Single-chromosome transcriptional profiling reveals chromosomal gene expression regulation

Marshall J Levesque & Arjun Raj

We report intron chromosomal expression FISH (iceFISH), a multiplex imaging method for measuring gene expression and chromosome structure simultaneously on single chromosomes. We find substantial differences in transcriptional frequency between genes on a translocated chromosome and the same genes in their normal chromosomal context in the same cell. Correlations between genes on a single chromosome pointed toward a *cis* chromosome-level transcriptional interaction spanning 14.3 megabases.

The transcription of a gene's DNA into RNA is thought to be controlled largely by the interaction of regulatory proteins with DNA sequences proximal to the gene itself. At the same time, genes are organized by the thousands into chromosomes, thus raising the possibility that the structure or organization of chromosomes influences transcription[1,2]; however, little is known about how organization at the chromosome-length scale affects gene expression. Here we begin to address this question with a method based on RNA FISH[3,4] called iceFISH that enabled us to generate transcriptional profiles of 20 genes simultaneously along individual copies of human chromosome 19 in single cells.

We took advantage of the fact that introns typically degrade rapidly after being spliced out of nascent RNA. Labeling the intron with 16 short, fluorescently labeled oligonucleotides[4] (**Fig. 1a** and **Supplementary Fig. 1**) enabled us to measure whether the gene was actively transcribing[5] and, if it was, to identify the three-dimensional coordinates of the gene[6–8]. Control experiments confirmed that the intron spot marked the site of transcriptionally active genes (**Supplementary Figs. 2–7**). Note that even genes considered constitutively active transcribe RNA in 'bursts' thought to arise from random aspects of the transcriptional process[9–12]. The overall transcription rate is proportional to the probability of finding such a spot for each gene (**Supplementary Discussion**).

To concomitantly measure overall chromosome structure, we designed probes targeting the introns of 20 genes along chromosome 19 (**Supplementary Fig. 1** and **Supplementary Table 1**). This yielded an average resolution of 3 megabases (Mb) with a minimum of 360 kilobases (kb), but we were also able to distinguish loci separated by just 30 kb (**Supplementary Fig. 6**). To measure all 20 genes' transcriptional statuses simultaneously, we labeled each gene's introns with a particular 'pseudocolor', which is a distinct code for each gene consisting of either two or three (out of five) spectrally distinguishable fluorophores[13,14] (**Fig. 1a**). To assign gene identity, we looked for colocalization of two or three spots in each fluorescence channel (**Supplementary Fig. 8**). In human foreskin fibroblasts, we could discern two clearly separated chromosomes (**Fig. 1b**) 78% of the time (**Supplementary Fig. 9**). On average, we found 6 ± 2 expressing genes (out of the 20 labeled) per chromosome. The positions of these genes were more spread out than expected[15] (**Supplementary Fig. 6** and **Supplementary Discussion**). We found that using 16 singly colored probes did not change the spot detection efficiency (Online Methods), nor did pseudocoloring incur

