Single-allele analysis of transcription kinetics in living mammalian cells

Sharon Yunger[1], Liat Rosenfeld[2], Yuval Garini[2] & Yaron Shav-Tal[1]

We generated a system for *in vivo* visualization and analysis of mammalian mRNA transcriptional kinetics of single alleles in real time, using single-gene integrations. We obtained high-resolution transcription measurements of a single cyclin D1 allele under endogenous or viral promoter control, including quantification of temporal kinetics of transcriptional bursting, promoter firing, nascent mRNA numbers and transcription rates during the cell cycle, and in relation to DNA replication.

Stable genomic integrations of exogenous polymerase II–transcribed genes in mammalian cells have been used to examine transcriptional kinetics *in vivo*[1]. Such integrations typically culminate in the formation of multicopy tandem gene arrays[2] that cannot provide a true picture of the endogenous state of single-copy genes. Transcriptional kinetics have been directly measured in *Dictyostelium* sp.[3] and *Escherichia coli*[4], but single alleles in living human cells have not been visualized directly. Here we describe a method for probing and quantifying transcription kinetics of single alleles in living mammalian cells.

We used HEK-293 cells with a single Flp recombinase target (*FRT*) genomic locus for Flp-in homologous recombination of a single copy of the coding region of the human cyclin D1 gene (*CCND1*) driven by its well-characterized promoter (*CCND1pr*) or a CMV promoter (*CMVpr*) (**Fig. 1a**, **Supplementary Fig. 1a** and **Supplementary Note**). To visualize mRNA transcription in real time, we inserted repeats of MS2 protein–binding sequence into the extensive regulatory 3′ untranslated region (UTR) of *CCND1* to generate *CCND1-MS2* constructs. As nascent transcripts emerged from the polymerase they were bound by MS2-GFP fusion proteins[5,6].

We isolated two comparable stable clones each containing a single copy of *CCND1-MS2* in the same genomic context but regulated by either *CCND1pr* or *CMVpr*. We verified accurate single-allele integration and expression by loss of *lacZ* expression, PCR and DNA sequencing of genomic DNA, gain of hygromycin resistance and expression of hemagglutinin (HA) tag–labeled cyclin D1 by immunofluorescence (**Supplementary Figs. 1** and **2**). Semiquantitative reverse transcription (RT)-PCR demonstrated that integrated *CCND1pr-CCND1-MS2* was expressed to a similar extent as the endogenous *CCND1* alleles and responded similarly to serum stimulation, indicating the addition of a single active allele. Flow cytometry analysis showed that the cell-cycle profile was similar to that of the parental cell line (**Supplementary Fig. 2e**,**f**).

We detected active single transcription sites by RNA fluorescence *in situ* hybridization (FISH) with a Cy3-labeled probe to the MS2 sequence repeats in the entire *CMVpr-CCND1-MS2* cell population but only in 43% of the *CCND1pr-CCND1-MS2* cells (**Supplementary Fig. 3a**,**b**). Using transient expression of MS2-GFP, we observed active transcription sites as well as nucleoplasmic single mRNA-protein complexes (mRNPs) released from transcription sites (**Supplementary Fig. 3c**,**d** and **Supplementary Video 1**). To verify that polymerase II indeed actively transcribed at these sites, we used fluorescence recovery after photobleaching (FRAP)[5,6].

