SUPPLEMENTARY MATERIAL

Origin distributions in open and closed chromatin, as defined by Hi-C Eigenvector (HCE) analysis. As an independent measure of any global relationships among origins and epigenetic attributes in GM06990, we compared origin distributions to domains defined by the Hi-C Eigenvector (HCE), which measures the physical interaction between genomic sequences and, among other things, reflects the relative degree of chromatin compaction in various regions: positive values (HCE+) are associated with open chromatin and negative values (HCE-) with closed chromatin (Lieberman-Aiden et al. 2009b). When replication timing profiles for selected murine and human cell lines were compared to a number of epigenetic features, HCE values displayed the most significant correlations by far (Ryba et al. 2010). Indeed, when we compared the global replication timing (TR50) values versus HCE values for GM06990 (both for 1 Mb windows), the correlation coefficient was also extremely high (-0.74; Table SVIIIA). Note that the negative sign results from our convention that high TR50 values correspond to late replication, which tends to occur in regions with the lowest HCE values (Table SVIIIA).

On the other hand, the correlation between origin *efficiencies* (RDs) and HCE values in the genome as a whole is poor (-0.04; Table SVIIIA6), even though there is a modest but steady downward trend in RD as replication time increases, and in HCE- compared to HCE+ (Fig. S3). The poor correlation is due to the high variability of RD levels (especially in early-replicating regions) as shown in Fig. 1B. The correlation between origin *density* and HCE values was better but modest (0.26; Table SVIIIA). Note that *density* here corresponds to the combined length of bubbles (in bp) divided by 1 Mb (i.e., the same window used to calculate HCE). The likely explanation for these observations is that origin densities are highest in early-replicating regions, drop in mid-replicating regions and increase again in late-replicating regions (Figure S4 and Table SIV). Conversely, HCE values tend to monotonically decrease from relatively large positive values in early-replicating regions to relatively large negative values in late-replicating regions (not shown). Indeed, when the HCE sign is disregarded, the correlation coefficient between origin densities and the absolute values of HCE rises to 0.5 (Table SVIIIA4). Similarly, correlation coefficients between HCE and origin density in early-, mid- and late-replicating regions are

0.58, 0.19, and -0.48, respectively (Table SVIIIA). These values are consistent with comparisons between origin and epigenetic factor densities in early- and late-replicating regions summarized in Table SVIIIB-F).

To determine the nature of origin distributions in loci that do not conform to the simplest paradigm (i.e., HCE+ regions residing in late-replicating domains and HCE- regions in early-replicating domains), we further divided early-, mid-, and late-replicating regions into HCE+ or HCE- categories. The genomic coverage of HCE+ and HCE- regions is roughly the same, occupying 1,439 and 1,443 Mb, respectively. Of the 123,412 bubble-containing fragments that unambiguously map to either of these regions, 53% reside in HCE+ regions, while 47% reside in HCE- regions (Table SIIE). Similarly, of the *initiation zones* that unambiguously map to either HCE region (25,904), 53% reside in HCE+ regions, with the mean number of fragments and the size of zones being greater (3.9 and 23.5 kb, respectively) than in HCE- regions (3.3 and 16.6 kb). Of the 47,061 *isolated* bubble-containing fragments, 55% reside in HCE- regions, with the mean size of fragments being smaller than those in HCE+ regions (5.8 kb and 6.3 kb, respectively). While we consistently find a bias toward higher densities and efficiencies of origins in open chromatin environments, nearly half of the mapped bubble-containing fragments are located in regions defined by a more closed chromatin conformation, as defined by negative HCE.

Interestingly, the characteristic edge-to-edge distances between origins (either isolated or zonal) are lowest for early-replicating HCE+ and late-replicating HCE- regions, with average values of 56 kb and 36.3 kb, respectively (not shown). Origin-containing loci are spaced much further apart in early HCE- and late HCE+ regions, with average values of 303.1 kb and 133 kb, respectively. Surprisingly, the characteristic inter-origin distances in late-replicating heterochromatic regions are comparable to those of early-replicating euchromatic regions. However, as shown in Fig. S3, origin efficiencies are lower in late-replicating HCE- regions relative to early-replicating HCE+ regions.

We also compared overlaps between origins and DNase I HSS, H3K36me3, H3K27me3, H3K4me3, and CTCF marks both in HCE+ and HCE- regions, as well as with the six HCE+ and HCE- subdivisions of early-, mid-, and late-replicating domains (Table SVII). In general, the results are

consistent with these same comparisons for the early-, mid-, and late-replicating regions as a whole (Table SVIID-F).

Large zones and high levels of H3K36me3 are detected in a late-replicating section of chromosome X. Fig. S6 displays a moderately late-replicating section of the X chromosome in which the overlaps among initiation zones, transcriptional activity, and H3K36me3 modifications are extremely high. Interestingly, the largest zones in late-replicating regions, which correspond to outliers in the rightmost boxplot in Fig. 3D, are located on the X chromosome. Given that the GM06990 cell line was isolated from a female, this is most likely due to X chromosome inactivation, wherein the transcriptionally-silent X is replicated late in S-phase, driving a moderately high TR50 value, while active genes on the other X chromosome are decorated with activating marks such as H3K36me3. Indeed, the variance of the replication timing is relatively high in this region (Fig. S6).

SUPPLEMENTARY ANALYTICAL METHODS

The one-fragment rule for defining zones. Replication initiation sites in mammalian cells often cluster into zones. We define a zone as two or more adjacent bubble-containing fragments. However, many obvious zones display single-fragment gaps (as in Fig. 1). These could represent *bona fide* gaps (as in the DHFR initiation zone (Dijkwel et al. 2002)), particularly if the fragments are small. Alternatively, they could result from lack of saturation either in the biological or sequencing samples. To assess the latter possibility, we plotted a distribution of the number of negative fragments in the genome that lie between any two bubble-containing fragments, and developed a simple null model, which assumed that the probability of an EcoRI fragment containing a bubble or not containing a bubble is independent of the bubble-containing status of its neighbors. We first estimated the probability p of a genomic EcoRI fragment containing a bubble to be the fraction of EcoRI fragments containing bubbles (0.1599365 excluding chr Y). Using these assumptions, the probability of finding *two* bubble-containing fragments separated by n non-bubble-containing fragments is given by $p^2 * (1-p)^n$. The null model, $p^2 * (1-p)^n$, was scaled and then overlaid on the distribution of the actual data (Fig. SI). This resulted in striking agreement between the empirical data and the null model for $n > 1$. However, when the number of non-

bubble containing EcoRI fragments between two bubbles was one, there was a ~3.6-fold enrichment of actual data over the null model (Fig. S2). This finding suggested either that: 1) many zones are interrupted by regions that don't initiate or do so at a much reduced frequency; or 2) single fragment interruptions result from lack of complete saturation in the isolation or sequencing of bubble-containing fragments. Since both effects are possible, a one-fragment rule therefore was adopted to define a *zone* as a cluster of adjacent bubble-containing EcoRI fragments, no two of which are separated by more than one negative fragment.

Estimation of TR50 and variance about the TR50 from GM06990 replication timing data. *FASTQ* files for timing data (Hansen et al. 2010) corresponding to the six replication timing windows for GM06990 (G1b, S1, S2, S3, S4 and G2) were uniquely mapped to the NCBI human genome assembly v36 (hg18), using BWA version 0.5.7. Sequence reads that mapped to chrY were removed. Individual enriched-read wig files were created and normalized for each timing sector to 4 M mapped reads by multiplying by a scale factor, similar to the approach taken by Hansen et al. (Hansen et al. 2010). This factor was calculated by dividing 4 M by the number of mapped reads for each individual timing sector. The *TR50* for a given genomic region was defined as the time at which 50% of the DNA in that region was replicated. It was assumed that 100% of the DNA in a region was replicated by the end of G2 phase. Cumulative reads within each EcoRI fragment for which there was timing data were first calculated using the normalized enriched-read profile data as follows: G1b ($C_0 = R_0$), S1 ($C_1 = R_0+R_1$), S2 ($C_2 = R_0+R_1+R_2$), S3 ($C_3 = R_0+R_1+R_2+R_3$), S4 ($C_4 = R_0+R_1+R_2+R_3+R_4$) and G2 ($C_5 = R_0+R_1+R_2+R_3+R_4+R_5$), where R_0, R_1, R_2, R_3, R_4, and R_5 are the number of reads in normalized G1b, S1, S2, S3, S4 and G2 read enrichment profile data, respectively. Replication times were indexed as follows: 0-1, 1-2, 2-3, 3-4 4-5, and 5-6 for G1b, S1, S2, S3, S4, and G2 intervals, respectively. We then determined the interval during which 50% of the DNA in that sector was replicated (i.e., the TR50) by linear interpolation. Regions with TR50 values in the first two intervals were classified as early-replicating, those in the second two intervals as mid-replicating, and those in the last two intervals as late-replicating.

For comparison with HCE data for GM06990 (Lieberman-Aiden et al. 2009a), TR50 values were determined in non-overlapping 1Mb windows across the genome, using the same procedure. TR50 values were also determined across the genome in 50 kb windows in sliding steps of 1 kb. These smoothed TR50 curves are assumed to predict the overall trend of replication.

Variance about the TR50 in a region measures the spread of replication-timing data relative to TR50 over that region. For a region with timing data, the variance about TR50 was defined as:

$$Variance = \left(\frac{\sum_{i=0}^{5} \left((i - TR50)^2 * R_i \right)}{\sum_{i=0}^{5} R_i} \right), \quad \sum_{i=0}^{5} R_i \neq 0 , \qquad (1)$$

where i is the replication time index from 0 to 5, *TR50* was calculated as specified earlier, and R_i is the normalized number of reads for replication time index i in a given region. Variances about the TR50s were determined for individual EcoRI fragments, non-overlapping 1Mb windows over the genome; as well as for 50 kb windows over the genome, in sliding steps of 1 kb.

Constructing a matching transcriptome for the GM06990 cell line. Sequence reads for the two cDNA libraries for the GM06990 cell line propagated in our laboratory were merged. TopHat version 1.0.13 (Trapnell and Salzberg 2009) and Bowtie version 0.12.5.0 (Langmead et al. 2009) were used to uniquely align RNA-Seq reads to NCBI human genome assembly v36 (hg18). The parameters used were *min-intron-length* 30, and a reference transcript annotation file (wgEncodeGencodeManualV3.gtf; transcript annotation of the human genome NCBI36). The RNA-Seq read alignments were converted to SAM format with TopHat (Trapnell and Salzberg 2009). Cufflinks version 0.8.3 (Trapnell et al. 2010) was then used to assemble transcripts from the aligned reads in SAM format, and to estimate their relative abundances based on the same reference transcript annotation file. Abundances were reported in *FPKM* (fragments per kilobase of exon model per million mapped fragments). Out of a possible 37,576 genes/isoforms, 21,322 putative isoforms were identified and their FPKM values were reported along with their confidence limits. The distribution of background FPKM values for the 21,322 putative isoforms displayed a very good fit to an exponential null model:

$$P(x) = c * \lambda^{x} , \quad (2)$$

in which c and λ are constants and x is FPKM. Best fits were observed with the parameters c = 135.9110 and λ = 0.1352. In total, 10,393 genes were found to have significant FPKM values after p-values were adjusted to FDR values, and an FDR cutoff of 5% was applied. The 5' UTRs, 3' UTRs, and exons corresponding to these 10,393 significant transcripts were identified from the Cufflinks output. In order to filter out genes that contained a small fraction of 5' UTR, 3' UTR, and exons with high FPKM values, which could represent spurious pile-ups of the mapped reads, we calculated the fractional coverage of expressed exons and the 5' and 3' UTRs of known genes. The plot of the distribution of the fraction of the gene that is expressed (excluding introns) displayed a bimodal distribution characteristic of expressed and non-expressed genes, with a dip near 0.5. Thus, we classified a gene as transcribed if ≥ 60% of its sequence (excluding introns) was expressed (i.e., exceeded the 5% FDR FPKM-based cutoff).

SUPPLEMENTAL FIGURE AND TABLE LEGENDS

Figure S1. Raw replication timing data for segments of the GM06990 genome demonstrating different TR50s and variances about the TR50s. IGV screen shots of raw replication timing data compiled by Hansen and coworkers and collected from six S-phase fractions sorted based on DNA content (Hansen et al. 2010). Four different ~1 Mb genomic windows are shown that display different relationships among raw replication timing data, TR50s, and variances about the TR50s. Panel A. An early-replicating region (i.e., low TR50) with low variance. Panel B. a late-replicating region (high TR50) displaying low variance. Panel C. A generally early-replicating region (moderately low TR50) displaying high variance. Panel D. A generally late-replicating region (moderately high TR50) displaying high variance.

