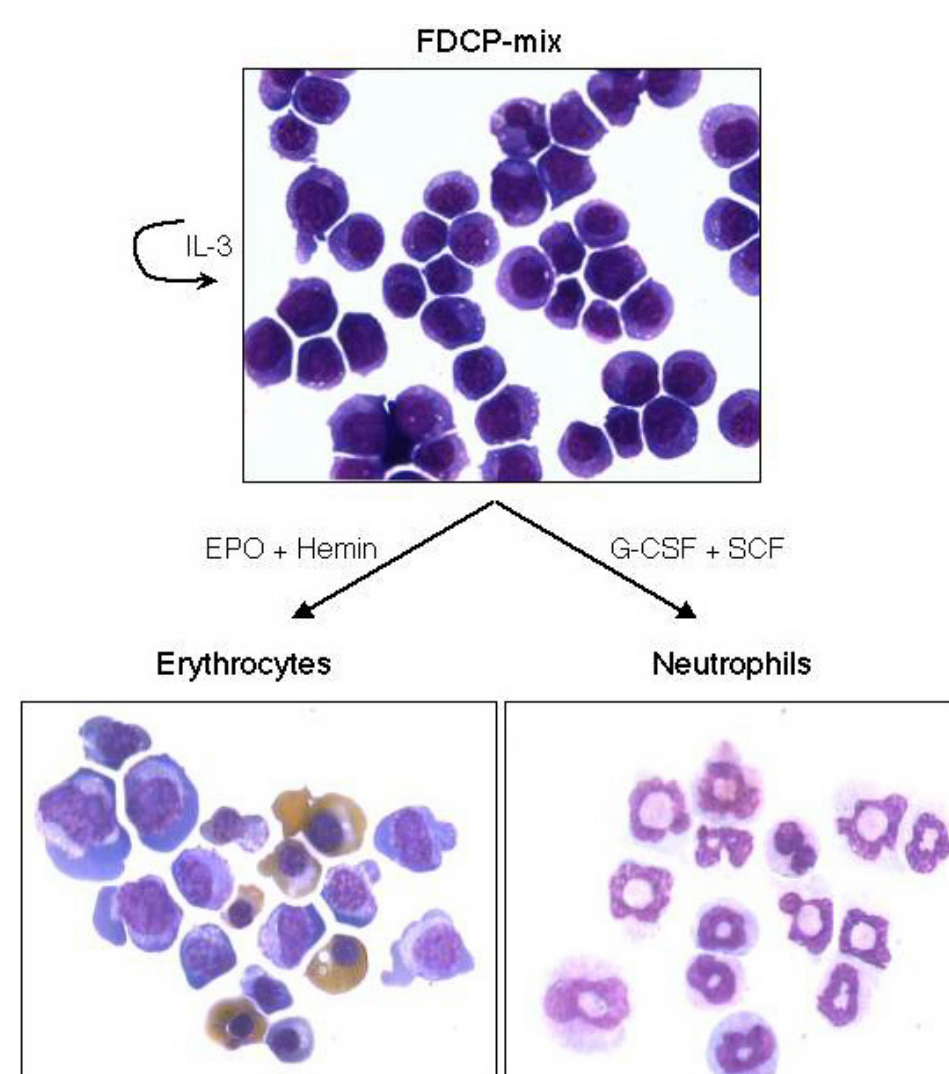


Steven T. Kosak<sup>1</sup>, Sam V. Alworth<sup>2</sup>, David Scalzo<sup>1</sup>, Fusheng Li<sup>1</sup>, Stephanie Palmer<sup>1</sup>, James S.J. Lee<sup>2</sup>, and Mark Groudine<sup>1</sup>. <sup>1</sup>Division of Basic Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, Washington 98109, USA. <sup>2</sup>SVision LLC, Bellevue, WA 98006.

## Abstract

To fully understand how such basic nuclear processes as transcription are regulated, it is necessary to determine their spatial organization in the nucleus. For example, many gene loci—such as  $\beta$ -globin—are specifically positioned within nuclei and looped away from their chromosome territory (CT) according to their lineage-specific activity. Additionally, expressed genes from a single CT have been shown to share transcription factories, and the regulation of genes significant in T cell differentiation is in part due to their physical interaction. These observations have rekindled interest in a long-established question in the study of nuclear organization: do chromosomes have defined positions relative to each other? Although examples of cell-specific organizations of particular chromosomes have been recently demonstrated in vertebrates, a common basis for non-random chromosome organization has yet to be elucidated. Using a novel software tool that performs spatial pattern recognition (SVCell™ recognition software), we have studied the simultaneous organization of all chromosomes during cellular differentiation in a murine model of hematopoiesis. We show that homologous chromosomes have a significant propensity to be proximal during differentiation, and that this organization is related to the distribution of co-regulated genes along chromosomes. These data indicate that proximity, both in the form homologous association and gene distribution, may be the basis for organizing the genome during cellular differentiation.

