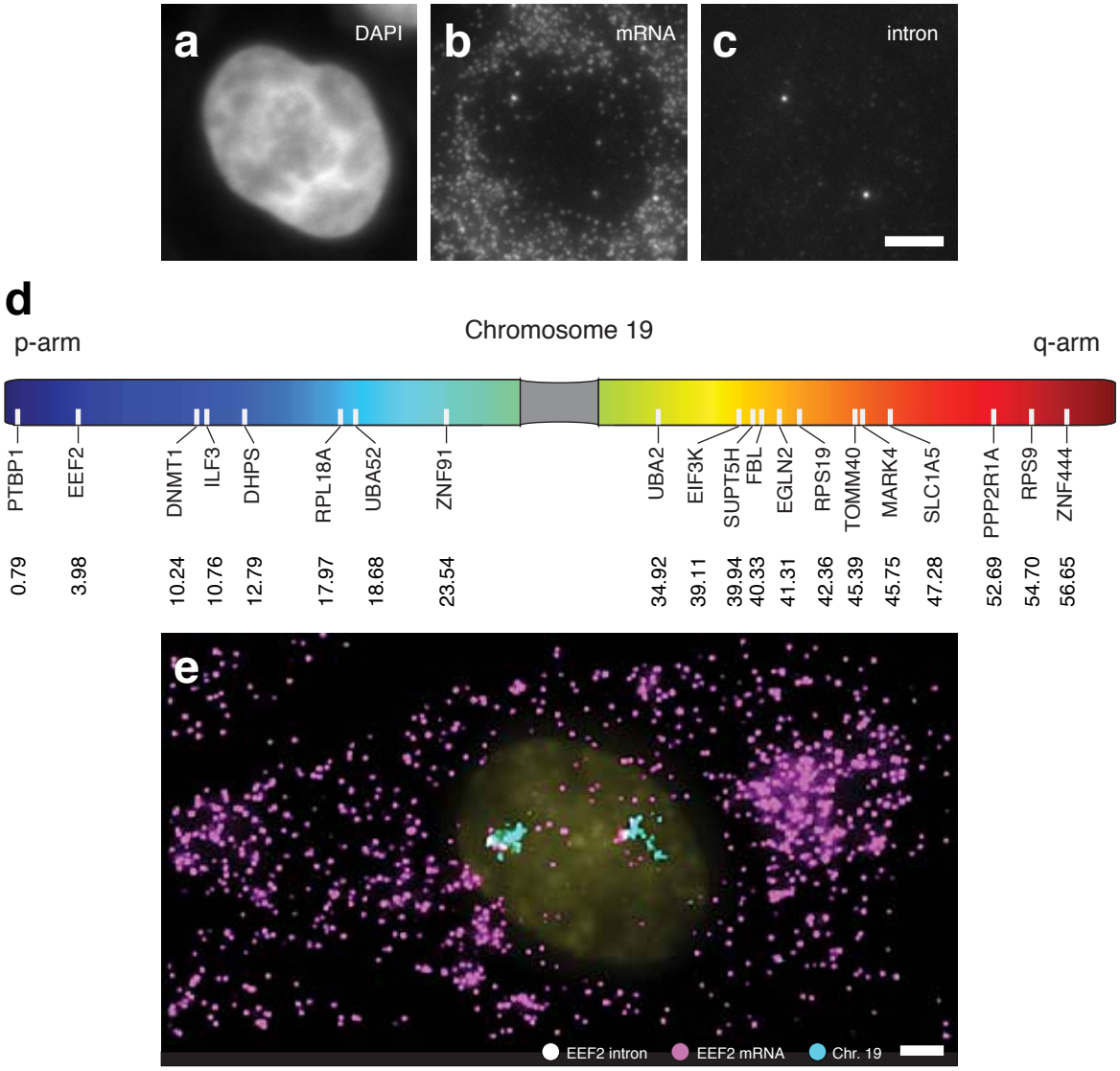
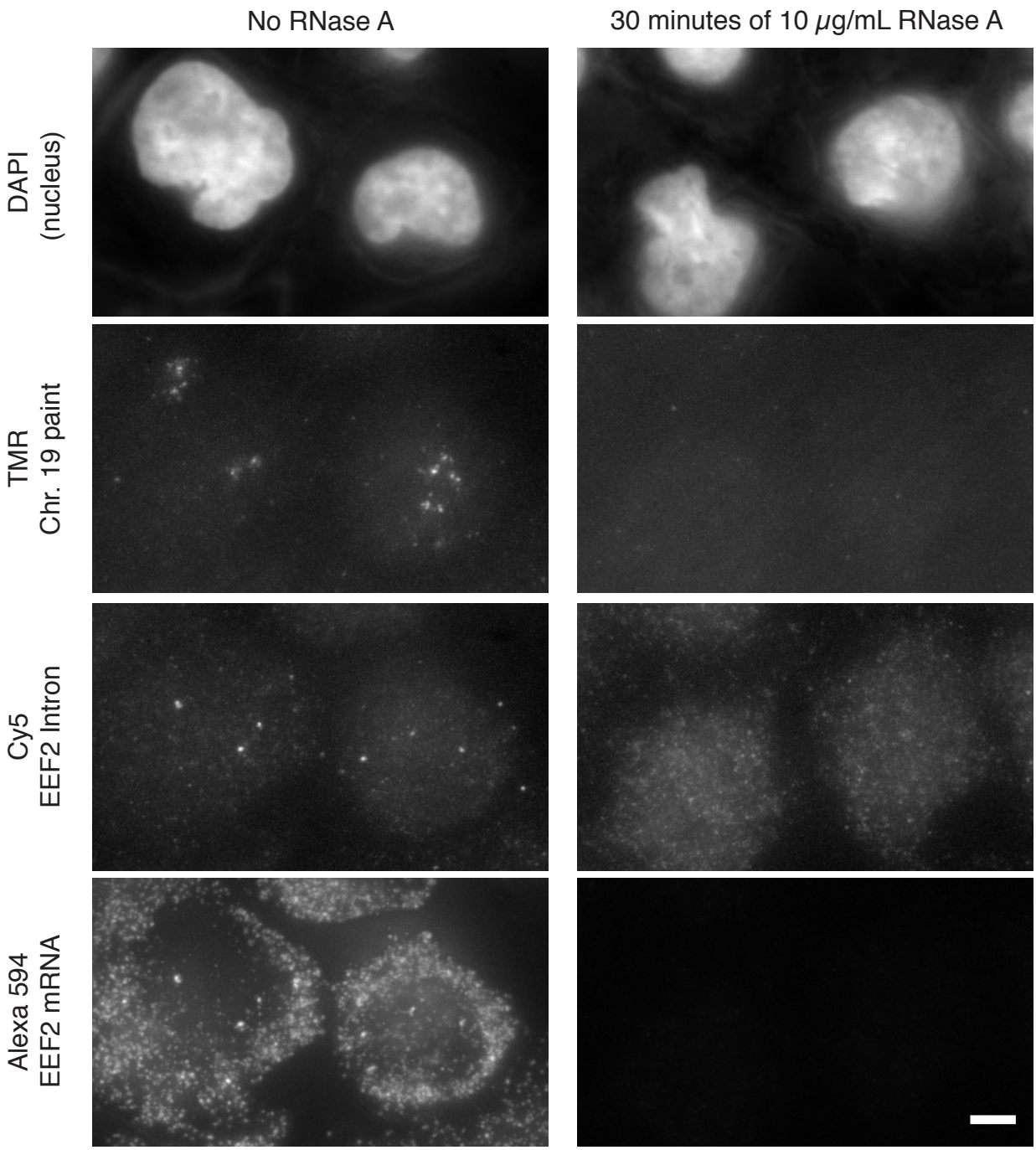
Supplementary Figure 1



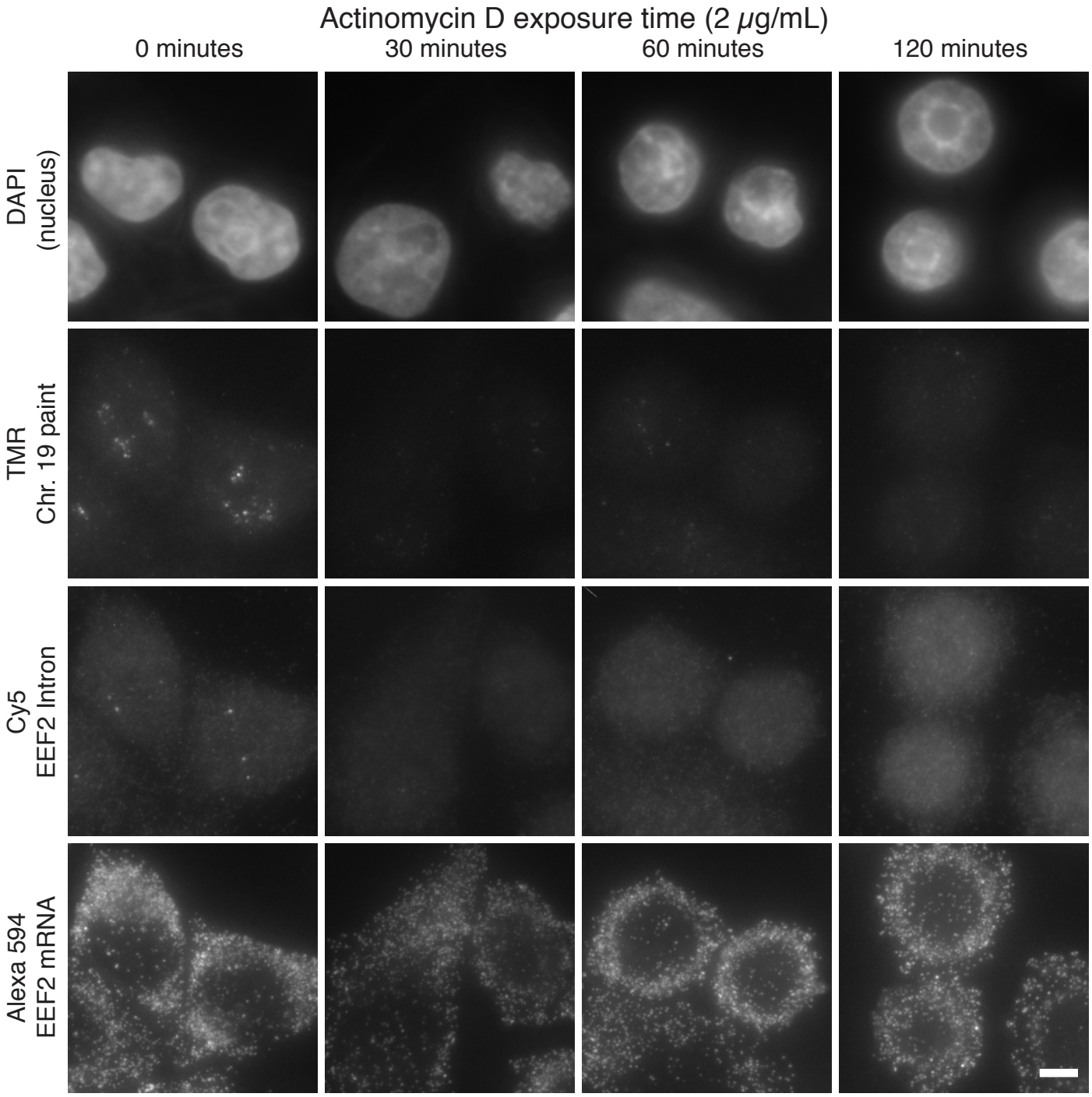
Supplementary Fig. 1. Targeting 20 introns enables chromomal paints. **a**. DAPI stain of the cell's nucleus. **b**. RNA FISH targeting mRNA from the EEF2 gene (labeled with Alexa594 dye). **c**. RNA FISH targeting EEF2 introns (labeled with Cy3). **d**. Genomic locations of all genes on chromosome 19 whose introns we targeted. We selected many potential genes based on high abundance of their mRNA (determined by RNA-seq and RT-PCR), then narrowing our focus to a subset of those with high intron spot frequencies as measured by RNA FISH. Ultimately, the set of genes we picked had spot frequencies ranging between 10% and 80% of chromosomes exhibiting a spot; see Fig. 2 for some representative numbers. The smallest genomic distance between loci was 0.36 megabases (between TOMM40 and MARK4). **e**. 17 gene intron chromosome 19 "paint" (cyan, labeled with Cy3), intron of EEF2 (white, labeled with Atto 647N), EEF2 mRNA (magenta, labeled with Alexa 594), and the nucleus (yellow), labeled with DAPI. All images are maximum z-projections of a three-dimensional z-stack. Scale bars are 5μm long.

