Endothelial cell differentiation is encompassed by changes in long range interactions between inactive chromatin regions

Henri Niskanen¹, Irina Tuszynska²,†, Rafal Zaborowski²,†, Merja Heinäniemi³, Seppo Ylä-Herttuala¹,4,* , Bartek Wilczynski²,* and Minna U. Kaikkonen¹,*

¹A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland, ²Institute of Informatics, Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw 02-097, Poland, ³School of Medicine, University of Eastern Finland, FI-70211 Kuopio, Finland and ⁴Heart Center and Gene Therapy Unit, Kuopio University Hospital, FI-70211 Kuopio, Finland

Received August 18, 2017; Revised October 21, 2017; Editorial Decision November 21, 2017; Accepted November 27, 2017

ABSTRACT

Endothelial cells (ECs) differentiate from mesodermal progenitors during vasculogenesis. By comparing changes in chromatin interactions between human umbilical vein ECs, embryonic stem cells and mesendoderm cells, we identified regions exhibiting EC-specific compartmentalization and changes in the degree of connectivity within topologically associated domains (TADs). These regions were characterized by EC-specific transcription, binding of lineage-determining transcription factors and cohesin. In addition, we identified 1200 EC-specific long-range interactions (LRIs) between TADs. Most of the LRIs were connected between regions enriched for H3K9me3 involving pericentromeric regions, suggesting their involvement in establishing compartmentalization of heterochromatin during differentiation. Second, we provide evidence that EC-specific LRIs correlate with changes in the hierarchy of chromatin aggregation. Despite these rearrangements, the majority of chromatin domains fall within a pre-established hierarchy conserved throughout differentiation. Finally, we investigated the effect of hypoxia on chromatin organization. Although hypoxia altered the expression of hundreds of genes, minimal effect on chromatin organization was seen. Nevertheless, 70% of hypoxia-inducible genes situated within a TAD bound by HIF1α suggesting that transcriptional responses to hypoxia largely depend on pre-existing chromatin organization. Collectively our results show that large structural rearrangements establish chromatin architecture required for functional endothelium and this architecture remains largely unchanged in response to hypoxia.

INTRODUCTION

The circulatory system is the first organ system to develop in the growing embryo as it is needed for the rapid delivery of oxygen and nutrients to tissues. Endothelial cells (ECs), which constitute the luminal layer of blood vessels, differentiate from their mesodermal progenitors during a process called vasculogenesis. ECs act as functional local oxygen sensors within the vascular system and during hypoxia they modulate vascular tone and induce vascular remodeling and angiogenesis. Recent studies have uncovered extensive chromatin reorganization during cellular differentiation (1). We and others have shown that different cell types are characterized by differences in the location of active and inactive chromatin compartments (1–3). These compartments are further divided into single or series of topologically associating domains (TADs) (4), large fraction of which are conserved between cell types and in response to extracellular signals (5). However, the chromatin interactions within and between TADs have been shown to exhibit extensive changes between cell types and in response to senescence and hormone induced gene regulation (1,6–8). The most extensive changes have been documented during the heat shock response in Drosophila, where a dramatic increase in inter-TAD interactions correlated with genome-wide repression of transcription (9).

Despite recent advances in describing local changes in the units of chromosome organization, our understanding of the long range interactions between these units and their

---

*To whom correspondence should be addressed. Tel: +358 40 355 2413; Email: minna.kaikkonen@uef.fi
Correspondence may also be addressed to Seppo Ylä-Herttuala. Tel: +358 40 355 2075; Email: seppo.ylaherttuala@uef.fi
Correspondence may also be addressed to Bartek Wilczynski. Tel: +48 22 5544 577; Email: bartek@mimuw.edu.pl
†These authors contributed equally to the paper as second authors.

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
role in driving cellular identity is incomplete. Here we analyze chromatin interactions during EC differentiation by comparing public Hi-C maps from H1 human embryonic stem cells (ESC) and mesendoderm stem cell (MDC) to our data generated from human umbilical vein endothelial cells (HUVECs). In addition to previously reported changes in chromatin compartments and TADs, we identified 1200 EC-specific long-range interactions (LRIs) between TADs. These regions were highly enriched for repressive chromatin marks and depletion of active transcription. We also present evidence that domain aggregation into hierarchy correlates with LRIs and the hierarchy in differentiated ECs largely pre-exists in progenitor cells. Finally, we also demonstrate how the response to hypoxic stimulus is largely determined by pre-existing chromatin organization.

MATERIALS AND METHODS

Cell culture

HUVECs used in this study were either obtained from the maternity ward of Kuopio University Hospital by the approval of the Kuopio University Hospital Ethics committee and used at passage 6 or purchased from Life Technologies and used at passage 8. HUVECs were maintained in endothelial cell growth medium (EGM; 0.1% human epidermal growth factor, 0.1% Gentamicin-Amphotericin-B, 0.4% bovine brain extract, 2% FBS; Lonza) on cell culture flasks coated with 10 µg/ml Gentamicin-Amphotericin-B, 0.4% bovine brain extract, 2% FBS; Lonza) on cell culture flasks coated with 10 g/ml fibronectin (Sigma, St Louis, MO, USA) and 0.05% gelatin in phosphate-buffered saline. To induce hypoxia, cells were incubated in Ruskinn Invivo2000 hypoxia workstation (The Baker Company) in the presence of 1% O2 and 5% CO2 for 8 h.

ChIP-seq libraries

ChIP-seq libraries were prepared from two biological replicates as previously described (3). Briefly, ∼10 M cells were formaldehyde-fixed for 10 min, scraped into lysis buffer (10 mM Tris–HCl pH 7.4, 10 mM NaCl, 5 mM MgCl2, formaldohyde-fixed for 10 min, scraped into lysis buffer) and TADs were identified as previously described (3). Briefly, ChIP-seq libraries were prepared from two biological replicates of control or hypoxia-treated HUVECs according to the protocol described (3). Briefly, 5–10 millions of nuclei were collected from each sample, run-on reactions were performed in the presence of Br-UTP and RNA was extracted with Trizol LS Reagent (Life Technologies, Carlsbad, CA, USA). RNA was DNase treated and fragmented using TURBO DNase and RNA fragmentation reagents (Life Technologies), purified using P-30 columns (Bio-Rad, Hercules, CA, USA) and dephosphorylated using PNK (Enzymatics, Y904L). Labeled RNA fragments were captured using anti-BrdU beads (sc-32323 AC, SantaCruz Biotech, Santa Cruz, CA, USA). Poly(A)-tailing and reverse transcription were performed as described (3). Libraries were amplified using 12–16 cycles, size-selected (180–350 bp) from 10% TBE gels (Life Technologies) and sequenced using Illumina HiSeq 2000.

Tethered conformation capture

For two replicates, 90 million cells were grown in normoxic (20% O2) or in hypoxic (1% O2) conditions (Ruskin In Vivo2000 hypoxia incubator) for 8 h. Tethered conformation capture (TCC) was performed as described (3). Digestions were performed using MboI (New England Biolabs), DNA was sheared with Bioruptor Next-Gen (Diagenode) for 30–35 cycles, library preparations continued as described (3) and products were amplified for 15 cycles, size-selected (225–550 bp) from 10% TBE gels (Life Technologies) and sequenced using Illumina HiSeq 2000 at EMBL Genomics Core Facility (Heidelberg, Germany).