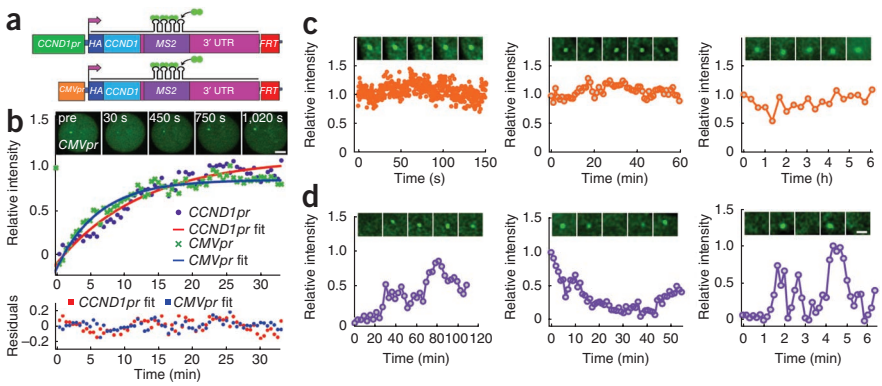

Figure 1 | Real-time transcription kinetics of single *CCND1* alleles. (**a**) Scheme of the two constructs. *HA* sequence encodes the hemagglutinin tag, and *MS2* encodes an array of MS2 protein binding sites. Green spheres represent MS2-GFP fusion protein binding to the MS2 stem-loops. (**b**) Live-cell fluorescence-image movie frames collected before and at indicated times after photobleaching of the transcription site (top; scale bar, 5 μm) and FRAP recovery curves. Goodness of fit evaluated by checking for a random distribution of residuals around 0 is plotted at the bottom. (**c**,**d**) MS2-GFP signal intensity profiles at transcription sites of three *CMVpr-CCND1-MS2* cells (**c**; imaging frequency: 300 ms (left), 1 min (middle) and 20 min (right)) and three *CCND1pr-CCND1-MS2* cells (**d**; imaging frequency: 3 min (left); 1 min (middle); 10 min (right)). Transcription site movie frames from random times during interphase show periods of transcription activity and inactivity (top). Scale bar, 1 μm.

[1]The Mina and Everard Goodman Faculty of Life Sciences and Institute of Nanotechnology, Bar-Ilan University, Ramat Gan, Israel. [2]The Department of Physics and Institute of Nanotechnology, Bar-Ilan University, Ramat Gan, Israel. Correspondence should be addressed to Y.S.-T. (shavtaly@mail.biu.ac.il).

RECEIVED 19 JANUARY; ACCEPTED 28 JUNE; PUBLISHED ONLINE 18 JULY 2010; DOI:10.1038/NMETH.1482

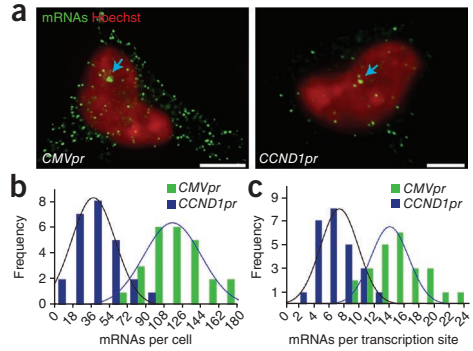

Figure 2 | mRNA quantification at transcription sites. (**a**) Images show deconvolved 3D stacks of images from RNA FISH experiments (with MS2-Cy3 probe, pseudocolored green) in *CMVpr-CCND1-MS2* (left) and *CCND1pr-CCND1-MS2* (right) cells. Transcription sites (arrows) and cytoplasmic mRNAs (green) were identified. Hoechst staining of DNA is pseudocolored red. Scale bars, 5 μm. (**b**) Distribution of cellular *CCND1-MS2* mRNAs (n = 25 cells). (**c**) Distribution of *CCND1-MS2* mRNAs at the transcription sites. ANOVA (P = 0.000; F = 76.262, d.f. = 3, 90; n = 25).

Recovery of fluorescence signified generation of new transcripts at the transcription site (**Fig. 1b** and **Supplementary Note**).

We followed the kinetic behavior of ongoing transcription using four-dimensional time-lapse imaging. We always found *CMVpr*-driven *CCND1-MS2* in an active transcriptional state (~10 h), independent of imaging rates (**Fig. 1c** and **Supplementary Fig. 4**). MS2-GFP signal intensity at the transcription site was relatively constant, and we identified no bursts or intervals of inactivity (**Supplementary Video 2**), demonstrating transcriptional potency of the viral promoter. In contrast, endogenous *CCND1pr*-driven transcription was not constant. Single-allele measurements showed gene-active periods of up to 200 min with relatively constant amount of mRNA as reported by the MS2-GFP signal, whereas periods of gene inactivity between active states averaged 22 min (range, 12–36 min). Transitions between 'on' and 'off' states were gradual; gene 'turning on' and gene 'shutdown' lasted an average of 35 min and 27 min, respectively (**Fig. 1d**, **Supplementary Figs. 4** and **5** and **Supplementary Videos 3–5**).