Supplementary Figure 2



Supplementary Fig. 2. RNase A eliminates intron spots. We exposed HeLa cells to 10 μg/mL of RNase A for 30 minutes after fixation and before hybridization (right panels), with control cells exposed to the same procedure but without the addition of RNase. The top row contains images of the DAPI nuclear stain. The second row contains images in which we labeled the introns of all genes except EEF2, thus painting active genes in chromosome 19, and in the third row we labeled the introns of EEF2. In the fourth row, we labeled EEF2 mRNA. All images are maximum intensity projections of a z-stack of fluorescence images. The scale bar is 5μm long and applies to all images depicted. These results show that RNase A eliminates all the FISH signals we observed in the cells. This shows that the signals we are detecting are due to binding to RNA and not DNA, showing that our probes are detecting RNA from actively transcribing genes and not the DNA of gene itself.

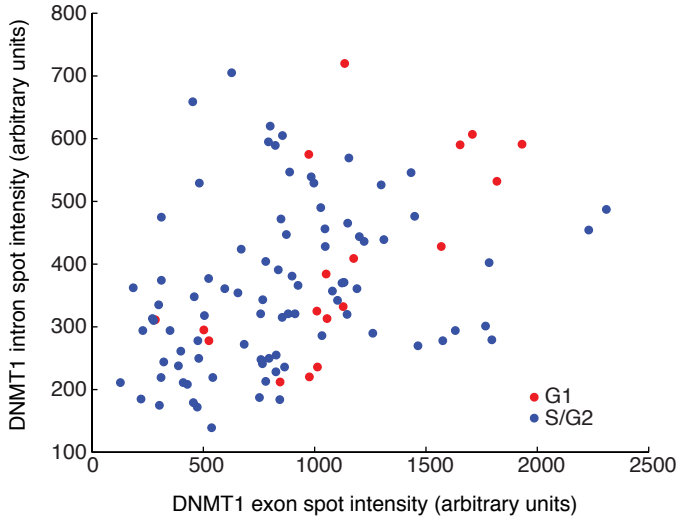
Supplementary Figure 3



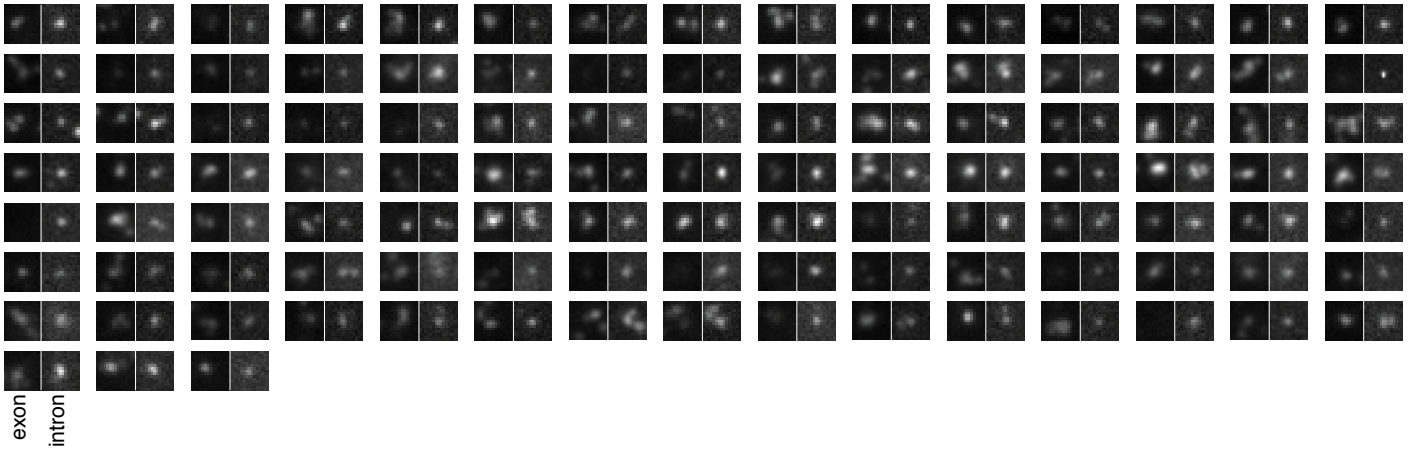
Supplementary Fig. 3. Inhibition of transcription by Actinomycin D results in rapid degradation of intronic RNA spots. We exposed HeLa cells to 2 μg/mL of actinomycin D for varying amounts of time as indicated. The top row contains images of the DAPI nuclear stain. The second row contains images in which we labeled the introns of all genes except EEF2, thus painting active genes in chromosome 19, and in the third row we labeled the introns of EEF2. In the fourth row, we labeled EEF2 mRNA. All images are maximum intensity projections of a z-stack of fluorescence images. The scale bar is 5μm long and applies to all images depicted. We found that virtually all the intronic RNA disappeared or greatly diminished in intensity after 30 minutes of Actinomycin D exposure. In the case of EEF2 intron, we found absolutely no spots; in the case of the chromosome paint, we did occasionally see dim spots even at later time points, although they were considerably dimmer. These may be residual RNA still attached to RNA polymerases stalled by Actinomycin D. We observed mature mRNA at all timepoints, indicating that the cells were still alive and that the treatment did not affect the RNA itself. Altogether, our results show that intronic RNA degrades rapidly; thus, the presence of an intronic RNA spot indicates that the targeted gene is transcriptionally active.

Supplementary Figure 4

a



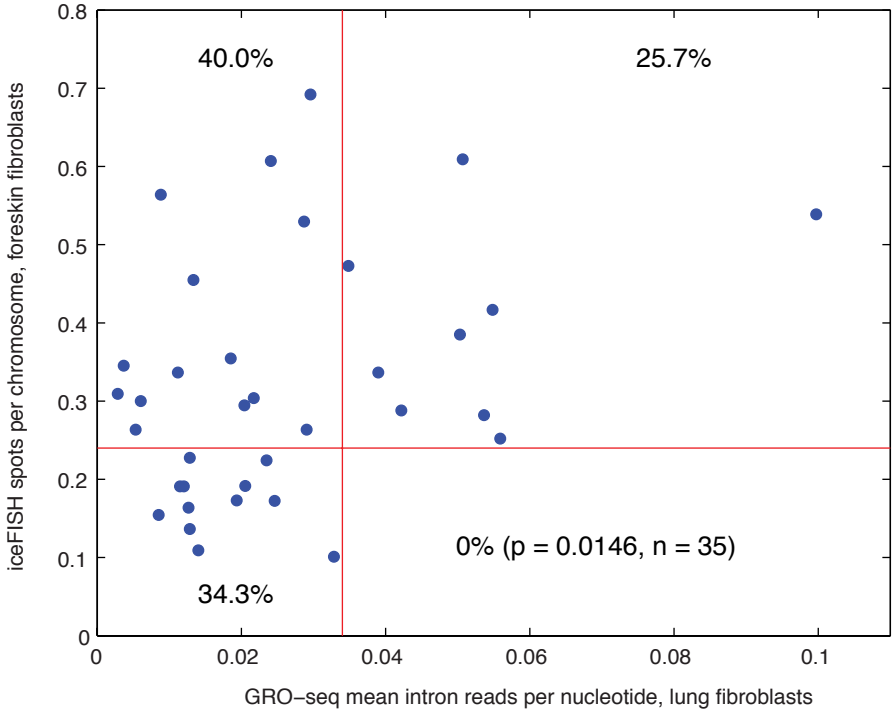
b



Supplementary Fig. 4. Analysis of exon vs. intron spot intensity. **a**. We show intron vs. exon spot intensity, defined as maximum pixel intensity of the spot minus the nearby background pixel intensity. Blue spots correspond to S/G2 phase cells (as scored by Cyclin A2 mRNA levels); red spots correspond to G1 cells. We analyzed a total of 105 spots. The correlation between intron and exon intensity is 0.378. **b**. Images depicting all the exon and intron spot pairs we analyzed (exon on left, intron on right). The slight shift between the exon and intron spot results from slight registration shifts in the images between fluorescence channels. We found that only 2 out of 105 intron spots displayed a lack of corresponding exon spot.

These results further establish that the presence of an intron spot corresponds to active transcription of the gene. We almost never observe an intron spot wthout any corresponding exon spot. Moreover, when there is no visible intron spot, we never find a bright transcription site in the exon channel (data not shown). Also, these results provide further evidence that the intron spots are truly located at the site of transcription because the intron spots strongly colocalize with bright exon spots that researchers have shown to represent nascent transcripts emanating from the site of transcription (see Levsky and Singer Science 2002, Vargas et al. PNAS 2005, Chubb et al. Current Biology 2006, Raj et al. PLoS Biology 2006).

