SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells

A. Rasim Barutcu  
*University of Massachusetts Medical School*

Bryan R. Lajoie  
*University of Massachusetts Medical School*

Andrew J. Fritz  
*University of Vermont*

*See next page for additional authors*

Follow this and additional works at: http://escholarship.umassmed.edu/cellbiology_pp

Part of the Biochemistry Commons, Cell Biology Commons, Computational Biology Commons, Developmental Biology Commons, Genomics Commons, Structural Biology Commons, and the Systems Biology Commons

Repository Citation
Barutcu, A. Rasim; Lajoie, Bryan R.; Fritz, Andrew J.; McCord, Rachel P.; Nickerson, Jeffrey A.; van Wijnen, Andre J.; Lian, Jane B.; Stein, Janet L.; Dekker, Job; Stein, Gary S.; and Imbalzano, Anthony N., "SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells" (2016). *Cell and Developmental Biology Publications*. 197.  
http://escholarship.umassmed.edu/cellbiology_pp/197

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Cell and Developmental Biology Publications by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells

A. Rasim Barutcu,1 Bryan R. Lajoie,2 Andrew J. Fritz,3 Rachel P. McCord,4 Jeffrey A. Nickerson,1 Andre J. van Wijnen,5 Jane B. Lian,3 Janet L. Stein,3 Job Dekker,2,6 Gary S. Stein,3 and Anthony N. Imbalzano1

1 Department of Cell and Developmental Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA; 2 Program in Systems Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA; 3 Department of Biochemistry, University of Vermont College of Medicine, Burlington, Vermont 05405, USA; 4 Department of Biochemistry and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, USA; 5 Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, Minnesota 55905, USA; 6 Howard Hughes Medical Institute, Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA

The packaging of DNA into chromatin plays an important role in transcriptional regulation and nuclear processes. Brahma-related gene-1 SMARCA4 (also known as BRG1), the essential ATPase subunit of the mammalian SWI/SNF chromatin remodeling complex, uses the energy from ATP hydrolysis to disrupt nucleosomes at target regions. Although the transcriptional role of SMARCA4 at gene promoters is well-studied, less is known about its role in higher-order genome organization. SMARCA4 knockdown in human mammary epithelial MCF-10A cells resulted in 176 up-regulated genes, including many related to lipid and calcium metabolism, and 1292 down-regulated genes, some of which encode extracellular matrix (ECM) components that can exert mechanical forces and affect nuclear structure. ChIP-seq analysis of SMARCA4 localization and SMARCA4-bound super-enhancers demonstrated extensive binding at intergenic regions. Furthermore, Hi-C analysis showed extensive SMARCA4-mediated alterations in higher-order genome organization at multiple resolutions. First, SMARCA4 knockdown resulted in clustering of intra- and inter-subtelomeric loci, demonstrating a novel role for SMARCA4 in telomere organization. SMARCA4 binding was enriched at topologically associating domain (TAD) boundaries, and SMARCA4 knockdown resulted in weakening of TAD boundary strength. Taken together, these findings provide a dynamic view of SMARCA4-dependent changes in higher-order chromatin organization and gene expression, identifying SMARCA4 as a novel component of chromatin organization.

[Supplemental material is available for this article.]

Organization of chromatin is essential for many biological processes. Packaging of the DNA around the nucleosomes acts to tightly condense the genome (Cutter and Hayes 2015). At the same time, the cell has to regulate the accessibility of the chromatin to many enzymes for the regulation of gene expression, DNA replication, and repair. Maintaining a balance between tight packaging and accessibility of the chromatin is an important function of the eukaryotic nucleus. This balance is achieved by multiple specialized protein complexes that dynamically alter chromatin structure in an ATP-dependent manner (Clapier and Cairns 2009). Four families of ATP-dependent chromatin remodelers exist: SWI/SNF, ISWI, INO80, and CHD (for review, see Varga-Weisz 2001; Flaus and Owen-Hughes 2011). The ATPase subunits of each family have a conserved helicase-like ATPase domain that uses the energy from ATP hydrolysis to evict, reposition, or modify nucleosomes (Sala et al. 2011; Narlikar et al. 2013). Different families of remodelers work in cells in a dynamic and orchestrated way to fine tune DNA accessibility (Morris et al. 2014).

The mammalian mating-type switching and sucrose non-fermenting (SWI/SNF) complexes contain one of two distinct ATPase subunits, SMARCA2 (also known as Brahma [BRM]) or SMARCA4 (Muchardt and Yaniv 1993; Wang et al. 1996). SMARCA4 is thought to be dispensable, as SMARCA2 null mice can properly develop to adulthood (Reyes et al. 1998), although this finding has recently been questioned (Thompson et al. 2015). On the other hand, SMARCA4 is essential, as SMARCA4 null mice are embryonic lethal, and SMARCA4 heterozygous mice show developmental defects and are prone to mammary tumor formation (Bultman et al. 2000, 2008). SMARCA4 has been shown to be involved in many developmental processes and in transcriptional regulation, DNA repair, cell cycle control, and cancer (Trotter and Archer 2008; King et al. 2012). The role of SMARCA4 in gene regulation is contextual, as it can activate some promoters while repressing others. In addition, extensive dysregulation and mutations of SMARCA4 have been implicated...
in many different cancer types, making SMARCA4 a potential therapeutic target for cancer (Kadoch et al. 2013; Shain and Pollack 2013).