Fluorescence intensity at the transcription sites comprises ongoing transcription elongation and mRNA release. FRAP curves showed full recovery, indicating that mRNA transcription and release was being monitored in real time. The MS2-GFP signal recovery rates were proportional to the rates of polymerase elongation (**Supplementary Note**). The time required to reach steady state after photobleaching was different between the two different promoters, with slower kinetic rates for the *CCND1pr*–driven gene: *CCND1pr-CCND1-MS2* half-time of recovery ($t_{1/2}$) = 10 min, full recovery = 38 min; versus *CMVpr-CCND1-MS2* $t_{1/2}$ = 5.5 min, full recovery = 23 min (**Fig. 1b**). Also, mRNA fluorescence MS2-GFP intensity was greater in *CMVpr-CCND1-MS2* cells than in *CCND1pr-CCND1-MS2* cells. These data suggest that different amounts of mRNAs were being produced per time unit, either because of differences in polymerase recruitment or changes in polymerase activity rates.

We therefore quantified the nascent mRNA molecules present on both the *CMVpr* and *CCND1pr*-driven transcription sites. First, we acquired three-dimensional (3D) stacks of total cell volumes after RNA FISH experiments (**Fig. 2a**, **Supplementary Fig. 6a**,**b** and **Supplementary Video 6**) and determined the distinct and uniform *CCND1-MS2* single mRNA signals based on previous studies demonstrating that cellular mRNPs contain a single mRNA molecule[7–9]. *CCND1pr-CCND1-MS2* cells contained 41 ± 30 mRNAs per cell, and *CMVpr-CCND1-MS2* cells contained 114 ± 40 mRNAs per cell (**Fig. 2b**). We then quantified the intensity of the transcription site signal in these single cells and divided it by the single-mRNP signal intensity. Quantification showed that 7 ± 4 mRNA molecules were associated with the *CCND1pr-CCND1-MS2* transcription sites, compared to 14 ± 4 mRNAs at *CMVpr-CCND1-MS2* sites (**Fig. 2c**

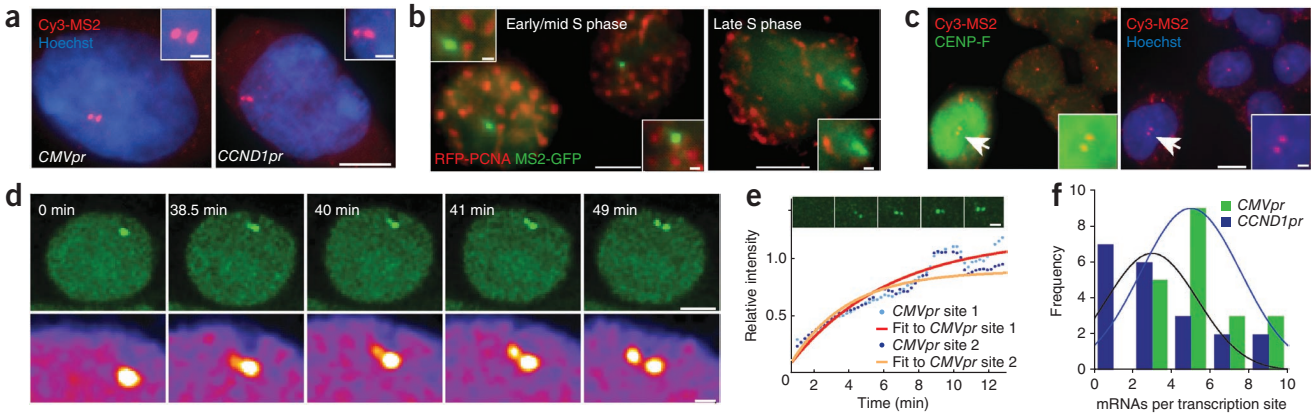

Figure 3 | Transcriptional activity during the cell cycle. (**a**) Fluorescence images revealing two adjacent transcription sites (RNA FISH with a Cy3-MS2 probe) in a *CMVpr-CCND1-MS2* (left) and *CCND1pr-CCND1-MS2* (right) cell. (**b**) Stained nuclei from a *CMVpr-CCND1-MS2* cell with one transcription site (green) in early to mid-S phase (left), and two sites in late S phase (right). RFP-PCNA fusion marks replication foci. (**c**) Fluorescence image showing MS2 FISH signal (red; arrows) and CENP-F labeling (green) of *CMVpr-CCND1-MS2* cells in G2 phase (left). Same field (right) shows Hoechst staining (blue) and RNA FISH (red). (**d**) Frames from a time-lapse movie showing transcription site duplication. Magnification of the transcription sites is shown in the lower images (gray intensity levels were pseudocolored using the ImageJ 'fire' look-up table). (**e**) FRAP frames (13-s interval after bleach), recovery curves and fits of replicated transcription site doublets, site 1 (left) and site 2 (right). (**f**) Quantification of number of *CCND1-MS2* transcripts on replicated sites (n = 20). ANOVA post-hoc Bonferroni test showed a significant difference between the intensities of the single and duplicated sites (P = 0.000 (F = 76.262, d.f. = 3,90), n = 20). Scale bars, 5 μm (insets in **a**–**c**, **e** and magnifications in **d**, 1 μm).