Preprocessing of GRO-seq, RNA-seq and ChIP-seq data

Raw data from public datasets and new samples were processed identically. Poly-A tails introduced in the library preparation were trimmed out using using HOMER 4.7 (http://homer.salk.edu/homer) (10) software (homertools trim) and reads shorter than 25 bp were discarded. Quality of reads was controlled using FastQC tool (11) and bases with poor quality were trimmed using FastX toolkit (http://hannolab.cshl.edu/fastx-toolkit) available in Galaxy platform (12) so that minimum 97% of all bases in one read have to have a minimum phred quality score of 10. Alignments to hg19 genome were performed using bowtie-0.12.7 (13) allowing up to two mismatches and up to three locations per read and the best alignment was reported. Reads aligning to rRNA (abundant sequences as annotated by iGenomes) and blacklisted regions (unusually low or high mappability, defined by ENCODE) were excluded from analysis. For RNA-seq experiments alignments were performed using TopHat (14) while allowing up to two mismatches and reporting only one alignment for each read. ChIP-seq data was processed similarly as GRO-seq data, except that all reads passing quality filtering were used in the alignments.

Preprocessing of TCC/Hi-C data

Raw data from public datasets and new samples were processed identically. Paired-end reads were filtered and aligned separately. Before alignments, reads were checked for restriction enzyme sites (MboI or HindIII) and sequences after the site were removed to improve mappability using HOMER 4.7 software (homertools trim), reads...
shorter than 25 bp were discarded and quality was controlled using FastQC and FastX tools, similarly as in GRO-seq data. Alignments to hg19 genome were performed using bowtie-0.12.7 allowing up to two mismatches and up to three locations per read and the best alignment was reported. Paired-end reads were connected when preparing HOMER Hi-C tagdirectories using standard settings. These settings included considering read pairs with exact same ends were only once (-tbp1 option) and removing paired-end reads that are likely continuous genomic fragments or re-ligation events, i.e. reads separated less than 1.5× the estimated sequencing insert length (-removePEbg option). Settings also included excluding reads from regions with unusually high tag density by removing reads from 10 kb regions which contain more than five times the average number of reads (-removeSpikes 10000 5 option). Finally, only the reads where both ends of the paired-end read have a restriction site within the fragment length estimate 3′ to the read were included (-restrictionSite GATC or AAGCTT and -both options) and reads were removed if their ends form a self-ligation with adjacent restriction sites (-removeSelfLigation).

Analysis of GRO-, RNA- and ChIP-seq data
All analysis with GRO-, RNA- and ChIP-seq data was performed using HOMER. TagDirectories were prepared from preprocessed data or from public .bam or .bed files with fragment length set to 75. When preparing bedGraph and bigWig visualization files, sequencing experiments were normalized to 10⁷ uniquely mapped reads. Differentially regulated genes were detected from GRO- and RNA-Seq data using edgeR v3.2.2 (15) and thresholds of RPKM > 0.5, FDR < 0.1 and fold change > 1.7 were used. Gene ontology analysis was performed using HOMER ‘findGO.pl’ program. ChIP-seq peaks were detected using the ‘findPeaks.pl’ tool in HOMER software, with default settings for ‘style factor’ option: requiring 4-fold enrichment over appropriate input sample and 0.001 FDR significance. Analysis of GRO-seq and ChIP-seq data from oxidized phospholipid (oxPAPC), interleukin 1 beta (IL1β), tumor necrosis factor alpha (TNFα) stimuli were performed as previously described (16) and from ChIP-seq peaks only those present in treatment conditions were selected as induced peaks. Average fold changes for TADs and the fractions of regulated genes were calculated using fold changes in gene expression for genes with RPKM > 0.5 in basal or treatment conditions. Hypoxia-responsive eRNAs were detected by combining all HUVEC H3K4me1, H3K4me2, H3K4me3 and H3K27ac peaks, which were further filtered to contain only intergenic enhancers (at least 3 kb upstream of TSS and 10 kb downstream of TTS) exhibiting enrichment of H3K4me1 over H3K4me3, GRO-seq RPKM > 0.3 and 1.7-fold increase or decrease in normalized GRO-seq tags in response to hypoxia. Enhancers without a change in eRNA expression in response to hypoxia were defined by having fold change below 1.3. Hypoxia-responsive H3K27ac (17) and H3K4me1 (18) histone modifications were analyzed and regions showing strongest changes in response to hypoxia were selected by using log2 fold change cutoff 1 for H3K27ac and cutoff 4 for H3K4me1 data. To assess CTCF binding in MCF-7 cells (19), closest features tool from BEDOPS software (20) was used to select CTCF peaks closest to hypoxia-inducible factor 1 alpha (HIF1α) peaks (found in both HUVEC and MCF-7 datasets (21)) or TSS of hypoxia-regulated genes that were shared in HUVECs and MCF-7 cells (22).

Normalization of Hi-C/TCC data
For the analysis performed using HOMER, data was normalized as described (23). Background models, which take into the account the sequencing depth and linear distances between the interacting regions, were created for 10, 25, 50 and 100 kb resolutions. For generation of long range contact maps, we have utilized the iterative correction software by (24) with binning into 50 kb (used for differential intra- and inter-TAD contacts) and 150 kb (used for hierarchical domain aggregation). These bin sizes were used to avoid bias due to comparison of our libraries with MboI-cutter (4-cutter) and the comparison data generated with HindIII (a 6-cutter) (1).

Analysis of interactions, compartmentalization and TADs
Unless stated otherwise, analysis of Hi-C and TCC data was performed for data where replicates were pooled. Analysis of individual interactions with 10 kb resolution was performed as described (3, 23). Principal component analysis (PCA)-based detection of active and inactive chromatin compartments were performed using HOMER tool ‘runHiCpca.pl’ using 25 kb resolution and 50 kb superResolution options and correlation differences between two experiments were calculated using ‘getHiCcorrDiff.pl’ program. Detection of differential compartments between different cell types and treatment conditions were detected using ‘findHiCCompartments.pl’ tool and requiring minimum 80 difference in PC1 values and correlation difference of 0.4. Analysis of TADs is based on directionality index and it was performed using HOMER command ‘findHiCDomains.pl’ using a resolution of 50 and 100 kb superResolution. Overlap of TAD boundaries was determined using bedTools (25) intersect, while allowing one bin mismatch while still considering boundaries identical.

We have used DiffTAD software (https://www.biorxiv.org/content/early/2016/12/13/093625) on 50-kb resolution contact maps to annotate Homer-derived TADs with significantly differential contact densities using the permutation test approach (P < 0.05). The same maps were used to find long-range inter-TAD contacts with the hypergeometric test (P < 0.1 * 10^-10) (scripts for the inter-TAD contact detection are available here: https://github.com/regulomics/long_inter). As the first step of our method the new matrix M (with the shape of NxN where N is the number of domains) is calculated with M[i,j] representing the sum of Hi-C map's values calculated for the pair of domains i and j. Next, for each pair of domains in the new matrix P-value was calculated based on the hypergeometric test (see formulas below), where k is the sum of contacts for the particular pair of domains, M is a sum of all contacts between all domains, i.e. the sum of all values in the matrix, n and N are the sum of contacts of first or second domains from the ana-
lyzed pair of domains with all other domains (sum of values in the particular column or row of the new matrix).

\[ p(k, M, n, N) = \binom{n}{k} \binom{M-n}{N-k} \binom{M}{N} \]

\[ P - value = 1 - \sum_{i=0}^{k-1} p(i, M, n, N) \]

This procedure was followed for ESC, MDC and HUVEC (hypoxia and normoxia) contact matrices. The long range contacts were called for each contact map separately using the same P-value threshold and the specific interactions were identified by removing any interactions that were overlapping with contacts detected in other cell type.

In the analysis of intra-TAD interactions in response to hypoxia, HOMER was used to prepare differential contact matrices between the two conditions using 50 kb resolution. Geometric means for the differential of observed/expected interactions were calculated from the sub-matrices that position TADs in following three groups: upregulated, unchanged or downregulated according to average log2 fold change of genes within the domain, >0.3, between 0.3 and −0.3 or <−0.3, respectively. To analyze the dynamics of individual interactions in response to hypoxia, HOMER software ‘annotateInteractions.pl’ was used to compute pairwise feature enrichments from significant interactions detected at 10 kb resolution which connected hypoxia upregulated TSS to other features of interest.