Through interactions with many different protein partners, SMARCA4 is involved in nuclear structure and in mediating specific long-range chromatin interactions (Trotter and Archer 2008; Euskirchen et al. 2011; Imbalzano et al. 2013a). The organization of chromatin occurs in a hierarchical manner. Chromosomes are positioned in distinct volumes forming the chromosome territories (Cremer et al. 2006), which consist of open (A-type) and closed (B-type) genomic compartments (Lieberman-Aiden et al. 2009). The genomic compartments are further folded into sub-megabase-scaled structures called topologically associating domains (TADs) (Dixon et al. 2012; Nora et al. 2012, 2013), where local looping interactions between promoters and enhancers occur (Sanyal et al. 2012; Symmons et al. 2014; Tang et al. 2015; Smith et al. 2016).

SMARCA4 regulates inter-chromosomal interactions between tissue-specific promoters during myogenesis (Harada et al. 2015) and is required for looping at many gene loci, including the beta and alpha-globin genes (Kim et al. 2009a,b), the Igf1 locus (Bossen et al. 2015), and the class II major histocompatibility complex gene (CIITA) (Ni et al. 2008). SMARCA4 binds to poised developmental enhancers in embryonic stem cells (Hu et al. 2011; Rada-Iglesias et al. 2011) and B-cells (Bossen et al. 2015) and colocalizes with pluripotency factors (Ho et al. 2009), suggesting important roles in enhancer function. Furthermore, previous work classifying genome-wide interactions according to their histone modifications and transcription factor binding revealed SMARCA4 enrichment at open chromatin regions, indicating a possible structural role for SMARCA4 (Lan et al. 2012). In addition, SMARCA4 knockdown affects nuclear size (Hill et al. 2004) and the integrity of nuclear shape via a mechanism independent of cytoskeletal connections (Imbalzano et al. 2013b). Recently, SMARCA4 was shown to be involved in the lncRNA-dependent assembly of nuclear bodies (Kawaguchi et al. 2015).

Taken together, apart from its chromatin remodeling activity at the regulatory regions of target genes, emerging evidence suggests a possible important role for SMARCA4 in maintaining the structural integrity of the nucleus by regulating global chromatin structure (Imbalzano et al. 2013a). To date, very little is known about the role of SMARCA4 in global higher-order genome organization. To gain insight into the role of SMARCA4 in nuclear organization at a genome-wide level, we performed RNA-seq and Hi-C seq data set (GSM935588) (The ENCODE Project Consortium 2012) and determined that 27% of all SMARCA4 bound sites are intergenic regions (Fig. 1D). A significant number of these genes were associated with cell adhesion, including many proteoglycans, integrins, and laminins. This result suggests a role for SMARCA4 in ECM biology and nuclear structure. On the other hand, up-regulated genes were associated with calcium signaling and lipid metabolism, including pathways related to “regulation of cholesterol biosynthesis by SREBP,” “fatty acids,” and “eicosanoids,” which are the byproducts of fatty acid oxidation (Fig. 1E).

Furthermore, we assessed the transcriptional changes of polyadenylated long noncoding RNAs (lncRNAs) using the GENCODE v19 lncRNA gene annotation (Trapnell et al. 2013). We identified 88 down-regulated and 64 up-regulated poly(A) lncRNAs upon SMARCA4 knockdown (Supplemental Table S2). This suggests a widespread role of SMARCA4 in the transcriptional regulation of both coding and noncoding genes. Although the majority of the differentially expressed lncRNAs were unannotated, we observed down-regulation of several well-known lncRNA genes, including XIST, NEAT1, and MALAT1, and up-regulation of the imprinted lncRNA H19. We validated the RNA-seq results by performing qRT-PCR on 23 coding and noncoding genes of interest. There was a significant correlation between qRT-PCR and RNA-seq results (Spearman’s ρ = 0.65, P = 4.17 × 10−8) (Supplemental Fig. S2C,D).

Taken together, these results suggest that SMARCA4 positively regulates genes associated with ECM, supporting the role for SMARCA4 in ECM biology (Xu et al. 2007; Stankunas et al. 2008; Saladi et al. 2010), and negatively regulates genes related to lipid metabolism and calcium signaling in mammary epithelial cells. Moreover, SMARCA4 also regulates the expression of many lncRNAs.

**SMARCA4 ChIP-seq analysis reveals extensive binding to intergenic and intronic regions in MCF-10A cells**

To confirm that SMARCA4-dependent alterations in gene expression observed by RNA-seq are directly related to SMARCA4 binding, we performed ChIP-seq analysis in wild-type MCF-10A cells (Fig. 2; Supplemental Fig. S3A,B). We identified 15,046 SMARCA4 ChIP-seq peaks. The binding profile of SMARCA4 demonstrated that SMARCA4 binds to defined locations in the genome (Fig. 2A). Annotation of SMARCA4 ChIP-seq peaks revealed that 60% of the binding sites were localized in promoters, introns, and exons (gene bodies), whereas 40% of the sites were bound to intergenic regions (Fig. 2B). Consistent with the annotation of SMARCA4 peaks, normalized SMARCA4 binding signal across all human genes was highest at promoter regions (Fig. 2C). Because SMARCA4 was enriched at the promoters, we intersected the SMARCA4 peaks with a publicly available MCF-10A Pol II ChIP-seq data set (GSM935588) (The ENCODE Project Consortium 2012) and determined that 27% of all SMARCA4 bound sites were also bound by Pol II (Supplemental Fig. S3C).