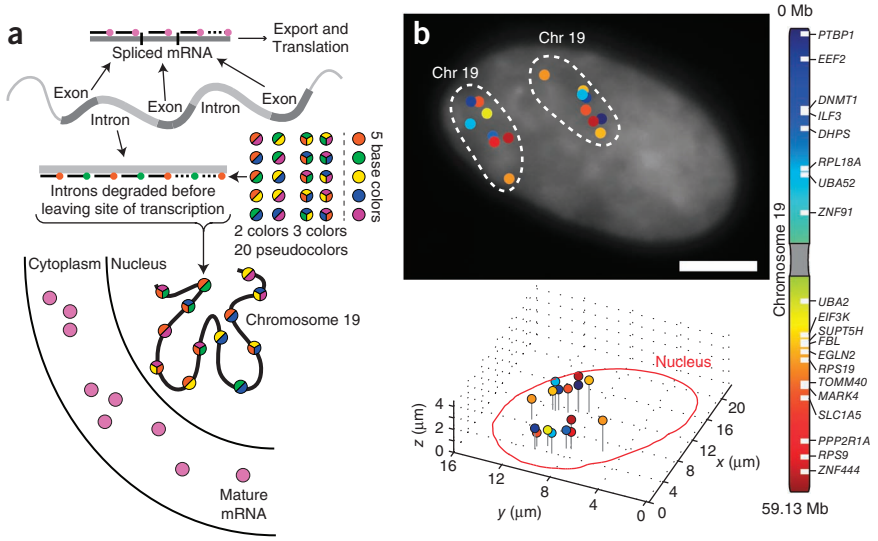

Figure 1 | Unique identification of 20 loci on chromosome (chr) 19 by RNA FISH targeting introns in human foreskin fibroblasts. (**a**) Pseudocoloring scheme for labeling the site of transcription by targeting gene introns with a series of labeled oligonucleotide probes. (**b**) Transcriptional activity and location of the 20 genes computationally identified from images from each fluorescence channel. Scale bar, 5 μm.

Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania, USA. Correspondence should be addressed to M.J.L. or A.R. (rajlaboratory@gmail.com).

RECEIVED 17 NOVEMBER 2012; ACCEPTED 18 JANUARY 2013; PUBLISHED ONLINE 17 FEBRUARY 2013; DOI:10.1038/NMETH.2372

a significant rate of spot misidentification (**Supplementary Fig. 10**). We ensured that the cells we analyzed were in the G0/G1 stage of the cell cycle by colabeling cyclin A2 mRNA, which is abundant during the S, G2 and M phases of the cell cycle[16] (**Supplementary Fig. 11**).

By grouping actively transcribing genes into territories corresponding to each chromosome, we constructed transcriptional profiles showing which of our 20 genes were on or off per chromosome. Researchers largely believe that gene transcription depends on both chromosome-extrinsic *trans* factors (such as transcription factors) and local *cis* factors on the DNA (typically within 1 Mb of the gene itself). Our method enabled us to examine the possibility that nonlocal mechanisms at the chromosome scale may also regulate transcription.

Translocations provide a means to search for such possibilities: although they disrupt the large-scale structure of a chromosome, the cell's *trans* environment and local *cis* DNA regulatory code remain unchanged for most genes on the translocated chromosome. For example, HeLa cells contain two intact copies of chromosome 19 and one copy that is split into two pieces fused to parts of other chromosomes[17]: one, denoted t(6;19), consists of the first 17–20 Mb of chromosome 19 fused to part of chromosome 6; and the other, denoted t(13;19), consists of the remaining 40–43 Mb of chromosome 19 translocated onto a portion of chromosome 13 (**Fig. 2a** and **Supplementary Fig. 12**). Our iceFISH data recapitulated these genetic rearrangements (**Fig. 2b**). We found that most genes on t(13;19) were up to fivefold more transcriptionally active than those on the normal copies of chromosome 19 (**Fig. 2c** and **Supplementary Fig. 13**), consistent with the existence of chromosome-specific transcriptional regulation that the translocation may have disrupted. Intron spot intensities were roughly the same on all the chromosomes we examined (**Supplementary Fig. 14**), suggesting that transcriptional hyperactivation results from an increased probability of a gene being active (**Supplementary Discussion**).

We also found that the transcriptional frequencies of two genes from chromosome 13 (*DIAPH3* and *MZT1*) were roughly twofold higher on t(13;19) than on the normal copies of chromosome 13 (**Fig. 2c** and **Supplementary Fig. 15**), suggesting that this translocation resulted in hyperactivation of all genes on t(13;19) irrespective of location. Meanwhile, transcription of the chromosome 19 genes on t(6;19) was similar to that of the normal copies (**Fig. 2c**), suggesting that translocations do not necessarily lead to transcriptional changes. Reports of such effects[18] are not widespread because of the averaging effects of most assays. We also found that these transcriptional differences did not appear to correlate with differences in spatial chromosome conformation (**Supplementary Figs. 6**, **16** and **17**).

