Supplementary Methods

for "Genomic encoding of transcriptional burst kinetics" by Larsson A. et al.

Derivation of primary mouse fibroblasts. Primary fibroblasts were derived from female and male adult CAST/EiJ x C57BL/6J or C57BL/6J x CAST/EiJ mice (approval by the Swedish Board of Agriculture, Jordbruksverket: N343/12; no randomization or blinding of animals was performed). Derivation was performed by skinning, mincing and culturing tail explants in fibroblast medium (DMEM high glucose (Invitrogen 41965-039), 10% ES cell FBS (Gibco Ref. 16141-079), 1% Penicillin/Streptomycin (Invitrogen 15140-114), 1% Non-essential amino acids (Invitrogen 11140-035), 1% Sodium-Pyruvate (Invitrogen 11360-039), 0.1mM b-Mercaptoethanol (Sigma). After removal of explants, the culture was passaged twice to attain a pure fibroblast culture. Cells were either 1) dissociated and diluted to very low cell density in medium, and cells were manually picked under 10x or 20x magnification using a thin glass capillary in picking volumes of 0.5µl medium, or 2) dissociated and distributed by FACS into 96-well plates (using the core facility at the Department of Cell and Molecular Biology, Karolinska Institutet) or into 384-well plates using BD Influx sorter. For single-cell RNA-seq, female fibroblasts were used. Cells for RNA-seq were lysed in 4µl Smart-seq2 lysis buffer (1.9µl 0.2-0.4% TritonX-100; 0.1µl Recombinant RNase inhibitor (TaKaRa 40U/µl), 1µl 10µM Smarter oligo-dT, 1µl 10mM dNTPs and 0.1µl ERCC spike-in RNA (Thermo Fischer Scientific, cat no. 4456740), or in 2.3 ul lysis buffer (0.95µl 0.2% TritonX-100; 0.05µl Recombinant RNase inhibitor (TaKaRa 40U/µl), 0.05µl 100µM Smarter oligo-dT, 0.4µl 25mM dNTPs, 0.825µl RNase-free water and 0.025µl ERCC spike-in RNA) per reaction and snap-frozen on dry ice. Smart-seq2 cDNA libraries were prepared as described below.

Culturing of mouse embryonic stem cells. The mESCs (129/CAST) with and without a CRISPR-induced deletion of the Sox2 super-enhancer on the CAST allele was constructed in the Bing Ren lab[23]. Single cells of the 129/CAST cross (wild type) and with the Sox2 enhancer deletion on the CAST

allele was grown as described[29] and later FACS sorted with BD Influx cell sorter (BD Biosciences) individually into 384-well plates containing 2.3 ul of the single-cell RNA-seq lysis buffer. Plates were immediately frozen until further processed.

Single-cell RNA-sequencing: Reverse transcription. Single-cell Smart-seq2 libraries were prepared as earlier described[30,31]. Briefly, cell lysates were incubated for 3 min at 72°C in a thermal cycler, and then put on ice. Reverse transcription (RT) was performed by the addition of 5.6µl Smart-seq2 RT mix (0.5µl SuperScriptII (Invitrogen Cat. 18064-014); 2µl 5x SuperScriptII buffer; 0.5µl 100mM DTT; 2µl 5M betaine; 0.1µl 1mM $MgCl_2$; 0.25µl Recombinant RNase inhibitor (TaKaRa 40U/µl Ref. 2313A); 0.1µl 100µM LNA strand switch primer; 0.15µl water, per reaction) and incubation (90 min 42°C; 10 "strand-switch" cycles of (2 min 50°C; 2 min 42°C) 10min 70°C; 4°C) in a thermal cycler. For Smart-seq2 libraries with a unique molecular identifier, we instead used the following oligo-dt primer: 5Biosg//idSp//idSp//idSp/ACGAGCATCAGCAGCATACGATTTTTTTTTTTTTTTTTTTTTTTTTTTTVN, and template switching oligo: Biotin-AGAGACAGATTGCGCAATGHHHHHHrG+GG, both purchased from Eurogentech.

