SUPPLEMENTARY MATERIAL

# Multi-Mappability Signal Aggregations

We used the multi-mappability signal profiles generated by MUSIC (See Methods) and aggregated the signal on different regions. Figure S1 shows the aggregation plots of multi-mappability signal over protein coding gene promoters (TSS +/- 2.5 kbs), randomly selected set of introns and exons, and randomly selected regions on the human genome. Promoters, exons, and the regions downstream of TSS into the first exon show significantly higher mappability compared to random regions. In addition, introns show slightly higher mappability compared to random regions and exons show are much more mappable than random regions. Transcription start sites and mid points of exons show almost the same amount of average multi-mappability, 1.2 reads.

# Motivational Example for Multiscale Decomposition

Figure S2 illustrates the 4 level multiscale decomposition of histone modification read depth signal profile. The minima and maxima of the decomposition are indicated on the signal. As the scale level increases, the large scale structures are revealed in the signal. The 3 broadly enriched regions are revealed at the 4th scale as 3 “blobs”, which can be identified as the regions between local minima regions. The smaller scales contain more detailed structure in the signal.

# Statistics on ERs Identified by Different Methods

Table S1 shows the count and coverage for the enriched regions (ERs) for H3K36me3 datasets identified by different methods.

# Fraction of CTCF Peaks with CTCF Motifs

Table S2 shows the fraction of CTCF peaks (for each compared method) for which there is a known CTCF motif within the 150 bp vicinity of the summit reported by the methods.

# Multi-mappability Signal Aggregation on MACS Specific Troughs

For comparing the regions that are merged by MACS and not merged by MUSIC, i.e., the troughs in the signal specific to MACS, we used the H3K36me3 ERs identified by each method. We first overlapped the ERs that are called in both. Then we subtracted the overlapping MACS ERs from the overlapping MUSIC ERs to generate a set of ERs that are not called by MACS but are called by MUSIC. These remaining regions are used as the troughs specific to MACS. Then, to evaluate the mappability of these MACS specific troughs, we aggregated the multi-mappability signal on these regions. For aggregation, we use the center point of each trough and extended the region 2500 base pair to each side. Finally, we generated set of control regions by translating all the MACS specific troughs 10000 base pairs to right. Figure S6 shows the aggregation plots on the MACS specific troughs and for the control regions. It can be seen that MACS specific troughs show a significant decrease in the mappability.