Figure S2. The RDs for technical and biological replicates are better correlated for origins firing in early S-phase than for those firing in mid- or late S-phase. Scatter plots of RDs for bubble-containing fragments in early (A-C), mid- (D-F), and late-replicating (G-I) genomic regions between pairs of technical replicates for the B1 bubble library (B1A and B1B; Panels A, D, and G), between the B1A

and B2 bubble libraries (Panels B, E, and H), and between B1B and B2 (Panels C, F, and I). The indicated R^2 values are the corresponding Pearson correlation coefficients. RDs were included in the plot and in the calculation of R^2 if they passed the 0.1% FDR cutoff in at least one of the two replicates being compared.

Figure S3. Origin efficiencies and transcription levels both are lower in late-replicating DNA and in DNA with decreased chromatin accessibility. RDs of bubble-containing fragments (panel A) and RNA transcript levels (panel B; our cDNA sequencing data) for early-, mid-, and late-replicating regions. These were further subdivided into HiC eigenvector (HCE) positive and negative loci. Positive and negative HCE values correspond to open and closed chromatin regions, respectively. The decreasing trend in both origin efficiency and RNA level appears to be driven more by replication timing than HCE sign, suggesting that timing may be a slightly better measure of chromatin accessibility. Note that the horizontal black line in each box corresponds to the *median* of values, whereas the data in Table SIIE&F report the *means* (averages).

Figure S4. Origin end-to-end distances in late-replicating regions tend to be smaller than those in early-replicating regions. Distributions of end-to-end distances between origins in early (A-C) and late (D-F) replicating regions, calculated for all origins (A and D), isolated origins only (B and E), and zones only (C and F). Numbers in the upper right hand are the mean/median distances. Zones were treated as units in these calculations (i.e., end-to-end distances were measured between the centers of either the isolated or the outermost fragments of zones). Origins are generally less efficient and initiation zones are smaller in late-replicating regions, but inter-origin distances tend to be smaller.

Figure S5. The densities of H3K36me3, H4K4me3, and CTCF sites increase with increasing number of fragments in zones. Boxplots of epigenetic factor broadpeak site densities per bubble-containing fragment, averaged over the fragments in a zone as a function of the number of fragments in that zone. Note that *density* here corresponds to the combined length of bubbles (in bp) within each EcoRI fragment divided by its length (in bp). Panel A. H3K36me3. Panel B. H3K4me3. Panel C. CTCF broadpeak sites.

Figure S6. X-inactivation likely rationalizes high origin activity, large initiation zones, high transcription levels, and the H3K36me3 histone modification in a moderately late-replicating region. IGV screen shot of a 729 kb region on chrX, showing: RDs for bubble-containing fragments, initiation zone calls, smoothed TR50s calculated for 50 kb sliding windows in steps of 1 kb, smoothed TR50 variances calculated for 50 kb sliding windows in steps of 1 kb, HCE values calculated for non-overlapping 1 Mb windows, H3K36me3 read profiles at narrowpeak sites, H3K36me3 narrowpeak site densities in 1 Mb windows, H3K36me3 read profile at broadpeak sites, H3K36me3 broadpeak site densities in 1 Mb windows, cDNA overlapping read profiles, and RefSeq gene annotations. High transcriptional and H3K36me3 levels likely arise from the *active* X which is likely early-replicating, while the relatively late-replicating phenotype (based on TR50) is likely driven by the *inactive* X chromosome. Thus, the variance in replication timing about TR50 is relatively high at this locus.

TABLE LEGENDS

Table SI. RDs of bubble-containing EcoRI fragments in the integrated EBV genome. All of the EcoRI fragments in the *circular* EBV genome are listed, along with their RDs in the three sequencing datasets. The 5' end of a 4,167 bp fragment is arbitrarily designated position one, but the viral genome was linearized at some unknown position by integration into the GM06990 genome. The leftmost (49,622 bp) EcoRI fragment on the portion of the EBV map contains the viral lytic origin (*oriP),* including the DS and FR elements, near its left end. *B1A* and *B1B*: RDs for sequencing replicates. *B1* and *B2*: RDs for biological replicates. *GM*: combined RDs. Note that RD cutoffs were established independently for each of the three replicates as well as for the combined RDs, as indicated in the last row.

Table SII. Size and signal strength of bubble-containing EcoRI fragments. Panel A. Distribution of bubble-containing fragments within the isolated and zonal fragment categories, and numbers of zonal fragments with and without the inclusion of single negative fragments. Panel B. Distribution of bubble-containing fragments among the three replication timing domains. Panel C. Size distributions of: origins (whether isolated or zonal); zonal fragments with or without the inclusion of

single negative fragments; and as a function of their replication timing windows shown for all fragments, whether isolated or zonal, and for all origins whether isolated or zonal. Panel D. Numbers of origins, whether isolated or zonal, and the numbers of zonal fragments distributed among the three timing windows. Panel E. Number and average signal strength (RDs) of bubble-containing fragments in HCE+ and HCE- domains in the whole genome, or as a function of origin firing times. Panel F. Numbers and average signal strengths of 5' UTRs, 3' UTRs and exons that overlap HCE+ and HCE- domains in the indicated categories of EcoRI fragments. Panel G. Numbers of isolated origins and zones that overlap HCE+ and HCE- regions. Panel H. Size distributions of isolated bubble-containing fragments and zones in HCE+ and HCE- regions.

Table SIII. Significance of origin distributions within early-, mid-, and late-replication timing windows. The genome was divided into early-, mid-, and late-replicating regions, with these categories representing 23.7%, 35.3%, and 32.3% of the genome, respectively. The distributions of bubble-containing fragments within each timing sector were then determined to be 31.9%, 31.9%, and 36.2%, respectively. To determine the significance of the actual bubble distributions in each timing sector, all genomic EcoRI fragments (bubbles and non-bubbles) were randomly permuted 10,000 times within non-overlapping 1 Mb windows across the genome, as described in Methods. The low p-value and high z-score for the early-firing bubble distribution suggest a large and *bona fide* increase in bubble density in this compartment. Also note the ~20% lower number of bubbles in mid-S-phase regions relative to mean random, suggesting a relative depletion in initiations in these timing transition regions (Desprat et al 2009).

Table SIV. Statistical relationships among bubble-containing fragments, transcription units, and replication time signatures (in support of Figure 4). Distributions and genomic coverage by RefSeq genes (non-transcribed and transcribed, the latter determined from the cDNA library prepared from GM06990 for this study; see Methods). Panels A-C: Data for the entire genome, for bubble-containing fragment calls made for the combined sequencing datasets, and for bubble calls from the individual origin libraries, respectively.

Table SV. Significance of overlaps between bubble-containing fragments and nine gene annotation categories (in support of Figure 4). Significance analyses were performed as described in the legend to Table SIII for all bubble-containing fragments (Panel A), and for early-firing (Panel B), mid-firing (Panel C), and late-firing fragments (Panel D). The nine transcription-related categories listed in the first column of each panel correspond to each sector of the pie-chart in Fig. 4. The normalized values in the right three columns were calculated as described in the main text. The two highest fold-changes are highlighted.

Table SVI. Significance of overlaps between bubble-containing factors and epigenetic factors. The significance of the overlaps between bubble-containing fragments and the indicated epigenetic factor sites were performed as described in the legend to Table SIII for all bubble-containing fragments (Panel A), and for early-firing (Panel B), mid-firing (Panel C), and late-firing (Panel D) bubble-containing fragments. The normalized values in the right three columns were calculated as described in the main text.

Table SVII. Origin-epigenetic factor overlap summaries. Basic overlap statistics between bubble-containing fragments and six different epigenetic factor datasets and as a function of replication time and HCE status. Data are presented for the whole genome (Panels A-C), and for early-, mid-, and late-firing bubble-containing fragments (Panels D-F). The latter three categories are further broken down into HCE+ and HCE- categories in Panels G-L. Reading from left to right, the values correspond to: the number of bubble-containing fragments that overlap one or more particular sites; the number of bubbles in the selected regions (e.g., whole genome, early-replicating, or early-replicating HCE+ or early-replicating HCE-); the percentage of bubbles that overlap the factor sites; the number of epigenetic factor sites that overlap bubble-containing fragments; the number of epigenetic factor sites in the selected regions (e.g., early-replicating), and the percentage of epigenetic factor sites that overlap bubble-containing fragments.

Table SVIII. Correlation coefficients between origin densities (isolated or zonal) and epigenetic factors calculated for 1 Mb intervals. Summaries of correlation coefficients between origin

densities and HCE values, TR50 values, and the indicated epigenetic factor site densities calculated in 1 Mb windows. Columns, from left to right, are: Pearson correlation coefficients, p-values calculated for the Pearson correlation coefficient, Spearman correlation coefficients, and Kendall tau correlation coefficients (Becker et al. 1988).

Table SIX. Significance of NS/Bubble overlaps and G-quadruplexes distributions. Panel A. The numbers and percentages of NS in each human cell line that overlap GM06990 bubble-containing fragments (NS data from Besnard et al. 2012). The p-values, fold-changes over mean random, and Z-scores are shown. Significance analyses were performed as described in Methods. Panel B. The numbers and percentages of GM06990 bubble-containing fragments that overlap NS in each of the four human cell lines, along with significance values. Panel C. The numbers, percentages, and significance of G-quadruplex/bubble overlaps.

REFERENCE LIST

Desprat R, Thierry-Mieg D, Lailler N, Schildkraut CL, Thierry-Mieg and Bouhassira EE. Predictable dynamic program of timing of DNA replication in human cells. *Genome Res.* **19**:2288-2299.

Diffley JF, Cocker JH, Dowell SJ, Rowley A. 1994. Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell* **78**: 303-316.

Dijkwel PA, Wang S, Hamlin JL. 2002. Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies. *Mol. Cell Biol.* **22**: 3053-3065.

Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, Dorschner MO, Gartler SM, Stamatoyannopoulos JA. 2010. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc. Natl. Acad. Sci. U. S. A* **107**: 139-144.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**: R25.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. 2009a. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289-293.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J. 2009b. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289-293.

Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM. 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**: 761-770.

Trapnell C and Salzberg SL. 2009. How to map billions of short reads onto genomes. *Nat. Biotechnol.* **27**: 455-457.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**: 511-515.

Figure S1

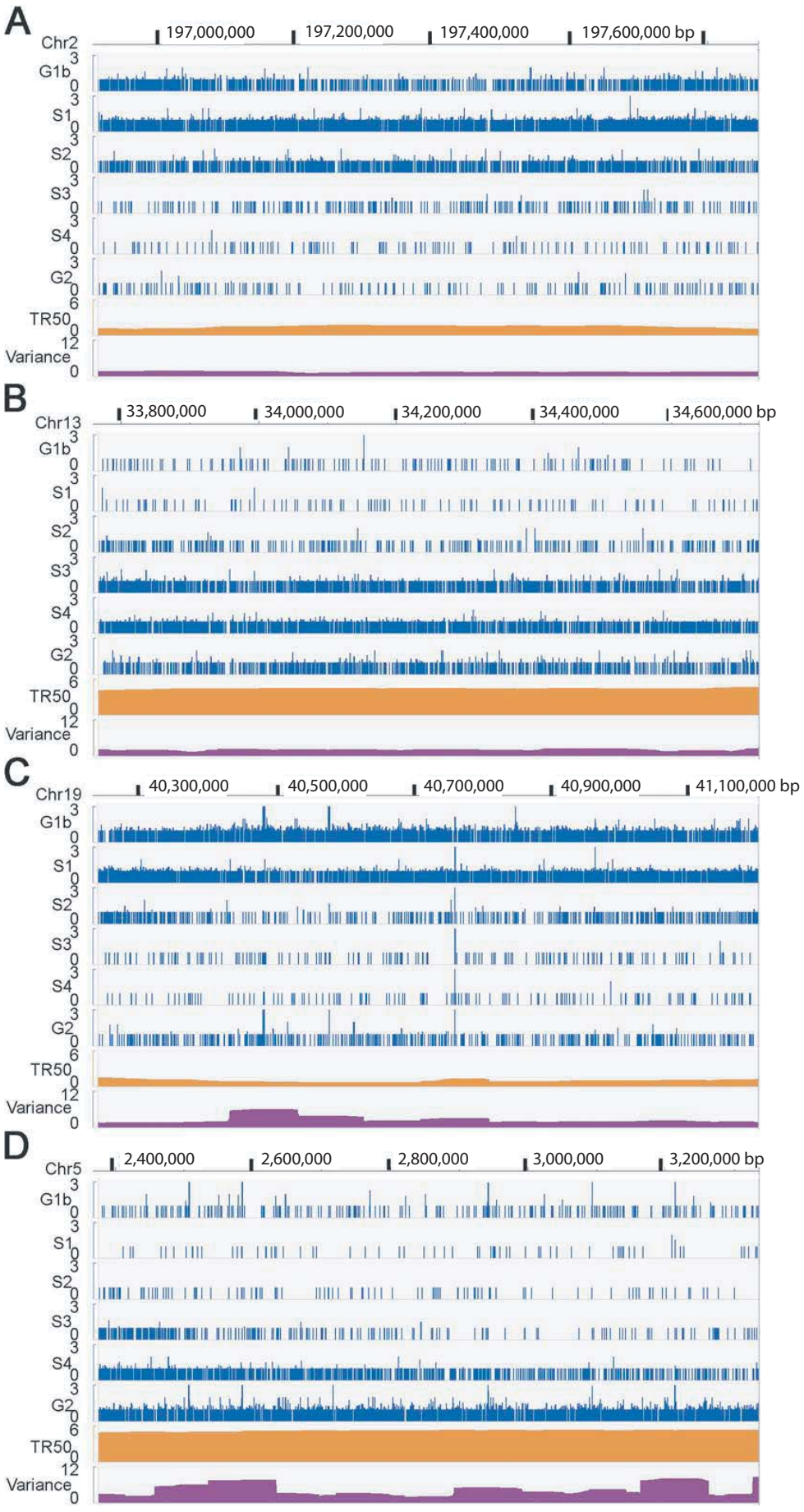
A
Chr2
197,000,000
197,200,000
197,400,000
197,600,000 bp
G1b
S1
S2
S3
S4
G2
TR50
Variance
B
Chr13
33,800,000
34,000,000
34,200,000
34,400,000
34,600,000 bp
C
Chr19
40,300,000
40,500,000
40,700,000
40,900,000
41,100,000 bp
D
Chr5
2,400,000
2,600,000
2,800,000
3,000,000
3,200,000 bp

