

H3K4me3 Breadth Is Linked to Cell Identity and Transcriptional Consistency

B erence A. Benayoun, Elizabeth A. Pollina, Duygu Ucar, Salah Mahmoudi, Kalpana Karra, Edith D. Wong, Keerthana Devarajan, Aaron C. Daugherty, Anshul B. Kundaje, Elena Mancini, Benjamin C. Hitz, Rakhi Gupta, Thomas A. Rando, Julie C. Baker, Michael P. Snyder, J. Michael Cherry, and Anne Brunet*

*Correspondence: anne.brunet@stanford.edu
<http://dx.doi.org/10.1016/j.cell.2015.10.051>

(Cell 158, 673–688, July 31, 2014)

Our paper reported that broad H3K4me3 domains in a given cell type are associated with genes that are important for the identity/function of that cell type and that they are associated with increased transcriptional consistency, but not increased expression.

It has come to our attention that we made a programming error in the code used to generate Figure S1J. When the code is corrected, the top 5% broadest H3K4me3 domains display a statistically significant increased expression compared to the rest of the distribution (see corrected Figure S1J below). In addition, if one uses the rank-based Spearman correlation instead of the Pearson correlation we had used, H3K4me3 breadth exhibits a positive correlation with gene expression. Thus, the correct conclusion is that broad H3K4me3 domains are, on average, more expressed than non-broad H3K4me3 domains. This error does not affect our conclusions that H3K4me3 breadth is associated with cell identity and transcriptional consistency. However, we acknowledge that the increased transcriptional consistency of genes marked by broad H3K4me3 domains could be due to their increased average expression, as normalized transcriptional variability and average expression have been observed to be anti-correlated.

The corrected Figure S1J is shown below. The text changes are as follows, with additions in bold and deletions in bracketed italics:

Summary:

“Indeed, genes marked by the broadest H3K4me3 domains exhibit enhanced transcriptional consistency **and** [rather than] increased transcriptional levels.”

Page 674, second paragraph of Results:

“H3K4me3 breadth **quantiles** did not **linearly** correlate with mRNA levels (Figures 1D and 1E, **Pearson correlation**). **However, H3K4me3 breadth showed positive rank correlation with mRNA levels (R~0.19-0.31, Spearman correlation). In addition, the top 5% broadest H3K4me3 domains were more highly expressed on average compared to the rest of the distribution (Figure S1J)** [even when comparing the most extreme example to the rest of the distribution (Figure S1J)]. Thus, broad H3K4me3 domains are present in many cell types across taxa but cannot be explained as simple readouts of promoter complexity [or high expression levels].”

Figure 1 title:

“Breadth Is an Evolutionarily Conserved Feature [that Is Not Predictive of Expression Levels]”

After we identified this programming error, we had the entire manuscript and lines of code independently scrutinized. This process identified the following inadvertent errors that do not affect our conclusions but that we would like to correct.

In Figures 1E and S6A, there was an incorrect attribution of datasets (C2C12 myotubes for myoblasts and H1 hESC population for single cells). The corrected Figures 1E and S6A are shown below. The conclusions are not changed.

In Figures 7D, 7E, 7G, and S7L, the statistical analyses were done using two different tests (one-sided one-sample and one-sided two-sample Wilcoxon tests), but only one set of p values was reported in the original panels, and the corresponding statistical tests were not appropriately described. Results from both tests are shown in updated Figures 7D, 7E, 7G, and S7L. Upper p values (7D, 7G), black lines (7E, S7L): one-sided two-sample Wilcoxon tests for increased variability between genes with maintained versus changed H3K4me3 breadth. Lower p values (7D, 7G), gray lines (7E, S7L): one-sided one-sample Wilcoxon tests for increased variability between genes with changed H3K4me3 breadth versus the expectation of no change in variability. The overall conclusions are not changed.

We sincerely regret these errors and apologize for any inconvenience they may have caused. We would also like to thank Wei Li and Kaifu Chen from the Baylor College of Medicine for alerting us to the discrepancy between the Pearson and Spearman correlation results and for helping us to identify the error in Figure S1J.