and **Supplementary Table 1**). Other studies have also identified several mRNAs per active gene[3,10,11]. Transcript numbers should correlate with the maximum number of RNA polymerases actively engaged on the gene (from MS2 to the 3′ UTR; **Supplementary Fig. 6c**), hence we conclude that *CCND1pr-CCND1-MS2* engaged fewer polymerases.

We used this quantification to examine the difference in the FRAP rates measured above to determine whether the elongation rates in the two clones differed (**Fig. 1b**). To describe the FRAP data, we used a model based on a rate equation that expresses the increase in the fluorescence signal at the transcription site resulting from the transcription of the MS2 region and the decrease in fluorescence resulting from release of completed transcripts (**Supplementary Note**). Considering the number of quantified polymerases (**Supplementary Table 1**) that contribute to the fluorescent signal buildup and fitting the solution of the rate equation to the FRAP data, the elongation rates in the two clones were 0.31–0.78 kb min^{-1}, in the lower range of previous measurements[1] (**Supplementary Discussion**). Rather than different elongation rates, this analysis demonstrated more frequent *CMVpr* firing compared to *CCND1pr* firing, every 22 s versus 52 s, resulting in calculated polymerase spacing of 237 nucleotides and 335 nucleotides, respectively (**Supplementary Table 1**). Thus, *CMVpr*-driven expression followed the kinetics of overexpression, which in fact resulted from engagement of more polymerases per time unit.

Next, we examined the transcriptional profiles of the *CMVpr*- and *CCND1pr*-driven genes during the cell cycle. We detected nuclei with two active transcription sites in close proximity (doublets), which could correspond to transcription sites on sister chromatids, in both *CCND1pr-CCND1-MS2* and *CMVpr-CCND1-MS2* clones (**Fig. 3a** and **Supplementary Fig. 7a**). To examine the possibility of transcriptional activity occurring after DNA replication, we labeled cells with specific markers of cell-cycle S phase (proliferating cell nuclear antigen; PCNA) and G2 phase (centromere protein F, CENP-F). We found transcription site doublets only in cells that had replicated their DNA, as judged by a display of peripheral replication foci, a phenotype of late S phase[12] and cells expressing nuclear CENP-F (**Fig. 3b**,**c**). Synchronization to late S phase increased doublet frequency (**Supplementary Fig. 7b**,**c**). The doublet state persisted until mitosis, whereas daughter cells had one site only (**Supplementary Fig. 7d** and **Supplementary Video 7**). We describe the nuclear distribution and dynamics of the transcription sites in **Supplementary Figure 8** and **Supplementary Video 8**.

Live-cell imaging in unsynchronized cells containing one transcription site showed the sudden appearance of a second transcription site, next to the 'mother site' (**Fig. 3d**, **Supplementary Fig. 9** and **Supplementary Videos 9** and **10**), which we could follow for ~1 h. Typically, the second transcription site emerged and distanced itself from the initial transcription site (1.8 ± 0.3 μm between sites) over ~10 min, allowing detection of chromatid separation. As replication can precede chromatid separation by several hours[13,14], we followed the 'mother site' for 3 h before site duplication (**Supplementary Fig. 10a**,**b**) and saw no transcription shutdown during this time (we cannot rule out that shutdown occurred in the frame intervals (8–10 min)). However, faster imaging rates (every 15 s) showed that 'mother-site' transcription persisted uninterrupted for at least 40 min (**Supplementary Fig. 9a**). These data suggest an interplay between replication and transcription. Simultaneous FRAP analysis of the replicated sites showed that both recovered with similar kinetics (**Fig. 3e** and **Supplementary Video 11**), albeit slower than before replication. Quantified mRNA numbers engaged with the duplicated transcription sites were: *CMVpr-CCND1-MS2*, 5 ± 3 mRNAs per site and *CCND1pr-CCND1-MS2*, 3 ± 2 mRNAs per site (**Fig. 3f** and **Supplementary Fig. 10c**), implying a drastic reduction in promoter firing after replication compared to the rate before replication (**Fig. 2c**). Our approach demonstrates the high-resolution of single-molecule analysis, as changes in *CCND1* mRNA amounts were not detectable by RT-PCR (**Supplementary Fig. 11**), whereas RNA FISH suggests a decline during S to G2 phases[15].