By performing de novo motif analysis on SMARCA4 peaks, we identified enriched motifs for MEF2A/C, USF2, SMAD2/4, TP53, and SP1 (Fig. 2D). These factors may be functionally related to SMARCA4.
because SMARCA4 is a coregulator for many of these factors. MEF2C has been shown to interact with SMARCA4 and is required for the activity-dependent recruitment of SMARCA4 to its target regions (Zhang et al. 2015). Moreover, USF1/2 can bind a SMARCA4-associated factor, SMARCD3, and recruit other BAF sub-units including SMARCA4 (Wang and Sul 1997). In addition,
Figure 2. (A) Example of a ChIP-seq genome browser view of SMARCA4 binding and the input control, the shSCRAM and shSMARCA4 RNA-seq on Chr 5, and a zoom-in on the VCAN (Versican) gene, which is regulated by SMARCA4, in the lower panel. The y-axis represents the normalized tag densities relative to hg19 genomic coordinates. (B) Distribution of SMARCA4 ChIP-seq peak annotation for genic and intergenic regions. (C) Normalized SMARCA4 ChIP-seq signal intensity plot for all human UCSC genes ±2 kb. SMARCA4 binding is enriched at the promoter regions. (D) Top five sequence motifs associated with the SMARCA4 peaks. (E) SMARCA4 peak density within ±20 kb of the TSS of significantly down-regulated (blue), or up-regulated genes (red). (F) Distribution of SMARCA4 ChIP-seq signal across the MCF-10A enhancers. SMARCA4 binding is not uniformly distributed across the enhancers, as 109 super-enhancers display higher (log$_2$ ~1.5-fold) levels of SMARCA4 binding. (G) SMARCA4 signal is greater over super-enhancers (red) than typical enhancers (green). (H) Distribution of SMARCA4-bound super-enhancers in genic and intergenic regions.

SMARCA4 regulates chromatin interactions
SMARCA4 is a direct interaction partner for SMAD proteins and TP53 (Xi et al. 2008; Naidu et al. 2009).

Next, we asked whether SMARCA4 binding was associated with differentially expressed genes. To address this, we analyzed the SMARCA4 peak frequency at the promoters of up- and down-regulated genes. We observed ~10-fold increase in the frequency of SMARCA4 binding at the promoters of genes that are down-regulated upon SMARCA4 knockdown, whereas there was minimal association of SMARCA4 with the promoters of genes that are up-regulated when SMARCA4 is knocked down (Fig. 2E). This result suggests a more direct role for SMARCA4 in the positive regulation of gene expression in MCF-10A cells.

Recently, super-enhancers, a novel type of regulatory regions having an unusual enrichment of transcription factors that are in close proximity, were described (Whyte et al. 2013). Super-enhancers are mostly associated with developmentally regulated genes. SMARCA4 is localized at super-enhancers in leukemic and in normal B-cells (Shi et al. 2013; Bossen et al. 2015). Using the approach previously published (Whyte et al. 2013), we identified 109 SMARCA4-bound super-enhancers in the MCF-10A genome (Fig. 2F). SMARCA4 signal intensity was greater at super-enhancers than at typical SMARCA4 ChIP-seq peak regions (Fig. 2G). Annotation of the super-enhancers revealed ~60% localization at intergenic regions and ~40% localization at gene bodies, but not at promoters (Fig. 2H).

Overall, we identified SMARCA4 binding mostly (60%) at promoters and gene bodies. In the promoter regions, there may be individual single (stand-alone) high-density SMARCA4 peaks, but since they are not proximal to other high-density SMARCA4 peaks, they are not defined as super-enhancers. This result may explain why the super-enhancers are mostly (60%) enriched at intergenic regions, whereas the typical SMARCA4 ChIP-seq peaks are mostly (60%) at promoters and gene bodies. Interestingly, the localization pattern of SMARCA4-bound super-enhancers is opposite of the annotation of typical SMARCA4 peaks (intergenic versus promoter) (Fig. 2B), suggesting differential regulatory functions for regular SMARCA4 peaks and SMARCA4-bound super-enhancers.

Hi-C analysis of SMARCA4 knockdown and control MCF-10A cells

To identify the genome-wide changes in higher-order chromatin structure following SMARCA4 knockdown, we performed Hi-C on doxycycline-induced shSCRAM and shSMARCA4 MCF-10A cells (Fig. 3A,B; Supplemental Table S3). With an average depth of approximately 115 million reads per biological replicate and with two biological replicates, we achieved up to 40-kb resolution genome-wide. The Hi-C heatmaps, which are normalized by total number of reads, revealed hierarchical higher-order chromatin structures, such as genomic compartments and TADs, at increasing resolutions (Fig. 3A,B). There was a high correlation ($R^2 > 0.90$) between two independent biological replicates at multiple resolutions (Supplemental Fig. S4A–J). The shSCRAM and shSMARCA4 Hi-C data sets exhibited high correlation ($R^2 > 0.95$) with the previously published wild-type MCF-10A Hi-C data set (Barutcu et al. 2015). The scaling plot curves of genomic interaction frequencies along genomic distance showed similar trends of decay and an increase at distances larger than ~200 Mb, which we previously showed to be a characteristic of the MCF-10A genome (Supplemental Fig. S4C; Barutcu et al. 2015). The Hi-C data sets displayed similar cis/trans interaction ratios (an average of ~25%) (Supplemental Fig. S4K,L).
Genome-wide Hi-C interaction heatmaps showed that, consistent with previous Hi-C studies and the notion of chromosome territories (Cremer et al. 2006), the intra-chromosomal interactions, which are visualized as dark boxes along the diagonal, were more frequent than inter-chromosomal interactions (Fig. 3). Moreover, we identified large blocks of inter-chromosomal interactions between Chr 3 and Chr 5, Chr 3 and Chr 9, and Chr 6 and Chr 19, which represent previously known translocations in the MCF-10A genome (Supplemental Fig. S5; Marella et al. 2009; Barutcu et al. 2015). Taken together, these analyses reflect the high quality and reproducibility of the Hi-C data.