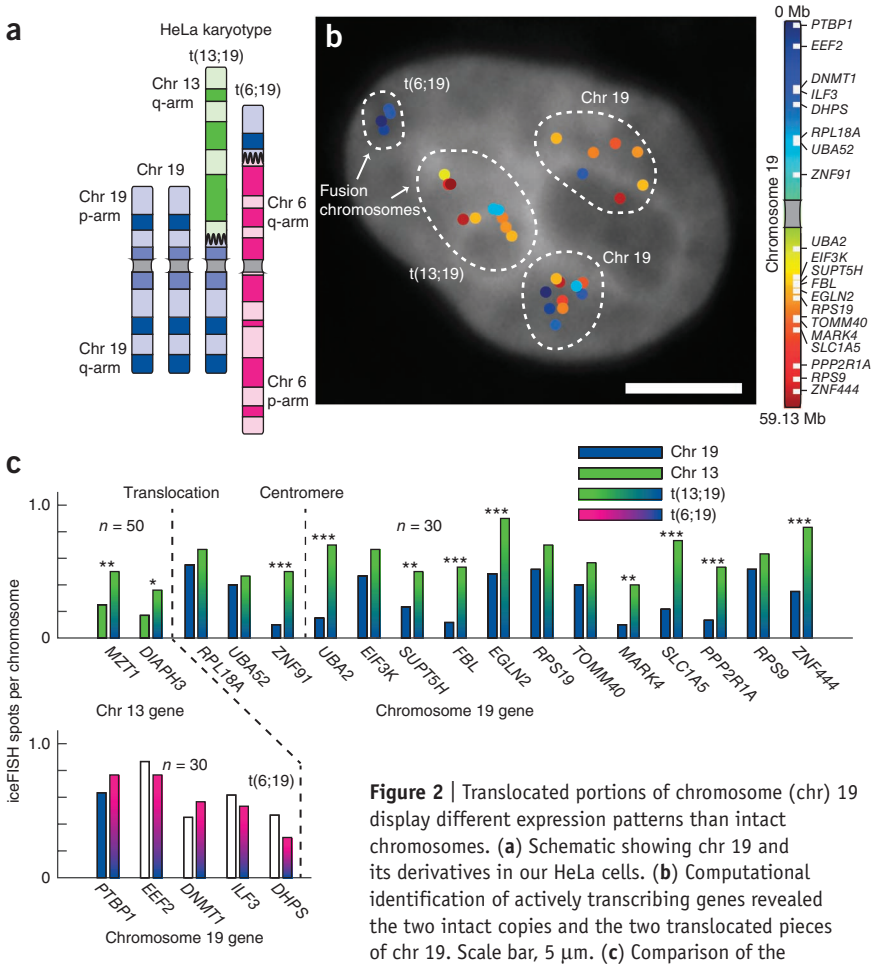

Figure 2 | Translocated portions of chromosome (chr) 19 display different expression patterns than intact chromosomes. (**a**) Schematic showing chr 19 and its derivatives in our HeLa cells. (**b**) Computational identification of actively transcribing genes revealed the two intact copies and the two translocated pieces of chr 19. Scale bar, 5 µm. (**c**) Comparison of the transcriptional activity of the genes on the translocated fragments (as measured by frequency of observing a transcription site per chr) to the activity of the intact copies of chr 19. We similarly analyzed two genes on chr 13 (*MZT1* and *DIAPH3*), as described in **Supplementary Figure 15**. *P* values for the difference in frequency (using a binomial distribution test) $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

We next looked for evidence of interactions governing the transcription of genes within a single chromosome by examining whether the transcriptional status of one gene in our panel affected the transcriptional status of another gene on the same chromosome. Such an interaction would manifest itself as a deviation from independence, with positive correlations signifying that the two genes 'A' and 'B' would have a likelihood greater than chance to be actively transcribing at the same time on the same chromosome, and with anticorrelations indicating that transcription of genes A and B would be mutually exclusive.