Our real-time kinetic analysis enables single-allele visualization and analysis during the cell cycle *in vivo*. The endogenous promoter had alternating periods of gene activity as seen in transcriptional bursting (**Supplementary Discussion**), and the viral promoter had higher and constant expression. This system will allow additional observations and analysis of the kinetic behavior of single endogenous genes as well as examination of possible interplay between the transcription and replication machineries.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank R. Pestell (Thomas Jefferson University) for the cyclin D1 promoter, C. Cardoso (Technische Universität Darmstadt) for RFP-PCNA and R. Drummer (Bar-Ilan University) for statistical analysis. This work was supported by grants to Y.S.-T. by the Israel Cancer Research Fund (Research Career Development Award), the Israel Ministry of Health, the Israel Cancer Association, the Alon Fellowship and the Jane Stern Lebell Family Fellowship of Bar-Ilan University; and to Y.S.-T. and Y.G. by the Israel Science Foundation Bikura grant. We thank the Israel Science Foundation for the fluorescence live-cell imaging microscopes.

AUTHOR CONTRIBUTIONS

S.Y. generated the cell system and performed the experiments; L.R. and Y.G. performed FRAP and diffusion experiments, and model analysis; and Y.S.-T. designed the project and wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Hager, G.L., McNally, J.G. & Misteli, T. *Mol. Cell* **35**, 741–753 (2009).
2. Darzacq, X., Singer, R.H. & Shav-Tal, Y. *Curr. Opin. Cell Biol.* **17**, 332–339 (2005).
3. Chubb, J.R., Trcek, T., Shenoy, S.M. & Singer, R.H. *Curr. Biol.* **16**, 1018–1025 (2006).
4. Golding, I., Paulsson, J., Zawilski, S.M. & Cox, E.C. *Cell* **123**, 1025–1036 (2005).
5. Darzacq, X. *et al. Nat. Struct. Mol. Biol.* **14**, 796–806 (2007).
6. Boireau, S. *et al. J. Cell Biol.* **179**, 291–304 (2007).
7. Fusco, D. *et al. Curr. Biol.* **13**, 161–167 (2003).
8. Shav-Tal, Y. *et al. Science* **304**, 1797–1800 (2004).
9. Vargas, D.Y. *et al. Proc. Natl. Acad. Sci. USA* **102**, 17008–17013 (2005).
10. Femino, A.M., Fay, F.S., Fogarty, K. & Singer, R.H. *Science* **280**, 585–590 (1998).
11. Osheim, Y.N., Miller, O.L. Jr. & Beyer, A.L. *Cell* **43**, 143–151 (1985).
12. Sporbert, A. *et al. Mol. Cell* **10**, 1355–1365 (2002).
13. Azuara, V. *et al. Nat. Cell Biol.* **5**, 668–674 (2003).
14. Mesner, L.D., Hamlin, J.L. & Dijkwel, P.A. *Proc. Natl. Acad. Sci. USA* **100**, 3281–3286 (2003).
15. Guo, Y., Stacey, D.W. & Hitomi, M. *Oncogene* **21**, 7545–7556 (2002).

ONLINE METHODS

Plasmids and cloning. The cyclin D1 endogenous promoter[16] was obtained from R.G. Pestell (Thomas Jefferson University). RFP-PCNA was obtained from C. Cardoso (Technische Universität, Darmstadt)[12]. The pcDNA5/FRT plasmid was manipulated by the following steps to insert the gene of interest. To generate the plasmid backbone for further cloning, the *CMVpr* was removed with BamHI and BglII from the pcDNA5/FRT plasmid. Then an adaptor containing the coding sequence of the HA tag that will join N-terminally in frame with the cyclin D1 protein was inserted (adaptor 1, **Supplementary Table 2**).