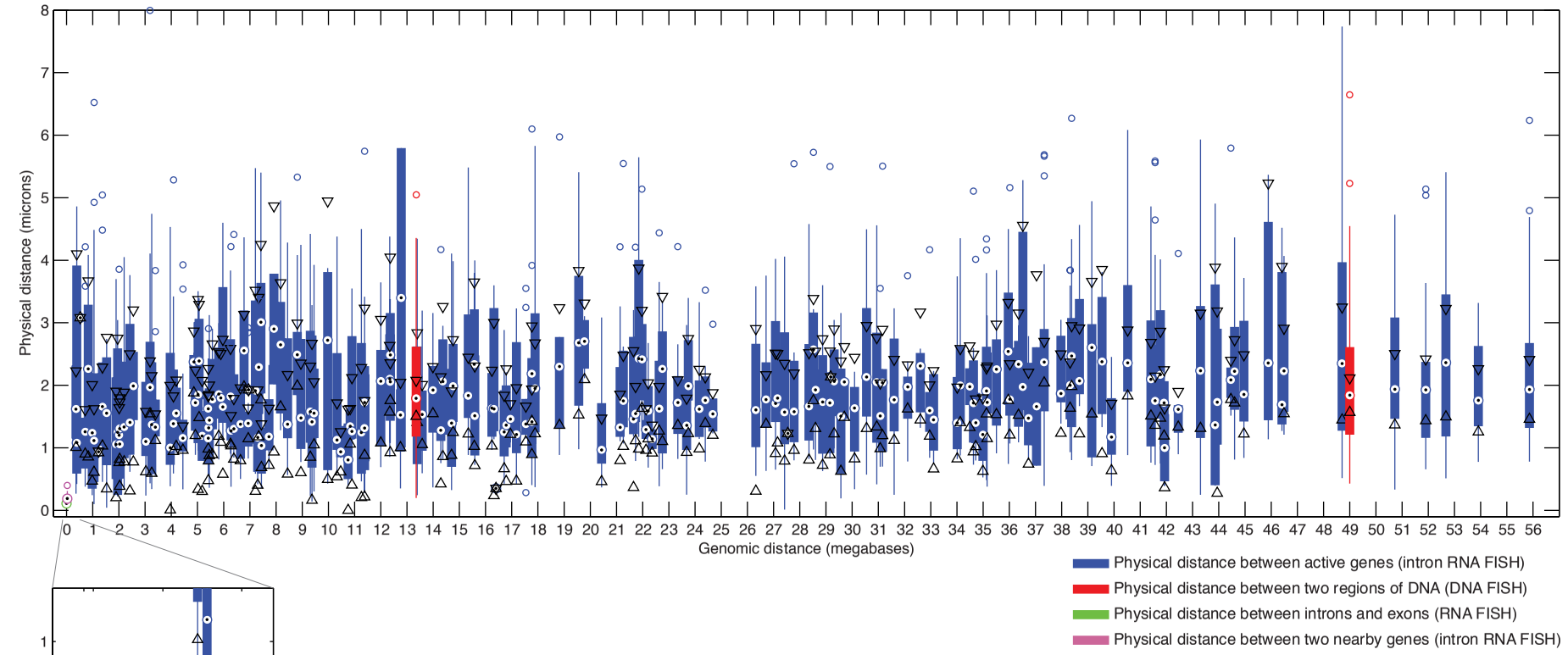
Supplementary Figure 5

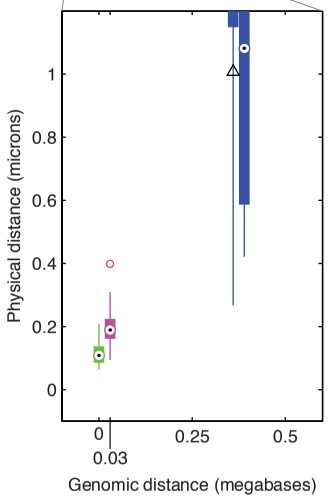


Supplementary Fig. 5. Comparison of iceFISH spot frequency with GRO-seq data (Core et al. Science 2008). In order to compute the GRO-seq mean intron reads, we measured the average number of reads within the first 3000 intronic bases in the transcript (which was roughly the same region that we probed by iceFISH in our study). We computed the frequency as discussed at length in the main text. We also included data from a further 15 genes on chromosome 19 not studied in the main paper to provide better sampling in the high GRO-seq regime.

Our observation is that above a certain number of GRO-seq reads, the iceFISH spot frequency was always relatively high. We also point out that there is a large set of points that we have not shown that are at the origin of the graph: they have no GRO-seq reads and generally do not show intronic FISH spots. (We have not included these data as we were not interested in intron probes that do not show spots and so did not systematically analyze such probes.) We believe there are many reasons that the correlation we observe is rather noisy: 1. these two sets of data arise from different cell lines and represent completely different treatments and procedures; 2. comparing iceFISH signals from gene to gene is not necessarily valid (nor are those from GRO-seq); and 3. our data are solely from G1 cells, whereas GRO-seq is from a mixed population. As such, we believe that the fact that these data show that high GRO-seq signals correspond to high iceFISH spot frequency is a valid means by which to measure transcriptional activity.

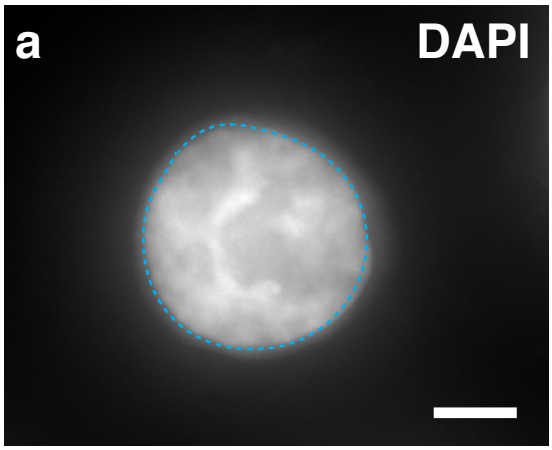
Supplementary Figure 6

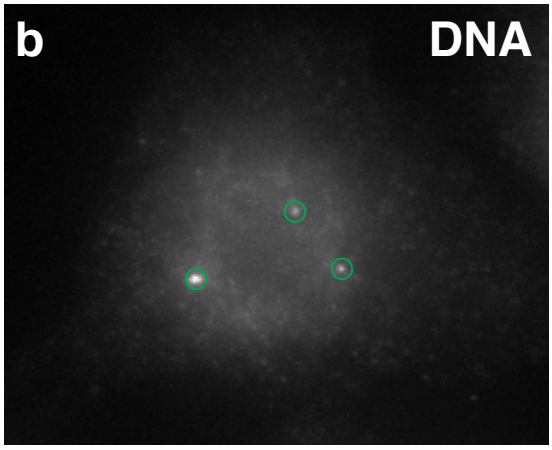


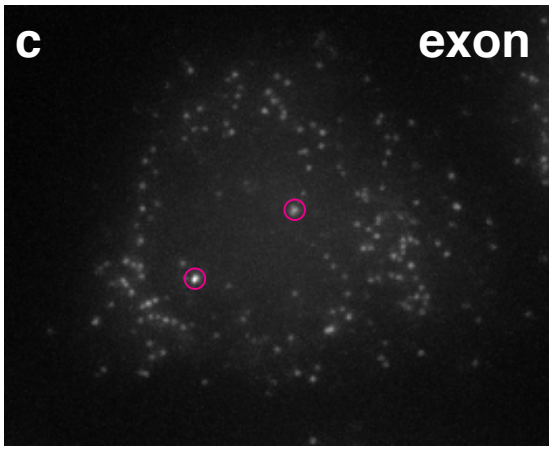


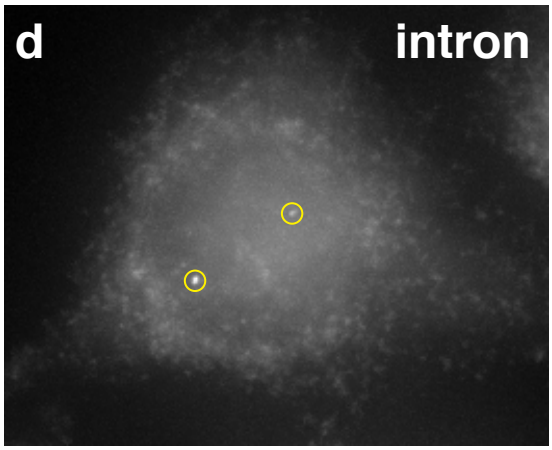
Supplementary Fig. 6. Comparison of genomic and physical distances between actively transcribing pairs of genes. Each box plot depicts a statistical analysis of the distances between a particular pair of actively transcribing genes, and we placed the box plot at a location on the x-axis representing the genomic distance between the pair of genes. The spot in the center of the box corresponds to the median distance, the box itself corresponds to 25th and 75th percentiles, and the whiskers reflect the range of the data (with the open circles reflecting data points deemed outliers). The open triangles provide comparison intervals: two medians are different at a 5% significance level if the intervals represented by these triangles do not overlap. The red box plots correspond to distances measured by DNA FISH (the experimental details of which are described in the methods). There is no particular difference between the DNA FISH results and the distances between transcriptionally active loci at the length scales we examined, although our results indicate that genetically proximal transcriptionally active loci (<0.5 megabases) are more spread out physically than pure DNA FISH measurements that do not distinguish between transcriptional status (Mateos-Langerak et al. PNAS 106:3812-3817. 2009). However, a simple model demonstrates that our data are compatible with these previous results (see dicussion point in supplement). The inset shows the physical distance between loci separated by 30 kilobases (EIF3K and ACTN4; magenta) and introns and exons of the same gene (DNMT1; green). These data show that intron spots are very close to the site of transcription as measured by exonic probes (we believe the measurements are almost to within our spatial discrimination limit), and that genetically very proximal genes do indeed come very close to each other, showing that the spread out characteristics that we observe are not just an artifact of our method.