We found that most pairwise interactions on single chromosomes did not show significant deviations from independence (**Fig. 3** and **Supplementary Figs. 18** and **19**). However, one pair of genes, *RPS19* and *ZNF444* (separated by 14.3 Mb), showed an anticorrelation ($R = -0.40 \pm 0.08$; $P = 3.99 \times 10^{-5}$, Fisher exact test). One explanation for this anticorrelation is fluctuations in a potential *trans*-acting factor, such as a transcription factor, that activated *RPS19* and inactivated *ZNF444* in some cells while activating *ZNF444* and inactivating *RPS19* in others. Such a *trans* factor would, however, also affect the copy of the gene on the

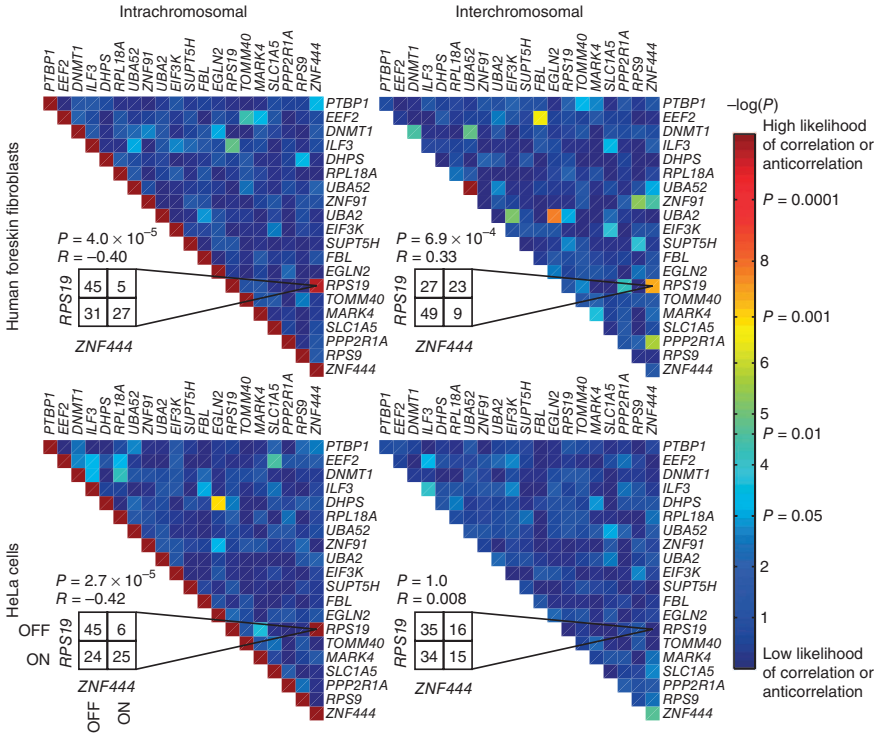

Figure 3 | Identification of a pair of genes showing an intrachromosomal but not interchromosomal expression relationship. The heat maps show the deviation from independence of the intra- and interchromosomal transcriptional activity of all pairs of genes we measured in human foreskin fibroblasts ($n = 54$ cells) and the two intact copies of chromosome 19 in HeLa cells ($n = 50$ cells). A smaller P value (Fisher exact test) indicates a more significant deviation from independence. The grid shows the number of chromosomes in which *RPS19* and *ZNF444* are transcriptionally active or inactive (designated ON and OFF as marked in lower left; markings apply to all boxes).

other chromosome 19 (ref. 19). Thus, we looked for an anticorrelation between *RPS19* on one chromosome and *ZNF444* on the other copy of chromosome 19 in the same cell. We found that the interchromosomal interaction between the genes was qualitatively different, having a mild and less statistically significant positive correlation ($R = 0.33 \pm 0.09$; $P = 6.90 \times 10^{-4}$, Fisher exact test), indicating that the interaction between these genes is indeed a *cis* effect confined to the chromosome itself. The lack of anticorrelation between the chromosome 19 copies (both between the pair of genes and between each gene and its copy; **Fig. 3**) also precludes the possibility of genetic imprinting. Notably, we also found the exact same pattern of interactions when examining the two intact copies of chromosome 19 in HeLa cells (**Fig. 3**). The interaction between *RPS19* and *ZNF444* disappeared, however, on t(13;19) (**Supplementary Fig. 20**), suggesting that whatever mechanism is responsible requires a fully intact chromosome 19. This interaction did not, however, depend on the distance between any pair of genes (**Supplementary Fig. 21**) or on the distance of the chromosome from the nuclear periphery (**Supplementary Fig. 17**).