RT-PCR amplification of the full human *CCND1* coding region (888 bp) containing also 378 bp from the 3′ UTR region, was performed using specific primers to the *CCND1* gene, with added NotI and XhoI restriction sites. A second adaptor was inserted into the 3′ UTR of the cyclin D1 gene between AflII and HindIII endogenous restriction sites, for cloning of the 24 copies of the MS2 sequence repeats (adaptor 2, **Supplementary Table 2**). Then the 24 MS2 sequence repeats (1,308 bp) were added using BamHI and BglII sites.

Insertion of the endogenous promoter and 3′ UTR. Owing to cloning limitations for ligating of the cyclin D1 endogenous promoter (*CCND1pr*) or the *CMVpr* directly into the *FRT* plasmid, the shuttle pUC19 plasmid was used. First, an adaptor with suitable restriction sites was ligated into pUC19 using EcoRI and HindIII sites (adaptor 3; **Supplementary Table 2**). Then the full length sequence of the *CCND1pr* (1,903 bp) was inserted into pUC19 using EcoRI and HindIII sites. RT-PCR of the entire 3′ UTR region (2,711 bp) of cyclin D1 was performed with primers that contain suitable restriction sites (BglII and XhoI), using a proofreading polymerase (Fermentas), and inserted using BglII and XhoI sites. The *CCND1pr* was inserted into the *FRT/HA-CCND1/24MS2/3′UTR* plasmid using AgeI and NruI sites (see **Supplementary Table 2** for primer sequences).

Insertion of the *CMVpr* and 3′ UTR. The *CMVpr* was inserted into pUC19 to gain suitable restriction sites using BglII and HindIII sites, and then moved to the *FRT/HA-CCND1/24MS2/3′UTR* plasmid with the 3′ UTR using AgeI and NruI sites.

Ligations were performed with concentrated T4 DNA ligase (5 U μl^{-1}; New England Biolabs). Recircularization of vectors was prevented by shrimp alkaline phosphatase (Fermentas) treatment.

Cell culture and stable cell line generation. HEK-293 Flp-in cells were obtained from Invitrogen and maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% FBS (HyClone Laboratories). Transfections were performed either by electroporation (10–15 μg DNA; Gene pulser Xcell; BioRad) or by calcium-phosphate precipitation (for transient MS2-GFP expression). Stable integration into the *FRT* site in the HEK-293 Flp-in cells was performed using the *FRT-HA-CCND1-MS2* plasmids together with the pOG44 plasmid expressing the Flp recombinase. Stable selection was performed with 100 μg ml^{-1} hygromycin (Sigma-Aldrich). Single colonies were picked and screened for *CCND1-MS2* expression. For *LacZ* detection, fixed cells were incubated for 6 h at 37 °C with X-gal solution (5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$ and 2 mM $MgCl_2$). For cell-cycle synchronization, cells were incubated with 2 mM thymidine (Sigma) for 24 h, or 1 mM etoposide for 14 h. RFP-PCNA in replication foci were observed after fixation in methanol[12].

Fluorescence microscopy, live-cell imaging and data analysis. Widefield fluorescence images were obtained using the Cell^R system based on an Olympus IX81 fully motorized inverted microscope (60× PlanApo objective, 1.42 numerical aperture (NA) or 100× objective, 1.40 NA) fitted with an Orca-AG charge-coupled device (CCD) camera (Hamamatsu) driven by the Cell^R software. For time-lapse imaging, cells were plated on glass-bottom tissue-culture plates with collagen coating (MatTek) in medium containing 10% FCS at 37 °C. The microscope is equipped with an on-scope incubator, which includes temperature and CO_2 control (Life Imaging Services). For long-term imaging of transcription-site activation, several cell positions were chosen and recorded by a motorized stage (Scan IM; Märzhäuser).

A laser scanning confocal microscope (Olympus FV1000) with a 60×/1.35 NA PlanApo objective was also used. eGFP fluorescence was detected using an Aragon laser (488 nm, 3 mV output) with 1–2% of laser power and 0.041 × 0.041 μm pixel size. Cells were maintained at 37 °C in a 5% CO_2 humidified atmosphere using a stage adapted incubator in conjunction with an objective heater (Tokai).