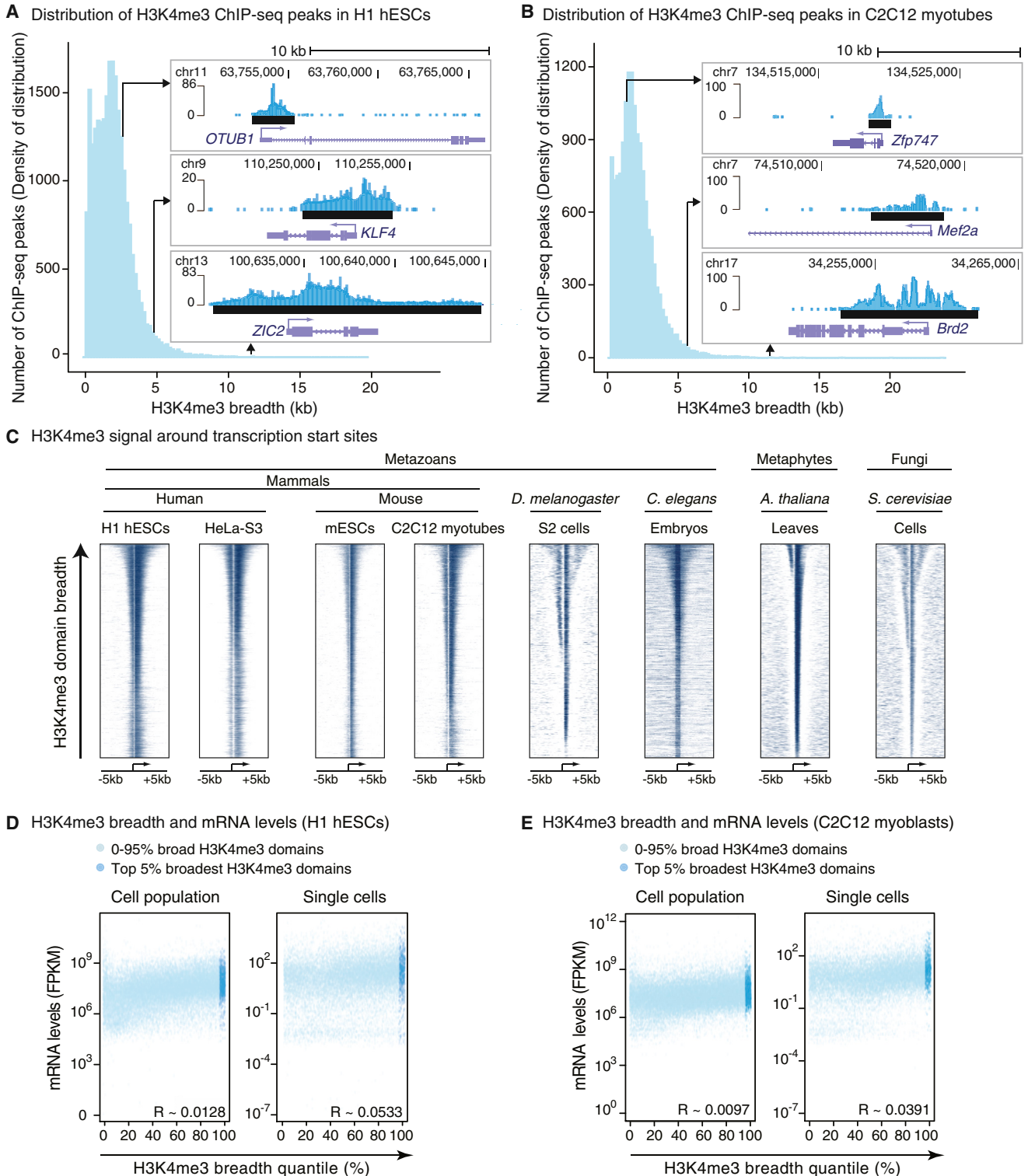


Figure 1. H3K4me3 Breadth Is an Evolutionarily Conserved Feature