Figure S2

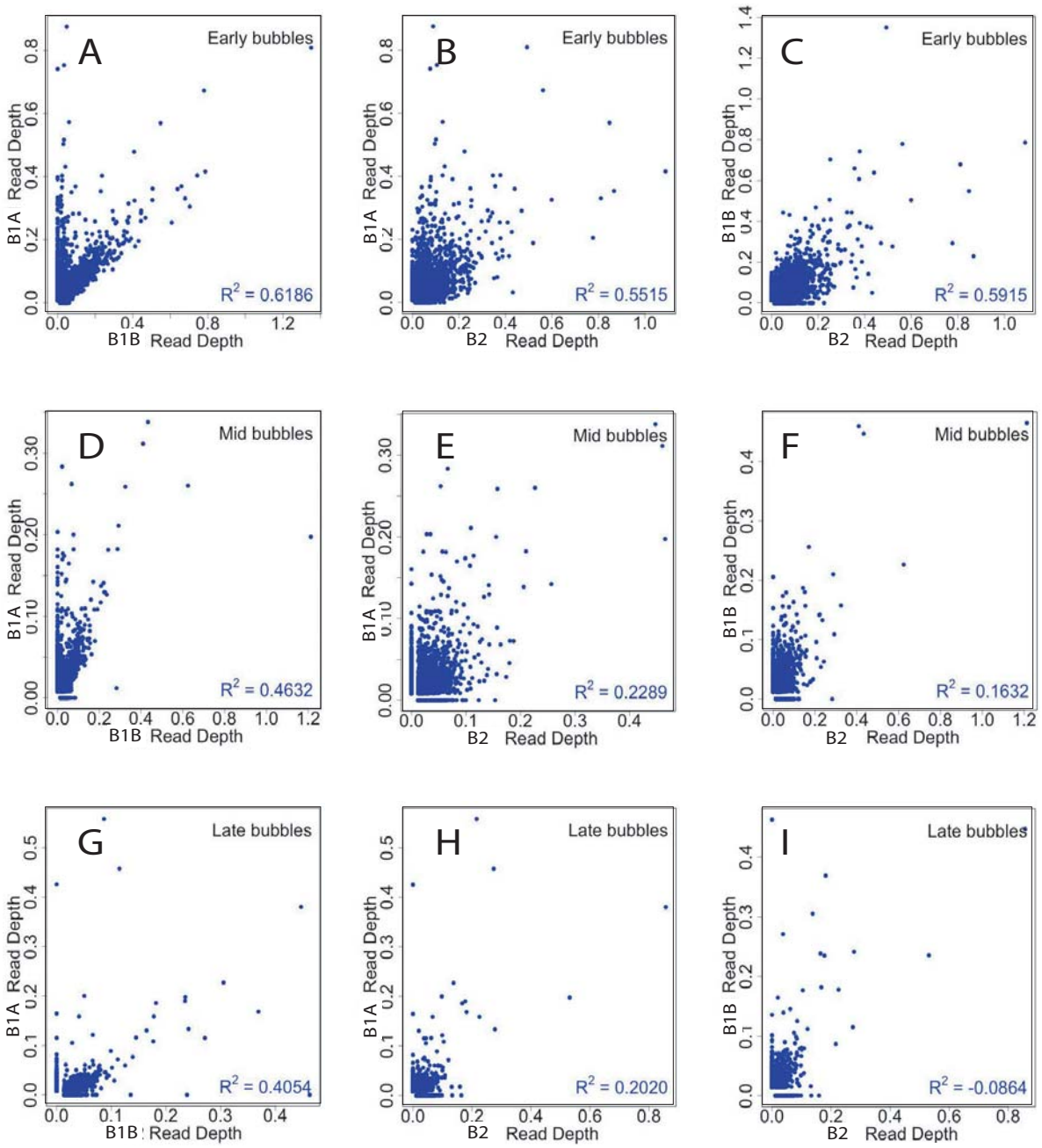
A
Early bubbles
B1A Read Depth
B1B Read Depth
R² = 0.6186
B
Early bubbles
B1A Read Depth
B2 Read Depth
R² = 0.5515
C
Early bubbles
B1B Read Depth
B2 Read Depth
R² = 0.5915
D
Mid bubbles
B1A Read Depth
B1B Read Depth
R² = 0.4632
E
Mid bubbles
B1A Read Depth
B2 Read Depth
R² = 0.2289
F
Mid bubbles
B1B Read Depth
B2 Read Depth
R² = 0.1632
G
Late bubbles
B1A Read Depth
B1B Read Depth
R² = 0.4054
H
Late bubbles
B1A Read Depth
B2 Read Depth
R² = 0.2020
I
Late bubbles
B1B Read Depth
B2 Read Depth
R² = -0.0864

Figure S3

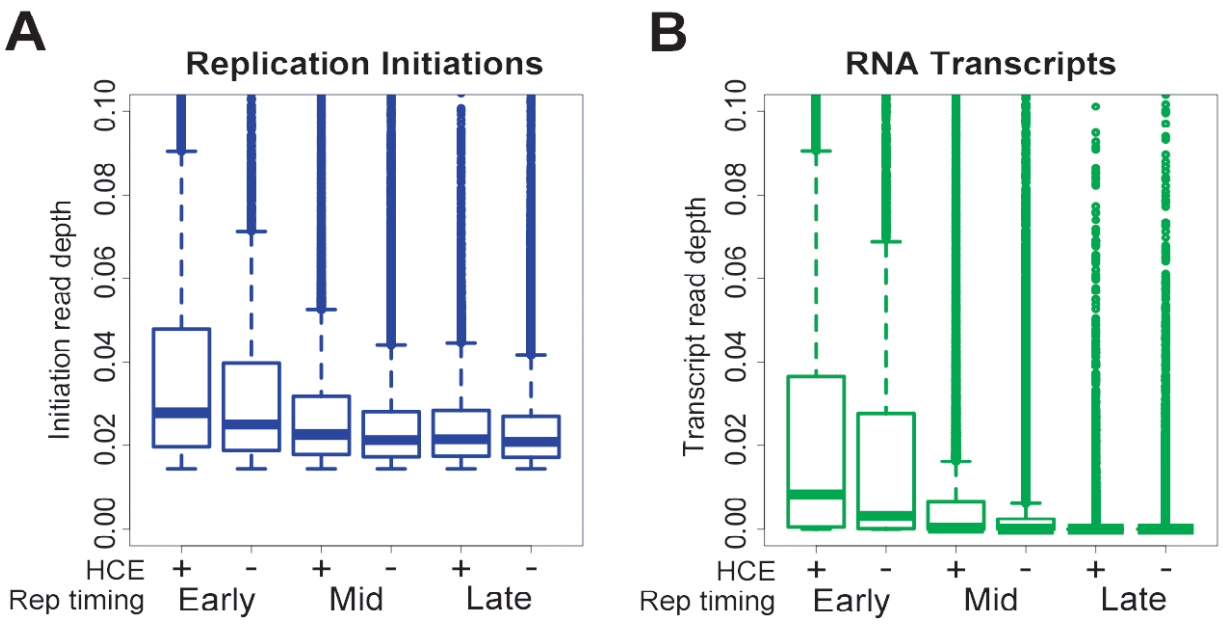
A
Replication Initiations
Initiation read depth
0.00
0.02
0.04
0.06
0.08
0.10
HCE + - + - + -
Rep timing Early Mid Late
B
RNA Transcripts
Transcript read depth
0.00
0.02
0.04
0.06
0.08
0.10
HCE + - + - + -
Rep timing Early Mid Late

Figure S4

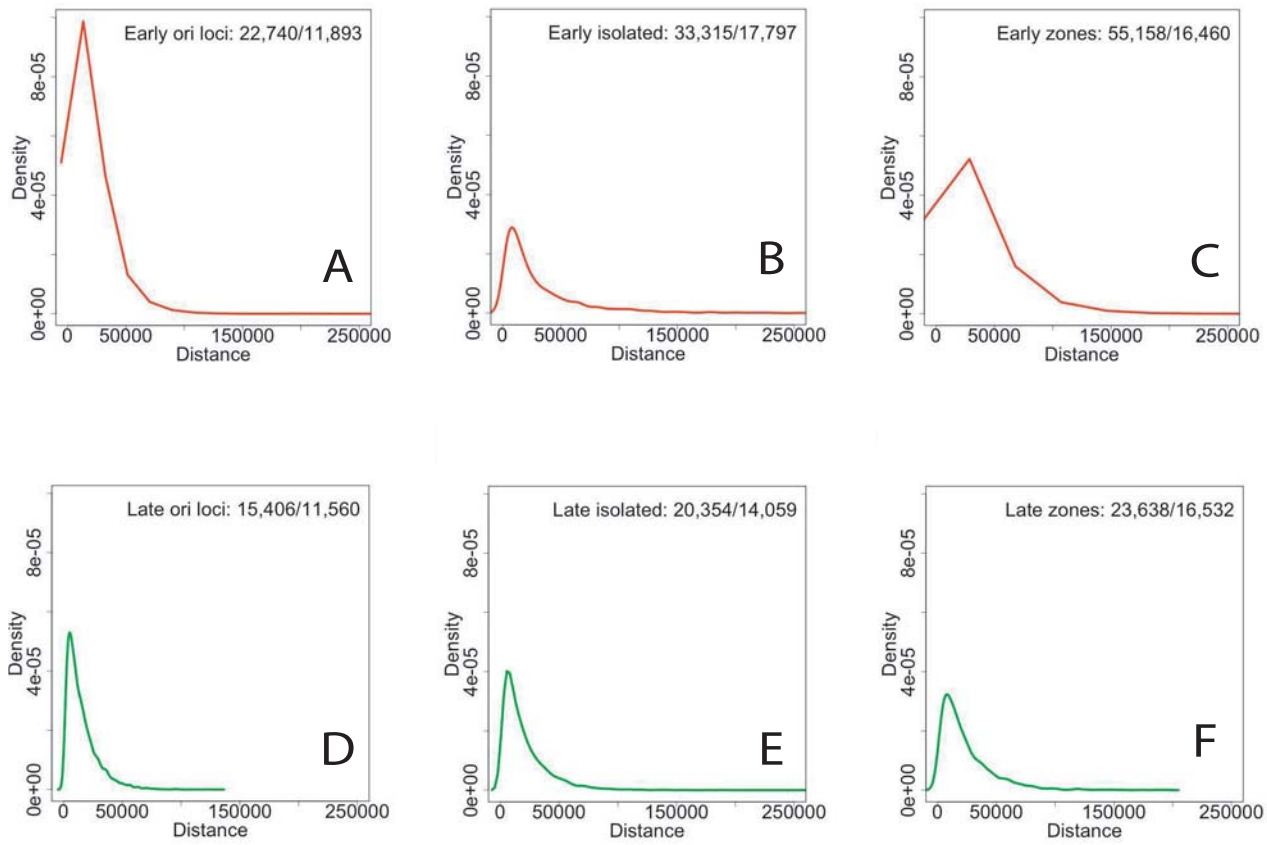
Early ori loci: 22,740/11,893
A
Early isolated: 33,315/17,797
B
Early zones: 55,158/16,460
C
Late ori loci: 15,406/11,560
D
Late isolated: 20,354/14,059
E
Late zones: 23,638/16,532
F
Density
0e+00
4e-05
8e-05
0
50000
150000
250000
Distance