iceFISH provides a complement to chromosome conformation assays[20] that look for interactions at the DNA (rather than transcriptional) level. We believe that iceFISH and similar tools will allow us to determine the prevalence of these chromosome-level regulatory phenomena and uncover their underlying mechanisms.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank Biosearch Technologies (especially M. Beal and R. Cook) for providing many of the reagents used in the iceFISH assay; J. Morrissette for her lab's assistance in performing the cytogenetic analysis; L. Cai, H. Youk, D. Muzzey, J. Gore, H. Maamar and O. Padovan-Merhar for insightful comments; G. Nair for discussions about statistics; and the lab of P. Sharp (Massachusetts Institute of Technology) for providing the HeLa cells. We gratefully acknowledge a University of Pennsylvania University Research Foundation pilot grant, the US National Institutes of Health Director's New Innovator Award (1DP2OD008514) and a Burroughs-Wellcome Fund Career Award at the Scientific Interface for supporting our work.

AUTHOR CONTRIBUTIONS

M.J.L. and A.R. conceived of the project, performed the analyses and wrote the paper. M.J.L. performed the experiments.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Published online at http://www.nature.com/doifinder/10.1038/nmeth.2372. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Fraser, P. & Bickmore, W. *Nature* **447**, 413–417 (2007).
2. Cremer, T. & Cremer, C. *Nat. Rev. Genet.* **2**, 292–301 (2001).
3. Femino, A.M., Fay, F.S., Fogarty, K. & Singer, R.H. *Science* **280**, 585–590 (1998).
4. Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A. & Tyagi, S. *Nat. Methods* **5**, 877–879 (2008).
5. Fremeau, R.T., Lundblad, J.R., Pritchett, D.B., Wilcox, J.N. & Roberts, J.L. *Science* **234**, 1265–1269 (1986).
6. Gribnau, J. *et al. EMBO J.* **17**, 6020–6027 (1998).
7. Xing, Y., Johnson, C.V., Dobner, P.R. & Lawrence, J.B. *Science* **259**, 1326–1330 (1993).
8. Vargas, D.Y. *et al. Cell* **147**, 1054–1065 (2011).
9. Golding, I., Paulsson, J., Zawilski, S.M. & Cox, E.C. *Cell* **123**, 1025–1036 (2005).
10. Chubb, J.R., Trcek, T., Shenoy, S.M. & Singer, R.H. *Curr. Biol.* **16**, 1018–1025 (2006).
11. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y. & Tyagi, S. *PLoS Biol.* **4**, e309 (2006).
12. Suter, D.M. *et al. Science* **332**, 472–474 (2011).
13. Levsky, J.M., Shenoy, S.M., Pezo, R.C. & Singer, R.H. *Science* **297**, 836–840 (2002).
14. Lubeck, E. & Cai, L. *Nat. Methods* **9**, 743–748 (2012).
15. Mateos-Langerak, J. *et al. Proc. Natl. Acad. Sci. USA* **106**, 3812–3817 (2009).
16. Eward, K.L., Van Ert, M.N., Thornton, M. & Helmstetter, C.E. *Cell Cycle* **3**, 1057–1061 (2004).
17. Macville, M. *et al. Cancer Res.* **59**, 141–150 (1999).
18. Harewood, L. *et al. Genome Res.* **20**, 554–564 (2010).
19. Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S. *Science* **297**, 1183–1186 (2002).
20. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. *Science* **295**, 1306–1311 (2002).

ONLINE METHODS

Cell culture, fixation and fluorescence *in situ* hybridization. We grew primary human foreskin fibroblasts (ATCC CRL 2097) or HeLa cells in Dulbecco's modified Eagle's medium with Glutamax (DMEM, Life Technologies) supplemented with penicillin/streptomycin and 10% FBS. We enriched for G0/G1-phase cells through a double-thymidine block (2 mM thymidine in medium) procedure, which arrested cells at the beginning of S phase. We released the cells and let them go through S, G2, M, G1, S, G2 and M and then fixed them when they were in G1. We let the cells go through one complete cell cycle after the block to minimize any potential transcriptional or structural effects due to the block itself. To fix the cells, we followed the protocol of Raj *et al.*[4] Briefly, we fixed the cells for 10 min at room temperature using 4% formaldehyde/10% formalin in 1× phosphate buffered saline solution (PBS) and followed the fixation by two rinses in 1× PBS, after which we permeabilized the cells with 70% EtOH and stored them at 4 °C at least overnight.