**SMArcA4 knockdown results in extensive gain and loss of long-range chromatin interactions and altered telomeric associations**

To map the SMARCA4-mediated alterations in higher-order chromatin structure, we compared the genome-wide interactions as previously described (Fig. 4A; Barutcu et al. 2015; Crane et al. 2015). At multiple resolutions (10-Mb to 40-kb resolution genome-wide), SMARCA4 knockdown resulted in the disruption of existing interactions and the emergence of novel specific contacts throughout each chromosome (Fig. 4B; Supplemental Fig. S6). Mapping the high-confidence interactions that are depleted/enriched upon SMARCA4 knockdown revealed specific regions potentially important for transcriptional regulation (Fig. 4D,E). For instance, a zoom-in view of the differential interactions at the promotor regions of MALAT1 and NEAT1 lncRNAs, which are downregulated upon SMARCA4 knockdown, displayed several differentially interacting regions (Fig. 4D,E).

Although there were several alterations in long-range interactions at high resolution, the positioning of chromosomes relative to each other was similar following SMARCA4 knockdown (Supplemental Fig. S7A,B). At the chromosomal scale, when only the significant differential interactions were considered, SMARCA4 knockdown resulted in a higher frequency (Wilcoxon rank-sum test, \( P < 1 \times 10^{-16} \)) of *trans* interactions between the chromosomes (Supplemental Fig. S7C). Moreover, SMARCA4 knockdown resulted in a lower frequency (Wilcoxon rank-sum test, \( P = 1.2 \times 10^{-16} \)) of *cis* interactions both within and across the different chromosomal arms on each chromosome (Supplemental Fig. S7D).

Interestingly, among the subtelomeric regions of the chromosomes, we observed a systematic pattern of increased interactions in shSMARCA4 cells, both in *cis* and in *trans* (Fig. 4F). In other words, the subtelomeric regions of each chromosome displayed a striking enrichment of interactions with each other upon SMARCA4 knockdown compared with the shSCRAM control (Fig. 4F). Quantification of the subtelomeric interactions suggested a significant increase in both intra- and inter-chromosomal associations in shSMARCA4 cells compared to shSCRAM (Fig. 4G). However, this was not the case when interactions in the same regions were randomized (Supplemental Fig. S7E). To validate the Hi-C results, we performed DNA-FISH by probing the intra-chromosomal telomeric interactions of Chr 1 and Chr 4. Consistent with our Hi-C findings, SMARCA4 knockdown resulted in a significant increase in telomeric proximity of the p and the q chromosomal arms of Chr 1 and Chr 4 (Fig. 4H,I). Taken together, these results indicate a novel role for SMARCA4 in telomere structure and suggest that disruption of SMARCA4 levels results in altered three-dimensional organization of telomeric regions of the genome.

**SMArcA4 occupancy is enriched at open compartment regions**

Each chromosome territory is composed of megabase-scale genomic compartments that are either A-type (i.e., open, gene rich) or B-type (i.e., closed, gene poor). The frequency of interactions within one compartment is greater than the frequency of interactions between compartments (Lieberman-Aiden et al. 2009). We asked whether SMARCA4 knockdown resulted in any compartment changes. To address this, we binned the genome at 250-kb nonoverlapping intervals and compared the type of compartmentalization for each bin (Fig. 5A). The majority of the compartmentalization was similar in shSCRAM and shSMARCA4-MCF-10A cells, with 42% of the genome consisting of constitutive A-type compartments and 54% consisting of constitutive B-type compartments (Fig. 5B). Upon SMARCA4 knockdown, a total of 2% of the genome altered its compartmentalization from A-type to B-type and 2% showed alteration from B-type to A-type (Fig. 5B).

Compartmentalization of the genome is correlated with gene expression (Barutcu et al. 2015; Dixon et al. 2015). To understand the link between compartment switching and gene expression upon SMARCA4 knockdown, we plotted the shSMARCA4/shSCRAM log2 fold change RNA-seq expression levels of the genes that were located either within unchanged compartments or within compartments that switched (Fig. 5C). The genes located in regions with a compartment switch from A-type to B-type upon SMARCA4 knockdown showed significantly lower expression levels than the genes within unaltered compartment regions. The differences in gene expression within B-type to A-type compartment-switching regions (shSMARCA4 to shSCRAM), although statistically significant (Wilcoxon rank-sum test; \( P = 0.02 \)), showed much more similar expression levels upon SMARCA4 knockdown (Fig. 5C). Nevertheless, these results suggest a prominent correlation between differential compartmentalization and gene expression.

Next, we asked whether the regions with SMARCA4 binding were associated with compartment switching. Of all SMARCA4-bound sites, 76% were located within the constitutive open A-type compartments. In contrast, 21% of SMARCA4 ChIP-seq peaks were found in the closed B-type compartment regions (Fig. 5D), and ~3% of SMARCA4 peaks were in regions showing compartment-switching (Fig. 5D). We then assessed the percentage of genomic compartment-switching regions that were either bound or unbound by SMARCA4. We observed that SMARCA4 was bound to ~80% of constitutive A-type compartments (A to A) and ~50% of constitutive B-type compartments (B to B) (Fig. 5E). The frequency of SMARCA4 binding to altered compartment regions was similar, because 75% of “A to B” and 55% of “B to A” compartment-switching regions showed SMARCA4 binding, suggesting a similar degree of localization at compartment-switching regions. This suggests that SMARCA4 is therefore not likely to be directly mediating the compartment switch.