Figure S5

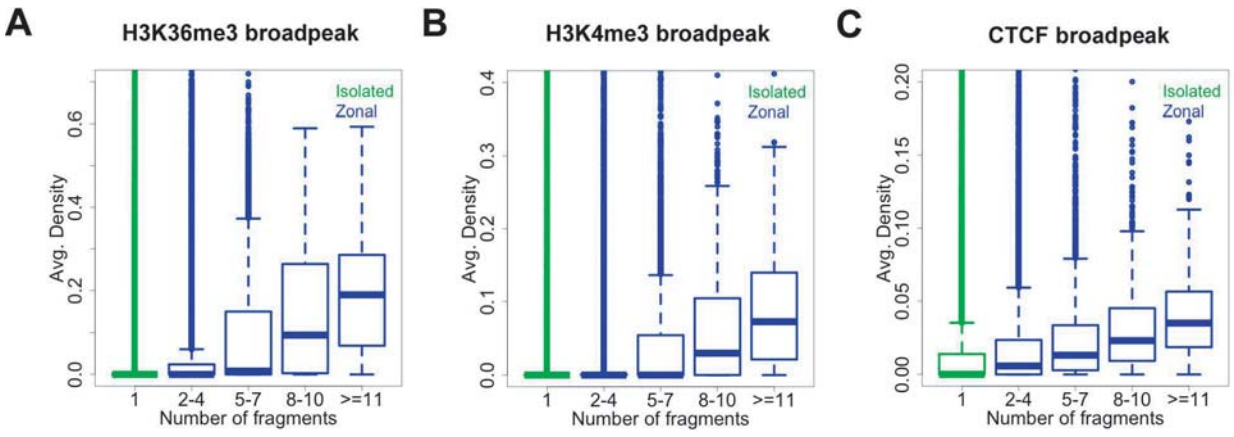
A
H3K36me3 broadpeak
Isolated
Zonal
Avg. Density
0.0
0.2
0.4
0.6
1
2-4
5-7
8-10
>=11
Number of fragments
B
H3K4me3 broadpeak
Isolated
Zonal
Avg. Density
0.0
0.1
0.2
0.3
0.4
1
2-4
5-7
8-10
>=11
Number of fragments
C
CTCF broadpeak
Isolated
Zonal
Avg. Density
0.00
0.05
0.10
0.15
0.20
1
2-4
5-7
8-10
>=11
Number of fragments

Figure S6

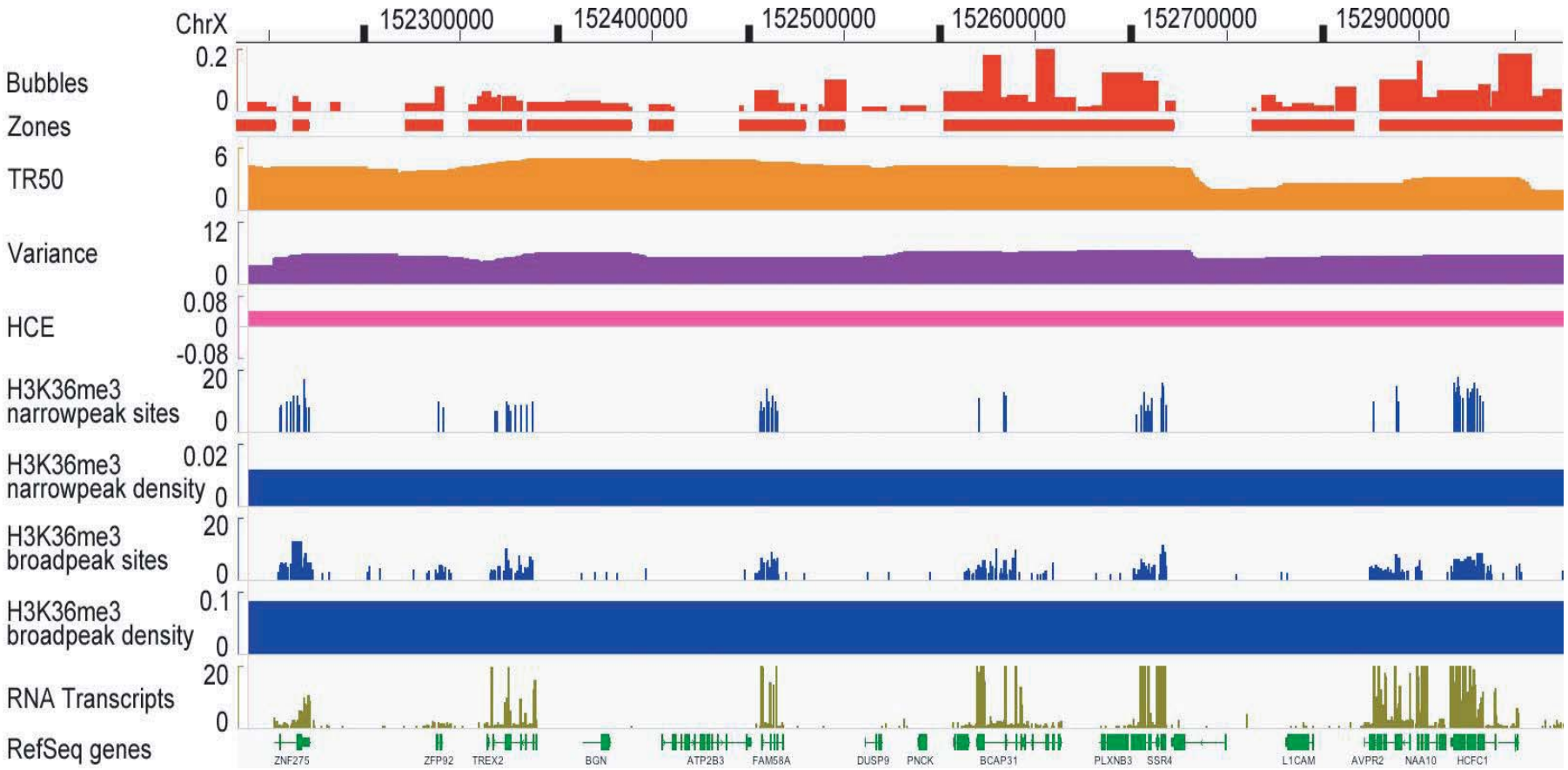
ChrX
152300000
152400000
152500000
152600000
152700000
152900000
Bubbles
0.2
0
Zones
TR50
6
0
Variance
12
0
HCE
0.08
0
-0.08
H3K36me3
narrowpeak sites
20
0
H3K36me3
narrowpeak density
0.02
0
H3K36me3
broadpeak sites
20
0
H3K36me3
broadpeak density
0.1
0
RNA Transcripts
20
0
RefSeq genes
ZNF275
ZFP92
TREX2
BGN
ATP2B3
FAM58A
DUSP9
PNCK
BCAP31
PLXNB3
SSR4
L1CAM
AVPR2
NAA10
HCFC1

Table SI – Bubble-containing EcoRI fragments in the integrated EBV genome

EBV fragment sizes in base pairs			B1A Read Depth	B1B Read Depth	B2 Read Depth	Combined Read Depth
5'-end	3'-end	Size				
1	4167	4167	1.0114	1.6955	0.8609	3.8228
4168	7320	3153	0.4138	0.3382	0.0859	0.8169
7325	56947	49623	0.0181	0.0211	0.0197	0.0576
56952	63183	6232	1.3674	1.4237	0.6551	3.4403
63188	64442	1255	0.2248	0.3023	0.0282	0.5739
64447	70766	6320	0.8332	1.3339	1.6357	3.8536
70771	79204	8434	3.0855	4.2626	2.4135	9.6673
79209	79838	630	0.2819	0.6344	0.1147	0.8761
79843	80939	1097	0.0693	0.1058	0.0265	0.1106
80944	82961	2018	0.1839	0.2819	1.1633	0.4510
82966	113858	30893	0.1597	0.0762	0.1671	0.3838
113863	125760	11898	1.3287	1.6525	2.4147	5.3575
125765	131225	5461	1.1530	1.7489	0.9004	3.9296
131230	160879	29650	3.1135	3.8814	2.1473	9.1970
160880	172764	11885	2.1003	1.8219	1.2762	5.1623
Read Depth Cutoffs			0.0081	0.0139	0.0123	0.0144

Table SII – Size and signal strength of bubble-containing EcoRI fragments

A. Statistics based on the one-fragment joining rule	Frag #	Percent
Total number of bubble-containing fragments (excluding negative RIs)	123297	100.0%
Number of *isolated* bubble-containing fragments	47005	38.1%
Number of *zonal* fragments (excluding negative EcoRIs)	76292	61.9%
Number of *zones* (excluding negative fragments)	25835	
Mean number of EcoRIs in *zones* (excluding negative EcoRIs)	2.95	
Total number of *origins* if *zones are treated as units*	72840	
Total number of bubble-containing fragments (including negative RIs)	140685	100.0%
Number of *isolated* bubble-containing fragments	47005	33.4%
Number of *zonal* fragments (including negative EcoRIs)	93680	66.6%
Number of *zones* (including negative fragments)	25807	
Mean number of EcoRIs in *zones* (including negative EcoRIs)	3.63	
Total number of *origins* if *zones are treated as units*	72812	

B. Classifications based on timing data	Frag #	Percent
Total number of bubble-containing fragments (excluding negative RIs)	123297	100.0%
Number of early-firing bubble-containing fragments	39236	31.8%
Number of mid-firing bubble-containing fragments	39313	31.9%
Number of late-firing bubble-containing fragments	44527	36.1%
Number of bubble-containing fragments with no timing data	221	0.2%

C. Origin size distributions	Mean (bp)	Median(bp)
Isolated bubble-containing fragments	6031	5643
Zones	20235	15755
Zonal fragments (excluding negative EcoRIs)	6113	5448
Zonal fragments (including negative EcoRIs)	5580	4840
Early-firing bubble-containing fragments	6773	5888
Mid-firing bubble-containing fragments	5952	5488
Late-firing bubble-containing fragments	5615	5345
Bubble-containing fragments with no timing data	872	733
Early-firing isolated fragments	6675	6037
Early-firing zones	27054	19871
Mid-firing isolated fragments	5971	5617
Mid-firing zones	18076	14988
Late-firing isolated fragments	5791	5535
Late-firing zones	16028	14061

D. Statistics associated with timing data	Frag #	Percent
Number of early-firing EcoRI fragments (excluding negative EcoRIs)	39236	100.0%
Number of early-firing *isolated* bubble-containing fragments	9686	24.7%
Number of early-firing *zonal* fragments (excluding negative EcoRIs)	29550	75.3%
Number of early-firing *zones* treated as units (excluding negative EcoRIs)	8321	46.2%
Mean number of EcoRIs in early-firing *zones* (excluding negative EcoRIs)	3.55	
Total number of early-firing *origins* if *zones* are treated as units	18007	100.0%
Total number of early-firing EcoRI fragments (including negative EcoRIs)	45514	100.0%
Number of early-firing *isolated* bubble-containing fragments	9686	21.3%
Number of early-firing *zonal* fragments (including negative EcoRIs)	35828	78.7%
Number of early-firing *zones* treated as units (including negative EcoRIs)	8313	46.2%
Mean number of EcoRIs in early-firing *zones* (including negative EcoRIs)	4.31	
Total number of early-firing *origins* if *zones* are treated as units	17999	100.0%
Total number of mid-firing EcoRI fragments (excluding negative EcoRIs)	39313	100.0%
Number of mid-firing *isolated* bubble-containing fragments	17448	44.4%
Number of mid-firing *zonal* fragments (excluding negative EcoRIs)	21865	55.6%
Number of mid-firing *zones* treated as units	8279	32.2%
Mean number of EcoRIs in mid-firing *zones* (excluding negative EcoRIs)	2.64	
Total number of mid-firing *origins* if *zones* are treated as units	25727	100.0%
Total number of mid-firing EcoRI fragments (including negative EcoRIs)	45209	100.0%
Number of mid-firing *isolated* bubble-containing fragments	17448	38.6%
Number of mid-firing *zonal* fragments (including negative EcoRIs)	27761	61.4%
Number of mid-firing *zones* treated as units	8287	32.2%
Mean number of EcoRIs in mid-firing *zones* (including negative EcoRIs)	3.35	
Total number of mid-firing *origins* if *zones* are treated as units	25735	100.0%
Total number of late-firing EcoRI fragments (excluding negative EcoRIs)	44527	100%
Number of late-firing isolated bubble-containing fragments	19792	44.5/%
Number of late-firing zonal fragments (excluding negative EcoRIs)	24735	55.5%
Number of late-firing zones treated as units (excluding negative EcoRIs)	9233	31.8%
Mean number of EcoRIs in late-firing zone (excluding negative EcoRIs)	2.68	
Total number of late-firing *origins* if *zones* are treated as units	29025	100.0%
Total number of late-firing EcoRI fragments (including negative EcoRIs)	49876	100.0%
Number of late-firing *isolated* bubble-containing fragments	19792	39.7/%
Number of late-firing *zonal* fragments (including negative EcoRIs)	30084	60.3%
Number of late-firing *zones* treated as units (including negative EcoRIs)	9228	31.8%
Mean number of EcoRIs in late-firing *zones* (including negative EcoRIs)	3.26	
Total number of late-firing *origins* if *zones* are treated as units	29020	100.0%