For the analysis of the hierarchy of topological domains, we have used contact maps normalized into 150 kb domains to obtain a stable hierarchy reconstruction between replicates. The hierarchical aggregation was done using the Sherpa software (https://github.com/Regulomics/sherpa) with up to 11 levels of aggregation starting from the 150 kb bins. Similarity of hierarchical domains was measured using Jaccard index, which is calculated by dividing the number of overlapping domain boundaries with the total number of non-overlapping boundaries in cell type A and B.

Enrichment for RAD21 peaks and super-enhancers within TADs and compartments

Permutation tests were performed using R/Bioconductor package regioneR (26) to study the associations between RAD21 peaks or super-enhancers (27) with TADs and differential compartments. Numbers of overlaps between regions A (TADs of interest or 25 kb bins with differential compartmentalization) and regions B (RAD21 peaks or HUVEC super-enhancers) were quantified with ‘numOverlaps’ function. Randomization function ‘resampleRegions’ was used to select from universe of all TADs or from 25 kb bins of the whole genome. A total of 10 000 permutations were performed to estimate enrichment Z-scores and P-values.

Analysis of LDTFs and HUVEC-specific active compartments

To study the role of LDTFs in regulation of genes residing in HUVEC-specific active compartments, RNA-seq data from (28) was downloaded and processed as described earlier. Quantification of gene expression was performed using HOMER ‘analyzeRNA.pl’ and fold changes in response to LDTF overexpression was calculated from normalized log2 tag counts for genes residing in HUVEC-specific compartments or random gene sets. Test gene set consisted of 1458 genes residing in differential active compartments (HUVEC versus H1-ESC) and random sets of same size were selected from all genes residing in active compartments in HUVECs. Fold changes from 20 different random sets were pooled together for visualization. To analyze the enrichment of TFs binding to specific active compartments, regioneR software was used as described above, except that regions were divided to specific active compartments, all active compartments, all inactive compartments and specific inactive compartments. To test enrichment in all active or inactive compartments, same size sets of 25 kb bins were selected as in the specific sets and averages from 50 random selections were displayed. Whole genome was used as the background from which 1000 permutations were performed to estimate enrichment Z-scores. In visualizing the expression of transcription factors (TFs) in HUVEC-specific compartments, 1672 genes categorized as TFs in FANTOM5 SSTAR catalog (29) were selected and filtered for those overlapping with specific regions and for active transcription (RPKM > 0.5) in at least one of the three cell types (MDCs, H1-ESCs, MDCs or HUVECs) according to RNA-seq data.

Analysis of connectivity to pericentromeric heterochromatin

Normalized Hi-C matrices and correlation matrices were produced with HOMER using 100 kb resolution. H3K9me3 ChIP-seq data from HUVECs was also analyzed using 100 kb bins and the regions exhibiting strong enrichment for H3K9me3 in proximity to centromeric regions were selected for analysis (Supplementary Table S4). Chr9 was excluded from the analysis due to the large amount of gaps around the centromere in the hg19 genome assembly.

For each 100 kb bin along a chromosome, the connectivity to pericentromeric region was analyzed by taking the average value of log2 (observed/expected interactions) for the position of pericentromeric region from normalized Hi-C matrix. BedGraph files for visualization were prepared using the correlation matrix. For analyzing connectivity of PC1 negative compartments to pericentromeric regions, inactive compartments were selected for each cell type. The average value for the connectivity to pericentromeric regions was calculated for all inactive compartments within each chromosome and values for each chromosome in ESCs, MDCs and HUVECs were reported.

Association between LIRIs and pericentromeric regions

Permutation test was performed to assess the enrichment of pericentromeric regions within the LIRIs connecting inactive TADs. Sets of random interactions were prepared from
the list of PC1 negative TADs for each chromosome, while keeping the size of the set similar as the number of observed LRIs. Randomization process was repeated 1000 times and the number of interactions in which a TAD overlaps pericentromeric region was calculated for each set. Distribution of the overlaps for the random interactions was used to estimate the enrichment $Z$-score for the overlap in observed interactions.

**Association between changes in hierarchy, compartmental switching and epigenetic features**

Regions ongoing hierarchical changes between ESC and HUVEC were determined as domains with at most 50% overlap between the cell-types and at least 100 kbp in length. In case of overlapping reorganized regions we have reported all included reorganized subregions as the result. Permutation tests were performed using regioneR similarly as described earlier. Reorganized domains of certain level were used as a test set and association with differential compartments and regions showing differential H3K4me1 (log2 fold change cutoffs 3 and $-3.5$), H3K27ac (log2 fold change cutoffs 3 and $-3$) and H3K27me3 (log2 fold change cutoffs 2 and $-2$) histone modifications was analyzed. Randomization function ‘resampleRegions’ was used to select from universe of all domains of same level and 1000 permutations were performed to estimate enrichment $Z$-scores and $P$-values.

For assessment of the enrichment of ‘within-hierarchy’ LRIs, we have calculated the number of LRIs falling within the SHERPA domains at all levels. The statistical significance of the enrichment was assessed using binomial distribution estimated from the total number of LRIs at any given level the number of within-hierarchy domains falling into randomly reshuffled domains.

**RESULTS**

**Endothelial-specific units of chromosome organization are characterized by cell type specific transcription and cohesin binding**

We performed TCC experiments from two replicates of HUVEC cells under normoxia and 8h hypoxia (1% O2), generating a total of nearly 1 billion reads (Supplementary Table S1). This extended the sequencing depth from our previous study and allowed us to investigate the effect of cellular differentiation and external stimulus on EC chromatin architecture (3). To assess the changes in genome compartmentalization during endothelial differentiation, we compared our data to recent Hi-C data from H1 human ESC and MDC (1). We performed PCA-based segmentation to decompose each chromosome into active and inactive compartments, in which interactions within each compartment type are enriched and contacts between them depleted (2). To further identify TADs (Supplementary Table S1) within these compartments, we used the directionality index to quantify the degree of upstream or downstream interaction bias for a genomic region (4). In line with previous findings, we saw extensive switching between the active and inactive chromatin compartments and in the intra-TAD interaction frequency (Figure 1A) (1,3). Altogether around $\sim$2000 compartments underwent a switch from inactive to active (PC1 negative to PC1 positive) or active to inactive (PC1 positive to PC1 negative), representing 9–14% of compartments (Figure 1A and Supplementary Table S2). The compartment changes were highly reproducible as strong correlation between the replicates ($R = 0.924 \sim 0.815$) and previously published HUVEC in situ Hi-C dataset (35) ($R = 0.704 \sim 0.636$) was detected (Supplementary Figure S1A and B). Of the total of $\sim$3400 TADs, $\sim$1000 TADs were enriched, i.e. exhibiting more intra-TAD interactions in HUVECs versus progenitors, and $\sim$200 depleted in intra-TAD interactions (Supplementary Table S3, Figure 1A-B, Supplementary Figure S1D, S2A-D). As expected, activated compartments were more likely to contain enriched TADs and inactivated compartments depleted TADs (Supplementary Figure S1C). We observed that both of these alterations in chromatin structure correlated with changes in gene expression and histone marks (Figure 1B and C; Supplementary Figure S1D and E). Interestingly, however, the compartment changes were more strongly segregated according to histone marks and gene expression. Accordingly, the genes within EC-specific active compartments were enriched for endothelial-specific functions, such as cell migration, cell adhesion and vasculature development while genes in EC-specific inactive compartments were involved in functions related to nervous system development, stem cell division, metabolic functions and cell adhesion (Supplementary Figure S1F). In addition, we found multiple TFs, such as myocyte-specific enhancer factor (MEF), nuclear receptor subfamily 2 (NR2) and Kruppel-like factor (KLF) family TFs, residing in the EC-specific active compartments majority of which were induced in ECs (Supplementary Figure S3A). In contrast majority of TFs residing in EC-specific inactive compartments were repressed in ECs, as exemplified by stem cell regulators Nanog homeobox (NANOG), POU Class 3 Homeobox 1 (POU3F1) and orthodenticle homeobox 2 (OTX2).