Taken together, these results show that SMARCA4 knockdown affects compartment organization in the genome. Although this change is confined to a subset of the genome, it is associated with differential gene expression.

**SMArcA4 is associated with TAD boundaries and TAD boundary strength**

Each compartment is composed of TADs, which are sub-megabase-scale structures constituting a confined nuclear microenvironment for the proper association and regulation of promoters and enhancers (Nora et al. 2013). MCF-10A shSCRAM and shSMARCA4 interaction maps at 40-kb resolution revealed these
Figure 4. (A) Genome wide interaction heatmap at 2.5 Mb resolution showing the differences between interactions that are gained and lost upon SMARCA4 knockdown. The chromosomes are stacked from top-left to bottom-right in order (Chr 1, Chr 2, ..., Chr 22, and Chr X). (B) A zoom-in of Chr 11 at 250-kb resolution showing all differential interactions. (C) The interactions that are altered with significance (see Methods). (D) A further zoom-in view of a genomic region on Chr 11 (Chr 11: 60000001–81750000) (top) showing the differential interactions where the MALAT1 and NEAT1 loci reside (Chr 11: 64750339–65807685). (E) RNA-seq tracks from shSMARCA4 and shSCRAM cells showing a reduction of expression in NEAT1 and MALAT1 IncRNA genes upon SMARCA4 knockdown. (F) A zoom-in of the inter-chromosomal interactions between Chr 1 and Chr 2 through Chr 5, with arrows indicating the enriched telomeric interactions in the shSMARCA4 cells. (G) Quantification of the interactions among subtelomeric ends for shSCRAM and shSMARCA4 Hi-C data sets. The subtelomeric end shows significantly (Student’s t-test: P < 0.01 for Chr 1 and P < 0.05 for Chr 4) higher frequency of interactions in shSMARCA4 cells compared to control cells. (H) DNA-FISH images of shSMARCA4 and shSCRAM cells showing the intra-chromosomal telomeric interactions of Chr 1 and Chr 4. (I) Box plot showing the quantification of the telomere distances in shSMARCA4 and shSCRAM cells, quantified as described in the methods. P-value: Student’s t-test.
sub-megabase-scale interaction domains on all chromosomes, suggesting that SMARCA4 knockdown does not result in a loss of TAD formation (Fig. 6A; Supplemental Fig. S8A,B). In order to quantify the TAD boundary scores and identify specific TADs, we generated insulation plots using a method that was described previously (Barutcu et al. 2015; Crane et al. 2015). We identified 2963 and 2796 TAD boundaries in shSCRAM and shSMARCA4 MCF-10A cells, respectively (Fig. 6B). The identified TADs showed known characteristics, such as increased density of genes and Pol II binding at the boundaries compared to the surrounding regions (Supplemental Fig. S8C,D). Despite the notion that TADs are stable across different cell types, species, and different biological contexts (Dixon et al. 2012, 2015; Nora et al. 2012, 2013), interestingly, SMARCA4 knockdown altered the localization of ∼14% of the TAD boundaries. The majority of TAD boundaries (83% of shSCRAM and 88% of shSMARCA4) were overlapping between the control and the SMARCA4 knockdown cells (Fig. 6B).

SMARCA4 plays a role in TAD boundaries and assessed whether SMARCA4 localization was enriched at TAD boundaries. Intersection of SMARCA4 ChIP-seq peaks with both the shSMARCA4 and shSCRAM TAD boundary definitions yielded similar results, in which ∼25% of all SMARCA4 binding was located at TAD boundaries (Fig. 6C). Surprisingly, we observed an enrichment of SMARCA4 binding at the boundaries when the frequency of SMARCA4 binding was
plotted around the TAD borders (Fig. 6D). Sixty-six percent of all TAD boundaries were bound by SMARCA4. A similar phenomenon was observed when SMARCA4-bound super-enhancers were plotted across the TAD boundaries (Fig. 6E).

The strength of a TAD boundary is a measure of the allowance of inter-TAD interactions across the boundary (Van Bortle et al. 2014; Li et al. 2015). Although our data indicate that the majority of TAD boundaries are similar, we wondered whether SMARCA4...
knockdown resulted in a change in TAD boundary strength. Interestingly, SMARCA4 knockdown resulted in an overall decrease (Wilcoxon rank-sum test; \( P = 3.7 \times 10^{-6} \)) in the overlapping TAD boundary strength, as shown by plotting the boundary scores for the overlapping and unique TAD boundaries for the shSMARCA4 and shSCRAM cells (Fig. 6F). Furthermore, to address whether this decrease is related to SMARCA4 binding, we compared the scores of the TAD boundaries that were either SMARCA4-bound or not. We observed a significant decrease (Wilcoxon rank-sum test; \( P = 0.003 \)) of TAD boundary strength at borders that lacked SMARCA4 binding (Fig. 6G).