Single-cell RNA-sequencing: cDNA amplification. cDNA amplification was performed by the addition of 15µl of Smart-seq2 PCR mix (12.5µl 2x KAPA HiFi HotStart ReadyMix (KAPA Biosystems Ref. KK2602); 0.25µl 10µM ISPCR primers; 2.25µl water, per reaction) and incubation (3 min 98°C; 18 cycles of (20 s 98°C; 15s 67°C; 6 min 72°C); 5 min 70°C; 4°C) in a thermal cycler. The cDNA was purified using 0.7:1 volume of AMPure XP beads (Beckman Coulter, Ref. A63882) according to the manufacturer's protocol (No. PT5163-1). The purified cDNA was inspected on an Agilent 2100 Bioanalyzer to determine cDNA concentration and size distribution, using Agilent High Sensitivity DNA chips (Ref. 5067-4626). For Smart-seq2 libraries with a unique molecular identifier, we instead used the following PCR primers: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTGCGCAATG and

ACGAGCATCAGCAGCATACGA.

Single-cell RNA-sequencing: Tagmentation and sequencing. Successful cDNA libraries were tagmented using transposase Tn5, either provided in the Nextera XT DNA kit (Illumina Cat. FC-131-1024, and in accordance with the provided protocol No. 15031942) or by our in-house-generated Tn5 (which produces sequencing libraries of indistinguishable characteristics to those generated using the Nextera XT kit)[32]. For tagmentation using our in-house Tn5, 1ng of cDNA in 5µl water was mixed with 15µl tagmentation mix (1µl of Tn5; 2µl 10x TAPS $MgCl_2$ Tagmentation buffer; 5µl 40% PEG8000; 7µl water, per reaction) and incubated 8 min at 55°C in a thermal cycler. Tn5 was inactivated and released from the DNA by the addition of 5µl 0.2% SDS and 5 min incubation at room temperature. Sequence library amplification was performed using 5µl Nextera XT Index primers (Illumina, Ref. 15032356) and 15µl PCR mix (1µl KAPA HiFi DNA polymerase (KAPA Biosystems Ref. KK202); 10µl 5x KAPA HiFi buffer; 1.5 µl 10mM dNTPs; 2.5µl water, per reaction), and incubation (3 min 72°C; 30 s 95°C; 10 cycles of (10 s 95°C; 30 s 55°C; 30 s 72°C); 5 min 72°C; 4°C) in a thermal cycler. More details about the tagmentation procedure and the synthesis of Tn5 are available in our recent publication. DNA sequencing libraries were purified using 1:1 volume of AMPure XP beads (Beckman Coulter, Ref. A63882) according to the manufacturer's protocol (No. PT5163-1, version PR0X3693), inspected on an Agilent 2100 Bioanalyser, using Agilent High Sensitivity DNA chips (Ref. 5067-4626), and DNA concentration was measured using a Qubit 2.0 Fluorometer (Invitrogen) with the Qubit dsDNA High Sensitivity Assay kit (Molecular Probes, Ref. Q32854). Pools of samples for multiplexing, with unique Illumina barcode for each cell, were prepared according to the Nextera XT DNA Sample Preparation Guide (No. 15031942 page 46, llumina). The multiplexed single-cell libraries were sequenced on either Illumina HiSeq 2000 and 3000 machines, from a single-end for 45 nucleotides in addition to the index reads (P5 and P7). The general depth was approximately 5M reads per cell, which we have determined as being a suitable

depth for obtaining sufficient read coverage across single-nucleotide polymorphisms to enable transcriptome-wide allelic expression analyses in the individual cells.