E. Signal strengths of bubble-containing fragments	Number	Avg. RD
All bubble-containing fragments overlapping HCE+	65009	0.0387
All bubble-containing fragments overlapping HCE-	58403	0.0259
Bubble-containing fragments overlapping HCE+ & HCE-	116	0.0379
Early-firing bubble-containing fragments overlapping HCE+	34752	0.0471
Early-firing bubble-containing fragments overlapping HCE-	4529	0.0372
Early-firing bubble-containing fragments overlapping HCE+ & HCE-	45	0.0534
Mid-firing bubble-containing fragments overlapping HCE+	23955	0.0300
Mid-firing bubble-containing fragments overlapping HCE-	15410	0.0264
Mid-firing bubble-containing fragments overlapping HCE+ & HCE-	52	0.0293
Late-firing bubble-containing fragments overlapping HCE+	6263	0.0243
Late-firing bubble-containing fragments overlapping HCE-	38283	0.0258
Late-firing bubble-containing fragments overlapping HCE+ & HCE-	19	0.0248

F. Signal strengths of transcripts overlapping bubbles in HCE+/-	Number	RNA RD
Transcripts overlapping non-bubble-containing fragments in HCE+	298333	0.0126
Transcripts overlapping non-bubble-containing fragments in HCE-	349691	0.0035
Transcripts overlapping bubble-containing fragments in HCE+	65009	0.0318
Transcripts overlapping bubble-containing fragments in HCE-	58403	0.0137
Transcripts overlapping early-firing bubble-containing fragments in HCE+	34752	0.0470
Transcripts overlapping early-firing bubble-containing fragments in HCE-	4529	0.0319
Transcripts overlapping mid-firing bubble-containing fragments in HCE+	23955	0.0173
Transcripts overlapping mid-firing bubble-containing fragments in HCE-	15410	0.0335
Transcripts overlapping late-firing bubble-containing fragments in HCE+	6263	0.0036
Transcripts overlapping late-firing bubble-containing fragments in HCE-	38283	0.0035

G. Distribution of isolated bubble-containing fragments/zones in HCE+/-	Number
Isolated bubble-containing fragments overlapping HCE+	21036
Isolated bubble-containing fragments overlapping HCE-	26025
Zones overlapping HCE+	13716
Zones overlapping HCE-	12188

H. Size distribution of origins in HCE+ and HCE-	Mean # of frags	Mean size (bp)
Isolated bubble-containing fragments in HCE+	1.00	6337
Isolated bubble-containing fragments in HCE-	1.00	5789
Zones in HCE+ (including negative EcoRIs)	3.93	23548
Zones in HCE- (including negative EcoRIs)	3.30	16585

Table SIII – Significance of number of early-, mid-, and late-firing bubble-containing fragments

	# Actual	Fold-change	Mean random	SD random	p-value	Z-score
Total number of bubbles	123279					
Number of early-firing bubbles	39236	1.53	25613	340	< 1e-04	40.01
Number of mid-firing bubbles	39313	0.79	49829	267	1.0000	-39.40
Number of late-firing bubbles	44527	0.97	46096	124	0.9990	-12.68

Table SIV – Statistical relationships among bubble-containing fragments, transcription units, and replication time signatures

A. Whole genome	
% of the genome represented by RefSeq genes	38.65 %
% of the genome represented by transcribed RefSeq genes (0.6 cutoff)	13.07 %
% of the genome represented by non-transcribed RefSeq genes (0.6 cutoff)	26.04 %
% of the genome that is early-replicating	23.67 %
% transcribed RefSeq genes that are early-replicating	7.63 %
% non-transcribed RefSeq genes that are early-replicating	6.08 %
% of the genome that is intergenic and early-replicating	10.21 %
% of the genome that is mid-replicating	35.27 %
% transcribed RefSeq genes that are mid-replicating	4.78 %
% non-transcribed RefSeq genes that are mid-replicating	11.26 %
% of the genome that is intergenic and mid-replicating	19.41 %
% of the genome that is late-replicating	32.34 %
% transcribed RefSeq genes that are late-replicating	0.56 %
% non-transcribed RefSeq genes that are late-replicating	8.16 %
% of the genome that is intergenic and late-replicating	23.65 %
% of genome with no timing data	1.42 %
% of genome with unmapped contigs	7.29 %

B. Bubble-containing fragment distributions	
Number of mapped bubble-containing fragments	123297
% of the genome represented by bubble-containing fragments	24.35 %
% of the genome represented by early-firing bubble-containing fragments	8.63 %
% of the genome represented by early-firing bubble-containing fragments that overlap transcribed genes	2.98 %
% of the genome represented by early-firing bubble-containing fragments that overlap non-transcribed genes	1.99 %
% of the genome represented by early-firing bubble-containing fragments that are intergenic	3.75 %
% of the genome represented by mid-firing bubble-containing fragments	7.60 %
% of the genome represented by mid-firing bubble-containing fragments that overlap transcribed genes	1.12 %
% of the genome represented by mid-firing bubble-containing fragments that overlap non-transcribed genes	2.36 %
% of the genome represented by mid-firing bubble-containing fragments that are intergenic	4.17 %
% of the genome represented by late-firing bubble-containing fragments	8.12 %
% of the genome represented by late-firing bubble-containing fragments that overlap transcribed genes	0.11 %
% of the genome represented by late-firing bubble-containing fragments that overlap non-transcribed genes	1.95 %
% of the genome represented by late-firing bubble-containing fragments that are intergenic	6.06 %

C. Individual GM06990 libraries	
Number of B1A bubbles	34184
% of the genome represented by B1A bubbles	7.28
Number of B1B bubbles	39971
% of the genome represented by B1B bubbles	7.83
Number of B2 bubbles	36225
% of the genome represented by B2 bubbles	7.84

Table SV – Significance of association between bubble-containing fragments and transcription units

A. ALL GENOMIC BUBBLE-CONTAINING FRAGMENTS

Association type		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % Actual	Fold-change	% Random
1	Fragment in tx gene	16771	1.03	16222	96.88	< 1e-04	5.66	13.60%	1.03	13.20%
2	Tx gene(s) in fragment	751	1.77	424	17.04	< 1e-04	19.22	0.60%	1.77	0.30%
3	Frag overlaps 5' tx gene(s)	4013	1.80	2229	37.51	< 1e-04	47.56	3.30%	1.80	1.80%
4	Frag overlaps 3' tx gene(s)	2885	1.39	2070	36.44	< 1e-04	22.36	2.30%	1.39	1.70%
5	Fragment in non-tx gene	27401	0.91	30159	118.86	1.0000	-23.20	22.20%	0.91	24.50%
6	Non-tx gene(s) in fragment	798	1.21	659	22.07	< 1e-04	6.31	0.60%	1.21	0.50%
7	Frag overlaps 5' non-tx gene(s)	2951	1.24	2383	40.48	< 1e-04	14.02	2.40%	1.24	1.90%
8	Frag overlaps 3' non-tx gene(s)	2235	1.04	2159	38.98	0.0255	1.96	1.80%	1.03	1.80%
9	Intergenic	65492	0.98	66992	132.71	1.0000	-11.30	53.10%	0.98	54.30%
	Total	123297		123298				100.00%		100.00%

B. EARLY-FIRING BUBBLE-CONTAINING FRAGMENTS

Association type		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % Actual	Fold-change	% Random
1	Fragment in tx gene	10981	1.46	7509	105.71	< 1e-04	32.84	28.00%	0.95	29.30%
2	Tx gene(s) in fragment	635	2.19	289	14.31	< 1e-04	24.14	1.60%	1.43	1.10%
3	Frag overlaps 5' tx gene(s)	2907	2.27	1277	31.71	< 1e-04	51.37	7.40%	1.49	5.00%
4	Frag overlaps 3' tx gene(s)	2211	1.88	1175	29.68	< 1e-04	34.90	5.60%	1.23	4.60%
5	Fragment in non-tx gene	6423	1.25	5127	92.32	0.0010	14.03	16.40%	0.82	20.00%
6	Non-tx gene(s) in fragment	392	1.72	227	13.53	< 1e-04	12.16	1.00%	1.13	0.90%
7	Frag overlaps 5' non-tx gene(s)	1423	1.66	857	26.23	< 1e-04	21.58	3.60%	1.08	3.30%
8	Frag overlaps 3' non-tx gene(s)	1067	1.47	728	24.18	< 1e-04	14.02	2.70%	0.96	2.80%
9	Intergenic	13197	1.57	8421	148.42	< 1e-04	32.18	33.60%	1.02	32.90%
	Total	39236		25610				100.00%		100.00%

C. MID-FIRING BUBBLE-CONTAINING FRAGMENTS

Association type		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % Actual	Fold-change	% Random
1	Fragment in tx gene	5173	0.67	7680	102.78	1.0000	-24.39	13.20%	0.85	15.40%
2	Tx gene(s) in fragment	103	0.91	113	9.64	0.8754	-1.07	0.30%	1.16	0.20%
3	Frag overlaps 5' tx gene(s)	985	1.19	831	27.60	< 1e-04	5.58	2.50%	1.50	1.70%
4	Frag overlaps 3' tx gene(s)	613	0.78	788	25.76	0.9990	-6.79	1.60%	0.99	1.60%
5	Fragment in non-tx gene	10699	0.77	13880	101.48	1.0000	-31.35	27.20%	0.98	27.90%
6	Non-tx gene(s) in fragment	258	0.89	289	15.07	0.9795	-2.07	0.70%	1.13	0.60%
7	Frag overlaps 5' non-tx gene(s)	1086	0.96	1128	30.02	0.9296	-1.39	2.80%	1.22	2.30%
8	Frag overlaps 3' non-tx gene(s)	877	0.83	1053	28.87	0.9991	-6.11	2.20%	1.06	2.10%
9	Intergenic	19519	0.81	24067	136.48	1.0000	-33.32	49.70%	1.03	48.30%
	Total	39313		49829				100.00%		100.00%

D. LATE-FIRING BUBBLE-CONTAINING FRAGMENTS

Association type		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % Actual	Fold-change	% Random
1	Fragment in tx gene	606	0.65	930	27.85	1.0000	-11.67	1.40%	0.67	2.00%
2	Tx gene(s) in fragment	13	0.69	19	3.93	0.9540	-1.49	0.00%	0.68	0.00%
3	Frag overlaps 5' tx gene(s)	121	1.06	115	9.47	0.2631	0.68	0.30%	1.09	0.20%
4	Frag overlaps 3' tx gene(s)	61	0.59	103	8.84	0.9995	-4.74	0.10%	0.61	0.20%
5	Fragment in non-tx gene	10236	0.94	10916	72.37	0.9998	-9.40	23.00%	0.97	23.70%
6	Non-tx gene(s) in fragment	148	1.15	128	9.73	0.0282	2.02	0.30%	1.20	0.30%
7	Frag overlaps 5' non-tx gene(s)	442	1.18	375	17.10	< 1e-04	3.94	1.00%	1.22	0.80%
8	Frag overlaps 3' non-tx gene(s)	291	0.81	358	16.66	0.9999	-4.03	0.70%	0.84	0.80%
9	Intergenic	32609	0.98	33152	105.80	0.9990	-5.13	73.20%	1.02	71.90%
	Total	44527		46096				100.00%		100.00%

Table SVI – Significance of origin-factor overlaps

A. All bubble-containing frags Overlaps with:		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % actual	Fold-change	% random
	Total bubble-containing fragments	123297		121538						
1	GM Ctcf rep1 broadpeak	45788	1.13	40652	132.57	< 1e-04	38.74	37%	1.11	33%
2	GM Ctcf rep1 narrowpeak	17994	1.26	14273	88.76	< 1e-04	41.92	15%	1.24	12%
3	GM DNaseI HSS rep1 broadpeak	35931	1.29	27948	111.61	< 1e-04	71.53	29%	1.27	23%
4	GM DNaseI HSS rep1 narrowpeak	21905	1.41	15515	89.80	< 1e-04	71.16	18%	1.39	13%
5	GM DNaseI HSS rep2 broadpeak	39059	1.22	31976	115.96	< 1e-04	61.08	32%	1.20	26%
6	GM DNaseI HSS rep2 narrowpeak	20953	1.39	15067	88.94	< 1e-04	66.18	17%	1.37	12%
7	GM H3k27me3 rep1 broadpeak	44096	1.00	43923	128.56	0.0935	1.34	36%	0.99	36%
8	GM H3k27me3 rep1 narrowpeak	4638	1.14	4084	50.78	< 1e-04	10.92	4%	1.12	3%
9	GM H3k27me3 rep2 broadpeak	38723	0.99	38933	123.94	0.9532	-1.70	31%	0.98	32%
10	GM H3k36me3 rep1 broadpeak	39879	1.19	33605	120.00	< 1e-04	52.28	32%	1.17	28%
11	GM H3k36me3 rep1 narrowpeak	14549	1.29	11322	79.65	< 1e-04	40.51	12%	1.27	9%
12	GM H3k36me3 rep2 broadpeak	45081	1.17	38489	123.44	< 1e-04	53.40	37%	1.15	32%
13	GM H3k4me3 rep1 broadpeak	19284	1.44	13436	86.76	< 1e-04	67.41	16%	1.41	11%
14	GM H3k4me3 rep1 narrowpeak	12912	1.55	8343	69.53	< 1e-04	65.71	11%	1.53	7%
15	GM H3k4me3 rep2 broadpeak	20650	1.41	14677	90.27	< 1e-04	66.17	17%	1.39	12%