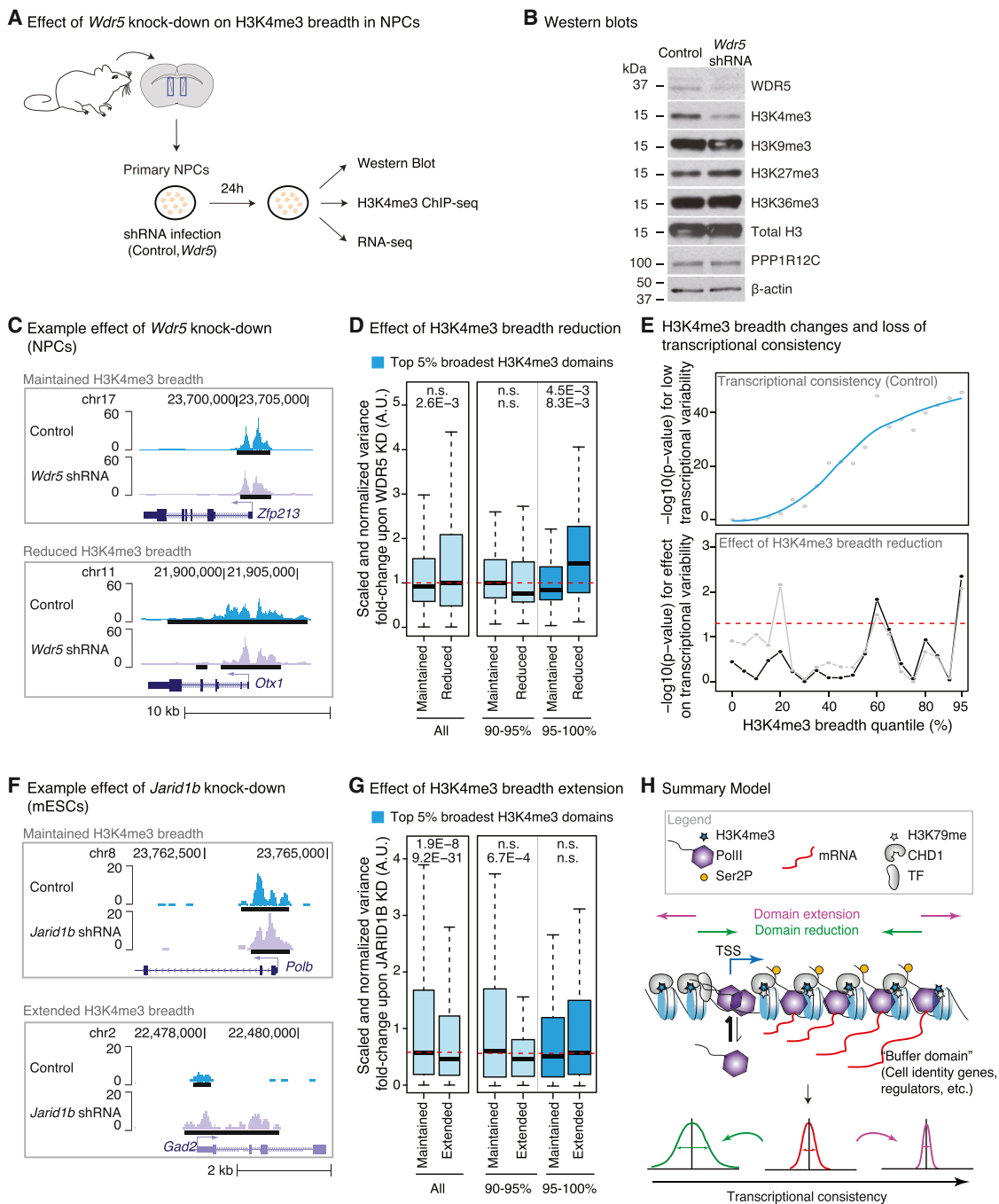


Figure 7. Experimental Perturbation of H3K4me3 Breadth Results in Changes to Transcriptional Consistency