To understand the mechanism behind the cell type-specific chromatin organization, we studied the effect of lineage determining TFs (LDTFs) in regulating gene expression in compartments undergoing a switch from inactive to active. Accumulating evidence suggests that a relatively small set of LDTFs are responsible for establishing enhancer-like open chromatin regions that are required for cell-type specific gene expression (10). To this end, we analyzed the expression of genes within EC-specific active compartments upon reprogramming of ESCs by LDTF overexpression (28). We observed that overexpression of EC-specific TFs ETS variant 2 (ETV2) and GATA binding protein 2 (GATA2) increased the transcription within the EC-specific active compartments whereas the overexpression of a TF not relevant for EC-differentiation runt related TF 1a (RUNX1a) caused much smaller differences (Supplementary Figure S3B and C). In line with this, we observed that EC-specific active compartments are enriched for GATA2 binding events (Supplementary Figure S3D) and also for JUN and FOS TFs. This supports the notion that LDTFs play a role in defining EC-specific gene expression patterns.

Recent studies have indicated that cohesin co-localizes with LDTFs and maintains the cell type-specific chromatin contacts originally established by the binding of LDTFs at
Figure 1. EC differentiation is accompanied by extensive changes in chromatin interactions. (A) Table representing the number of regions exhibiting compartmental switch and number of TADs enriched or depleted for intra-TAD interactions. Regions that are detected in both HUVEC versus H1-ESC and HUVEC versus MDC comparisons are shown in the overlap column. (B) UCSC Genome browser image displaying compartmentalization: black color indicates active compartments with positive PC1 values, and gray color inactive compartment with negative PC1 values. Transcriptional activity is represented as normalized GRO-seq tags at NR2F2 locus. Normalized contact difference of HUVEC versus H1-ESC interactions shown on top. (C) Violin plots depicting the fold changes in ChIP-seq, RNA-seq and GRO-seq signals at differential PCA compartments (top) and TADs with enriched or depleted intra-TAD interactions (bottom). Black dot represents the mean and whiskers standard deviation. (D) Z-scores derived from permutation tests are shown for enrichment of RAD21 peaks within TADs (enriched or depleted in intra-TAD interactions) and compartments (* = P < 0.0001, x = P < 0.0008 and + = P < 0.00014). (E) Z-scores from permutation tests showing enrichment of HUVEC super-enhancers within TADs and compartments (* = P < 0.0001). (F) Circos plot showing significant interactions (black lines; 10 kb resolution), GRO-seq (orange) and ChIP-seq signal for H3K27ac (green) for the region inside the two dashed lines in (B).
enrichments during cell division when most TFs are evicted from DNA (30). Thus the binding of cohesin is expected to reflect the sites of LDTF binding without the need to study each LDTF separately. To this end, we performed ChIP-Seq for the RAD21 subunit of cohesin and compared the location of binding to that seen in ESCs. In line with reported direct interactions between cohesin and CCCTC-binding factor (CTCF) and their collaboration in shaping the 3D genome, extensive overlap with CTCF binding was seen (Supplementary Figure S3E) (30–32). Accordingly, these coobound peaks showed localization at the borders of TADs (Supplementary Figure S3F) (33,34). However, only the cohesin without coobound CTCF were found significantly enriched for binding LDTFs and E1A binding protein p300 (EP300) coactivator (Supplementary Figure S3G) and associated within EC-specific active compartments but not TADs (Figure 1D). Similar enrichment of cohesin, but not CTCF, has previously been reported at super-enhancers (27) and indeed we also demonstrate that EC-specific active chromatin compartments and enriched TADs are enriched for HUVEC-specific super-enhancers (Figure 1E) that have been detected using H3K27ac ChIP-seq (27). We also detected increased connectivity between the super-enhancers and target genes within the enriched TADs, exemplified by nuclear receptor subfamily 2 group F member 2 (NR2F2) and cavelin 1 (CAV1) gene loci (Figure 1F and Supplementary Figure S1G). Furthermore, we observed that the cell-type specific RAD21 binding events were strongly associated with the cell type-specific active compartments (Supplementary Figure S3H) and TADs enriched for intra-TAD interactions (Supplementary Figure S3I). Taken together, these results suggest an important role for cohesin and LDTFs in driving the transcriptional activity of genes required for functional endothelium which lie within cell type-specific units of chromatin organization.

Changes in long range interactions encompass epigenetic changes within chromatin compartments

To study the structural dynamics in LRIs between TADs, we developed a method to compute $P$-values for each TAD–TAD interaction and kept only interactions above a stringent cutoff of $P < 10^{-10}$ (Figure 2A and Supplementary Figure S4A). This led to identification of $\sim 60 000–80 000$ LRIs in each cell type of which 60% were shared between the hESC/MDC and HUVECs, 90% between the two HUVEC replicates and 80% between our HUVEC dataset and the previously published dataset from in situ Hi-C (35) (Supplementary Figure S4B–D). Next we filtered the data to retain only the EC-specific LRIs common to both HUVEC datasets. This led to identification of 1200 LRIs with a median distance of 10 Mb (Figure 2A, Supplementary Figures S5 and Table S4). The vast majority of these interactions were located within regions enriched for constitutive heterochromatin mark H3K9me3 which was not present in the progenitor cells (Figure 2A–C, Supplementary Figures S4A, S6A and B). This is supported by the observation that the LRIs were found much more frequently between two TADs residing in inactive chromatin compartments than expected by chance even taking into consideration the over-representation of inactive domains among the ones in contact (Figure 2D). About 18–25% of the compartments undergoing switch during cellular differentiation were contacted by LRIs and a clear shift towards a more inactive compartment was seen during differentiation (Figure 2E). To further characterize the LRIs we compared the transcriptional activity and histone modifications between ECs and progenitor cells. Consistent with enrichment for inactive compartments, we found that transcription, binding of architectural proteins CTCF and RAD21 and active histone marks H3K27ac and H3K4me1 were depleted in EC-specific LRIs, whereas PRC2-associated H3K27me3 repressive mark was enriched (Figure 2F; Supplementary Figure S6A and B). In line with this, we identified several examples where the expression of a gene was repressed in ECs (Figure 2C, Supplementary Figure S6A and B). Similar results were seen when the TADs engaged in LRIs were compared to TADs outside LRIs (Supplementary Figure S4E).

To dissect the different subtypes of LRIs based on the activity of the interaction counterparts, we classified the LRIs based on their connectivity between TADs falling into two active or two inactive compartments, representing 12 and 77% of all the LRIs, respectively (Figure 2D). In line with the majority of LRIs belonging to the inactive-inactive category, depletion of active chromatin marks and enrichment of repressive ones was seen together with significant (hypergeometric test $P$-value $= 2E-10$) enrichment for compartments undergoing switch from active to inactive compartment in HUVECs (Supplementary Figure S4F–H). Genes residing in these ‘inactive TADs’ were involved in functions related to ectodermal germ layer, such as sensory perception, neuronal, keratinocyte and skin development, in line with the repression of alternative germ layer activities during differentiation (Supplementary Figure S4I). In contrast, the active-active LRIs were enriched for enhancer marks H3K27ac and H3K4me1 and transcription suggesting that cell type-specific enhancer activity might be involved in establishing the epigenomic status of these TADs which could result in stronger associations between the interacting TADs. Accordingly, the active-active LRIs were significantly ($P$-value $= 1E-4$) enriched for compartments undergoing a switch from inactive to active (Supplementary Figure S4F–H). Most highly enriched gene ontologies were related to chromatin organization, humoral response, regulation of lipoprotein lipase activity and vascular wound healing (Supplementary Figure S4J).