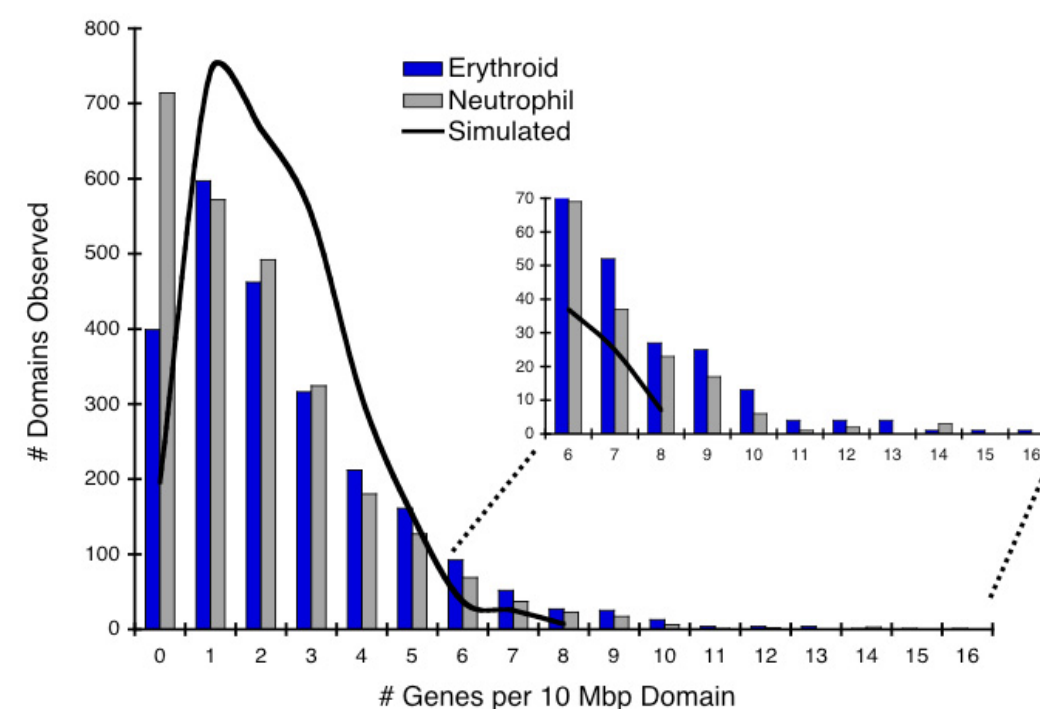
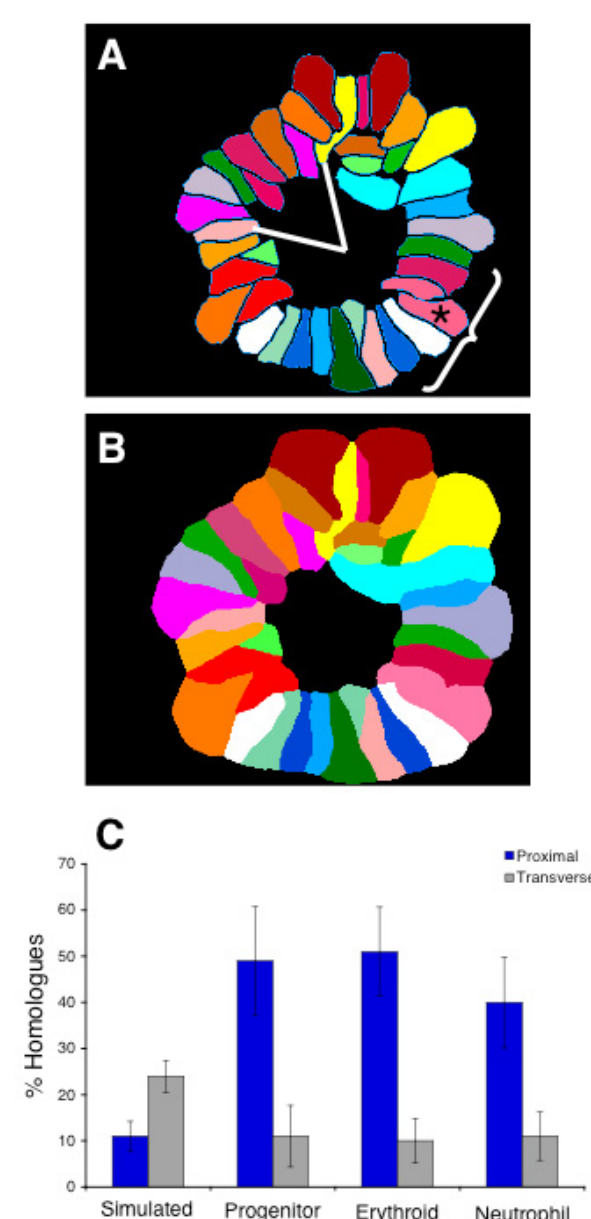
## Results