B. Early-firing bubbles Overlaps with:		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % actual	Fold-change	% random
	Total early-firing fragments	39236		25610						
1	GM Ctcf rep1 broadpeak	20598	1.68	12285	162.04	< 1e-04	51.30	53%	1.09	48%
2	GM Ctcf rep1 narrowpeak	10232	1.80	5691	84.59	< 1e-04	53.68	26%	1.17	22%
3	GM DNaseI HSS rep1 broadpeak	21941	1.80	12174	159.62	< 1e-04	61.19	56%	1.18	48%
4	GM DNaseI HSS rep1 narrowpeak	14551	1.93	7540	105.38	< 1e-04	66.53	37%	1.26	29%
5	GM DNaseI HSS rep2 broadpeak	23402	1.72	13595	173.30	< 1e-04	56.59	60%	1.12	53%
6	GM DNaseI HSS rep2 narrowpeak	13912	1.90	7312	103.02	< 1e-04	64.07	35%	1.24	29%
7	GM H3k27me3 rep1 broadpeak	13626	1.47	9299	144.75	< 1e-04	29.89	35%	0.96	36%
8	GM H3k27me3 rep1 narrowpeak	2075	1.54	1351	33.46	< 1e-04	21.65	5%	1.00	5%
9	GM H3k27me3 rep2 broadpeak	11170	1.44	7776	126.11	0.0004	26.91	29%	0.94	30%
10	GM H3k36me3 rep1 broadpeak	24679	1.65	14939	185.57	< 1e-04	52.49	63%	1.08	58%
11	GM H3k36me3 rep1 narrowpeak	10470	1.71	6108	84.20	< 1e-04	51.81	27%	1.12	24%
12	GM H3k36me3 rep2 broadpeak	26223	1.64	15958	199.39	< 1e-04	51.49	67%	1.07	62%
13	GM H3k4me3 rep1 broadpeak	13102	1.96	6696	96.80	< 1e-04	66.18	33%	1.28	26%
14	GM H3k4me3 rep1 narrowpeak	9020	2.06	4380	69.53	< 1e-04	66.74	23%	1.34	17%
15	GM H3k4me3 rep2 broadpeak	13738	1.94	7099	102.20	< 1e-04	64.96	35%	1.26	28%

C. Mid-firing bubbles Overlaps with:		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % actual	Fold-change	% random
	Total mid-firing fragments	39313		49829						
1	GM Ctcf rep1 broadpeak	15213	0.83	18299	153.48	0.9990	-20.11	39%	1.05	37%
2	GM Ctcf rep1 narrowpeak	5728	0.87	6559	84.31	0.9990	-9.86	15%	1.11	13%
3	GM DNaseI HSS rep1 broadpeak	11398	0.87	13067	149.80	0.9990	-11.14	29%	1.11	26%
4	GM DNaseI HSS rep1 narrowpeak	6232	0.91	6823	100.15	0.9990	-5.91	16%	1.16	14%
5	GM DNaseI HSS rep2 broadpeak	12825	0.84	15234	164.18	0.9990	-14.68	33%	1.07	31%
6	GM DNaseI HSS rep2 narrowpeak	5986	0.90	6633	97.62	0.9990	-6.62	15%	1.14	13%
7	GM H3k27me3 rep1 broadpeak	20446	0.83	24530	142.08	1.0000	-28.75	52%	1.06	49%
8	GM H3k27me3 rep1 narrowpeak	1939	0.87	2218	41.74	0.9990	-6.68	5%	1.11	5%
9	GM H3k27me3 rep2 broadpeak	18394	0.84	21971	128.80	1.0000	-27.77	47%	1.06	44%
10	GM H3k36me3 rep1 broadpeak	12750	0.80	15916	174.62	0.9990	-18.13	32%	1.02	32%
11	GM H3k36me3 rep1 narrowpeak	3674	0.78	4701	78.17	0.9990	-13.14	9%	0.99	9%
12	GM H3k36me3 rep2 broadpeak	15030	0.81	18557	185.84	0.9990	-18.98	38%	1.03	37%
13	GM H3k4me3 rep1 broadpeak	5217	0.91	5751	90.94	0.9990	-5.87	13%	1.15	12%
14	GM H3k4me3 rep1 narrowpeak	3358	0.98	3417	64.59	0.8924	-0.92	9%	1.25	7%
15	GM H3k4me3 rep2 broadpeak	5771	0.90	6411	96.40	0.9990	-6.64	15%	1.14	13%

D. Late-firing bubbles Overlaps with:		# Actual	Fold-change	Mean random	SD random	p-value	Z-score	*Normalized to fragment number* % actual	Fold-change	% random
	Total late-firing fragments	44527		46096						
1	GM Ctcf rep1 broadpeak	9968	0.99	10049	82.21	0.8565	-0.98	22%	1.03	22%
2	GM Ctcf rep1 narrowpeak	2034	1.01	2018	39.26	0.3401	0.42	5%	1.04	4%
3	GM DNaseI HSS rep1 broadpeak	2590	0.96	2687	48.99	0.9856	-1.97	6%	1.00	6%
4	GM DNaseI HSS rep1 narrowpeak	1121	0.98	1141	31.85	0.7558	-0.62	3%	1.02	2%
5	GM DNaseI HSS rep2 broadpeak	2828	0.91	3123	52.19	0.9990	-5.66	6%	0.94	7%
6	GM DNaseI HSS rep2 narrowpeak	1054	0.95	1112	31.61	0.9781	-1.84	2%	0.98	2%
7	GM H3k27me3 rep1 broadpeak	10022	0.99	10085	80.72	0.8085	-0.77	23%	1.03	22%
8	GM H3k27me3 rep1 narrowpeak	624	1.21	514	19.81	< 1e-04	5.54	1%	1.26	1%
9	GM H3k27me3 rep2 broadpeak	9154	1.00	9180	76.17	0.6479	-0.34	21%	1.03	20%
10	GM H3k36me3 rep1 broadpeak	2447	0.90	2727	49.36	0.9990	-5.67	5%	0.93	6%
11	GM H3k36me3 rep1 narrowpeak	404	0.79	509	20.65	0.9990	-5.06	1%	0.82	1%
12	GM H3k36me3 rep2 broadpeak	3824	0.97	3952	57.79	0.9931	-2.22	9%	1.00	9%
13	GM H3k4me3 rep1 broadpeak	963	0.99	977	29.36	0.6971	-0.47	2%	1.02	2%
14	GM H3k4me3 rep1 narrowpeak	532	0.99	538	21.33	0.6308	-0.29	1%	1.02	1%
15	GM H3k4me3 rep2 broadpeak	1137	0.98	1155	31.97	0.7320	-0.55	3%	1.02	3%

Table SVII – Summary of overlaps between bubble-containing fragments and epigenetic factors

A. All bubble containing fragments (123297)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	45788	123297	37.14	76466	216126	35.38
2	Overlaps with GM Ctcf rep1 narrowpeak	17994	123297	14.59	24836	60662	40.94
3	Overlaps with GM DNaseI hyper rep1 broadpeak	35931	123297	29.14	74840	160648	46.59
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	21905	123297	17.77	42511	84278	50.44
5	Overlaps with GM DNaseI hyper rep2 broadpeak	39059	123297	31.68	86715	198915	43.59
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	20953	123297	16.99	40143	81480	49.27
7	Overlaps with GM H3k27me3 rep1 broadpeak	44096	123297	35.76	127796	443030	28.85
8	Overlaps with GM H3k27me3 rep1 narrowpeak	4638	123297	3.76	7297	19303	37.80
9	Overlaps with GM H3k27me3 rep2 broadpeak	38723	123297	31.41	93933	342242	27.45
10	Overlaps with GM H3k36me3 rep1 broadpeak	39879	123297	32.34	124287	328303	37.86
11	Overlaps with GM H3k36me3 rep1 narrowpeak	14549	123297	11.80	42384	84906	49.92
12	Overlaps with GM H3k36me3 rep2 broadpeak	45081	123297	36.56	134106	354850	37.79
13	Overlaps with GM H3k4me3 rep1 broadpeak	19284	123297	15.64	33074	63090	52.42
14	Overlaps with GM H3k4me3 rep1 narrowpeak	12912	123297	10.47	35805	61792	57.94
15	Overlaps with GM H3k4me3 rep2 broadpeak	20650	123297	16.75	34856	67797	51.41

B. All bubble-containing fragments in HCE+ (65009)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	30003	65009	46.15	53979	134,340	40.18
2	Overlaps with GM Ctcf rep1 narrowpeak	13582	65009	20.89	19279	42,706	45.14
3	Overlaps with GM DNaseI hyper rep1 broadpeak	27983	65009	43.04	60579	121,309	49.94
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	17442	65009	26.83	34857	65,713	53.04
5	Overlaps with GM DNaseI hyper rep2 broadpeak	30474	65009	46.88	70438	150,183	46.90
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	16766	65009	25.79	33073	63,495	52.09
7	Overlaps with GM H3k27me3 rep1 broadpeak	28456	65009	43.77	89687	286,530	31.30

8	Overlaps with GM H3k27me3 rep1 narrowpeak	3436	65009	5.29	5472	13,800	39.65
9	Overlaps with GM H3k27me3 rep2 broadpeak	24712	65009	38.01	64430	218,916	29.43
10	Overlaps with GM H3k36me3 rep1 broadpeak	31599	65009	48.61	102274	253,264	40.38
11	Overlaps with GM H3k36me3 rep1 narrowpeak	12210	65009	18.78	36597	69,668	52.53
12	Overlaps with GM H3k36me3 rep2 broadpeak	34809	65009	53.54	109226	270,642	40.36
13	Overlaps with GM H3k4me3 rep1 broadpeak	15494	65009	23.83	26988	48,956	55.13
14	Overlaps with GM H3k4me3 rep1 narrowpeak	10527	65009	16.19	29665	49,371	60.09
15	Overlaps with GM H3k4me3 rep2 broadpeak	16431	65009	25.27	28180	51,933	54.26

C. All bubble-containing fragments in HCE- (58403)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	15837	58403	27.12	22578	81,792	27.60
2	Overlaps with GM Ctcf rep1 narrowpeak	4432	58403	7.59	5583	17,957	31.09
3	Overlaps with GM DNaseI hyper rep1 broadpeak	7997	58403	13.69	14362	39,354	36.49
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	4495	58403	7.70	7724	18,569	41.60
5	Overlaps with GM DNaseI hyper rep2 broadpeak	8634	58403	14.78	16388	48,749	33.62
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	4218	58403	7.22	7131	17,987	39.65
7	Overlaps with GM H3k27me3 rep1 broadpeak	15687	58403	26.86	38239	156,520	24.43
8	Overlaps with GM H3k27me3 rep1 narrowpeak	1207	58403	2.07	1830	5,504	33.25
9	Overlaps with GM H3k27me3 rep2 broadpeak	14051	58403	24.06	29587	123,335	23.99
10	Overlaps with GM H3k36me3 rep1 broadpeak	8331	58403	14.26	22196	75,071	29.57
11	Overlaps with GM H3k36me3 rep1 narrowpeak	2362	58403	4.04	5884	15,244	38.60
12	Overlaps with GM H3k36me3 rep2 broadpeak	10335	58403	17.70	25088	84,232	29.78
13	Overlaps with GM H3k4me3 rep1 broadpeak	3819	58403	6.54	6138	14,153	43.37
14	Overlaps with GM H3k4me3 rep1 narrowpeak	2405	58403	4.12	6197	12,423	49.88
15	Overlaps with GM H3k4me3 rep2 broadpeak	4250	58403	7.28	6727	15,882	42.36