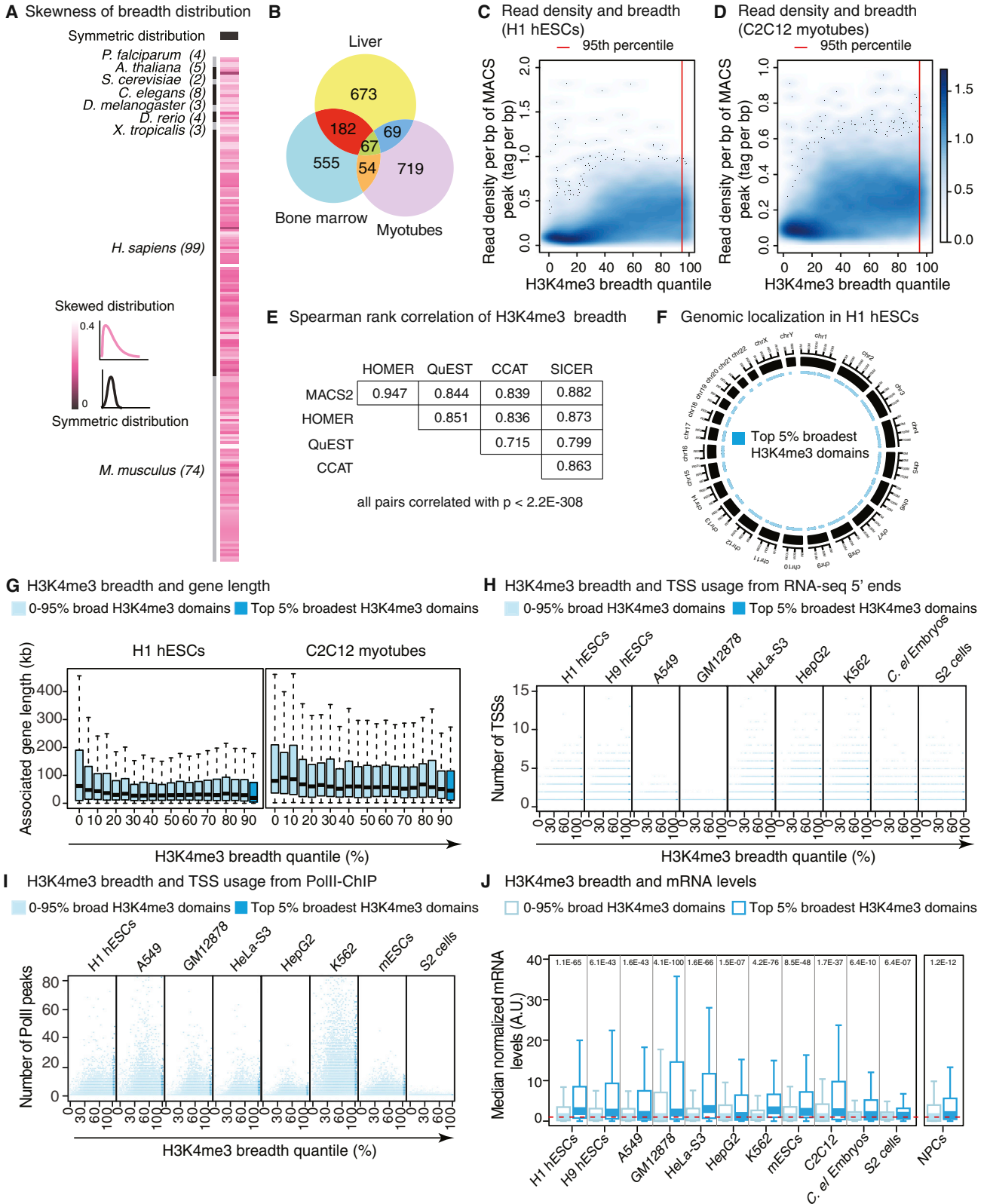
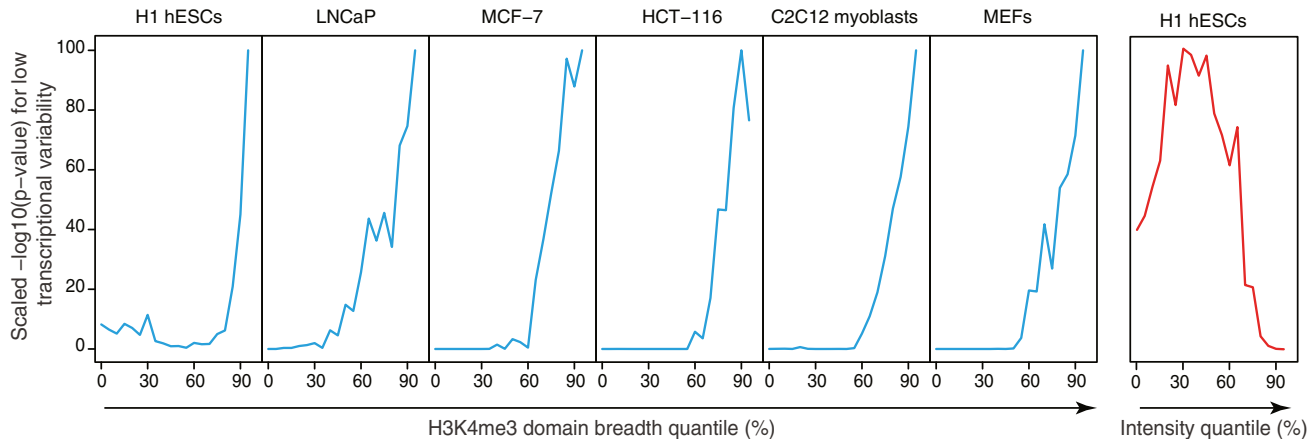
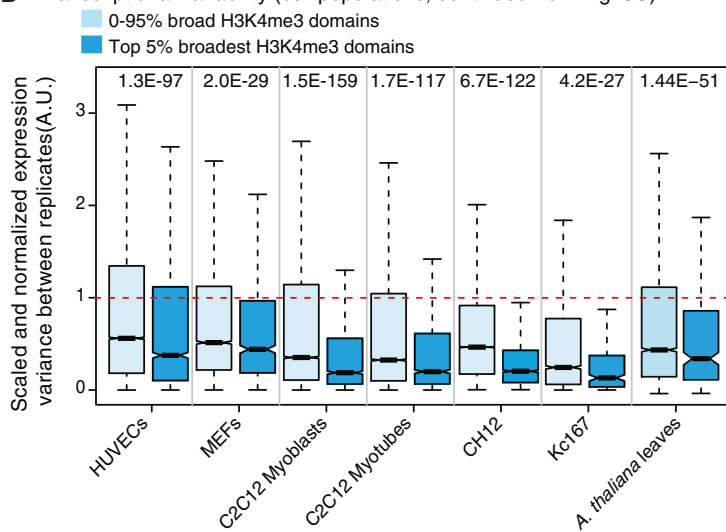