Single-molecule RNA FISH: Cellular preparation. Male primary mouse adult tail fibroblasts and male embryonic fibroblasts were seeded on 19mm #1 coverslips (VWR, Cat# 631-0155) coated with 0.2% gelatin (Sigma-Aldrich, Cat# G1890) or poly-L-ornithine (Sigma-Aldrich, Cat# P4957) and laminin (Fisher Scientific, Cat# CB-40232), respectively. The day after seeding, cells were 1) washed once in PBS 2) fixed for 10 minutes at room temperature (4% Formaldehyde (ThermoFisher Scientific, Cat# 28906), 1x PBS (HyClone, Cat# SH30378.02) and 3) washed three times in PBS. The PBS was finally replaced with 70% ice-cold ethanol (two times) and cells were stored at 4°C until hybridization.

Single-molecule RNA FISH: Probe hybridization. For hybridization of the probes, fixed cells were first washed once in 10% RNA wash buffer (10% Formamide (ThermoFisher Scientific Cat# AM9342), 2x SSC (ThermoFisher Scientific, Cat# AM9763)) followed by adding hybridization buffer (10% Dextran sulfate (Sigma-Aldrich Cat# D8906), 10% Formamide, 2x SSC, 1mg/ml E.coli tRNA (Sigma-Aldrich, Cat# 10109541001), 0.02% BSA (Ambion, Cat# AM2616), 10mM Ribonucleoside Vanadyl complex (New England Biolabs, Cat# S1402S)) containing 250nM of the corresponding probes (Biosearch Technologies, Custom assay with Quasar 570 dye). Segmentation of cells was facilitated by adding Vimentin antibody to tail fibroblasts (ThermoFisher Scientific, Cat# MA5-11883-A647, 1:50 dilution, Alexa Fluor 647) or SSEA-1 antibody to embryonic fibroblasts (BD Pharmingen, Cat#: 560120, 1:20 dilution, Alexa Fluor 647). Hybridization was carried out at 37°C for 16-20 hrs. After hybridization, cells were 1) washed twice in 10% RNA wash buffer at 37°C for 30 min (second wash with 0.5 μG/ml DAPI (Sigma-Aldrich, Cat# D9542)) and 2) washed three times in 2x SSC. The cells were finally mounted using ProLong gold (ThermoFisher Scientific, Cat# P10144) and imaged on a Nikon TiE microscope equipped with a 100x oil objective (NA 1.45).

Single-molecule RNA FISH: Imaging. For each tail fibroblast cell, a 2 μm thick z-stack of images with 0.2 μm of distance between individual planes was acquired. For embryonic fibroblasts, a total of 9 μm thick z-stack of images with 0.3 μm distance between individual planes was acquired. Imaged on a Nikon TiE microscope equipped with a 100x oil objective (NA 1.45).

Computational Supplemental Methods

Alignment of sequences libraries. We aligned reads using RNA-STAR[33] towards both the mm9 genome assembly of the C57 genotype and an in-house generated CAST mm9 assembly (replacing variable positions with the CAST based as cataloged in the Sanger mouse strain genome sequencing project). To filter out false positive SNPs before computing allelic expression levels, we requiring that each SNP occurred with both genotypes in the total set of sequenced fibroblasts. We used this filtered SNP set as previously described[12] and counted C57- and CAST-informative bases for each gene, using the samtools mpileup command. For analysis of mESCs from 129 and CAST crosses, we used the exact same strategy but analyzed 129 and CAST genomes.

Gene and allelic expression level quantification. We calculated RPKM values (reads per kilobase and million mapped reads) using rpkmforgenes[34] version 7 Feb 2014 with uniquely mapped reads; Ensembl annotation (from UCSC genome browser, last modification date 11 April 2012) for mouse, length compensation for unmappable positions[35] and the settings -fulltranscript -mRNAnorm -rmnameoverlap and -bothendsceil. Previous results indicate that one RPKM is close to one RNA molecule in primary fibroblast cells and mouse embryonic stem cells analyzed with Smart-seq2[12,13]. We calculated RNA molecules per gene and cell using the unique molecule identifiers (UMIs) present in the template switching oligonucleotide and sequenced in the scRNA-seq. We identified the subset

of reads that contained the template switching oligo sequence and collected UMIs per gene and cell. We collapsed all UMIs with an edit distance of one towards the UMI with most read coverage, to exclude over-counting of molecules from sequencing errors in the UMI sequences. For both RPKMs and molecules, we calculate the ratio of expression from the two alleles as the fraction of SNP-containing reads supporting each genotype multiplied by the total RPKM or total RNA molecules quantified for that gene and cell.