CTCF is an important component of TAD boundaries. When we intersected a previously published MCF-10A CTCF ChIP-seq data set (Ross-Innes et al. 2011) with SMARCA4 peaks, 10% of all SMARCA4 peaks and 12% of the SMARCA4 peaks that are located on a TAD boundary directly overlapped with CTCF, implying a crosstalk between SMARCA4 and CTCF at least for a subset of bound genomic regions (Supplemental Fig. S8E,F). The intersection of these two factors increased by twofold when the vicinity \( (±1 \, \text{kb}) \) of both SMARCA4 and CTCF peaks were considered. To assess whether the effects of SMARCA4 on TAD boundaries result from changes in the local chromatin structure around the CTCF sites, we analyzed publicly available SMARCA4-dependent accessibility and nucleosome positioning DNase I- and MNase-seq data sets (Tolstorukov et al. 2013; Morris et al. 2014; Stavreva et al. 2015). SMARCA4 perturbation affected nucleosome positioning within hundreds of base pairs of CTCF binding sites in mouse fibroblast cells (Supplemental Fig. S9A–C), showing decreased nucleosome occupancy immediately around the CTCF binding sites. SMARCA4 perturbation also altered chromatin accessibility in murine epithelial cells (Supplemental Fig. S9D). Taken together, these results demonstrate a role for SMARCA4, likely via mediating the local chromatin accessibility around the CTCF sites.

Discussion

Modification of the chromatin structure by ATP-dependent remodeling complexes is an essential process in transcriptional regulation. SMARCA4 affects both the activation and repression of many genes through its interactions with transcription factors and other cofactors (Trotter and Archer 2008). To understand the role of SMARCA4 in genome architecture, we characterized the SMARCA4-dependent alterations in gene expression and higher-order chromatin structure in mammary epithelial MCF-10A cells. RNA-seq analysis in the control and SMARCA4 knockdown cells showed an extensive down-regulation of genes related to ECM components and up-regulation of genes related to lipid metabolism and calcium signaling (Fig. 1). SMARCA4 is linked to calcium signaling (Lai et al. 2009; Nasipak et al. 2015). Literature also supports the idea that SMARCA4 (or the mammalian SWI/SNF complex) regulates lipid metabolism, as it was shown that SMARCA4 regulates PPAR\( \gamma \) expression, which in turn regulates lipid metabolism in differentiating adipocytes (Salma et al. 2004). In addition, SMARCD3, a subunit of the mammalian SWI/SNF complex, is required for activation of fatty acid and triglyceride synthesis in response to insulin and to feeding in liver (Wang et al. 2013). Interestingly, SMARCA4 binding was enriched at down-regulated response to insulin and to feeding in liver (Wang et al. 2013). Required for activation of fatty acid and triglyceride synthesis in re...

Growing evidence suggests that the shape of the nucleus, the stiffest organelle in the cell, might partly be affected by force-induced changes, a phenomenon known as nuclear mechanotransduction (Dahl et al. 2008). It is becoming well established that the cell surface adhesion receptors, such as integrins and cadherins, can exert mechanical forces to the nucleus and can potentially cause gene activation and/or chromatin reorganization (Wang et al. 2009). However, changes in nuclear shape can be induced from either external forces exerted by the cytoskeleton or via internal nuclear forces. Previous work showed that SMARCA4 knockdown results in nuclear shape alterations in MCF-10A cells (Imbalzano et al. 2013b). However, the disruption of the cytoplasmic filaments (actin, tubulin, and cytokeratins) did not alter SMARCA4-dependent structural changes observed in the MCF-10A cells (Imbalzano et al. 2013b). This implies that SMARCA4, apart from its chromatin remodeling function, might have additional roles in maintaining the structural integrity of the nucleus (Imbalzano et al. 2013a). Interestingly, in the present study, we determined that many of the genes that are down-regulated following SMARCA4 knockdown were associated with the extracellular matrix. These findings suggest an additional mechanism for SMARCA4-mediated regulation of nuclear integrity via regulation of ECM genes and possible alteration of cell surface connections and mechanotransducing forces to the nucleus.

Hi-C analysis of SMARCA4 knockdown and control MCF-10A cells revealed a significant enrichment of subtelomeric interactions in the shSMARCA4 cells (Fig. 4F–I). Literature evidence connecting SWI/SNF enzymes and telomeres is limited. SMARCA4 knockdown was shown to result in increased telomerase expression that leads to increased telomere lengths (Wu et al. 2014). Yeast SWI/SNF is required for telomeric silencing (Dror and Winston 2004), and telomeric expansion (Tomar et al. 2008). Mammalian cells lacking SMARCAL1, which is a SNF2 family ATPase but is not a subunit of the mammalian SWI/SNF complex, accumulate elevated levels of telomere-associated DNA damage (Poole et al. 2015). Moreover, knockdown of the mammalian SWI/SNF enzyme subunit, ARID1A, reduces telomeric repeat-containing RNA (TERRA) levels by at least twofold (Scheibe et al. 2013). Further experiments assessing whether SMARCA4 affects nucleosome positioning at or enhances/interferes with binding of specific proteins at telomeres to affect higher-order telomeric structures will shed further light on the relationship between SMARCA4 and telomere organization. Recently, Guidi et al. (2015) showed that the telomeres of yeast cells undergo spatial reorganization upon switching to different metabolic states. Because our RNA-seq analysis showed up-regulation of genes related to lipid synthesis and since our Hi-C analysis showed alterations in telomeric interactions in shSMARCA4 cells, it is tempting to speculate that the increased telomeric interactions (Fig. 4F–I) may be associated with changes in the metabolic state of the shSMARCA4 cells.