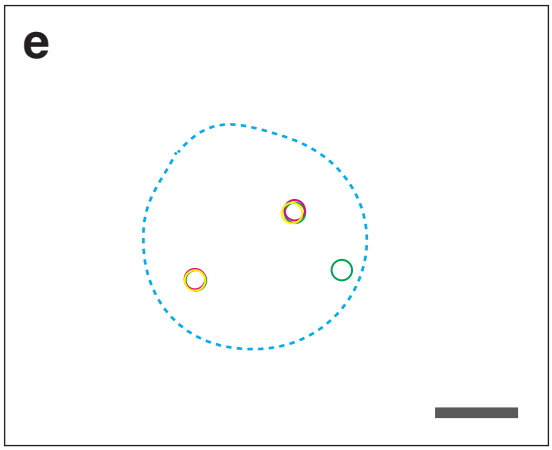
Supplementary Figure 7





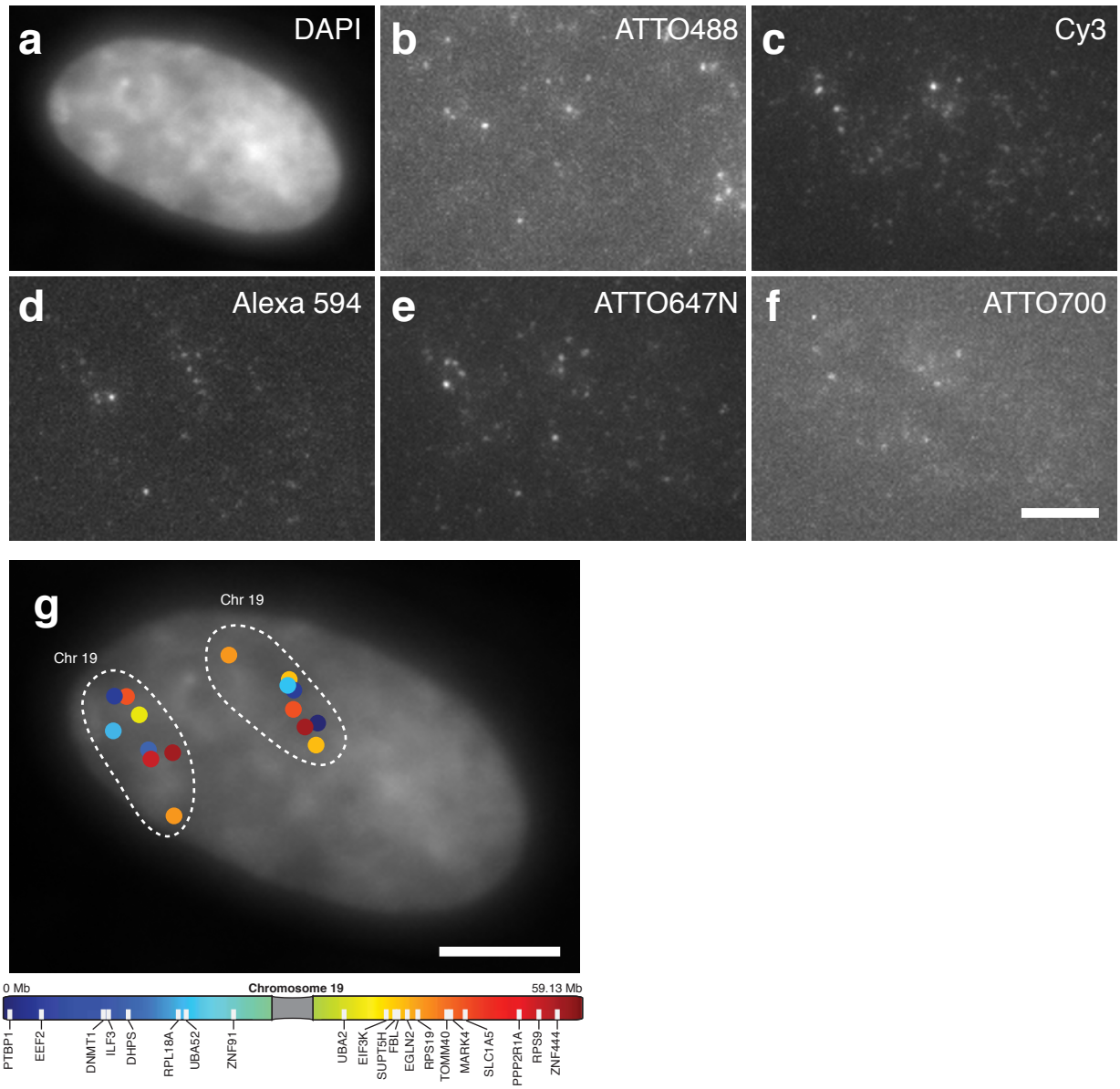






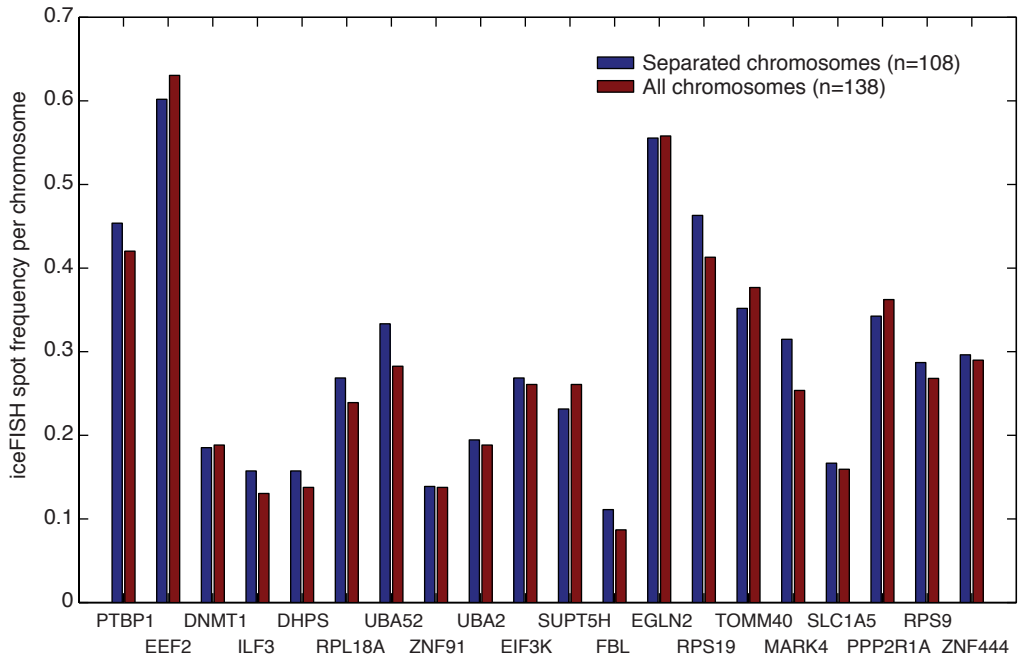
Supplementary Fig. 7. Colocalization of intron RNA FISH signals, exon RNA FISH signals, and DNA FISH signals. As we describe in the methods, we performed a protocol in which we combined DNA and RNA FISH, here using BAC probes targeting the DNA, RNA FISH probes targeting the exons of SLC1A5, and RNA FISH probes targeting the introns of SLC1A5 (**b-d**). We found that all the exon and intron signals colocalized with the DNA FISH signals (**e**) to within our detection limits. These results show that the intron RNA FISH signals we observed indeed remain at the site of transcription. Note that we observe three copies of the gene via DNA FISH, but only two of the three copies of the gene are transcriptionally active. This shows that our RNA FISH probes are not inadvertently binding to DNA. Another proof that our probes are not merely targeting DNA is the fact that both the exon and intron signals are brighter than what one would expect from a single RNA molecule. If the probes were bound to the gene's DNA, one would only see fluorescence intensity equivalent to that of a single RNA molecule. The scale bar is 5μm long.

Supplementary Figure 8



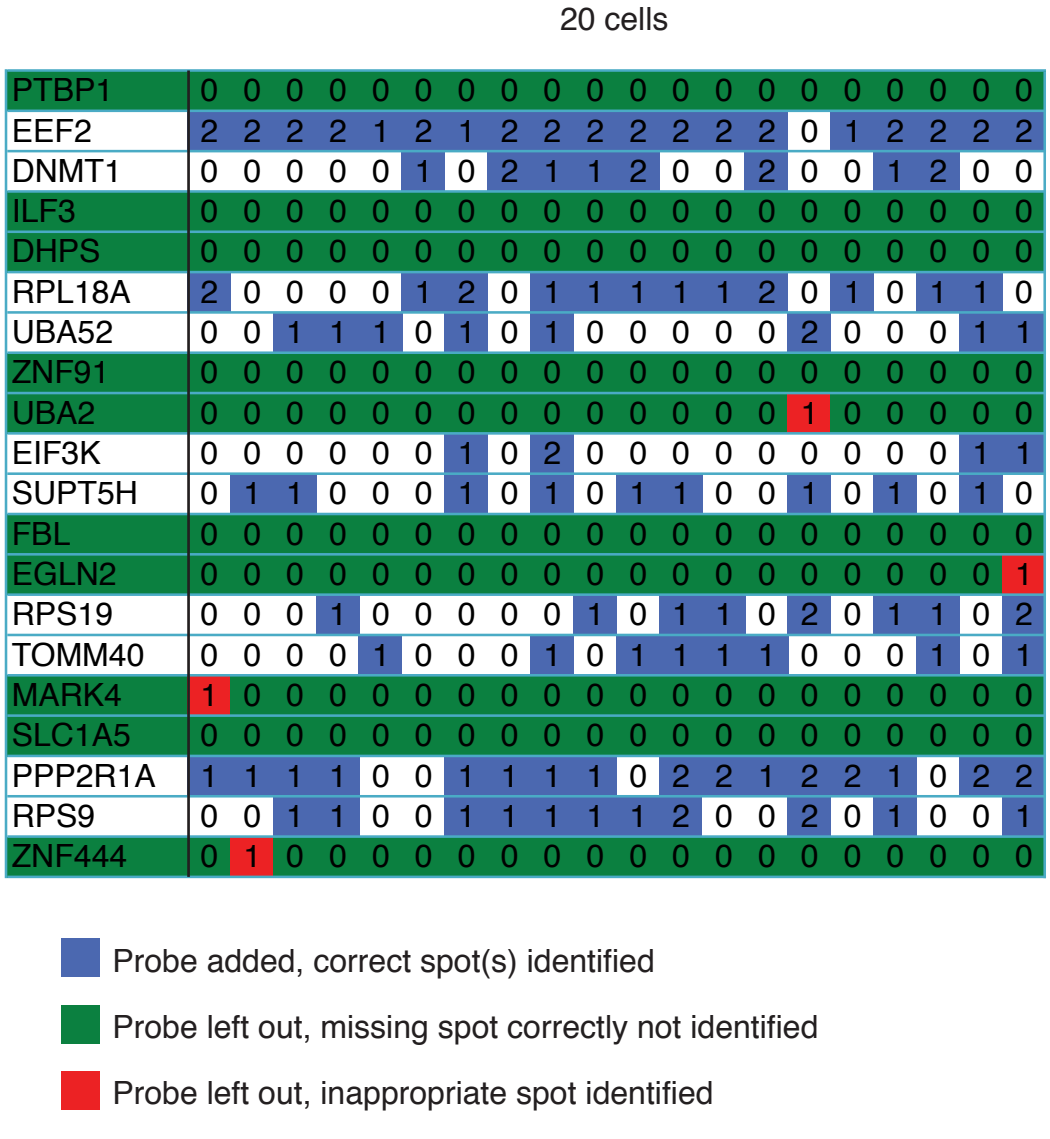
Supplementary Fig. 8. Images in individual fluorescence channels. **a-f.** Images for each fluorescence channels from the nucleus (labeled iwth DAPI in (a)) that we stained with probes labeled with the pseudocoloring scheme detailed in Supplementary Table 1. Along with probes targeting the introns, we also included probes targeting Cyclin A2 mRNA to determine position in cell cycle and SUZ12 mRNA as a fiducial marker. **g.** Computational identification of chromosome 19 gene positions. All images are maximum z-projections of a three-dimensional z-stack. All scale bars are 5μm long.

Supplementary Figure 9



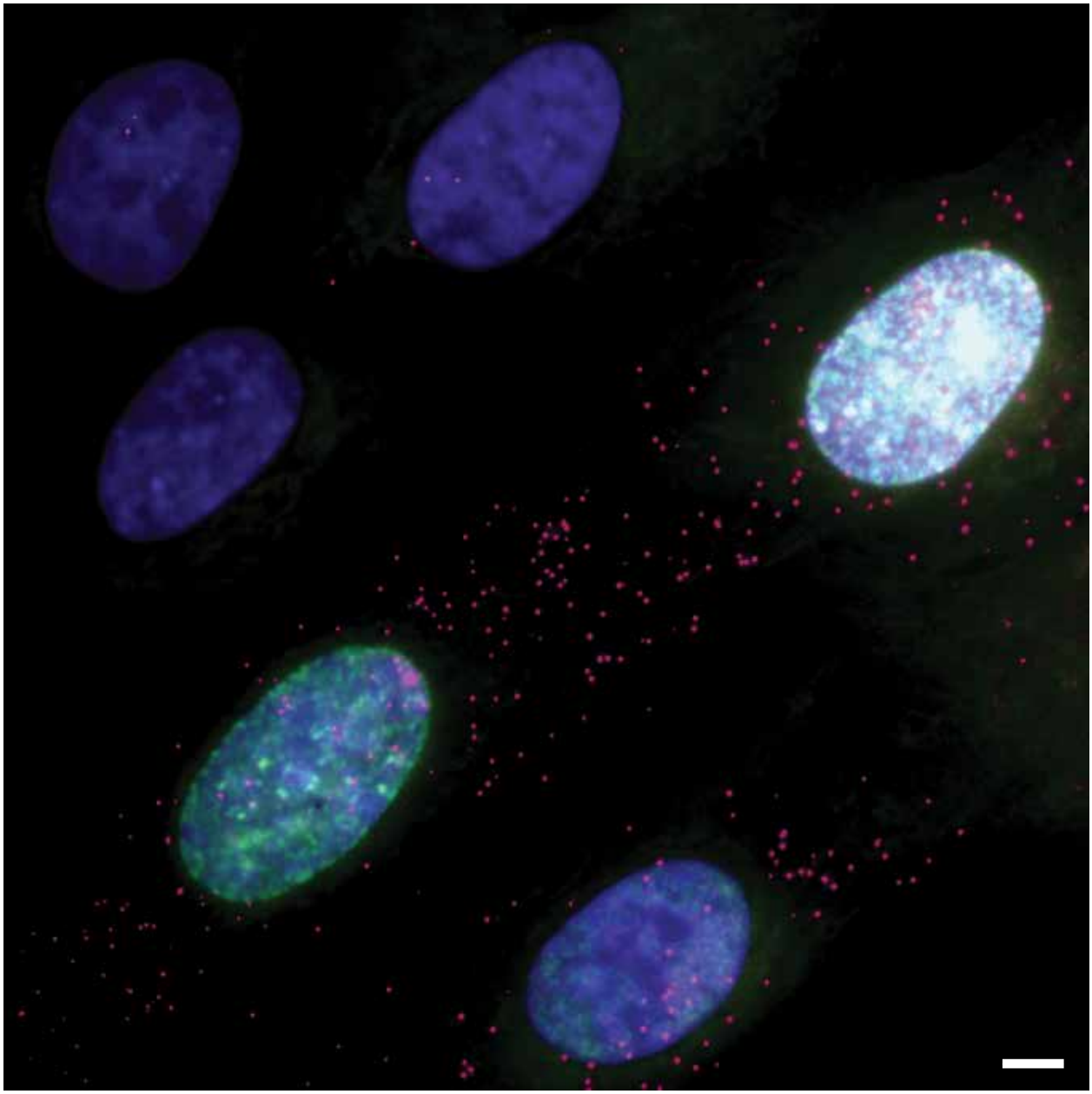
Supplementary Fig. 9. Chromosome identification does not bias spot frequency. In our analysis, we had to exclude chromosome in which both copies of chromosome 19 were too close for us to separate. In order to check whether this exclusion introduced any bias into our measurements, we measured iceFISH spot frequency for all chromosomes (including those which were excluded from transcriptional profiles). We found that the frequency was essentially identical to what we obtained from analyzing just the non-overlapping chromosomes, showing that our exclusion of overlapping chromosomes did not bias our results.