Two state model of stochastic gene expression. A two-state model of gene expression was introduced over twenty years ago to model stochastic gene expression[17]. In the two-state model, a promoter can either be in an ON or OFF state. Transcription occurs with a rate k_{syn} when the promoter is in the ON state, whereas no transcription occurs in the OFF state. In the OFF state, the promoter can be turned on with the rate k_{on}. Similarly, in the ON state, the time until the promoter turns off is exponentially distributed with rate k_{off}. Therefore, the two-state model of transcription has the states (i, n) where $i \in \{0, 1\}$ indicating if the promoter/gene is in the ON or OFF state, and *n* is the number of RNA transcripts present in the cell. Regardless of the promoter state, RNA transcripts are degraded with a constant rate λ. The model was introduced by Peccoud et. al., and more detailed information can be found in their original publication[17]. At steady state of this process, the stationary distribution is equivalent to the Poisson-beta distribution[15], which can be formally proved by considering their probability generating functions. For the Poisson-Beta distribution we let:

$$p|k_{on}, k_{off} \sim Beta(k_{on}, k_{off})$$

$$n|k_{syn}, p \sim Poisson(k_{syn} \cdot p).$$

The resulting marginal distribution $P(n|k_{syn}, k_{on}, k_{off})$ is the probability distribution for the amount of RNA transcripts *n* observed at steady state given the rates k_{on}, k_{off} and k_{syn}. For all the simulations of stochastic gene expression performed in this study, we generated cellular snapshots using the numpy package in Python. These parameters are further normalized by the degradation rate λ, e.g. $\frac{k_{on}}{\lambda}$, with the consequence that inferred parameters for each gene have the time-scale of degradation for the gene, but we normally drop the term in the denominator for the sake of brevity. An extended model, including a refractory state after gene activation[6], has been suggested to be more accurate. However, the resulting steady state distribution for the extended model is very close to the two-state model and to distinguish between these similar models, additional information such as multiple time measurements within the same cell is needed.

Maximum likelihood inference of transcriptional kinetics. For a given set of observations X, the maximum likelihood estimation of the parameters $\theta = (k_{on}, k_{off}, k_{syn})$ is found by maximizing the likelihood of the parameters given the observations, i.e.

$$\arg_{\max(\theta \; in \; \Theta)} \prod_{\{x \; in \; X\}} P(x|\, \theta)$$

Equivalently, we can maximize the log-likelihood:

$$\arg_{\max(\theta \; in \; \Theta)} \sum_{\{\mathrm{x \; in \; X}\}} \log(\mathrm{P(x|\theta)}$$

which in turn is equivalent to minimizing the negative log-likelihood:

$$\arg_{\min(\theta \; in \; \Theta)} \sum_{\{\mathrm{x \; in \; X}\}} -\ln(\mathrm{P(x|\theta)}$$

The probability distribution was computed as derived[36] and the negative log-likelihood was minimized using the scipy package in Python with the L-BFGS-B method with bounds $k_{on} = (10^{-3}, 10^{3})$, $k_{off} = (10^{-3}, 10^{3})$, $k_{syn} = (1, 10^{10})$.

Profile likelihood confidence intervals and hypothesis testing for significant differences. The profile likelihood[37] can be derived for any of the three parameters or combinations of parameters, and allows for confidence intervals to be derived. For example, in the case of k_{on}, we start by fixing k_{on} at the ML estimate and minimize the negative log-likelihood (ll) by letting k_{off} and k_{syn} vary freely, treating them as nuisance parameters. We then compute the minimum negative log-likelihood for fixed k_{on} values which are over and under the ML estimate. The negative log-likelihood as a function of k_{on} is quadratic. By standard theory[37],

$$q(k_{on}) = 2 \cdot \left(ll(k_{on}) - \min\left(ll(k_{on})\right)\right)$$

asymptotically approaches the Chi square distribution with $df = 1$, which allows for the construction of confidence intervals. The algorithm stops when $q(k_{on})$ reaches a predefined cutoff value on both sides of the ML estimate. For more detailed information regarding profile likelihood methodology, we recommend the textbook by Pawitan[37].