**Fig. 1. FDCP-mix experimental system.** Self-renewing FDCP-mix multipotential progenitors can be differentiated into the erythroid and neutrophil lineages with the appropriate cytokines (all cells have been hematoxylin and eosin stained).

## Fig. 2. Homologous chromosomes associate during differentiation.

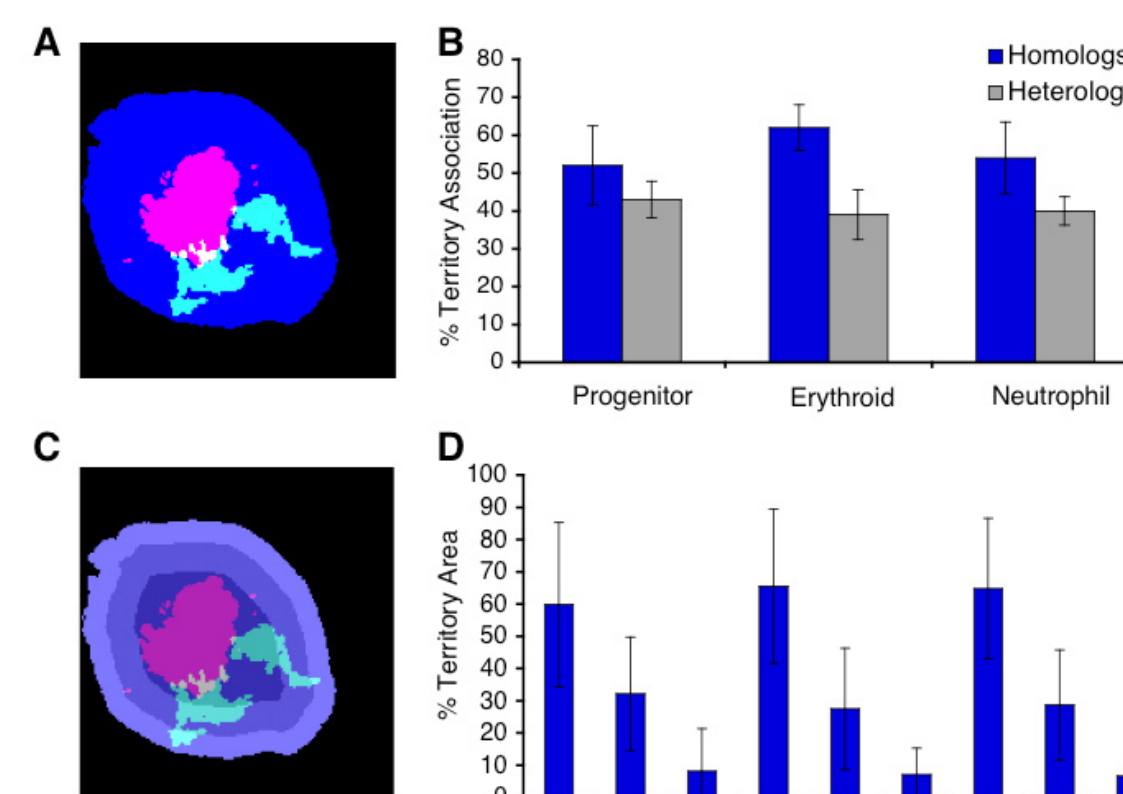
(A) We have studied chromosome organization by simultaneously detecting all of the chromosomes in lineage-specific rosettes using spectral karyotyping (SKY). We performed two types of analyses on at least 30 rosettes from each cell type and a simulated—random—rosette data set: 1) homologous sister chromatid pairs (homologous chromosomes) were assayed for proximity by determining the frequency of their being within two chromosomes of each other (\* indicates the assayed chromosome, the bracket identifies the region of proximity as three chromosomes on either side); 2) homologous chromosomes were scored for being transverse by determining their frequency of being across the center of a rosette in a 60° angle window (encompassing ~8 chromosomes). (B) Illustration of the distance constrained zone of influence (ZOI) operation performed by SVCell. Proximity of chromosome territories (CTs) is calculated automatically by performing a ZOI based region partition around each CT. The ZOI operation creates an unambiguous representation across which proximity transitions can be determined. If at least one pixel of two chromosomes' partitioned regions touch, then they are considered adjacent. (C) Spatial pattern rules were created in SVCell and used to measure proximal and transverse associations for all chromosomes in each of the rosettes for the three lineages and the simulated data set. The bar graph depicts these results as a mean of the chromosomal data for each lineage. Lines represent the standard deviation of the measurements for individual chromosomes.