In live-cell experiments, cells were typically imaged in four dimensions (three dimensions over time). For presentation of the movies, the four-dimensional (4D) image sequences were transformed into a time sequence by choosing the best focus (highest intensity) plane in each time point, using in-house–generated ImageJ scripts (US National Institutes of Health). To improve quality, movies were deconvolved using Huygens Essential II with the time series option (Scientific Volume Imaging). *CMVpr-CCND1-MS2, n* = 43; *CCND1pr-CCND1-MS2, n* = 26 cells. Tracking was performed using the tracking module of Imaris (Bitplane) or the ImageJ Spot Tracker plugin. Correction of cell movement during tracking was performed using the 'correct drift' option in Imaris (tracks the center of mass and cancels these movements). Bleaching correction was applied to time-lapse images.

Fluorescence *in situ* hybridization. RNA FISH was performed as previously described[17] using 40 ng of a Cy3-MS2 DNA probe[18]. For the quantification of the number of mRNAs on the transcription sites or in the total cell, 3D stacks (0.2-μm steps, 76 planes) of the total volume of the cells were collected from the *CMVpr-CCND1-MS2* or *CCND1pr-CCND1-MS2* cells. The 3D stacks were deconvolved and the specific signals of mRNPs were identified (Imaris). mRNP identification was performed in comparison to deconvolved stacks from native HEK-293 cells not containing the *D1-MS2* integration, which therefore served as background levels of nonspecific fluorescence. No mRNPs were identified in control cells. The sum of intensity of each mRNA particle and transcription sites was measured in the same cells using Imaris. The single mRNP intensities were pooled and the frequent value was calculated. The sum of intensity at the transcription site was divided by the frequent value of a single mRNP. This ratio provided the number of mRNAs associated with the transcription unit from the point of the MS2-region and onwards. As mRNAs should be associated with a polymerase, this number should

doi:10.1038/nmeth.1482

reflect the maximum number of polymerases engaged with this region. For quantification of the number of nascent mRNAs present on transcription sites or in the whole cell volumes, 25 cells containing single sites were analyzed; for quantification of nascent mRNAs on the replicated sites, 20 cells containing doublet sites were analyzed. For counting of single versus double transcription sites before and after synchronization we used 300 cells. For counting of active transcription sites in a cell population we used 400 *CMVpr-CCND1-MS2* cells and 425 *CCND1pr-CCND1-MS2* cells.

Immunofluorescence. After RNA FISH, cells were fixed for 20 min in 4% PFA and then permeabilized with 0.5% Triton X-100 for 3 min. After blocking with 5% BSA, cells were immunostained for 1 h with a primary antibody and after subsequent washes the cells were incubated for 1 h with a secondary antibody. Coverslips were stained with Hoechst and mounted in mounting medium. Primary antibodies used were rabbit anti-HA, rabbit anti–CENP-F, rabbit anti-fibrillarin (Abcam) and mouse anti-SC35 (Sigma). Secondary antibodies were FITC-labeled anti-mouse, Cy3-labeled anti-mouse (Jackson ImmunoResearch), Alexa Fluor 594–labeled anti-mouse and Alexa Fluor 594–labeled anti-rabbit (Molecular Probes).

Fluorescence-activated cell sorting (FACS) analysis. Cells were trypsinized and fixed with 70% ethanol (4 °C, overnight). After fixation, cells were centrifuged for 4 min at 200*g* and incubated for 30 min at 4 °C in 1 ml of PBS, centrifuged and resuspended in PBS containing 5 mg ml^{-1} propidium iodide and 50 μg ml^{-1} RNase A for 20 min at room temperature (25°C). Fluorescence intensity was analyzed using a BD Biosciences flow cytometer. Cell synchronization experiments were repeated 4 times. For cell-cycle analysis of reporter mRNA, cells were synchronized and cells from the same preparation were taken for FACS analysis and semiquantitative RT-PCR.

RNA extraction and RT-PCR. RNA extractions were carried out with Tri-Reagent (Sigma) or EZ-RNA Total RNA Isolation kit (Biological Industries). RNA samples were treated with DNA-free (Ambion) to remove traces of contaminating genomic DNA. cDNA was synthesized using the RevertAid First Strand cDNA Synthesis Kit and an oligo(dT) primer (Fermentas). Genomic PCR was performed on DNA isolated by the Puregene DNA Purification System. For primer sequences, see **Supplementary Table 2**.