In the case for burst size, we set $bs = \frac{k_{syn}}{k_{off}}$, fix bs, let k_{on} and k_{off} vary, and let k_{syn} be determined by $k_{syn} = bs \cdot k_{off}$. The profile likelihood is then computed as described above.

In figure 1b we simulate observations based on 140 cells with and calculate the mean relative width ((upper ci – lower ci)/point estimate) of 20 simulated sets of observations for parameter combinations throughout the whole parameters space (100 combinations in total), if we have any obtainable confidence intervals in that set and the calculated mean is lower than 10,000. These relative confidence interval widths (which we call the precision of the estimate) are presented in a contour plot in figure 1b with overlayed point estimates obtained by the maximum likelihood method.

For hypothesis testing, we fix the ratio theta between the two estimates, e.g. in k_{on} between alleles and let $k_{on_2} = 2^{\theta} \cdot k_{on_1}$. We test for significance by comparing the profile likelihood of the maximum likelihood estimate $\theta_A = \log_2(\frac{k_{on_2}}{k_{on_1}})$ to $\theta_H = 0$ (no change) by considering the value of $2 \cdot (\theta_H - \theta_A)$.

How the inference method compares to previous methods. A recent study developed an inference strategy from allelic scRNA-seq data[16] and used the moment estimator from Peccoud et. Al[17] after their histogram repiling method to account for various technical noise. They also apply a gene categorization procedure to prior to inference and their hypothesis testing and confidence intervals used a bootstrap approach. Our likelihood method allows us to develop a statistical modeling approach in which we can investigate the identifiability of the parameters through profile-likelihood instead of categorizing the genes prior to inference. We can determine the goodness-of-fit by using the probability distribution function directly. The hypothesis testing and confidence intervals procedures can directly use the likelihood and is arguably more appropriate than bootstrapping since the theoretical distribution is known. The moment estimator and bootstrap method for confidence intervals is known to produce negative values[15,16] and this is a further indication that the likelihood approach is more appropriate.

Inference of transcriptional kinetics in scRNA datasets with allelic resolution. Based on our previous allelic single-cell RNA-seq analysis in primary fibroblasts and mouse embryonic stem cells[12,13], we approximated the number of transcripts present in a cell by 1 rpkm ~ 1 transcript for rpkm data. Otherwise we used the UMI counts. The maximum likelihood inference was applied as described above to each gene. Genes were filtered based on whether the inferred kinetics fell within the predefined bounds defined above for the L-BFGS-B algorithm.

Calculating burst frequencies at an absolute time-scale using gene-specific MRNA degradation rates. As mentioned previously, the inferred parameters were initially in the time-scale of degradation, i.e. $\frac{k_{on}}{\lambda}$. Burst size is not affected by the time-scale since the two parameters cancel out the degradation term. There is however an interest in deriving the burst frequency on an absolute time-scale. This requires information of the decay rate of transcripts for the given genes. We used the data on half-lifes for genes expressed in mESCs ($t_{1/2}$) from Herzog et. Al[38]. Since we are interested in the decay rate (λ) we need to transform these values with the well-known formula

$$t_{1/2} = \ln(2)\,/\lambda.$$

We can the calculate the burst frequencies in the absolute timescale by

$$\frac{k_{on}}{\lambda} \cdot \lambda = k_{on}.$$

The timescale will be whichever timescale λ is in. The average waiting time until a burst is simply $\frac{1}{k_{on}}$.