To perform fluorescence *in situ* hybridization (FISH), we again followed the procedure of ref. 4 with some minor modifications. We prewashed with a wash buffer containing 10% formamide and 2× saline-sodium citrate (SSC) and then hybridized by adding the appropriate amount and type of probe (described below) in a buffer containing 10% formamide, 2× SSC and 10% dextran sulfate (w/v). We empirically determined the optimal concentration of each probe, which in most cases was roughly equivalent to the concentrations used in ref. 4. We hybridized our samples overnight in a humidified chamber kept at 37 °C, washed twice for 30 min with wash buffer at 37 °C (adding DAPI at a concentration of 50 ng/mL in the second wash) and then imaged in 2× SSC as described below.

In the case of the experiments involving actinomycin D, we incubated HeLa cells in 2 μg/mL of actinomycin D (Sigma) for 0, 30, 60 and 120 min (as described in **Supplementary Fig. 3**), after which we fixed the cells and performed FISH. We made sure to thoroughly mix the actinomycin D into the medium before adding it to avoid spatial inhomogeneity in the activity of the drug.

For the RNase experiments, we fixed and permeabilized the cells as just outlined, after which we aspirated the 70% EtOH, washed once with 1× PBS and then added 1× PBS with 10 μg/mL of RNase A (Sigma). We incubated the fixed cells at 37 °C for 30 min, washed with 1× PBS and then proceeded with FISH as outlined above. As a control, we performed the exact same procedure on cells in a neighboring well but did not add RNase A to the 1× PBS for the incubation (as described in **Supplementary Fig. 2**).

Imaging. We imaged all our samples on a Nikon Ti-E inverted fluorescence microscope using a 100× Plan-Apo objective (numerical aperture of 1.43) and a cooled CCD camera (Pixis 1024B from Princeton Instruments). We sequentially acquired three-dimensional stacks of fluorescence images in six different fluorescence channels using filter sets for DAPI, Atto 488, Cy3, Alexa 594, Atto 647N and Atto 700. Our exposure times were roughly 2–3 s for most of the dyes except for DAPI (which we exposed for ~100 ms) and Atto 700 (~5 s, owing to somewhat weaker illumination on our apparatus). The spacing between consecutive planes in our stacks was 0.3 μm. The filter sets we used were 31000v2 (Chroma), 41028 (Chroma), SP102v1 (Chroma), a custom set from Omega as described in ref. 4, SP104v2 (Chroma) and SP105 (Chroma) for DAPI, Atto 488, Cy3, Alexa 594, Atto 647N and Atto 700, respectively.

Image analysis. Once we acquired our images, we put them through an image analysis pipeline made up of custom semiautomated spot recognition software we wrote in MATLAB with the following series of steps.

1. We first identified candidate spots in the three-dimensional image by filtering the image with a Laplacian of Gaussian filter and taking the top 300 spots as candidates. In some cases, we also chose cells to analyze on the basis of their phase in the cell cycle. In those cases, we chose cells that had little or no cyclin A2 mRNAs. Our experiments in **Supplementary Figure 11** validate this approach.
2. For each candidate, we then fit the candidate to a Laplacian of Gaussian intensity profile, thereby giving us precise estimates of the center, width and intensity of the spot.
3. Using histograms of the intensities and widths, we manually selected a subset of the spots with qualities (uniform width, higher intensity) that were higher than background. This is similar in spirit to the procedure described in ref. 4, in which the experimenter chose a threshold to separate legitimate RNA spots from background spots. In this case, we erred on the side of including spots that may be background because our multicolor scheme for spot assignment provided us another means by which to discard background spots.
4. Once we had selected the spots, we ran software that found the fiducial markers (in this case, probes in all five RNA colors targeting *SUZ12* mRNA, which were present at an abundance of roughly 20–50 clear cytoplasmic spots per cell). In this manner, we could measure the displacements between different fluorescence channels in each cell individually. We then applied these shifts to align the computationally identified spots between the different fluorescence channels.
5. After alignment, we ran software that looked for colocalized spots corresponding to the particular pseudocoloring scheme we chose for the introns we targeted. We estimate that our software is roughly 75% accurate in assigning colocalized spots to particular genes at this stage.
6. We then went through a manual correction process in which we corrected mistakes the software made in identifying spots. Common issues were failure to detect dim (but clearly present) signals in one of the fluorescence channels and resolution of two spatially close fluorescent spots that the Laplacian of Gaussian filtering and candidate identification steps had labeled as a single spot.
7. Once we had correctly annotated the introns of the gene loci we had labeled, we examined cells manually to separate out individual chromosomes. We discarded cells in which the chromosomes overlapped because this made it difficult to assign gene spots to particular chromosomes.

