown as bars below the H3K27ac ChIP-seq tracks. MYC binding peaks are shown as yellow bars below the MYC ChIP-seq tracks. (**B**) 4C-seq of chromatin interactions at a randomly selected MYC bound superenhancer near the *PROC* gene (highlighted in pink) on Chromosome 2. Tracks show Hi-C chromatin loops, 4C-seq signal (red and blue tracks), difference between 4C-seq signal of High *MYC* and Low *MYC* cells (green), H3K27ac and MYC ChIP-seq signal. Gained Hi-C chromatin loop is highlighted in orange. Blue arrows show additional gained chromatin interactions identified using 4C-seq.

Figure 4 | Chromatin interactions are enriched at superenhancers and gained chromatin loops tend to connect MYC binding sites.

(**A**) Percentage of Hi-C chromatin loops occupied by MYC compared to randomly shuffled loop coordinates. Percentage of MYC bound chromatin loops are shown as orange dots. Box plots show 1000 iterations of MYC occupancy at random genomic loci of the same size and on the same chromosome as the actual chromatin loops. (**B**) Percentage of chromatin loop anchors occupied by MYC. (**C**) Percentage of Hi-C chromatin loops connecting *cis*-regulatory elements. (**D**) Percentage of *cis*-regulatory elements with Hi-C chromatin loops.

Figure 5 | MYC is enriched at both anchors of gained promoter-enhancer chromatin loops regardless of transcriptional activity.

(**A**) Promoter-enhancer chromatin loops are categorized as having promoter-proximal MYC binding or MYC binding at both loop anchors. (**B**) Percentage of lost (left), common (middle) and gained (right) promoter-enhancer chromatin loops occupied by MYC at the promoter proximal anchor or at both anchors. (**C**-**D**) Percentage of (**C**) MYC bound and (**D**) non MYC bound transcripts with Hi-C chromatin loops. Transcripts were stratified into top (yellow), middle (not shown) and bottom (grey) tertiles based on transcript expression after *MYC* overexpression.

Figure 6 | Chromatin contacts can be disrupted through CTCF siRNA knockdown.

(A) siControl and siCTCF Aggregate Peak Analysis (APA) plots at 5kb resolution showing the non-normalized aggregate signal at "Common" and "Gained" chromatin loops previously identified in High *MYC* cells. P: Peak signal at the centre pixel. (**B**) Fold change of non-normalized peak signal enrichment (P) at common and gained loops between siControl and siCTCF cells. (**C**) Same APA analysis as in (**A**), but with SIQHiC normalization applied. (**D**) Fold change of SIQHiC-normalized peak signal enrichment (P) at common and gained loops between siControl and siCTCF cells. (**E**) Genome browser view showing siControl and siCTCF virtual 4C signal of interactions at the *MYC* promoter (red and blue tracks respectively). The difference between the siControl and siCTCF virtual 4C signals are shown with and without SIQHiC normalization (green tracks). (**F**) Proposed model of overexpressed MYC accumulating at canonical promoter binding sites, binding to spatially adjacent superenhancers and forming protein-protein interactions with other transcription factors to stabilize chromatin interactions within the domain.