Interestingly, one third of TADs engaged in LRIs fell within negative or positive compartments without a notable difference in PCA between the progenitors and ECs, despite clear changes in gene expression (Supplementary Figure S4K). For example, active transcription (RPKM $> 0.5$) was detected in inactive TADs in 235 of H1-ESC genes majority (77%) of which, however, were silenced in in HUVECs. To test if LRIs could serve as indicators of gene expression change during differentiation, we studied the gene expression changes in TADs associated with or absent of EC-specific LRIs in relation to PCA value. Interestingly, higher level of gene repression and activation was seen within TADs engaged in LRIs within regions of PCA change (dPC1 $> 20$; Figure 2G) but even at regions with no change (Figure 2H). This suggests that LRIs could serve as an important predictor of gene expression change during...
Figure 2. HUVEC-specific LRIs are predominantly formed between regions of inactive chromatin. (A) Left: Hi-C interactions (blue) and -log10 $P$-value for LRIs (red) for H1-ESCs, MDCs and HUVEC datasets are depicted in heatmap for a representative region on chr17: 41,510,000 – 77,100,000. Middle:
differentiation instead or along with PCA-based compartment change.

HUVEC-specific LRIs connect inactive compartments and pericentromeric heterochromatin

To further explore the mechanism responsible for the establishment of the bulk of LRIs i.e. those within inactive compartments, we investigated the epigenomic characteristics of these domains. We found that TADs that reside in the pericentromeric heterochromatin (PCH) (Supplementary Table S4) were strikingly more prominent within the HUVEC-specific LRIs, exemplified by the left arm of chr10 (Figure 3A). In line with this, permutation analysis revealed that there was a clear enrichment (P-value < 0.001) for regions of PCH within HUVEC-specific LRIs (Figure 3B). Hi-C correlation matrices for the left arm of chr10 showed that the pericentromeric region is isolated from the active and inactive compartments in H1-ESCs, whereas the connections with inactive compartments and PCH are starting to emerge in MDCs and become strikingly evident in HUVECs (Supplementary Figure S7A and Figure 3C). This increase in connectivity between PCH and inactive compartments was observed genome-wide (Figure 3D). Regions that exhibited a raise in connectivity to PCH also showed a strong accumulation the H3K9me3 mark in HUVECs, whereas in ESCs strong H3K9me3 signal was only seen at the pericentromeric region (Figure 3C). Furthermore, regions that establish connections to PCH were transcriptionally repressed in HUVECs, whereas transcription was clearly visible in H1-ESCs (Figure 3C and Supplementary Figure S7B). An interesting exception to the rule was chr19 where connections between inactive compartments and pericentromeric regions were already present in H1-ESCs and strongly enriched for H3K9me3 in all cell types (Figure 3D and Supplementary Figure S7C). These regions coincided with clusters of KRAB-ZNF genes, which were actively transcribed both in H1-ESCs and HUVECs in spite of their heterochromatic nature. Of note, this same chromosome has recently been shown to present a distinct nuclear subcompartment with specific pattern of histone modifications with both activating chromatin marks, such as H3K36me3, and heterochromatin-associated marks H3K9me3 and H4K20me3 (35).

To confirm the findings were not specific to ECs, we also identified the LRIs for IMR90 fibroblast cells. In line with EC data, the majority of the 5000 fibroblast-specific LRIs were enriched for connection between inactive compartments and presence of H3K9me3 (Supplementary Figure S7D and E). Using the DamID map of Lamin B1 interactions in fibroblasts (36), we further demonstrate that these regions are largely overlapping with lamina-associated domains (LADs), regions of condensed chromatin that are bound by the nuclear lamina (Supplementary Figure S7F). These results support the view that cell type-specific long range interactions mainly reflect silencing of chromatin domains and the formation of heterochromatin in differentiated cells.

Changes in hierarchy of domains and correlation with LRIs

To study how differentiation affects higher-order interactions between domains and if LRIs could be explained by structural rearrangements within the hierarchy of domains, we developed a method called Simple Hierarchical Pearson Correlation (SHERPA) that detects optimal patterns of chromatin domain aggregation based on correlation of contact profiles. This led to identification of 11–12 levels of aggregation hierarchy within each cell type which are named sequentially SHERPA-domain 1, 2, 3, etc (Figure 4A and Supplementary Table S5). The levels 1–2 showed highest consistency with TADs and represents mostly regions of size <1 Mb (Figure 4A). Since the process of SHERPA domain detection is fundamentally different from the directionality index method, the overlap of the exact locations of domain boundaries as determined by SHERPA and directionality index method (4) applied in HOMER is below expectation (770/3400) when SHERPA is used at the high resolution of 50 kb. This is likely due to technical differences between the methods. However, when we aggregated the Hi-C maps to 150 kb before hierarchical analysis, the agreement between boundary locations defined by HOMER and SHERPA increased to 78% (2650). For this reason, we have focused on the SHERPA hierarchy obtained at the resolution of 150 kb not to focus on technical differences between computational approaches.

To investigate the changes in the architecture of higher-order chromosome folding during differentiation, we compared the hierarchy of SHERPA-domains between ECs and
Figure 3. Inactive regions are interacting with PCH in HUVECs. (A) Heatmap of HUVEC Hi-C interactions (in blue on left) and -log10 P-value (red) for LRIs within the left arm of chromosome 10. Light gray color on the right side of the heatmap represents active (PC1+) compartments and white inactive (PC1−) compartments. Compartments are also depicted by PC1 tracks on the top and on the left side of the heatmap (black = PC1 positive representing active compartment, gray = PC1 negative representing inactive compartment). (B) Permutation test for overlap of PCH and HUVEC-specific LRIs between PC1 (−) TADs. Number of observed overlaps is shown with a green line. Distribution of the overlaps of random interaction pairs and pericentromeric regions is shown by grey bars. (C) UCSC Genome browser image of the left arm of chr10. Active and inactive compartments shown as PC1 values from H1-ESCs, H1-MDCs and HUVECs. Connectivity to left pericentromeric region across the chr10 using 100 kb resolution is displayed as correlations of normalized interaction ratios. Normalized GRO-seq tags for H1-ESCs (green) and HUVECs (magenta) and H3K9me3 ChIP-Seq (blue) are also shown. (D) The average connectivity of PC1 (−) regions with PCH for each chromosome. Values are shown as log2 ratio of observed and expected interactions determined from normalized Hi-C matrix using 100 kb resolution. Chr9 was excluded from the analysis due to large gaps in the hg19 genome at the pericentromeric regions.
Figure 4. Analysis of aggregation of domains into hierarchy. (A) HUVEC Hi-C interaction matrix for chr14. TADs are shown as dark blue triangles above the diagonal and different levels of SHERPA hierarchy shown in colored triangles below the diagonal. Numbers along the axis represent the 150 kb bins. (B) Comparison of SHERPA domain boundaries between HUVECs and H1-ESCs or H1-MDCs shown as Jaccard index values for different domain levels. (C) Hi-C interaction matrices and different levels of SHERPA hierarchy shown for HUVECs (above the diagonal) and H1-MDCs (below the diagonal). Numbers along the axis represent the 150 kb bins. (D) Z-scores derived from permutation tests illustrating the association of reorganized SHERPA domains of different levels with regions undergoing compartmental changes (activated ΔPC1 > 80, inactivated ΔPC1 < −80). P-values < 0.001 for all except for...
the progenitor cells. The results show that 45–75% of domain boundaries are largely the same between cell types (Figure 4B), although some significant hierarchical rearrangements are also seen (Figure 4C). These rearrangements, especially at levels 1–6, were enriched for regions undergoing compartmental switch during differentiation suggesting that hierarchy is affected by the aggregation of active and inactive regions of chromatin (Figure 4D). In line with this, regions showing differential H3K4me1, H3K27ac and H3K27me3 histone modifications between HUVECs and H1-ESCs were enriched in domains undergoing changes in hierarchy (Figure 4E).