Figure S1. Broad H3K4me3 Stretches Are Present in Different Cell Types and Organisms but Are Independent of Signal Intensity, Promoter Architecture, Gene Length, and Genomic Location, Related to Figure 1

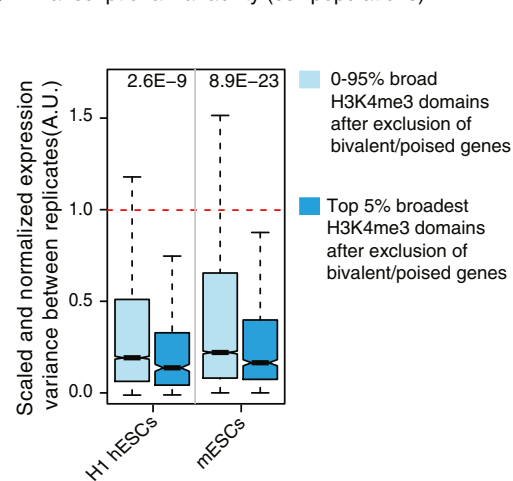
A H3K4me3 breadth and transcriptional variability (single cells)



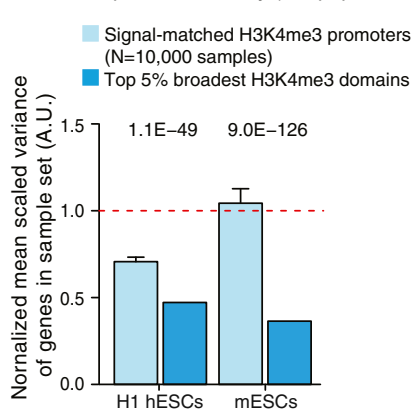
B Transcriptional variability (cell populations, continued from Fig. 6C)



C Transcriptional variability (cell populations)



D Transcriptional variability (cell populations)



E H3K4me3 breadth and transcriptional consistency (cell populations, nascent RNA)

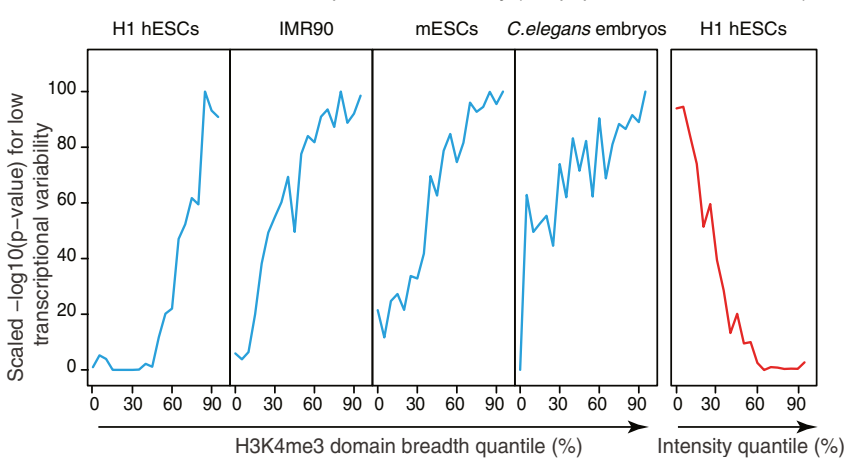


Figure S6. H3K4me3 Breadth Is Associated with Increased Transcriptional Consistency, Related to Figure 6