Supplementary Figure 10



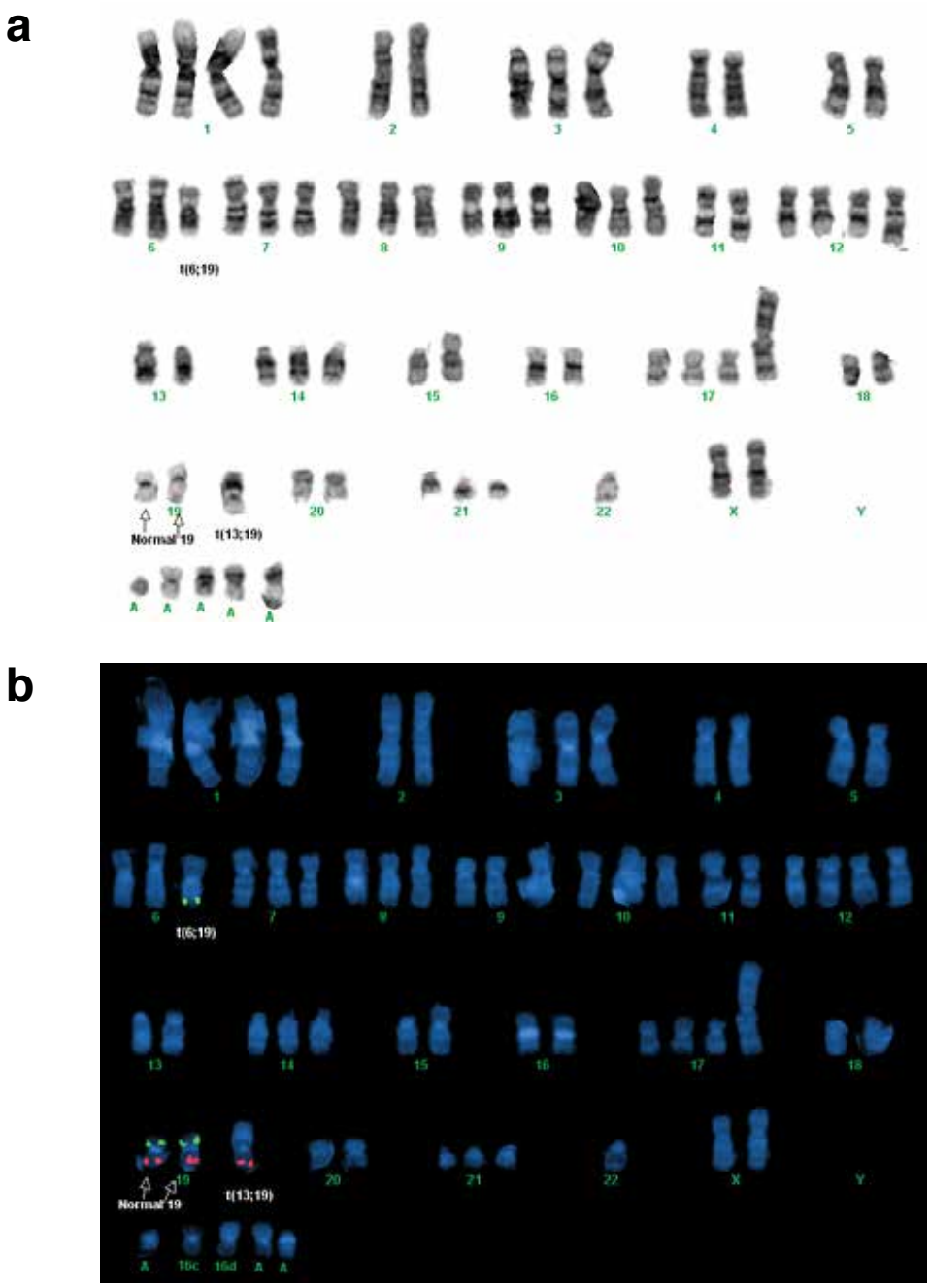
Supplementary Fig. 10. False identification rate. In order to assess the degree to which our intron spot color-coding scheme resulted in false positives, we left out 10 of the 20 probes (labeled in green) and ran the resultant images through our normal spot identification pipeline. In blue, we have indicated all the spots that we found that we ended up assigning to genes whose probes had been added to the hybridization (i.e., correct identifications). In red are spots that we misidentified in the sense that they correspond to genes that we hadn't added to the hybridization. We found that misidentified spots were relatively rare, with roughly 97% of spots we identified being assigned to genes that we had actually targeted in our hybridization. In order to minimize bias, we had another person in the lab randomly select 10 genes to leave out of the hybridization.

Supplementary Figure 11



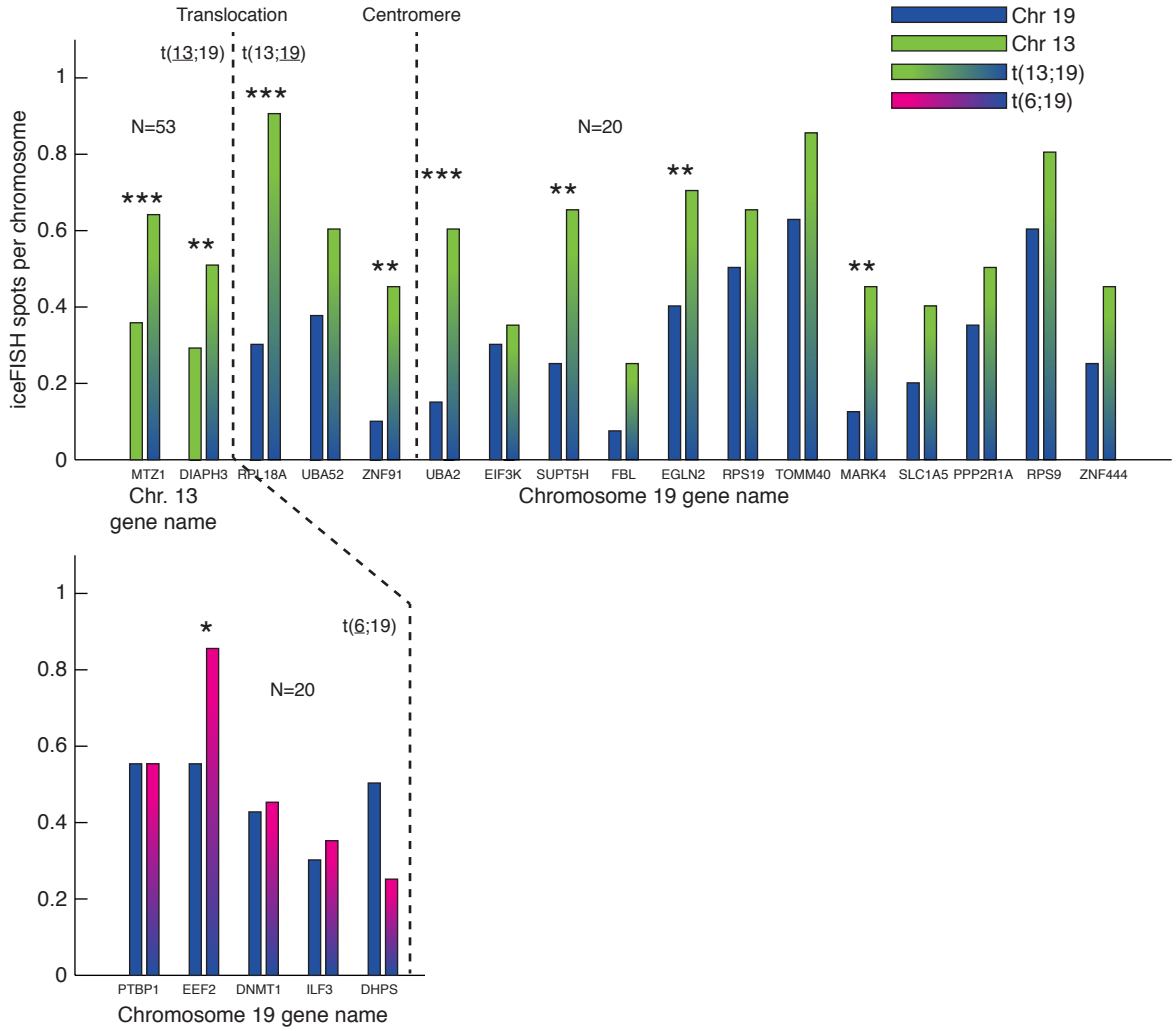
Supplementary Fig. 11. Labeling Cyclin A2 mRNA enables detection of cells in the S, G2, and M phases of the cell cycle. In order to isolate cells in G0/G1 for analysis, we performed RNA FISH targeting Cyclin A2 (magenta), which Eward et al. 2004 have shown to be present during S, G2, and M phases. We observed highly variegated expression, with some cells having having high levels of expression and some having very few mRNA. In order to demonstrate that the high expressing cells were indeed in S-M phase, we incubated the cells with Click-iT EdU 10μM for 5 minutes before fixation (green); the Click-iT EdU reagent incorporates into polymerizing DNA and produces a signal in cells undergoing DNA replication. We found that every cell displaying Click-iT EdU signal had high levels of Cyclin A2 mRNA, showing that Cyclin A2 mRNA provides a strong marker for cell-cycle. Sometimes a cell would have high levels of Cyclin A2 but would not be undergoing DNA replication; these cells are in G2, as we observed several double intron spots in these cells, indicating that those cells had already duplicated their DNA. In thus study, we were primarily interested in cells that had not undergone replication, so we selected cells with low levels of Cyclin A2. Note that we could not use the Click-iT EdU kit in combination with iceFISH because we found that incubating cells with the Click-iT reagent resulted in an abolition of transcription. We stained the nuclei with DAPI (purple); the scale bar is 5μm long.