We next compared the relationship between the hierarchy of SHERPA-domains and the LRIs. The majority of LRIs fell within mid-levels of SHERPA hierarchy (levels 4–7) in concordance with the LRI distance distribution and domain size distribution (Figure 4F). Overlay of the hierarchy with LRIs exhibited good concordance, suggesting that the aggregation of the domains could be described in a hierarchical framework (Figure 4G). In line with this, the HUVEC hierarchy was able to better predict the HUVEC-specific LRIs than random domains, namely 52% of all LRIs (1704/3252; Figure 4H). Interestingly, also the ESC hierarchy outperformed substantially the random domain hierarchy, suggesting that ESC hierarchy is predictive of EC-specific LRIs (Figure 4H). Taken together, our methodology revealed a hierarchical architecture of domains that can describe the coordination of long range interactions between domains.

**Hypoxia has subtle effects on chromatin dynamics**

We next characterized the dynamics of chromatin contacts after hypoxia signaling in HUVECs, which plays an important role in driving EC migration and angiogenesis. Altogether, 710 genes were found upregulated and 171 genes downregulated on nascent RNA level upon 8h hypoxia stimulus (Supplementary Table S6A; Fold change cutoff 1.7, 0.5 RPKM, FDR < 0.1). Interestingly, ~90% of hypoxia-regulated genes were found within active compartments and 12% of these has undergone a compartmental switch during differentiation. These include the well-established hypoxia-responsive genes vascular endothelial growth factor C (VEGFC), dual specificity phosphatase 6 (DUSP6), C-X-C motif chemokine ligand 8 (CXCL8), lysyl oxidase (LOX) and protein phosphatase 1 regulatory subunit 3C (PPPIR3C). However, we found that hypoxia stimulus itself has very limited effect on chromosome compartments as only three regions showed a change from inactive to active compartment (Figure 5A and Supplementary Table S6B). This is exemplified by a region encoding adenylate kinase 4 (AK4), which is located at a boundary of active and inactive compartment and which is highly induced by hypoxia (Figure 5B). Despite inability to detect compartmental switches, the change in the PC1 value did exhibit a clear correlation with changes in gene expression suggesting that the compartmentalization could contribute to the patterns of gene expression in response to stimuli (Figure 5C).

Previous studies have suggested that during differentiation and responses to extracellular stimuli the positioning of TADs remains largely unchanged whereas interactions within and between the TADs can undergo changes (1,6–9). Similarly, we found that 94% of TAD boundaries are conserved between normoxia and hypoxia (Figure 5B). Also, only a minor increase in intra-TAD interactions in TADs harboring hypoxia upregulated genes was detected (Figure 5D). Furthermore, we analyzed the connectivity of hypoxia-inducible TSS to hypoxia-responsive regulatory elements by using significant interactions detected with 10 kb resolution. As putative regulatory elements we selected HIF1α binding sites, hypoxia-responsive eRNAs and regions showing dynamic changes in H3K4me1 (18) and H3K27ac (17) histone modifications in hypoxic conditions. Interactions connecting hypoxia-responsive TSS to these regulatory regions were found to be enriched in the set of all interactions originating from these TSS, but these contacts were not more frequent in hypoxic conditions (Supplementary Figure S8A). We also utilized the available data on CTCF binding in response to hypoxia that was detected in MCF-7 cells (19). We focused on the CTCF sites residing up- or downstream of genes that were hypoxia-responsive both in HUVECs and MCF-7 cells or near the HIF1α binding sites found in both cell types, and found that there was no change in CTCF binding in response to hypoxia (Supplementary Figure S8B). Altogether, this data suggests that majority of cis-interactions originating from hypoxia-responsive promoters are pre-established, with few changes occurring in response to HIF activation. However, the majority of the hypoxia-activated genes resided within TADs that were bound by HIF1α outside the promoter region and increasing number of HIF1α peaks within a TAD correlated with higher level of gene induction and increase in the number of coregulated genes (Figure 5E and F). By taking into account if there is a HIF1α binding site in the same TAD as the hypoxia-regulated gene, we were able to situate 70% of hypoxia-inducible genes within an interacting hub with HIF1α. Interestingly, the correlation between the HIF1α binding and the fraction of upregulated genes and fold change of induction increased when moving to more active chromatin compartments (Figure 5G) suggesting that these regions are more prone to gene activation (as in Figure 5C) or regulation by HIF1α. To address if this is a common feature of TFs, we studied the correlation between the level of gene activation, PC1 values and binding of signal-dependent TFs nuclear factor kappa B (NF-κB), interferon regulatory factor 1 (IRF1) and nuclear factor (erythroid-

---

level 8, where \( P\)-value = 0.004 for activated and \( P\)-value = 0.1049 for inactivated. (E) Z-scores derived from permutation tests illustrating the association of reorganized SHERPA domains of different levels with regions showing differential H3K4me1, H3K27ac and H3K27me3 histone modifications in HUVECs versus H1-ESCs comparisons (\( P\)-values < 0.001). (F) Bar chart showing the fraction of HUVEC-specific LRIs that reside in various levels of SHERPA hierarchical domains. (G) Contact matrix showing Hi-C interactions and significance of LRIs overlapped with SHERPA-domains (triangles of different colors) for region chr9: 1 − 37,500,000 in HUVECs and numbers along the axis represent the 150 kb bins. (H) Bar charts showing numbers of HUVEC-specific LRIs falling within SHERPA-hierarchy determined for HUVECs, H1-ESCs or random hierarchy (\( P\)-values from binomial distribution).
Figure 5. Hypoxia induces minor changes in chromatin architecture of HUVECs. (A) Scatter plot showing the PC1 values in normoxia and hypoxia for 25 kb genomic windows. Three regions changing the most and encircled in red represent areas overlapping with AK4 gene. (B) UCSC Genome browser.
derived 2-like 2 (Nrf2) in ECs subjected to IL1β, TNFα and oxPAPC treatments (Figure 5H and Supplementary Figure S8C) (16). The results demonstrate that although strongest with HIF1α, all TFs show preferential correlation between the TF binding and gene regulation towards more active compartments. Altogether, this suggests that increasing level of chromatin accessibility is associated with more binding of TFs and greater change in gene expression.

Interestingly, we also detected 57 significant long range interactions specific for hypoxia and 64 LRIs specific for normoxia (Supplementary Table S6C). As expected, TADs engaged in hypoxia-specific LRIs harbored many hypoxia-induced genes exemplified by factors participating in metabolic adaptation to hypoxia, such as hexokinase 2 (HK2), 6-phosphofructo-2-kinase (PFKFB3) and aldolase (ALDOA). In contrast to cell type-specific LRIs, hypoxia-specific LRIs were enriched for active compartments and hypoxia-inducible gene expression and increase in H3K4me1 enhancer mark suggestive of their role in remodeling of active response to stimulus (Figure 5I and data not shown). These results support the view that chromatin interactions detected by Hi-C display only subtle changes in response to hypoxia stimulus which are mostly related to transcriptional changes in the regions carrying hypoxia induced genes.

DISCUSSION

Our work presents dynamic reorganization of the higher order chromatin structure during EC differentiation and hypoxic stress. We find that differentiation induces (i) switching between active and inactive chromatin compartments, (ii) alterations in intra-TAD interactions, (iii) appearance of LRIs between TADs and (iv) changes in the hierarchy of domains. The former two correlated with enrichment of cell type-specific gene expression, active epigenetic marks while being characterized by the binding of cohesin and LDTFs. These two groups of DNA-bound factors have been previously shown to establish super-enhancers often found at genes that define cell identity (27). To this end, we also discovered that EC super-enhancers are enriched in EC-specific active compartments and TADs. These results strongly support a role of super-enhancers and LDTFs in the control of cell type-specific chromatin structures and differentiation.