**Fig. 3. The erythroid and neutrophil transcriptomes are clustered.** We determined the linear genomic positions of the ~11,100 genes with unique GeneBank accessions from the Affymetrix GeneChip® MG-U74Av2. We compared the distribution of the lineage transcriptomes to the simulated data set by sliding window analysis, with a 10 megabase window moved in 1 megabase steps through the genome. A proportions test indicates a significant difference between the transcriptomes and the microarray gene distribution ( $P < 1.2e-05$ ), but not between the simulated and microarray ( $P < 0.22$ ).

**Table I. The association of homologues and co-regulated gene distribution are related.** A Kruskal-Wallis test (a one-way ANOVA by ranks) was performed by comparing the proximal homologous chromosome data set for each lineage to their transverse chromosome data sets, chromosomal distribution of co-regulated genes (for each cell type), overall chromosome size (as a ratio of the total genome), and gene density (as a ratio of total gene number). NS indicates a non-significant difference.

Cell Type	Transverse Chromosomes	Co-Regulated Genes	Chromosome Density	Chromosome Size
Progenitor	P<0.001	NS	P<0.01	P<0.01
Erythroid	P<0.001	NS	P<0.01	P<0.01
Neutrophil	P<0.001	NS	P<0.01	P<0.01



## Fig. 4. Evidence from interphase nuclei indicates the association of homologues.

(A) Erythroid nucleus counterstained with DAPI, and chromosome territories (CTs) 11 (red) and 2 (green) detected by FISH. (B) Bar graph of results for the analysis of CTs for chromosomes 2, 3, 4, 5, 6, 7, 11, 12, 14, 17, 19 (homologous association) and pairwise analysis of chromosomes 2-11, 3-6, 4-5, 7-19, 12-14, 17-19 (heterologous association) as a mean for all analyzed chromosomes in the three lineages. Lines indicate standard deviations. (C) Erythroid nucleus (from B) with the nuclear area divided into three shells, inner, middle and outer, of equal area (D) Bar graph of results from the three lineages for the analysis of positioning of CT areas (all chromosomes from B) in the three regions. Lines indicate standard deviation.

## Conclusion

### Fig. 5. Implications for the nuclear organization of gene activity.

At left is a SKY image of an interphase nucleus from a hematopoietic progenitor. Evidence supports the idea that a nuclear topology exists that ensures the transcriptional program (or transcriptome) that gives rise to or maintains a given cell type. Therefore, the SKY image may represent the appropriate nuclear topology for the progenitor's transcriptome. Analysis of the linear distribution of gene activity from the FDCP-mix progenitor to the erythroid and neutrophil lineages reveals a non-random clustering of genes. Co-regulated gene clusters from different genomic regions may be proximal in the nuclear volume, as depicted in the theoretical magnification at right, to facilitate their transcriptional regulation. Gene clusters may take advantage of protein concentrations (or in fact be the basis for them), which are exemplified by the various types of nuclear bodies found in the nucleus. This model for nuclear organization is supported by our evidence for the non-random association of chromosomes at the prometaphase rosette.