Supplementary Figure 12



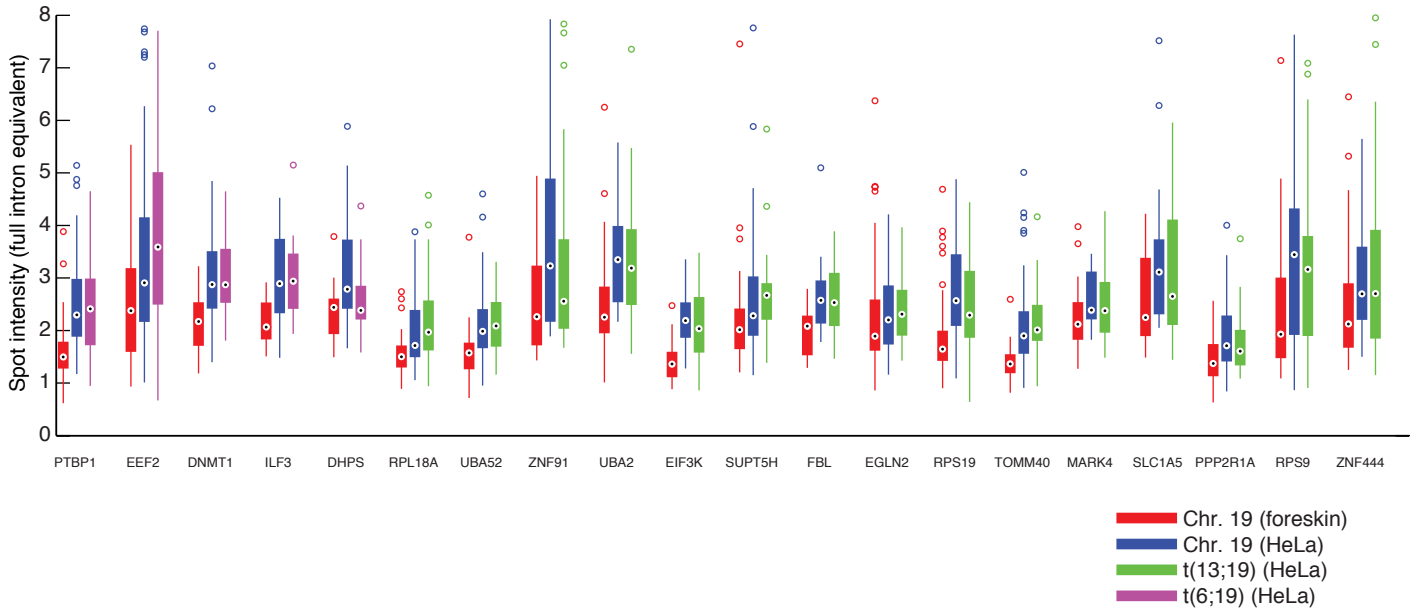
Supplementary Fig. 12. G-band and DNA FISH analysis confirms the translocated copies of chromosome 19 in our lab's HeLa cell line. a. We performed G-band analysis of a metaphase spread of our chromosomes to identify both the numbers of chromosomes in the cells as well as identify potential translocations. The numbers (and X and Y) correspond to the identities of the intact chromosomes. The specific translocations we were interested in for this study are fusions of portions of chromosome 19 to chromosome 13 and 6, denoted t(13;19) and t(6;19), respectively. **b**. We wanted to further verify the translocations of chromosome 19 via DNA FISH, which we performed upon the same chromosomes for the G-band analysis in **a**. We probed the chromosomes with probes against the p-arm of chromosome 19 (green) and the q-arm of chromosome 19 (red). This confirmed the results of our G-band analysis.

Supplementary Figure 13



Supplementary Fig. 13. Biological replicate of the data presented in Figure 3. We performed the exact same analysis as described in the legend of that figure. In this case, we analyzed 20 cells with the full set of 20 chromosome 19 probes and 53 cells in which we analyzed expression of MTZ1 and DIAPH3 on chromosome 13.

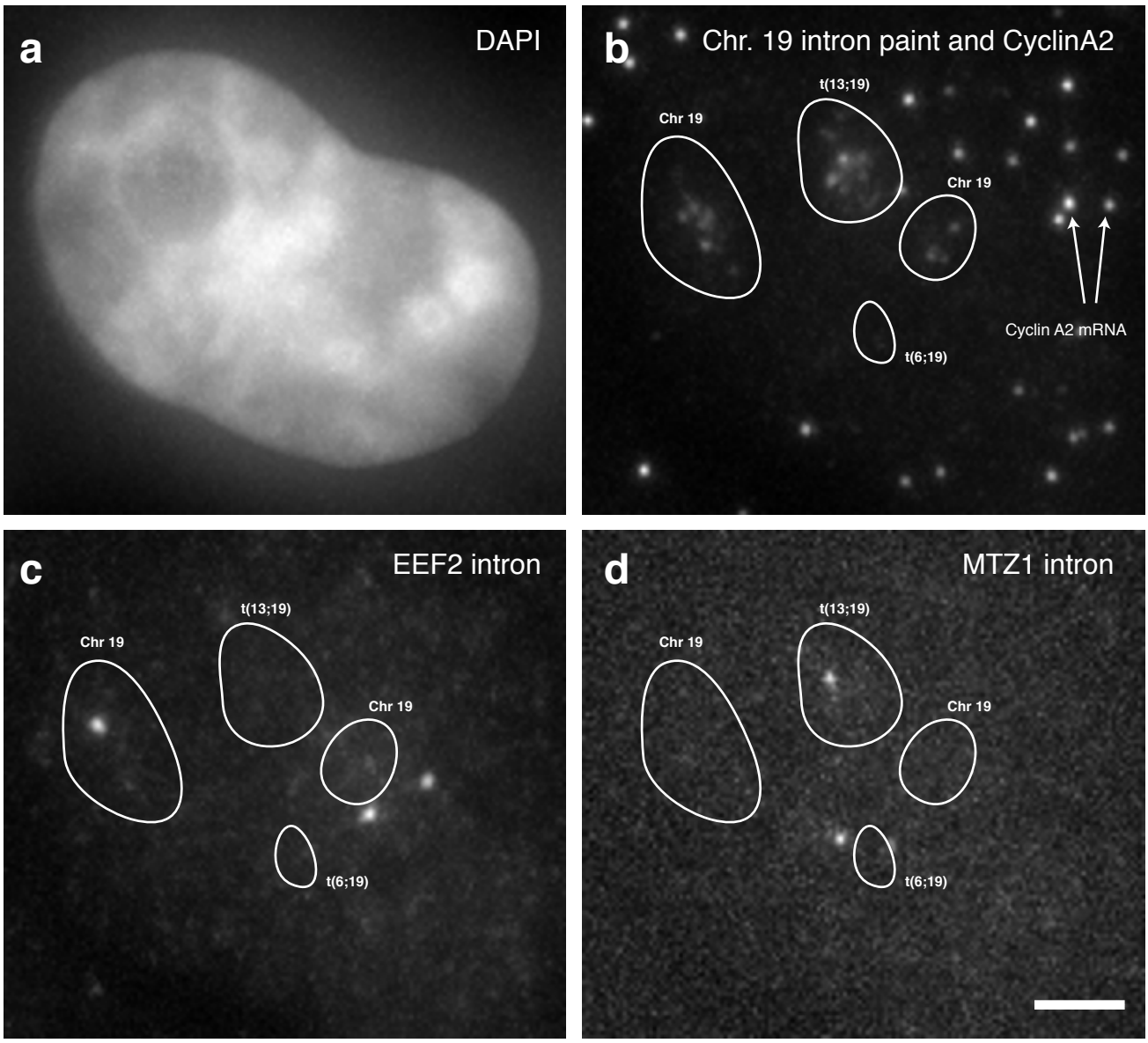
Supplementary Figure 14



Supplementary Fig. 14. Analysis of spot intensity. We computed spot intensity by finding the maximum pixel intensity in the center of a 3x3 pixel box around the center of the spot and subtracting the median local background pixel intensity around the spot. We then normalized this intensity for each color to number of probes on a cell-by-cell basis by calibrating to the intensities of the SuZ12 mRNA signals, which had 13 oligo-nucleotide probes in each fluorescence channel. A value of 1 for spot intensity here indicates and intensity equivalent to 1 complete set of intron probes bound. The box plots show spot intensity for each gene on Chr. 19 in foreskin fibroblasts and HeLa cells, as well as the t(6;19) and t(13;19) derivative chromosomes in HeLa cells. Center white spot represents the median, edges correspond to 25th and 75th percentile, and whiskers extend to extrema, with outliers plotted individually as circles.

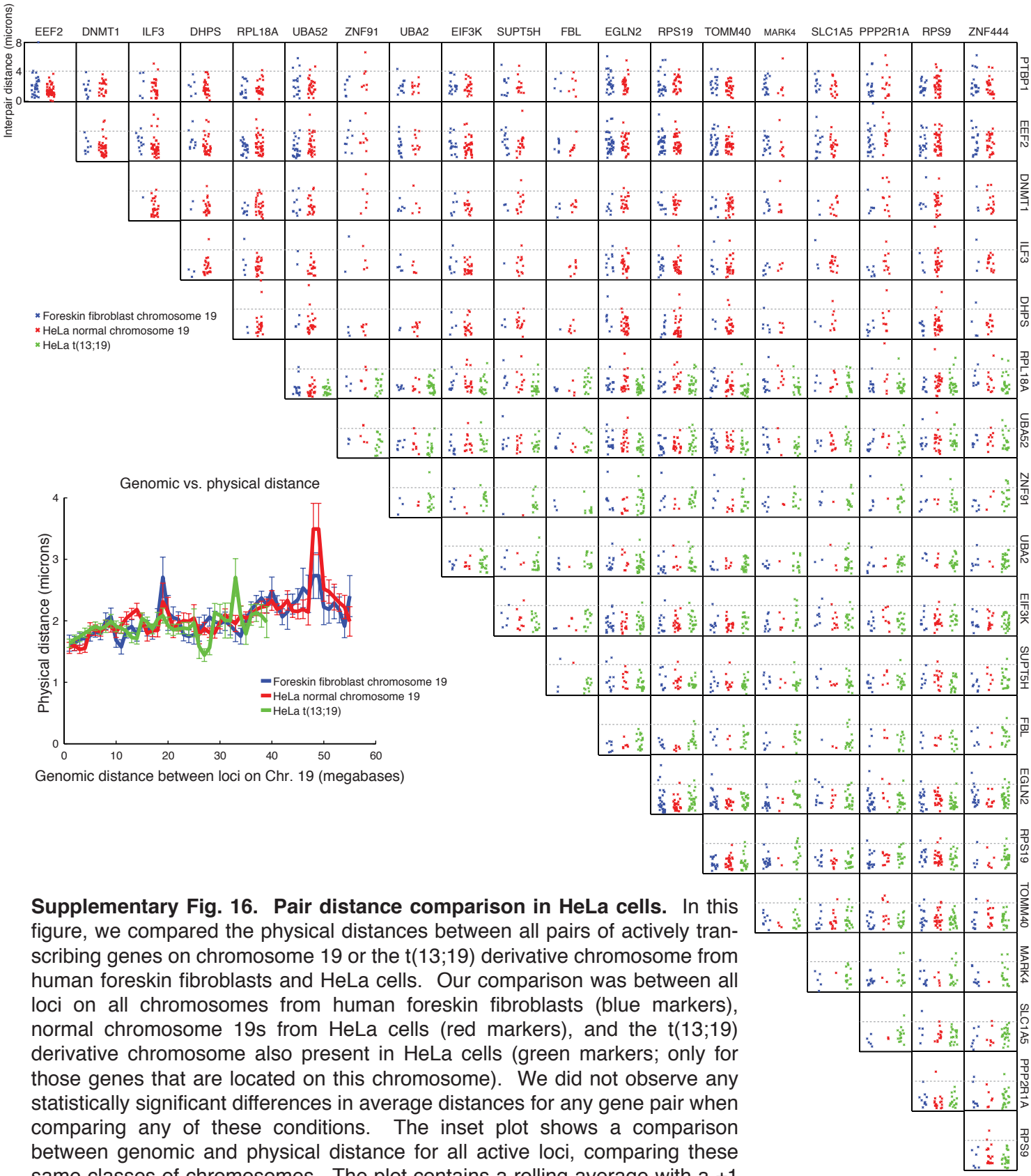
We found that the spot intensities varied from gene to gene, with most of the genes showing a median between 1 and 4 full intron equivalents. The foreskin spot intensities appeared to be somewhat lower than in HeLa cells. In HeLa cells, we found that spot intensities were essentially the same when comparing spots on the intact chromosome 19s and the translocated t(6;19) and t(13;19). This result implies that the changes in transcription of the chromosome 19 genes on t(13;19) arise via changes in transcriptional burst frequency but not changes in burst size. It also highlights the fact that the changes in spot frequency that we observe on t(13;19) are not likely to be the result of a relatively small increase in spot intensity leading to an increase in the number of spots above a putative detection threshold. Were that the case, then we would see many more low intensity spots on t(13;19), which would lower the mean intensity, which is not what we observe.