D. Early-firing bubble-containing fragments (39236)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	20598	39236	52.50	39657	84,276	47.06
2	Overlaps with GM Ctcf rep1 narrowpeak	10232	39236	26.08	15011	29,105	51.58
3	Overlaps with GM DNaseI hyper rep1 broadpeak	21941	39236	55.92	50255	92,916	54.09
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	14551	39236	37.09	30480	53,827	56.63
5	Overlaps with GM DNaseI hyper rep2 broadpeak	23402	39236	59.64	57416	111,912	51.30
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	13912	39236	35.46	28794	51,672	55.72
7	Overlaps with GM H3k27me3 rep1 broadpeak	13626	39236	34.73	43519	119,289	36.48
8	Overlaps with GM H3k27me3 rep1 narrowpeak	2075	39236	5.29	3642	7,839	46.46
9	Overlaps with GM H3k27me3 rep2 broadpeak	11170	39236	28.47	28661	85,511	33.52
10	Overlaps with GM H3k36me3 rep1 broadpeak	24679	39236	62.90	84365	192,769	43.76
11	Overlaps with GM H3k36me3 rep1 narrowpeak	10470	39236	26.68	33149	59,362	55.84
12	Overlaps with GM H3k36me3 rep2 broadpeak	26223	39236	66.83	88623	201,833	43.91
13	Overlaps with GM H3k4me3 rep1 broadpeak	13102	39236	33.39	23505	40,714	57.73
14	Overlaps with GM H3k4me3 rep1 narrowpeak	9020	39236	22.99	26243	42,043	62.42
15	Overlaps with GM H3k4me3 rep2 broadpeak	13738	39236	35.01	24377	42,571	57.26

E. Mid-firing bubble-containing fragments (39313)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	15213	39313	38.70	23739	86,185	27.54
2	Overlaps with GM Ctcf rep1 narrowpeak	5728	39313	14.57	7436	23,483	31.67
3	Overlaps with GM DNaseI hyper rep1 broadpeak	11398	39313	28.99	20637	57,005	36.20
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	6232	39313	15.85	10455	26,343	39.69
5	Overlaps with GM DNaseI hyper rep2 broadpeak	12825	39313	32.62	24820	72,902	34.05
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	5986	39313	15.23	9913	25,762	38.48
7	Overlaps with GM H3k27me3 rep1 broadpeak	20446	39313	52.01	61731	243,578	25.34
8	Overlaps with GM H3k27me3 rep1 narrowpeak	1939	39313	4.93	2755	9,029	30.51
9	Overlaps with GM H3k27me3 rep2 broadpeak	18394	39313	46.79	47135	191,378	24.63

10	Overlaps with GM H3k36me3 rep1 broadpeak	12750	39313	32.43	35082	123,340	28.44
11	Overlaps with GM H3k36me3 rep1 narrowpeak	3674	39313	9.35	8473	23,622	35.87
12	Overlaps with GM H3k36me3 rep2 broadpeak	15030	39313	38.23	38749	135,369	28.62
13	Overlaps with GM H3k4me3 rep1 broadpeak	5217	39313	13.27	8098	20,643	39.23
14	Overlaps with GM H3k4me3 rep1 narrowpeak	3358	39313	8.54	8434	17,427	48.40
15	Overlaps with GM H3k4me3 rep2 broadpeak	5771	39313	14.68	8837	23,003	38.42

F. Late-firing bubble-containing fragments (44527)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	9968	44527	22.39	13061	47,502	27.50
2	Overlaps with GM Ctcf rep1 narrowpeak	2034	44527	4.57	2389	8,379	28.51
3	Overlaps with GM DNaseI hyper rep1 broadpeak	2590	44527	5.82	3946	12,924	30.53
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	1121	44527	2.52	1575	4,480	35.16
5	Overlaps with GM DNaseI hyper rep2 broadpeak	2828	44527	6.35	4475	16,319	27.42
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	1054	44527	2.37	1435	4,413	32.52
7	Overlaps with GM H3k27me3 rep1 broadpeak	10022	44527	22.51	22544	84,000	26.84
8	Overlaps with GM H3k27me3 rep1 narrowpeak	624	44527	1.40	900	2,522	35.69
9	Overlaps with GM H3k27me3 rep2 broadpeak	9154	44527	20.56	18132	68,060	26.64
10	Overlaps with GM H3k36me3 rep1 broadpeak	2447	44527	5.50	4837	17,063	28.35
11	Overlaps with GM H3k36me3 rep1 narrowpeak	404	44527	0.91	761	2,285	33.30
12	Overlaps with GM H3k36me3 rep2 broadpeak	3824	44527	8.59	6729	21,963	30.64
13	Overlaps with GM H3k4me3 rep1 broadpeak	963	44527	2.16	1469	3,767	39.00
14	Overlaps with GM H3k4me3 rep1 narrowpeak	532	44527	1.19	1125	2,576	43.67
15	Overlaps with GM H3k4me3 rep2 broadpeak	1137	44527	2.55	1637	4,518	36.23

G. Early-firing bubble-containing fragments in HCE+ (34752)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	18597	34752	53.51	36187	75,207	48.12
2	Overlaps with GM Ctcf rep1 narrowpeak	9296	34752	26.75	13728	26,171	52.46
3	Overlaps with GM DNaseI hyper rep1 broadpeak	19735	34752	56.79	45420	82,814	54.85
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	13054	34752	37.56	27525	47,982	57.37
5	Overlaps with GM DNaseI hyper rep2 broadpeak	21028	34752	60.51	51865	99,726	52.01
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	12473	34752	35.89	25973	45,986	56.48
7	Overlaps with GM H3k27me3 rep1 broadpeak	12273	34752	35.32	39738	106,250	37.40
8	Overlaps with GM H3k27me3 rep1 narrowpeak	1900	34752	5.47	3361	7,174	46.85
9	Overlaps with GM H3k27me3 rep2 broadpeak	10005	34752	28.79	26049	75,818	34.36
10	Overlaps with GM H3k36me3 rep1 broadpeak	22265	34752	64.07	76390	171,808	44.46
11	Overlaps with GM H3k36me3 rep1 narrowpeak	9510	34752	27.37	30329	53,788	56.39
12	Overlaps with GM H3k36me3 rep2 broadpeak	23665	34752	68.10	80419	179,963	44.69
13	Overlaps with GM H3k4me3 rep1 broadpeak	11789	34752	33.92	21230	36,249	58.57
14	Overlaps with GM H3k4me3 rep1 narrowpeak	8140	34752	23.42	23678	37,561	63.04
15	Overlaps with GM H3k4me3 rep2 broadpeak	12352	34752	35.54	21967	37,777	58.15

H. Early-firing bubble-containing fragments in HCE- (4529)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	2030	4529	44.82	3530	9,208	38.34
2	Overlaps with GM Ctcf rep1 narrowpeak	951	4529	21.00	1303	2,982	43.70
3	Overlaps with GM DNaseI hyper rep1 broadpeak	2235	4529	49.35	4898	10,243	47.82
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	1514	4529	33.43	2999	5,944	50.45
5	Overlaps with GM DNaseI hyper rep2 broadpeak	2403	4529	53.06	5621	12,370	45.44
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	1456	4529	32.15	2859	5,774	49.52
7	Overlaps with GM H3k27me3 rep1 broadpeak	1365	4529	30.14	3813	13,226	28.83
8	Overlaps with GM H3k27me3 rep1 narrowpeak	176	4529	3.89	282	672	41.96
9	Overlaps with GM H3k27me3 rep2 broadpeak	1170	4529	25.83	2627	9,823	26.74

10	Overlaps with GM H3k36me3 rep1 broadpeak	2446	4529	54.01	8098	21,249	38.11
11	Overlaps with GM H3k36me3 rep1 narrowpeak	976	4529	21.55	2897	5,695	50.87
12	Overlaps with GM H3k36me3 rep2 broadpeak	2594	4529	57.28	8349	22,184	37.64
13	Overlaps with GM H3k4me3 rep1 broadpeak	1329	4529	29.34	2307	4,529	50.94
14	Overlaps with GM H3k4me3 rep1 narrowpeak	890	4529	19.65	2601	4,555	57.10
15	Overlaps with GM H3k4me3 rep2 broadpeak	1403	4529	30.98	2440	4,860	50.21

I. Mid-firing bubble-containing fragments in HCE+ (23955)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	9819	23955	40.99	15614	52,291	29.86
2	Overlaps with GM Ctcf rep1 narrowpeak	3843	23955	16.04	5024	14,833	33.87
3	Overlaps with GM DNaseI hyper rep1 broadpeak	7598	23955	31.72	14073	36,839	38.20
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	4068	23955	16.98	6867	16,805	40.86
5	Overlaps with GM DNaseI hyper rep2 broadpeak	8666	23955	36.18	17237	47,811	36.05
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	3963	23955	16.54	6646	16,528	40.21
7	Overlaps with GM H3k27me3 rep1 broadpeak	13369	23955	55.81	42561	157,393	27.04
8	Overlaps with GM H3k27me3 rep1 narrowpeak	1333	23955	5.56	1848	5,951	31.05
9	Overlaps with GM H3k27me3 rep2 broadpeak	12091	23955	50.47	32383	123,828	26.15
10	Overlaps with GM H3k36me3 rep1 broadpeak	8677	23955	36.22	24434	80,384	30.40
11	Overlaps with GM H3k36me3 rep1 narrowpeak	2557	23955	10.67	5983	15,531	38.52
12	Overlaps with GM H3k36me3 rep2 broadpeak	10114	23955	42.22	26792	87,914	30.48
13	Overlaps with GM H3k4me3 rep1 broadpeak	3431	23955	14.32	5334	13,071	40.81
14	Overlaps with GM H3k4me3 rep1 narrowpeak	2226	23955	9.29	5593	11,264	49.65
15	Overlaps with GM H3k4me3 rep2 broadpeak	3774	23955	15.75	5772	14,613	39.50

J. Mid-firing bubble-containing fragments in HCE- (15410)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	5413	15410	35.13	8151	34,053	23.94
2	Overlaps with GM Ctcf rep1 narrowpeak	1889	15410	12.26	2417	8,682	27.84
3	Overlaps with GM DNaseI hyper rep1 broadpeak	3818	15410	24.78	6600	20,268	32.56
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	2177	15410	14.13	3611	9,591	37.65
5	Overlaps with GM DNaseI hyper rep2 broadpeak	4176	15410	27.10	7619	25,214	30.22
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	2035	15410	13.21	3288	9,279	35.43
7	Overlaps with GM H3k27me3 rep1 broadpeak	7103	15410	46.09	19241	86,604	22.22
8	Overlaps with GM H3k27me3 rep1 narrowpeak	609	15410	3.95	910	3,087	29.48
9	Overlaps with GM H3k27me3 rep2 broadpeak	6329	15410	41.07	14807	67,850	21.82
10	Overlaps with GM H3k36me3 rep1 broadpeak	4091	15410	26.55	10707	43,191	24.79
11	Overlaps with GM H3k36me3 rep1 narrowpeak	1124	15410	7.29	2510	8,144	30.82
12	Overlaps with GM H3k36me3 rep2 broadpeak	4940	15410	32.06	12016	47,724	25.18
13	Overlaps with GM H3k4me3 rep1 broadpeak	1797	15410	11.66	2781	7,443	37.36
14	Overlaps with GM H3k4me3 rep1 narrowpeak	1141	15410	7.40	2859	6,198	46.13
15	Overlaps with GM H3k4me3 rep2 broadpeak	2009	15410	13.04	3084	8,436	36.56

K. Late-firing bubble-containing fragments in HCE+ (6263)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	1586	6263	25.32	2177	8,116	26.82
2	Overlaps with GM Ctcf rep1 narrowpeak	443	6263	7.07	527	1,907	27.64
3	Overlaps with GM DNaseI hyper rep1 broadpeak	649	6263	10.36	1085	3,275	33.13
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	319	6263	5.09	464	1,258	36.88
5	Overlaps with GM DNaseI hyper rep2 broadpeak	777	6263	12.41	1333	4,336	30.74
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	329	6263	5.25	453	1,312	34.53
7	Overlaps with GM H3k27me3 rep1 broadpeak	2813	6263	44.91	7387	25,704	28.74
8	Overlaps with GM H3k27me3 rep1 narrowpeak	203	6263	3.24	263	742	35.44
9	Overlaps with GM H3k27me3 rep2 broadpeak	2614	6263	41.74	5996	21,259	28.20

10	Overlaps with GM H3k36me3 rep1 broadpeak	654	6263	10.44	1447	4,893	29.57
11	Overlaps with GM H3k36me3 rep1 narrowpeak	142	6263	2.27	284	711	39.94
12	Overlaps with GM H3k36me3 rep2 broadpeak	1026	6263	16.38	5996	6,206	96.62
13	Overlaps with GM H3k4me3 rep1 broadpeak	272	6263	4.34	422	1,010	41.78
14	Overlaps with GM H3k4me3 rep1 narrowpeak	159	6263	2.54	391	792	49.37
15	Overlaps with GM H3k4me3 rep2 broadpeak	302	6263	4.82	437	1,206	36.24