The importance of nuclear structure–gene expression relationships is supported by our finding that 4% of genomic compartments are altered in a SMARCA4-dependent manner (Fig. 4B). The magnitude of the functional impact of SMARCA4 on the architectural landscape is illustrated by previous reports showing that comparisons of relative compartmental differences in distinct cell types ranged from 4% to 25% of the genome (Barutcu et al. 2015; Dixon et al. 2015). The compartmental changes are associated with gene expression (Fig. 5C). However, it is still debated whether compartmental change or chromosomal movement causes differential gene expression or vice versa. There is evidence supporting both viewpoints (Chuang and Belmont 2007; Therizols et al. 2014).
The formation of topologically associating domain boundaries is dependent on many factors (Van Borstel et al. 2014; Cuberas-Potts and Corces 2015), especially insulators such as CTCF and cohesin. Here, we report the rather remarkable observation that SMARCA4 peaks and super-enhancers were also enriched at TAD boundaries (Fig. 6D,E), and the depletion of SMARCA4 resulted in ~14% of TAD boundaries being altered and lower TAD boundary scores genome-wide (Fig. 6F). Furthermore, SMARCA4-bound boundaries exhibited a stronger boundary score than the boundaries not bound by SMARCA4 (Fig. 6G). Therefore, in addition to well-studied insulators such as CTCF and cohesin, maintenance of TAD boundaries is affected, directly or indirectly, by SMARCA4, and by extension, the mammalian SWI/SNF enzyme possibly via regulating CTCF and cohesin binding. A previous report showed that SMARCA4 and other subunits of the mammalian SWI/SNF complex bind near regions critical for genome organization (e.g., CTCF and lamin binding sites and DNA replication origins) (Euskirchen et al. 2011). Our analysis of publicly available (Tolstorukov et al. 2013; Morris et al. 2014) DNase I accessibility and nucleosome positioning data around the CTCF sites demonstrated a SMARCA4-dependent effect on local chromatin structure (Supplemental Fig. S9), similar to the effect of SMARCA4 knockdown that was previously noted around TSS of known genes (Tolstorukov et al. 2013). Further experiments investigating whether SMARCA4 is required for CTCF, as well as cohesin and mediator binding, would provide insight into the mechanism by which SMARCA4 affects TAD boundaries.

Despite affecting genome organization at multiple levels, the overall genome organization was largely preserved upon SMARCA4 knockdown. Recent studies knocking down other fundamental chromatin organizers, CTCF (Zuin et al. 2014), cohesin (Seitan et al. 2013), and histone H1 (Geeven et al. 2015) similarly showed little effect on overall genome organization as assayed by Hi-C. It is possible that the remaining SMARCA4 levels in the knockdown MCF-10A cells (Fig. 1A) resulted in more modest changes than otherwise might have been expected. It is also likely that there are redundant molecular mechanisms in the cell to ensure the integrity of the genome and the nucleus. In the shSMARCA4 MCF-10A cells, the SMARCA4 homolog SMARCA2 (also known as BRM) still exists, and we previously reported that MCF-10A cells are not viable when both SMARCA4 and SMARCA2 are simultaneously knocked down (Cohet et al. 2010). Total removal of SMARCA4 by CRISPR/Cas9 in the MDAMB-231 human metastatic breast cancer cell line and genetic ablation in primary myoblasts also resulted in cell death (Padilla-Benavides et al. 2013; Wu et al. 2015). These findings suggest that the importance of SMARCA4 in cell viability may preclude observing more severe effects on genome organization.

Taken together, we identify novel roles for SMARCA4 in regulating higher-order chromatin structure by affecting telomere organization, TAD boundary strength, and the frequency and specificity of long-range chromatin interactions.

Methods

Cell culture
MCF-10A cells expressing control shRNA and shRNA targeting SMARCA4 were generated as previously described (Cohet et al. 2010). The cells were maintained in monolayer in Dulbecco’s modified Eagle’s medium-F12 (DMEM/F12) (Invitrogen, 21041025) supplemented with 5% horse serum (Invitrogen, 16050122), 1% penicillin/streptomycin (Invitrogen, 15140122), 0.5 µg/mL hydrocortisone (Sigma, H-0888), 100 ng/mL cholera toxin (Sigma, C-8052), 10 µg/mL insulin (Sigma, I-1882), and 20 ng/mL recombinant human EGF (Peprotech, 100-15) as described previously (Debnath et al. 2003). The doxycycline induction was performed by the addition of 0.05 µg/mL doxycycline to the cells and incubating them for 3–4 d.

RNA-seq and analysis
RNA was isolated from MCF-10A cells at ~75% confluence using the TRIzol Reagent (Life Technologies #15596-026), including treatment with DNase I. The poly(A)-selected RNA-seq libraries were generated with TruSeq RNA Sample Preparation Kit v2, and single-end 100-bp sequencing was performed using a HiSeq 2000 instrument. RNA-seq analysis was performed by filtering and mapping the reads by Bowtie 2 (Langmead and Salzberg 2012), quantifying the transcripts by RSEM v1.2.7 (Li and Dewey 2011), and finding the differentially expressed genes (log2 fold change >1, P < 0.01) by DESeq2 (Love et al. 2014).

Preparation of Hi-C libraries
Hi-C was performed as previously described with minor modifications (Belton et al. 2012). The modification was in the biotin incorporation step, in which the mixture was incubated for 40 min at 37°C. The MCF-10A shSCRAM and shSMARCA4 samples displayed a range of 25%–50% biotin incorporation efficiency. At the end of Hi-C sample preparation, the libraries were sequenced using paired-end 100-bp reads with a HiSeq 2000 instrument.