Supplementary Figure 15



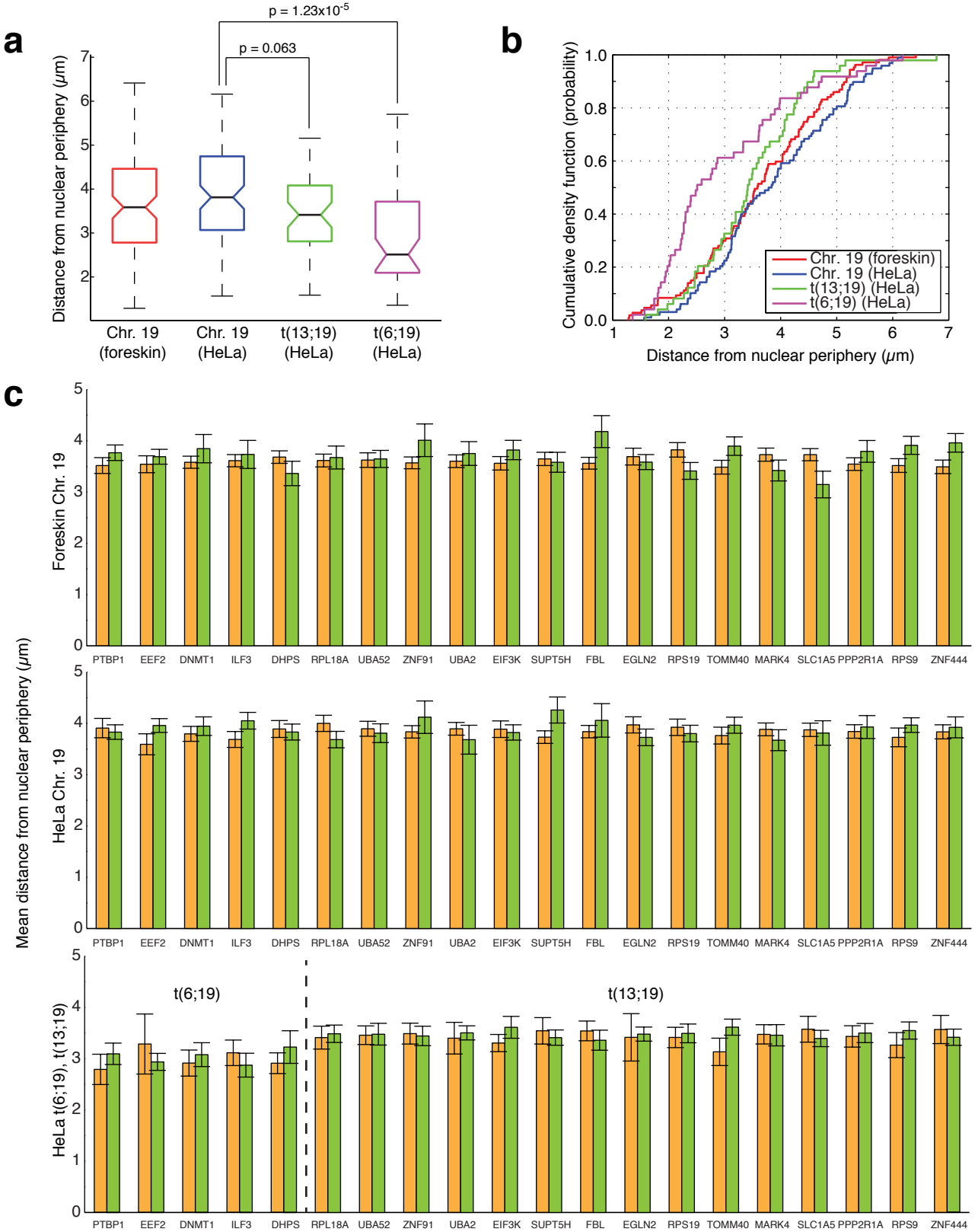
Supplementary Fig. 15. Chromosome 13 intron spot frequency. To measure the frequency of transcription of genes on the portion of chromosome 13 fused to chromosome 19 (denoted der(19)t(13;19)) in HeLa cells, we used a strategy in which we painted chromosome 19 with probes in one color (Cy3, panel b) to identify the chromosome territory. We also labeled Cyclin A2 mRNA in Cy3 to ensure analyzed cells were in G0/G1 phase of the cell cycle. We also labeled probes targeting EEF2's intron in Atto 647N (c), which is located on the portion of chromosome 19 that is part of der(6)t(6;19) and is not on der(19)t(13;19), thereby helping to identify those portions of chromosome 19 that are not fused to chromosome 13. Finally, we labeled the intron of the gene of interest on chromosome 13 (either MTZ1, in panel d, or DIAPH3, not depicted) with Alexa 594, and if the intron spot appeared in the chromosome territory identified by the chromosome 19 paint, we assigned it to der(19)t(13;19). We assigned spots for these chromosome 13 genes that appeared away from chromosome 19 as coming from one of the two normal copies of chromosome 13 in our HeLa cells. The scale bar is 5μm long.

Supplementary Figure 16



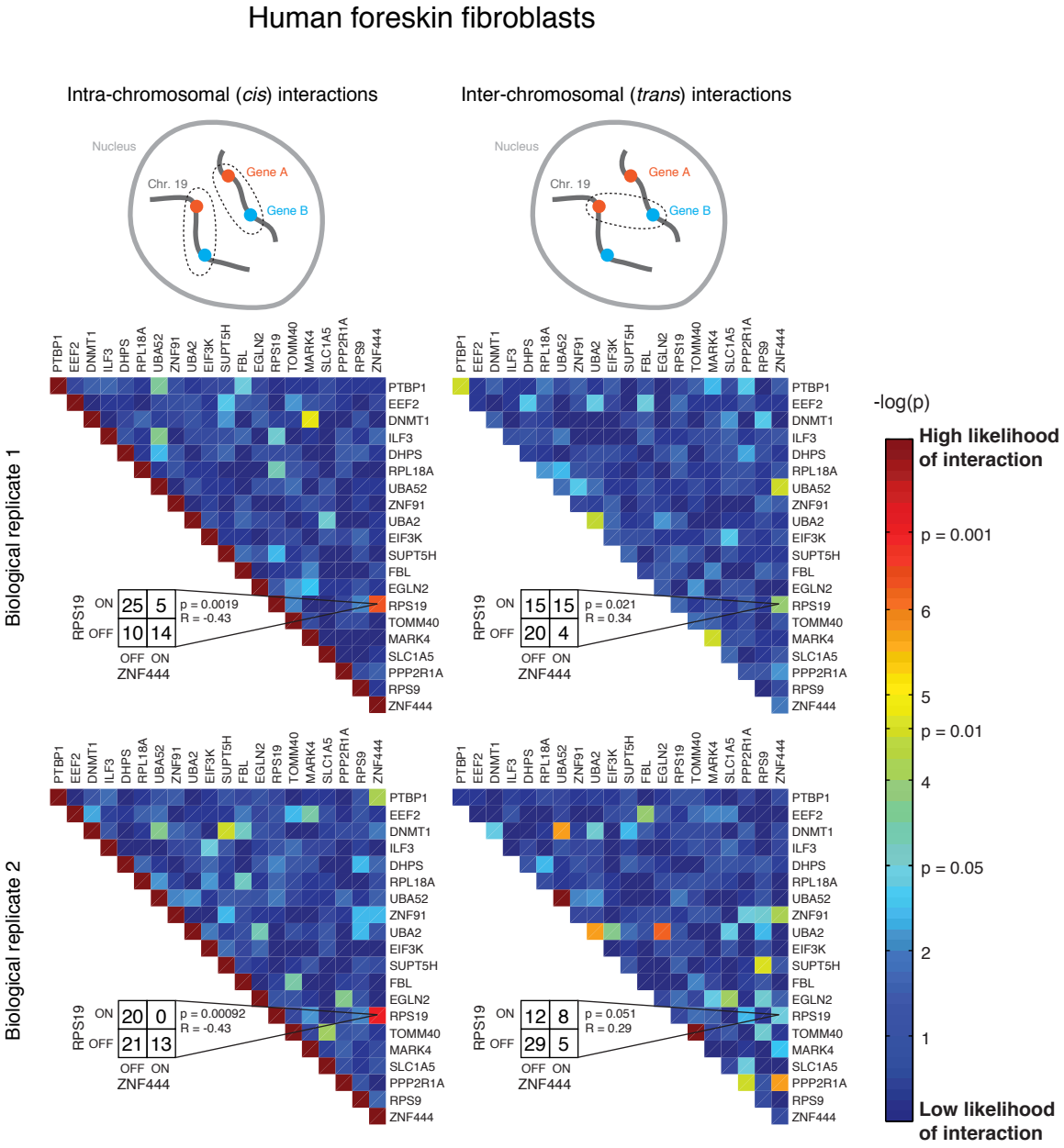
Supplementary Fig. 16. Pair distance comparison in HeLa cells. In this figure, we compared the physical distances between all pairs of actively transcribing genes on chromosome 19 or the t(13;19) derivative chromosome from human foreskin fibroblasts and HeLa cells. Our comparison was between all loci on all chromosomes from human foreskin fibroblasts (blue markers), normal chromosome 19s from HeLa cells (red markers), and the t(13;19) derivative chromosome also present in HeLa cells (green markers; only for those genes that are located on this chromosome). We did not observe any statistically significant differences in average distances for any gene pair when comparing any of these conditions. The inset plot shows a comparison between genomic and physical distance for all active loci, comparing these same classes of chromosomes. The plot contains a rolling average with a ±1 megabase window, with the error bars corresponding to the standard error of the mean. We observed no significant differences in the relationship between genomic distance and physical distance between active loci, suggesting that the increased expression on the t(13;19) derivative chromosome does not correlate with an overall increase in the physical distance between active loci.

Supplementary Figure 17



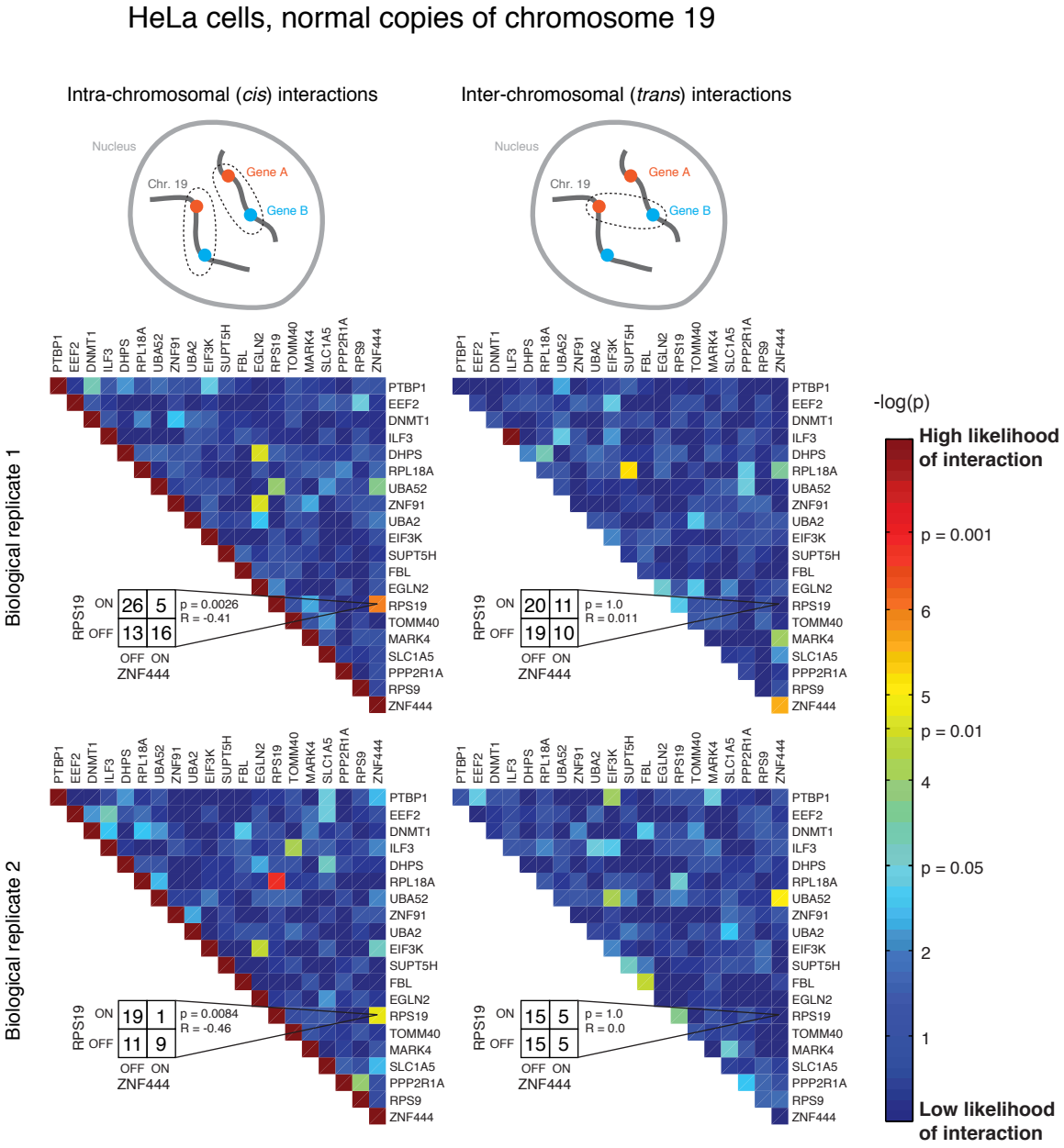
Supplementary Fig. 17. Analysis of distance of chromosomes from nuclear periphery. **a**. Box plot showing distance from nuclear periphery for Chr. 19 in foreskin fibroblasts and HeLa cells, as well as the t(6;19) and t(13;19) derivative chromosomes. Center line represents the median, edges correspond to 25th and 75th percentile, and whiskers extend to extrema. Medians are different with a $p < 0.05$ if the notched regions do not overlap. We computed p-values using the Kolmogorov-Smirnov test. **b**. The cumulative density function corresponding to A. **c**. Distance from nuclear periphery for chromosomes in which a given gene is actively transcribing (green) versus inactive (orange). Error bars reflect the standard error of the mean.