Previous studies have reported changes in the formation of chromatin loops within TADs or its subcompartments between human cell lines and during somatic cell reprogramming (35,37). The majority of such `loops` are <2 Mb apart and located at the TAD boundaries. Here, we present evidence that cellular differentiation is also characterized by emergence of LRIs between TADs. In striking contrast to chromatin loops (35), the appearance of LRIs was frequently accompanied by repressive chromatin environment. This is in line with the observation that compartment B interactions exhibit higher overall frequency (2). Moreover, a previous study employing 4C and Hi-C in mouse cells observed that the interactions between inactive regions are not as prevalent in pluripotent as in differentiated cells (38), suggesting that the organization of inactive chromatin is not completely established in progenitors. Here, we show that this also holds true genome-wide for the differentiation of human ECs. We demonstrate for the first time that these interactions between inactive regions largely contribute to the formation of cell-type specific LRIs between TADs that could serve as an important co-predictors of gene expression change during differentiation along with PCA-based compartment change. Although the phenomena detected by PCA and LRI detection are partly related, PCA by definition aggregates the similarity of complete interaction profiles while the LRIs are completely local. We do not find LRIs between all active TADs and neither for all inactive TADs, suggesting that they are not describing entirely same phenomenon. Supporting this, we demonstrate that one third of TADs engaged in LRIs fall within negative or positive compartments without a notable difference in PCA between the progenitors and ECs, despite clear changes in gene expression.

We also observed that TADs within HUVEC-specific LRIs were often connecting inactive regions to the PCH. These findings agree with recent observations that lineage commitment correlates with increased connectivity of pericentromere-associated domains (39) and chromatin compartments (40) to heterochromatin and that increased cellular maturation leads to heterochromatin localization to the nuclear periphery (41). Several studies have found that H3K9me3-marked heterochromatin can promote lamina-proximal positioning of chromatin domains (42,43). Supporting this, we observed enrichment for LRIs within inactive regions in IMR90 fibroblasts which were largely over-
lapping with LADs. Altogether, our data suggests that more connective heterochromatin is associated with mesendodermal lineage specification that could contribute to stable transcriptional silencing of genes residing in these domains. Future studies on the mechanisms by which heterochromatic regions are established and what units of chromatin topology demarcate these changes is needed to explain the changes in LRI and gene expression during differentiation.

Our results also demonstrate that the majority of LRIs fall within larger hierarchical structures that are comprised of multiple domains. The statistical analyses support the notion that these structures are already largely organized in progenitors and undergo limited rearrangement during differentiation. Recent study on TAD hierarchy concluded that only ~20% of TAD aggregates undergo structural reorganization upon differentiation (44). This is in agreement with our analysis of domain hierarchy derived with SHERPA, where we see that the vast majority of LRIs consistent with HUVEC hierarchy are also consistent with the hierarchy derived from the H1-ESC Hi-C data. Nevertheless, we can detect rearrangements of hierarchy at multiple levels that at least partly appear to be driven by compartmental switching. This is consistent with the model where large-scale hierarchy is pre-organized already in the stem cells and further changes in contacts within hierarchy are mostly determined by locally targeted changes in chromatin contacts as discussed above. Experimental validation of this model is beyond the scope of this work, but the statistical data strongly suggests such a scenario. Another important question is how such primary hierarchy arises. While they are far from being conclusive, the recent results from simulations (45) suggest that it might be enough to know the positions of active and inactive chromatin regions and the physical properties of chromosomes to fold the chromatin into a shape consistent with the experimental Hi-C maps. On the other hand, other recent studies suggests that chromatin contacts are more variable and frequently occur across TAD borders in single cells and TADs only emerge when averaged over a population of cells (46). Therefore, we should not expect each cell to fold its genome into the same hierarchical aggregation. Whether this holds true, warrants further investigations; nonetheless, our findings regarding LRIs occurring within the hierarchy of TADs should be interpreted in a similar, statistical sense, i.e. as events likely to occur more frequently than other contacts.

Finally, we interrogated the effect of hypoxia on the chromatin organization and coregulation of hypoxia-inducible genes genome-wide. In striking contrast to differentiation associated changes in chromatin contacts, we present evidence that hypoxia brings about subtle changes at the level of chromatin compartmentalization, intra-TAD interactions, connections between hypoxia-responsive regulatory regions and LRIs between TADs and these changes are largely limited to active compartments. This is in line with recent capture-C study focusing on 14 HIF-binding sites which demonstrated that HIF largely operates on pre-existing chromatin interactions (19). Previous studies have also demonstrated that coregulation of genes is prominent within TADs or even within chromatin compartments (3,7,47). In some cases coregulation has been explained by sharing of the same TFs as has been demonstrated for pluripotency factors, KLF1 and LSD1 (38,48,49). The observation that 70% of hypoxia upregulated genes fall within HIF1α containing TAD, provides evidence that HIF1α could also function in a similar fashion to mediate coordinated transcriptional control. Here, chromatin interactions could favor coregulation of distant genes by increasing the local amount of HIF1α required for their induction. This is supported by our observation that increasing level of active chromatin compartmentalization or accessibility correlates with more binding of TFs and greater change in gene expression. Interestingly, long-range chromosomal contacts at multigene complexes have also been shown to possess hierarchical structure with dominant and subordinate members (50). It remains to be studied whether HIF1α participates in the establishment or modulation of such hierarchy.

In conclusion, our results imply that transcriptional regulation during cellular differentiation and in response to stimuli should be considered in the context of the three-dimensional organization of the chromatin at several levels. Notably, we demonstrate that the appearance of EC-specific LRIs between TADs largely involves interactions between heterochromatic regions. Despite these rearrangements, the majority of chromatin domains fall within a pre-established hierarchy already present in progenitor cells. Finally, our results suggest a model in which lineage-specific chromatin interactions associated with compartment state changes are establishing the active regions of the genome that can respond to environmental stimuli. Shedding light into the combinatorial associations of lineage-determining and signal-dependent TFs at multiple genes within context of long range chromatin interactions or hierarchical chromatin hubs could help us gain further insight into the regulation of cell- and stimulus-specific gene expression patterns. This also has likely implications in human disease, where perturbations within any of the coregulated hubs could dysregulate the transcription of other members of the complex.

**DATA AVAILABILITY**

Public datasets used in this study are listed in Supplementary Table S7. Sequencing data from this study has been submitted to NCBI Gene Expression Omnibus under accession number GSE94872.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

We thank the Sequencing Service GeneCore Sequencing Facility (EMBL, http://www.genecore.embl.de) for NGS library sequencing and University of Eastern Finland Bioinformatics Center for server infrastructure. We are thankful to Assistant Professor Casey Romanoski (University of Arizona) for providing processed data files for RNA- and ChIP-Seq experiments from ECs subjected to II1β, TNFα and oxPAPC treatments (16).
REFERENCES

1. Dixon,J.R., Jung,L., Selvaraj,S., Shen,Y., Antosiewicz-Bourget,J.E., Lee,A.Y., Ye,Z., Kim,A., Rajagopal,N., Xie,W. et al. (2015) Chromatin architecture reorganization during stem cell differentiation. Nature, 518, 331–336.

2. Lieberman-Aiden,E., van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,J.P., Dorschner,M.O. et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science, 326, 289–293.

3. Kaikkonen,M.U., Niskanen,H., Romanoski,C.E., Kansanen,E., Kivela,A.M., Aaltonen,L.J., Heintz,S., Benner,C., Glass,C.K. and Yla-Herttuala,S. (2014) Control of VEGF-A transcriptional programs by pausing and genomic compartmentalization. Nucleic Acids Res., 42, 12570–12584.

4. Dixon,J.R., Selvaraj,S., Yue,F., Kim,A., Li,Y., Shen,Y., Hu,M., Liu,J.S. and Ren,B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature, 485, 378–380.

5. Jin,F., Li,Y., Dixon,J.R., Selvaraj,S., Ye,Z., Lee,A.Y., Yen,C.A., Schmitt,A.D., Espinoza,C.A. and Ren,B. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature, 503, 290–294.