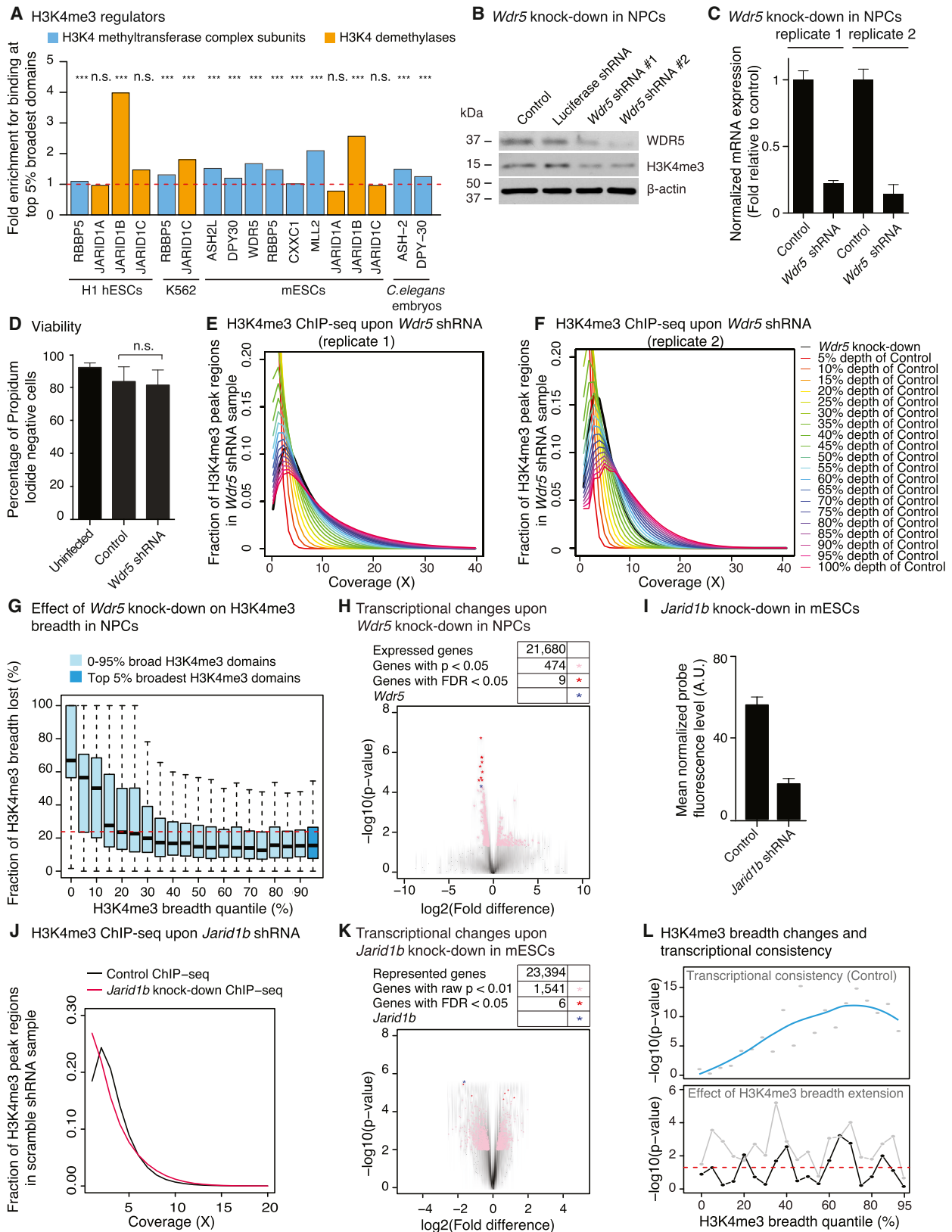


Figure S7. Effect of H3K4me3 Regulators on H3K4me3 Breadth and Transcriptional Consistency, Related to Figure 7