To determine the distance of the chromosome from the nuclear periphery, we first determined the average position of the spots of the chromosome in x and y and then found the Euclidean distance between this point and the nuclear periphery as outlined by our DAPI stain.

Characterization of error rate. To gain some sense of the rate of false positives, we performed a hybridization in foreskin

doi:10.1038/nmeth.2372

fibroblasts in which we left out 10 of the 20 genes comprising our iceFISH assay (randomly chosen by another member of the lab) and proceeded with our spot identification procedure as usual (**Supplementary Fig. 10**). We found that our rate of false identification was very low, with the vast majority (97%) of spots we assigned corresponding to genes that we had targeted in our assay.

We also probed a set of two genes (*RPS19*, *TOMM40*) one at a time with oligonucleotides labeled with a single dye rather than the combination of two or three dyes used in our pseudocoloring strategy. Our aim was to determine to what extent our pseudocoloring strategy would result in false negatives in spot identification. We found that the spot-per-chromosome frequencies measured with a singly colored probe alone were 0.56 and 0.27, whereas the spot frequencies measured by pseudocoloring were 0.57 and 0.25 for *RPS19* and *TOMM40*, respectively, in a total of 30 cells. Although statistical effects preclude a definitive statement, our results are consistent with our pseudocoloring strategy correctly identifying virtually all spots detectable by RNA FISH targeting introns.

Probe design. We designed 20 base oligonucleotide probes against introns using custom FISH design software (http://www.biosearchtech.com/stellarisdesigner/). 90% of probes worked without needing optimization. Where possible, we tried to design 16 oligonucleotides targeting the first intron of the gene; for cases in which we were unable to fit all the oligonucleotide probes onto the first intron, we targeted the second intron as well. We ordered the oligonucleotides from Biosearch Technologies, who synthesized the oligonucleotides with amine groups attached to the 3′ end. We coupled these 3′ ends to various organic dyes (including Atto 488 (Atto-Tec), Cy3 (GE), Alexa 594 (Invitrogen), Atto 647N (Atto-Tec) and Atto 700 (Atto-Tec)) as indicated in the text and in **Supplementary Table 1**. We purified the probes by HPLC[4].

Karyotyping of HeLa cells. We performed G-band analysis (karyotyping) on metaphase spreads of our HeLa cells, following standard procedures. The analysis indicated that our cells contained two intact copies of chromosome 19 and a full third copy of chromosome 19 split into two fragments and fused to other chromosomes (**Supplementary Fig. 12**). One fragment included the first half of the chromosome 19 p-arm and was fused to a large portion of chromosome 6. The second fragment was the remaining portion of chromosome 19 (half the p-arm through the centromere and entire q-arm), which was fused to the q-arm of chromosome 13. To conclusively demonstrate that chromosome 19 was split in this particular way, we performed a DNA FISH analysis on the same metaphase spreads that we performed the G-band analysis on. We used probes targeting loci within the 19p13 and 19q13 cytogenetic bands on chromosome 19, each labeled with a different fluorophore (Abbott Molecular). The results confirmed the results of the G-band analysis. We performed this analysis on ten cells, each of which showed the same genetic abnormalities, indicating that the cells did not vary much in this particular characteristic from cell to cell.