FRAP experiments. The 4D FRAP experiments were performed with the confocal system, with *z*-dimension intervals of 0.25 μm for a total of 10-μm stacks (40 *z*-dimension slices). Bleaching time was 250 ms for a circular region of interest using 100% laser power. After bleaching, image sequences were acquired at time intervals of 30 s for a total time of 45 min. Image sequences were then analyzed using Matlab (MathWorks) and ImageJ Spot Tracker plugin (http://bigwww.epfl.ch/sage/soft/spottracker/).

In the case of two sites, *z*-dimension steps of 0.2 μm for a total of 10 μm stacks (50 *z*-dimension slices) were captured. After photobleaching, image sequences were acquired every 13 s for a total time of 45 min. Image sequences were deconvolved and then analyzed using Imaris.

Bleach correction was applied to time-lapse images using

$$(I_s(t) - I_n(t) / I_s(t_0) - I_n(t_0)) / (I_n(t) / I_n(t_0))$$

in which t_0 and t are measurement times before and after bleaching, respectively. I_n is the fluorescence intensity of an arbitrary area in the nucleus. I_s is the fluorescence intensity at the transcription site.

Curves were analyzed using the 'reaction dominant' FRAP model[19] assuming that the kinetics of the recovery are governed by transcription (slow population) and that diffusion of free unbound MS2-GFP molecules is fast compared to binding and the FRAP experiment recovery time scale, and therefore does not affect the slow-time rate equations. According to the model, the data were fitted to the following equation:

$$I = C(1 - e^{-kt}).$$

in which C is the amplitude of the FRAP curve. By normalizing the equation to 1, C is eliminated and the data was fitted to $I = (1 - e^{-kt}) + const$, in which I is the relative intensity, k is dissociation rate and *const* is a postbleach background at the transcription site area at the first image after bleach ($n = 15$).

Model. A description of modeling of the data is available in **Supplementary Note**.

Transcription site diffusion. Diffusion measurements for the movement of single transcription sites were obtained with the confocal system. Measurements were performed on 4D image sequences imaged for a total time of ~30 min. Typically, 50 *z*-dimension slices of 320 × 320 pixels (0.35 μm steps) were used. Diffusion measurements of two transcribing sites were obtained with the Cell^R system (Olympus), and 4D image sequences measurements were obtained for a total time of ~12 min with typically 40 *z*-dimension slices (0.4 μm steps).

Transcription site movement data was imported into Matlab for calculation and modeling of the mean square displacement versus time and diffusion coefficients. Diffusion coefficients for three-dimensional Brownian motion of the sites were calculated by finding the average of the square displacements according to the classical diffusion solution: $<r^2> = 6Dt$ (ref. 20), in which $<r^2>$ is the mean square displacement (< > denotes average) of the site over time (t) and D is the diffusion coefficient. D was found by fitting $<r^2>$ as a function of t for the first 20 time intervals, and averaged from 16 single and 12 doublet transcription-site trajectories[21].

Statistical analysis. One-way ANOVA followed by post-hoc Bonferroni test was conducted for comparison between the quantified numbers of mRNAs associated with the two genes on single and double transcription sites. A significant difference was found ($F = 76.862$ (d.f. = 3, 90), $P = 0$). The post-hoc Bonferroni test revealed a significant difference between the quantified mRNAs generated from each promoter for single transcription sites ($n = 25$ each) and in when comparing the single sites to duplicated sites case for each promoter type ($n = 20$ each). For the synchronization experiments we present the mean values

and s.d. Comparison between the diffusion coefficients of the single and double transcription sites was done using the two tailed t-test (single-site, $n = 16$; double site, $n = 12$ ($t = 2.739$, (d.f. = 26), $P = 0.011$). Normal distribution can be assumed for all populations.

16. Albanese, C. *et al. J. Biol. Chem.* **270**, 23589–23597 (1995).
17. Chartrand, P., Bertrand, E., Singer, R.H. & Long, R.M. *Methods Enzymol.* **318**, 493–506 (2000).
18. Bertrand, E. *et al. Mol. Cell* **2**, 437–445 (1998).
19. Sprague, B.L., Pego, R.L., Stavreva, D.A. & McNally, J.G. *Biophys. J.* **86**, 3473–3495 (2004).
20. Saxton, M.J. & Jacobson, K. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 373–399 (1997).
21. Bronstein, I. *et al. Phys. Rev. Lett.* **103**, 018102 (2009).

doi:10.1038/nmeth.1482