Read mapping/binning/ICE correction
Supplemental Table S3 summarizes the mapping results and different classes of reads and interactions observed (Lajoie et al. 2015). The data were binned at 2.5 Mb, 1 Mb, 250 kb, 100 kb, and 40 kb nonoverlapping genomic intervals. In our Hi-C analyses, we utilized the iterative correction and eigenvector decomposition (ICE) method (Imakaev et al. 2012). The replicates showed high correlation (Pearson correlation; R² ranging from 0.69 to 0.92) at multiple scales (Supplemental Fig. S4). For the downstream analyses, sequences obtained from both biological replicates were pooled and ICE-corrected to serve as a combined data set.

Z-score calculation
We calculated the Z-scores by modeling the overall Hi-C decay with distance using a modified LOWESS method (α = 1%, IQR filter), as described previously (Sanyal et al. 2012). LOWESS calculates the weighted-average and weighted-standard deviation for every genomic distance and therefore normalizes for genomic distance signal bias.

Calculation of differential interactions
To capture the differences between shSCRAM and shSMARCA4 interactions, we used a method previously described (Barutcu et al. 2015). Briefly, we first transformed the Hi-C data into Z-score matrices for all four replicate data sets (shSMARCA4-R1, shSMARCA4-R2, shSCRAM-R1, and shSCRAM-R2). For each interaction, the mean sample:sample (between samples) Z-score difference was calculated from all pairwise combinations of the four data sets (shSMARCA4-R1–shSCRAM-R1, shSMARCA4-R1–shSCRAM-R2, shSMARCA4-R2–shSCRAM-R1, shSMARCA4-R2–shSCRAM-R2). The replicate:replicate Z-score difference (within samples) was also calculated for a random set of 500,000 interactions. These random replicate:replicate Z-score differences were then
used to build an expected distribution of Z-score differences. The resulting Z-score difference matrix was then derived by calculating for each bin the ratio of the mean of the set of four possible samples: sample Z-score differences minus the genome-wide mean of the replicate:replicate Z-score difference, divided by the genome-wide standard error of the replicate:replicate Z-score differences. All statistical analyses have been corrected for multiple correction tests.

**Compartment profiles**

To detect the genomic compartments, first, Pearson correlation of the Z-score matrices was calculated. In performing principal component analysis (Lieberman-Aiden et al. 2009; Zhang et al. 2012), the first principal component detects the patterns of increased and decreased interaction across the genome that appear as a plaid pattern in the heatmap. Each genomic region matches this prominent interaction pattern (positive eigenvector value) or its opposite (negative eigenvector value), and these represent the two spatially segregated compartments. The open, gene rich “A-type” compartment may end up with either a positive or negative eigenvector. To detect which compartment is the open “A-type” and which is the closed “B-type,” the genome-wide gene density was calculated to assign the “A-type” and “B-type” compartmentalization.

**Identification of TAD boundaries (insulation square analysis)**

TAD calling was performed as calculating the insulation score of each bin using the 40-kb resolution combined Hi-C data as previously described (Barutcu et al. 2015; Crane et al. 2015).

**ChIP-seq analysis**

The ChIP assay was performed as previously described (Lee et al. 2006). The chromatin was sheared by using a BioRuptor instrument on high setting, 30 sec on and 30 sec off, for 5 min for five cycles. The pull-down was performed using a SMARCA4 antibody (Santa Cruz #G-7). The pull-down and input control sequencing libraries were generated using the NEXTFlex Rapid DNA Sequencing Kit (Bioo Scientific #S144-02) and were sequenced by using single-end 100-bp reads with a HiSeq 2000 instrument. The adapters were clipped, trimmed from the sequencing reads, and the reads were aligned to the hg19 human genome using the Bowtie 2 tool (Langmead et al. 2009). The pull-down and input control libraries were generated using the NEXTflex Rapid DNA Sequencing Kit (Bioo Scientific #S144-02) and were sequenced by using single-end 100-bp reads with a HiSeq 2000 instrument. The adapters were clipped, trimmed from the sequencing reads, and the reads were aligned to the hg19 human genome using the Bowtie 2 tool (Langmead et al. 2009). The TAD assay was performed as previously described (Lee et al. 2015). Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. Genome Biol 16: 214.

**References**


SMARCA4 regulates chromatin interactions


Received November 5, 2015; accepted in revised form July 8, 2016.
# SMARCA4 regulates gene expression and higher-order chromatin structure in proliferating mammary epithelial cells


*Genome Res.* 2016 26: 1188-1201 originally published online July 19, 2016
Access the most recent version at doi:10.1101/gr.201624.115

<table>
<thead>
<tr>
<th>Supplemental Material</th>
<th><a href="http://genome.cshlp.org/content/suppl/2016/08/09/gr.201624.115.DC1.html">http://genome.cshlp.org/content/suppl/2016/08/09/gr.201624.115.DC1.html</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 96 articles, 30 of which can be accessed free at: <a href="http://genome.cshlp.org/content/26/9/1188.full.html#ref-list-1">http://genome.cshlp.org/content/26/9/1188.full.html#ref-list-1</a></td>
</tr>
<tr>
<td>Creative Commons License</td>
<td>This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <a href="http://genome.cshlp.org/site/misc/terms.xhtml">http://genome.cshlp.org/site/misc/terms.xhtml</a>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <a href="http://creativecommons.org/licenses/by-nc/4.0/">http://creativecommons.org/licenses/by-nc/4.0/</a>.</td>
</tr>
<tr>
<td>Email Alerting Service</td>
<td>Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.</td>
</tr>
</tbody>
</table>

To subscribe to *Genome Research* go to: http://genome.cshlp.org/subscriptions

© 2016 Barutcu et al.; Published by Cold Spring Harbor Laboratory Press