Click-iT EdU analysis of cell-cycle progression. To demonstrate that cyclin A2 mRNA was an accurate marker of position in the cell cycle, we used the Click-iT EdU Alexa Fluor 594 imaging kit (Invitrogen), which incorporates a targetable chemical into newly replicated DNA. In this case, we incubated foreskin fibroblasts with 10 μM Click-iT EdU reagent for 5 min before fixing the cells. We performed our FISH protocol on these cells using a cyclin A2 mRNA Cy3 probe and, after the hybridization and wash steps, followed the instructions provided with the kit for fluorescently labeling the incorporated EdU. We ultimately did not elect to use the Click-iT EdU kit directly in most of our experiments (and instead opted to use cyclin A2) because we found that performing the Click-iT EdU procedure interfered with our nascent RNA FISH detection, most likely either because of interference with transcription itself or because it made our spot detection less reliable owing to additional washing steps associated with the Click-iT procedure.

DNA FISH. We performed DNA FISH with BAC probes from Empire Genomics, using their reference hybridization protocol. In the human foreskin fibroblast cells we applied pairs of fluorescently labeled BAC clones from the human RPCI-11 library targeting human chromosome 19 at positions 2.8–4.5 Mb (268O21), 39.0–39.5 Mb (31D10) or 52.5–52.7 Mb (43N16). We denatured the DNA by immersing the cells in a solution containing 70% formamide and 2× SSC buffer at 80 °C for 5 min and then transferred them to a series of ethanol steps increasing to 70%, 85% and then 100% ethanol. We added 10 μL of BAC probes to the air-dried sample, applied a coverslip and incubated overnight in a humidified slide chamber. The next day we washed the sample with 0.4× SSC at 73 °C for 2 min, removed the coverslip, transferred it to room temperature 2× SSC for 1 min and then to 10 μL of 2× SSC with DAPI at 50 ng/mL, and applied a new coverslip. We performed imaging similar to that for our iceFISH probes with dye pairs red 5-ROX and TAMRA.

Combined DNA-RNA FISH. We performed a sequential DNA-RNA FISH in HeLa cells by first performing DNA FISH using BAC clones and then performing RNA FISH, both by following the protocols outlined above. We found that both the bright exonic transcription sites and the intron spots were considerably brighter than single mRNA spots, thus showing that the RNA probes were not simply targeting the DNA directly. We compared the location of *SLC1A5* exonic (Alexa 594) and intronic (ATTO647N) RNA to the location of DNA FISH probes using BAC clones RP11-687M15 (TAMRA).

Statistical analysis. In **Figure 3**, we looked for deviations from independence in the transcriptional frequencies of all pairs of genes we examined. We performed the Fisher exact test on each 2 × 2 table generated by counting the number of chromosomes in which gene A or B was transcriptionally active versus the number that were inactive. We reported the two-sided *P* value corresponding to the chance of obtaining a similar deviation from independence via random chance, with a smaller *P* value corresponding to a more significant result. In **Figure 3**, we show the results we obtained by analyzing a data set consisting of the combination of two independent biological replicates; we also performed the analysis on each individual biological replicate, as shown in **Supplementary Figures 18** and **19**. Note that we have not applied a multiple hypothesis correction in our presentation of the *P* values; however, our results would remain statistically

significant if we applied the crude correction of just multiplying our *P* values by 190, which is the number of pairs of genes we examined. We chose to convey the information in this manner because the number of hypotheses tested depends on the particular question being asked of the data. For instance, if one decides that, from the human foreskin fibroblast data, one wanted to focus on interactions between *RPS19* and *ZNF444*, then the *P* values for the specific hypothesis comparing these two genes in, say, HeLa cells, would not be subjected to this same correction. We leave such interpretative matters to the reader.

We also report the correlation coefficient between *RPS19* and *ZNF444*; although it is a somewhat imperfect measure of the lack of independence for this sort of data, it has the advantage of being familiar to many researchers. We obtained standard errors for the correlation coefficient by bootstrapping.

In **Figure 2**, we obtained *P* values for the difference in transcriptional frequency between the copy of the gene on the t(13;19) (or t(6;19)) chromosome and the copies of the gene on the normal copies of chromosome 19 by rejecting the null hypothesis in which the frequency of transcription was the same for all three copies. We did this by computationally generating the probability density function for the difference in transcriptional frequencies between two sets chosen to match our experimental data in size under the null hypothesis that the frequency is the same for both sets and then directly calculated the probability of finding our observed difference by chance.

doi:10.1038/nmeth.2372