6. Chandra,T., Ewels,P.A., Schoenfelder,S., Furlan-Magari,B., Wingett,S.W., Kirschner,K., Thuery,J.Y., Andrews,S., Fraser,P. and Reik,W. (2015) Global reorganization of the nuclear landscape in senescent cells. Cell Rep., 10, 471–483.

7. Le Dily,F., Bau,D., Pohl,A., Vicent,G.P., Serra,F., Soronellas,D., Castellano,G., Wright,R.H., Ballare,C., Filion,G. et al. (2014) Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes Dev., 28, 2151–2162.

8. Rafique,S., Thomas,J.S., Sproul,D. and Bickmore,W.A. (2015) Estrogen-induced chromatin decondensation and nuclear re-organization linked to regional epigenetic regulation in breast cancer. Genome Biol., 16, 145.

9. Li,L., Lyu,X., Hou,C., Takenaka,N., Nguyen,H.Q., Ong,C.T., Cubenas-Potts,C., Hu,M., Lei,E.P., Bosco,G. et al. (2015) Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. Mol. Cell, 58, 216–231.

10. Heinz,S., Benner,C., Spann,N., Bertolin,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell, 38, 576–589.

11. Krueger,F., Andrews,S.R. and Osborne,C.S. (2011) Large scale loss of data in low-diversity illumina sequencing libraries can be recovered by deferred cluster calling. PLoS One, 6, e16607.

12. Blankenberg,D., Gordon,A., Von Kuster,G., Coraor,N., Taylor,J., Nekrutenko,A. and Galaxy,T. (2010) Manipulation of FASTQ data with Galaxy. Bioinformatics, 26, 1783–1785.

13. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol., 10, R25.

14. Trapnell,C., Pachter,L. and Salzberg,S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics, 25, 1105–1111.

15. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26, 139–140.

16. Hogan,N.T., Whalen,M.B., Stolze,L.K., Hadeki,N.K., Lam,M.T., Springstead,J.R., Glass,C.K. and Romanoski,C.E. (2017) Transcriptional networks specifying homeostatic and inflammatory programs of gene expression in human aortic endothelial cells. Cell, 6, 22526.

17. Inoue,T., Kohro,T., Tanaka,T., Kakuk,M., Li,G., Poh,H.M., Mimura,I., Kobayashi,M., Taguchi,A., Maegima,T. et al. (2014) Cross-enhancement of ANGPTL4 transcription by HIF1 alpha and PPAR beta/delta is the result of the conformational proximity of two response elements. Genome Biol., 15, R63.

18. Mimura,I., Nangaku,M., Kanki,Y., Tsutsui,S., Inoue,T., Kohro,T., Yamamoto,S., Fujita,T., Shimamura,T. and Suyanjo,H. (2012) Dynamic change of chromatin conformation in response to hypoxia enhances the expression of GLUT3 (SLC2A3) by cooperative interaction of hypoxia-inducible factor 1 and KDM5A. Mol. Cell. Biol., 32, 3018–3032.

19. Platt,J.L., Salama,R., Smythies,J., Choudhry,H., Davies,J.O., Hughes,J.R., Ratcliffe,J.L. and Ratcliffe,J.L. (2016) Capture-C reveals preformed chromatin interactions between HIF-binding sites and distant promoters. EMBO Rep., 17, 1410–1421.

20. Nep,S., Kuehn,M.S., Reynolds,A.P., Haugen,E., Thurman,R.E., Johnson,A.K., Rynes,E., Maurano,M.T., Vierstra,J., Thomas,S. et al. (2012) BEDOPS: high-performance genomic feature operations. Bioinformatics, 28, 1919–1920.

21. Schodel,J., Oikonomopoulous,R., Ragozzin,J., Pugh,C.W., Ratcliffe,J.L. and Pohl,D.R. (2011) High-resolution genome-wide mapping of HIF-binding sites by ChiP-seq. Blood, 117, e207–e217.

22. Elvidge,G.P., Glenn,L., Appelhoff,R.J., Ratcliffe,J.L. and Giedlme,J.M. (2006) Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways. J. Biol. Chem., 281, 15215–15226.

23. Lin,Y.C., Benner,C., Mansson,R., Heinz,S., Miyazaki,K., Miyazaki,M., Chandra,V., Bossen,C., Glass,C.K. and Murre,C. (2012) Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. Nat. Immunol., 13, 1196–1204.

24. Imakaev,M., Fudenberg,G., McCord,R.P., Naumova,N., Goloborodko,A., Lajoie,B.R., Dekker,J. and Mirny,L.A. (2012) Iterative correction of H-Ci data reveals hallmarks of chromosome organization. Nat. Methods, 9, 999–1003.

25. Quinan,A.R. and Hall,M.I. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics, 26, 841–842.

26. Gel,B., Diez-Villanueva,S., Serra,E., Buschbeck,M., Peinado,M.A. and Malinvern,B. (2016) regioR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. Bioinformatics, 32, 289–291.

27. Huisz,D., Abraham,B.J., Lee,T.I., Lau,A., Saint-Andre,V., Sigova,A.A., Hoke,H.A. and Young,R.A. (2013) Super-enhancers in the control of cell identity and disease. Cell, 155, 934–947.

28. Elcheva,I., Brok-Volchanskaya,V., Kumar,A., Liu,P., Lee,J.H., Tong,L., Vodyanik,M., Stewart,D., Kyba,M. et al. (2014) Direct induction of haemoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nat. Commun., 5, 4372.

29. Abogesaisa,I., Shimojo,H., Sahin,S., Kondo,A., Harshbarger,J., Lizio,M., Hayashizaki,Y., Carninci,P., consortium,F., Forrest,A. et al. (2016) FANTOM5 transcriptome catalog of cellular states based on就行了。
physically links cohesin to chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 8309–8314.

32. Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J. et al. (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6456–E6465.

33. Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van I.W.F. et al. (2014) Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 996–1001.

34. Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Lander, E.S., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. et al. (2015) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.

35. Towbin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P. and Gasser, S.M. (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell*, **150**, 934–947.

36. Zhu, Y., Khanna, N., Alvikas, J. and Belmont, A.S. (2013) beta-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. *J. Cell Biol.*, **203**, 767–783.

37. Krijger, P.H., Di Stefano, B., de Wit, E., Limone, F., van Oevelen, C., de Boer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van I.W.F. et al. (2014) Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 996–1001.

38. Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Lander, E.S., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. et al. (2015) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.

39. Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Verstegen, M., Zhu, Y., Schell, Y., Vermeulen, C., de Laat, W. and Graf, T. (2016) Cell-of-origin-specific 3D genome architecture acquired during somatic cell reprogramming. *CellStem.*, **19**, 25–36.

40. Towbin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P. and Gasser, S.M. (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell*, **150**, 934–947.

41. Bian, Q., Khanna, N., Alvikas, J. and Belmont, A.S. (2013) beta-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. *J. Cell Biol.*, **203**, 767–783.

42. Fraser, J., Ferrai, C., Chiariello, A., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Lander, E.S., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. et al. (2015) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.

43. Zhu, Y., Gong, K., Denholtz, M., Chandra, V., Kamps, M.P., Alber, F. and Murre, C. (2017) Comprehensive characterization of neutrophil genome topology. *GenesDev.*, **31**, 141–153.

44. Toupin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P. and Gasser, S.M. (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell*, **150**, 934–947.

45. Briand, O., Kochar, S., Nunez, E., Cardamone, D., Hult, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G. et al. (2008) Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 19199–19204.

46. Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J. et al. (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6456–E6465.

47. Kuwano, Y., Kwon, Y.S., Nunez, E., Cardamone, D., Hult, K.R., Ohgi, K.A., Garcia-Bassets, I., Rose, D.W., Glass, C.K., Rosenfeld, M.G. et al. (2008) Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 19199–19204.