Supplementary Figure 18



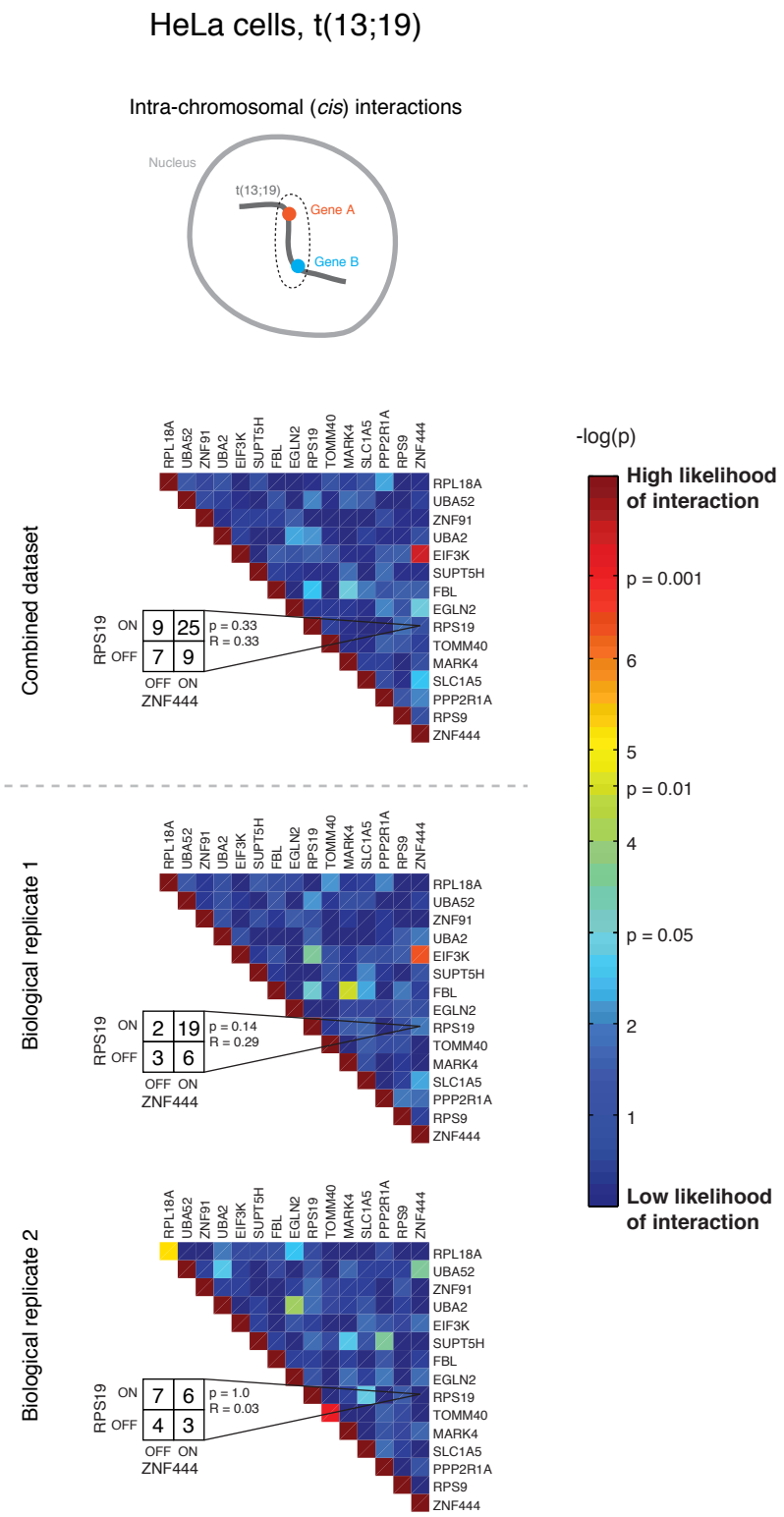
Supplementary Fig. 18. Replicates for foreskin correlations. Characterization of inter and intra-chromosomal interactions between gene pairs in independent biological replicates measured in human foreskin fibroblasts. The analysis is exactly the same as that performed in Fig. 3 and as described in the methods

Supplementary Figure 19



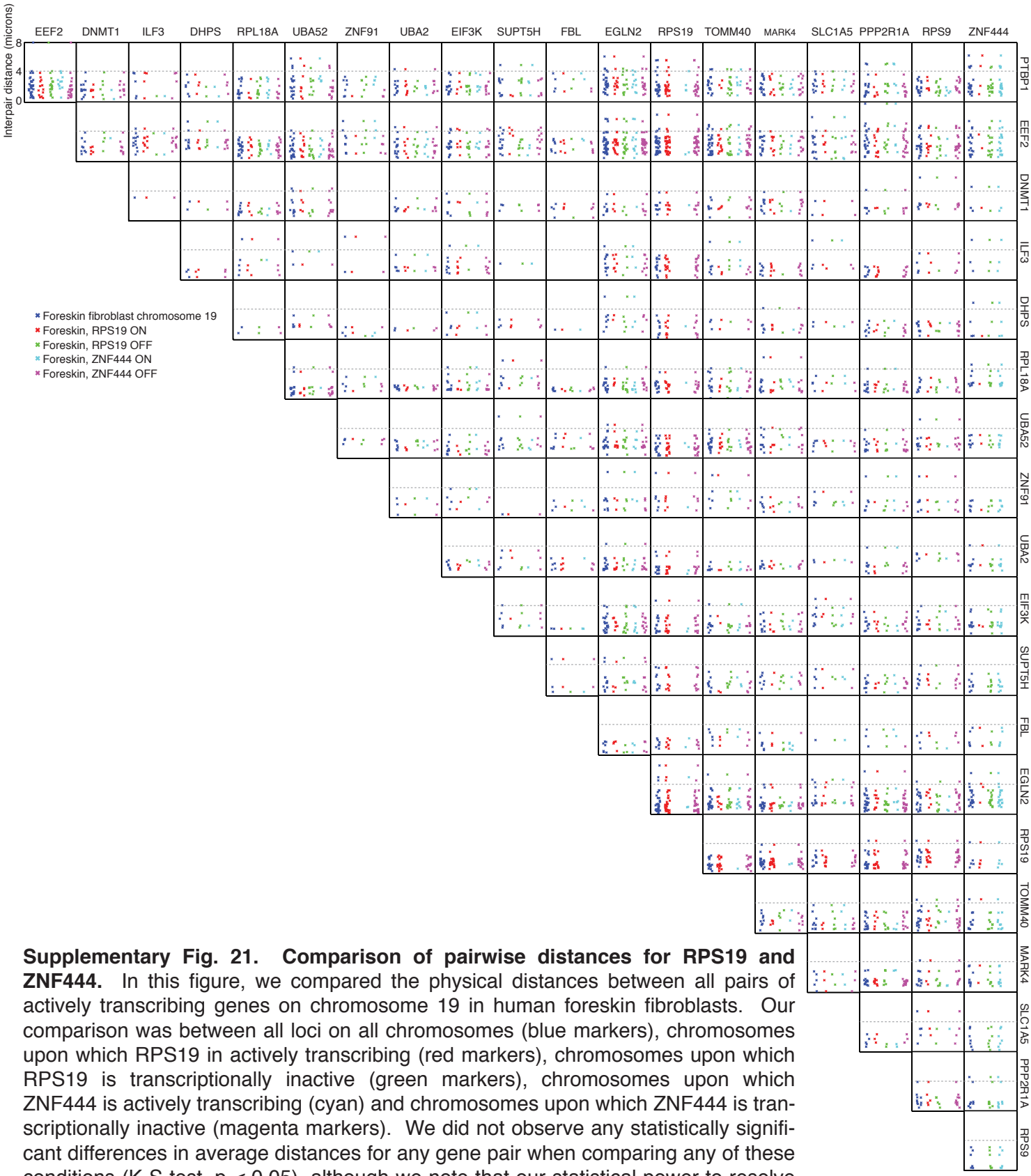
Supplementary Fig. 19. Replicates for HeLa correlations. Characterization of inter and intra-chromosomal interactions between gene pairs in independent biological replicates measured in HeLa cells. The analysis is exactly the same as that performed in Fig. 3 and as described in the methods.

Supplementary Figure 20



Supplementary Fig. 20. Replicates for t(13;19) correlations. Characterization of intra-chromosomal interactions between chromosome 19 gene pairs on the der(19)t(13;19) fusion chromosome in HeLa cells. The analysis is exactly the same as that performed in Fig. 3 and as described in the methods. We found no significant correlation between RPS19 and ZNF444, suggesting that the translocation disrupts the whatever the regulatory mechanism is that is responsible for the anti-correlation we obseved between this pair of genes on the normal copies of chromosome 19 in both human foreskin fibroblasts and HeLa cells.

Supplementary Figure 21



Supplementary Fig. 21. Comparison of pairwise distances for RPS19 and ZNF444. In this figure, we compared the physical distances between all pairs of actively transcribing genes on chromosome 19 in human foreskin fibroblasts. Our comparison was between all loci on all chromosomes (blue markers), chromosomes upon which RPS19 in actively transcribing (red markers), chromosomes upon which RPS19 is transcriptionally inactive (green markers), chromosomes upon which ZNF444 is actively transcribing (cyan) and chromosomes upon which ZNF444 is transcriptionally inactive (magenta markers). We did not observe any statistically significant differences in average distances for any gene pair when comparing any of these conditions (K-S test, $p < 0.05$), although we note that our statistical power to resolve these differences is in some cases rather low because of a low number of measurements. We also did not note any particular distance patterns or relationships, although some pairs did appear to display less distance variability than others.