L. Late-firing bubble-containing fragments in HCE- (38283)		# Bubble overlaps	# Bubbles	% Bubbles	# Factor overlaps	# Factors	% Factors
1	Overlaps with GM Ctcf rep1 broadpeak	8386	38283	21.91	10889	39,419	27.62
2	Overlaps with GM Ctcf rep1 narrowpeak	1592	38283	4.16	1863	6,480	28.75
3	Overlaps with GM DNaseI hyper rep1 broadpeak	1943	38283	5.08	2863	9,657	29.65
4	Overlaps with GM DNaseI hyper rep1 narrowpeak	804	38283	2.10	1114	3,227	34.52
5	Overlaps with GM DNaseI hyper rep2 broadpeak	2054	38283	5.37	3147	11,999	26.23
6	Overlaps with GM DNaseI hyper rep2 narrowpeak	727	38283	1.90	984	3,106	31.68
7	Overlaps with GM H3k27me3 rep1 broadpeak	7218	38283	18.85	15184	58,399	26.00
8	Overlaps with GM H3k27me3 rep1 narrowpeak	422	38283	1.10	638	1,785	35.74
9	Overlaps with GM H3k27me3 rep2 broadpeak	6549	38283	17.11	12150	46,873	25.92
10	Overlaps with GM H3k36me3 rep1 broadpeak	1794	38283	4.69	3391	12,178	27.85
11	Overlaps with GM H3k36me3 rep1 narrowpeak	262	38283	0.68	477	1,574	30.30
12	Overlaps with GM H3k36me3 rep2 broadpeak	2801	38283	7.32	4723	15,765	29.96
13	Overlaps with GM H3k4me3 rep1 broadpeak	693	38283	1.81	1050	2,765	37.97
14	Overlaps with GM H3k4me3 rep1 narrowpeak	374	38283	0.98	737	1,791	41.15
15	Overlaps with GM H3k4me3 rep2 broadpeak	837	38283	2.19	1202	3,317	36.24

Table SVIII – Correlation coefficients between origin densities (whether isolated or zonal) and epigenetic factors calculated for 1 Mb intervals

A. Non-epigenetic factor comparisons		Pearson CC	p-value	Spearman CC	Kendall CC
1	TR50 & origin density	-0.31	< 2.2e-16	-0.25	-0.17
2	TR50 & HCE (+/-)	-0.74	< 2.2e-16	-0.75	-0.55
3	Origin density & HCE (+/-)	0.26	< 2.2e-16	0.25	0.16
4	Origin density & abs HCE	0.50	< 2.2e-16	0.49	0.34
5	Avg RDs & TR50	0.09	6.47e-07	0.17	0.11
6	Avg origin RDs & HCE (+/-)	-0.04	0.01343	-0.05	-0.03
7	Isolated origin density & HCE (+/-)	-0.19	< 2.2e-16	-0.22	-0.15
8	Isolated origin density & abs HCE	0.26	< 2.2e-16	0.21	0.14
9	Zonal origin density & HCE (+/-)	0.35	< 2.2e-16	0.30	0.20
10	Zonal origin density & abs HCE	0.45	< 2.2e-16	0.46	0.32
11	Early-firing origin density & HCE (+/-)	0.58	< 2.2e-16	0.65	0.47
12	Mid-firing origin density & HCE (+/-)	0.19	< 2.2e-16	0.18	0.13
13	Late-firing origin density & HCE (+/-)	-0.48	< 2.2e-16	-0.53	-0.37

B. CTCF		Pearson CC	p-value	Spearman CC	Kendall CC
1	Rep1 broadpeak density & early-firing origin density	0.74	< 2.2e-16	0.73	0.53
2	Rep1 broadpeak density & mid-firing origin density	0.17	< 2.2e-16	0.26	0.18
3	Rep1 broadpeak density & late-firing origin density	-0.55	< 2.2e-16	-0.60	-0.43
4	Rep1 broadpeak density & origin density	0.58	< 2.2e-16	0.46	0.33
5	Rep1 broadpeak density & HCE (+/-)	0.60	< 2.2e-16	0.61	0.43
6	Rep1 narrowpeak & early-firing origin density	0.78	< 2.2e-16	0.76	0.57
7	Rep1 narrowpeak & mid-firing origin density	0.18	< 2.2e-16	0.26	0.19
8	Rep1 narrowpeak & late-firing origin density	-0.61	< 2.2e-16	-0.66	-0.49
9	Rep1 narrowpeak & origin density	0.58	< 2.2e-16	0.45	0.33
10	Rep1 narrowpeak & HCE (+/-)	0.64	< 2.2e-16	0.65	0.46

C. DNaseI HSS		Pearson CC	p-value	Spearman CC	Kendall CC
1	Rep1 broadpeak density & early-firing origin density	0.87	< 2.2e-16	0.86	0.67
2	Rep1 broadpeak density & mid-firing origin density	0.06	0.002223	0.22	0.16
3	Rep1 broadpeak density & late-firing origin density	-0.57	< 2.2e-16	-0.77	-0.57
4	Rep1 broadpeak density & origin density	0.59	< 2.2e-16	0.42	0.31
5	Rep1 broadpeak density & HCE (+/-)	0.63	< 2.2e-16	0.67	0.48
6	Rep2 broadpeak density & early-firing origin density	0.86	< 2.2e-16	0.86	0.68
7	Rep2 broadpeak density & mid-firing origin density	0.05	0.01678	0.20	0.14
8	Rep2 broadpeak density & late-firing	-0.57	< 2.2e-16	-0.77	-0.58

	origin density				
9	Rep2 broadpeak density & origin density	0.59	< 2.2e-16	0.41	0.30
10	Rep2 broadpeak density & HCE	0.64	< 2.2e-16	0.68	0.48
11	Rep1 narrowpeak & early-firing origin density	0.85	< 2.2e-16	0.86	0.66
12	Rep1 narrowpeak & mid-firing origin density	0.04	0.08435	0.19	0.14
13	Rep1 narrowpeak & late-firing origin density	-0.58	< 2.2e-16	-0.76	-0.57
14	Rep1 narrowpeak & origin density	0.55	< 2.2e-16	0.40	0.29
15	Rep1 narrowpeak & HCE (+/-)	0.64	< 2.2e-16	0.67	0.48
16	Rep2 narrowpeak & early-firing origin density	0.83	< 2.2e-16	0.85	0.66
17	Rep2 narrowpeak & mid-firing origin density	0.02	0.3945	0.18	0.12
18	Rep2 narrowpeak & late-firing origin density	-0.59	< 2.2e-16	-0.75	-0.57
19	Rep2 narrowpeak & origin density	0.53	< 2.2e-16	0.38	0.27
20	Rep2 narrowpeak & HCE (+/-)	0.64	< 2.2e-16	0.68	0.49

D. H3K27me3		Pearson CC	p-value	Spearman CC	Kendall CC
1	Rep1 broadpeak density & early-firing origin density	0.21	< 2.2e-16	0.34	0.23
2	Rep1 broadpeak density & mid-firing origin density	0.42	< 2.2e-16	0.42	0.30
3	Rep1 broadpeak density & late-firing origin density	-0.42	< 2.2e-16	-0.57	-0.41
4	Rep1 broadpeak density & origin density	0.36	< 2.2e-16	0.33	0.23
5	Rep1 broadpeak density & HCE (+/-)	0.46	< 2.2e-16	0.51	0.35
6	Rep2 broadpeak density & early-firing origin density	0.06	< 2.2e-16	0.21	0.14
7	Rep2 broadpeak density & mid-firing origin density	0.40	< 2.2e-16	0.43	0.30
8	Rep2 broadpeak density & late-firing origin density	-0.42	< 2.2e-16	-0.55	-0.39
9	Rep2 broadpeak density & origin density	0.23	< 2.2e-16	0.24	0.17
10	Rep2 broadpeak density & HCE (+/-)	0.38	< 2.2e-16	0.44	0.29
11	Rep1 narrowpeak & early-firing origin density	0.44	< 2.2e-16	0.45	0.32
12	Rep1 narrowpeak & mid-firing origin density	0.21	< 2.2e-16	0.30	0.21
13	Rep1 narrowpeak & late-firing origin density	-0.46	< 2.2e-16	-0.55	-0.40
14	Rep1 narrowpeak & origin density	0.38	< 2.2e-16	0.29	0.21
15	Rep1 narrowpeak & HCE (+/-)	0.48	< 2.2e-16	0.52	0.37

E. H3K4me3		Pearson CC	p-value	Spearman CC	Kendall CC
1	Rep1 broadpeak density & early-firing origin density	0.84	< 2.2e-16	0.85	0.66
2	Rep1 broadpeak density & mid-firing	0.03	< 2.2e-16	0.19	0.13

	origin density				
3	Rep1 broadpeak density & late-firing origin density	-0.54	< 2.2e-16	-0.72	-0.53
4	Rep1 broadpeak density & origin density	0.56	< 2.2e-16	0.40	0.29
5	Rep1 broadpeak density & HCE (+/-)	0.62	< 2.2e-16	0.67	0.48
6	Rep2 broadpeak density & early-firing origin density	0.84	< 2.2e-16	0.85	0.66
7	Rep2 broadpeak density & mid-firing origin density	0.03	0.147	0.19	0.13
8	Rep2 broadpeak density & late-firing origin density	-0.56	< 2.2e-16	-0.73	-0.54
9	Rep2 broadpeak density & origin density	0.56	< 2.2e-16	0.40	0.29
10	Rep2 broadpeak density & HCE (+/-)	0.62	< 2.2e-16	0.67	0.48
11	Rep1 narrowpeak & early-firing origin density	0.83	< 2.2e-16	0.84	0.65
12	Rep1 narrowpeak & mid-firing origin density	0.02	0.3726	0.17	0.12
13	Rep1 narrowpeak & late-firing origin density	-0.52	< 2.2e-16	-0.71	-0.53
14	Rep1 narrowpeak & origin density	0.56	< 2.2e-16	0.39	0.29
15	Rep1 narrowpeak & HCE (+/-)	0.61	< 2.2e-16	0.67	0.49

F. H3K36me3		Pearson CC	p-value	Spearman CC	Kendall CC
1	Rep1 broadpeak density & early-firing origin density	0.85	< 2.2e-16	0.82	0.63
2	Rep1 broadpeak density & mid-firing origin density	0.09	2.447e-05	0.22	0.16
3	Rep1 broadpeak density & late-firing origin density	-0.50	< 2.2e-16	-0.66	-0.47
4	Rep1 broadpeak density & origin density	0.60	< 2.2e-16	0.45	0.33
5	Rep1 broadpeak density & HCE (+/-)	0.63	< 2.2e-16	0.65	0.47
6	Rep2 broadpeak density & early-firing origin density	0.85	< 2.2e-16	0.82	0.63
7	Rep2 broadpeak density & mid-firing origin density	0.10	1.729e-06	0.23	0.17
8	Rep2 broadpeak density & late-firing origin density	-0.50	< 2.2e-16	-0.65	-0.46
9	Rep2 broadpeak density & origin density	0.61	< 2.2e-16	0.48	0.35
10	Rep2 broadpeak density & HCE (+/-)	0.63	< 2.2e-16	0.65	0.46
11	Rep1 narrowpeak & early-firing origin density	0.82	< 2.2e-16	0.79	0.61
12	Rep1 narrowpeak & mid-firing origin density	0.05	0.008073	0.20	0.14
13	Rep1 narrowpeak & late-firing origin density	-0.44	< 2.2e-16	-0.64	-0.48
14	Rep1 narrowpeak & origin density	0.60	< 2.2e-16	0.42	0.32
15	Rep1 narrowpeak & HCE (+/-)	0.57	< 2.2e-16	0.65	0.48

TABLE SIX. SIGNIFICANCE OF NS/BUBBLE OVERLAPS & G-QUADRUPLEX DISTRIBUTIONS

A. Distribution of nascent strands (NS) vis-à-vis bubbles

Cell line	Total NS sites	NS that overlap bubbles	Percent	Mean random	SD Random	p-value	Fold-change	Z-score
IMR-90 NS	256990	116378	45.3%	96,456	320.36	<1e-4	1.21	62.19
H9 NS	208520	96088	46.1%	79,671	280.87	<1e-4	1.21	58.45
iPS NS	246866	114332	46.3%	93,619	301.88	<1e-4	1.22	68.61
HeLa NS	233545	107844	46.2%	87,645	290.69	<1e-4	1.23	69.49

B. Distribution of bubble-containing fragments vis-à-vis nascent strands

Cell line	Total bubbles	Bubbles that overlap NS	Percent	Mean random	SD random	p-value	Fold-change	Z-score
IMR-90	123297	45391	36.8%	38719	120.44	<1e-4	1.17	55.40
H9	123297	40841	33.1%	34822	116.25	<1e-4	1.17	51.78
iPS	123297	45779	37.1%	38721	116.90	<1e-4	1.18	60.38
HeLa	123297	45361	36.8%	37755	115.95	<1e-4	1.20	65.60

C. Distribution of G-quadruplexes (G-Qs) vis-à-vis bubble-containing fragments

	Total G-Qs	G-Qs that overlap bubbles	Percent	Mean random	SD random	p-value	Fold-change	Z-score
	934355	369178	39.5%	308738	975.10	<1e-4	1.20	61.98
	Total bubbles	Bubbles that overlap G-Qs	Percent	Mean random	SD random	p-value	Fold-change	Z-score
	123297	79338	64.3%	75485	136.40	<1e-4	1.05	28.25