Analysis of core promoter elements effect on burst size. We retrieved core promoter annotations from the Eukaryotic Promoter Database (EPD)[39], and grouped genes by the presence of TATA and Inr elements within their core promoters. The EPD is a non-redundant database of RNA polymerase II promoters with experimentally validated transcription start sites. We used their promoter annotations for the TATA and Inr element. Next, we constructed a linear regression model that included the effect of gene length (log10), core promoters (encoded as dummy variables) and possible interactions on burst size (log10)

$$bs \sim\ gl * EPD_{TATA}\ + EPD_{TATA} * EPD_{INR}.$$

The linear regression was performed with the OLS function from the statsmodels package for Python. A similar analysis was made that instead considered mean expression and burst frequency (log10) as the dependent variable,

$$me \sim gl * EPD_{TATA} + EPD_{TATA} * EPD_{INR}$$

$$k_{on} \sim gl * EPD_{TATA} + EPD_{TATA} * EPD_{INR}.$$

For figures (Fig 2b,c and Extended Data Figure 3) we show the fitted values from the linear regression model, which we labeled as "Burst size (model)". Thus, the term "burst size" refers to the inferred gene-specific burst size while "burst size (model)" refers to the burst size predicted by the linear regression model which includes gene length and the existence of core promoter motifs.

Comparison of kinetics between cell types. The fold change of burst frequency and size of genes between mESCs (downloaded: GSE74155 from Chen et al.[41]) and Fibroblasts were divided into 100 bins based on mean expression fold changes and the median of each bin was calculated. The result is presented in Figure 3C. Boxplots were made to show the variability within the bins and is presented in Extended Data Figure 5a and 5b. The difference in mean expression fold changes between the top 100 genes in each direction for burst frequency and size was tested with the t-test and is presented in Extended Data Figure 5c and 5d.

Enhancer H3K27ac read density between cell types. We compared the kinetic parameters between the mESCs and fibroblasts cell types on the C57 allele and assessed whether their kinetics were significantly different using the profile likelihood method described above. The result is presented in Figure 3a and 3b (FDR < 5%). To assess if changes in burst frequencies between cell types (fibroblasts and mESCs) was correlating with enhancer activity, we re-analyzed published Chromatin immunoprecipitation and sequencing (ChIP-seq) data for H3K27ac in for fibroblasts and ES cells[40]. We detected H3K27Ac peaks for fibroblasts and ES cells with MACS 1.4 using input as control. The data read density was normalized between the two samples with MANorm[42].

To map genes to the enhancers in fibroblasts and mESCs, we utilized the enhancer-promoter units (EPUs) defined for these cell types in a large-scale epigenomics study of mouse tissues and cell types[40]. In the cited study, detected enhancer regions were assigned to a region of promoters which they are more likely to interact with based on a high correlation between enhancer and promoter activity for those regions within that cell type. See Shen et. al 2012[40] for details. We determined for each gene which enhancers they preferentially interact with using this EPU definition and was done for both cell types. We then summed magnitude of the obtained peaks (i.e. the read coverage across the enhancer regions) for each gene and cell type. We then computed the fold change in enhancer H3K27ac normalized read density between fibroblasts and ES cells. The particular commands used can be found at the bottom of this document.

The correlation between the fold changes in read densities were compared to fold changes in burst frequency and size by a rolling median (n=200). The result is presented in Figure 3d. The difference in normalized read density fold changes between the top 100 genes in each direction for burst frequency and size was tested with the t-test and is presented in Extended Data Figure 5e and 5f.

Strain variable SNPs in enhancer regions. We compared the kinetic parameters between the CAST and C57 alleles in fibroblasts and assessed whether their kinetics were significantly different using the profile likelihood method described above (FDR<5%). The result was presented in Figure 4a and 4b and histograms of expression distributions over cells were shown for specific cases in Extended Data Figure 8. We assessed whether genes with significant bursting kinetic differences between the CAST and C57 alleles had increased genetic variation within their regulatory regions. To this end, SNPs between CAST/EiJ and C57BL/6J was retrieved[43] and we required SNPs to have the quality label PASS. To map genes to the enhancers in fibroblasts and mESCs, we utilized the enhancer-promoter units defined for these cell types in the large-scale epigenomics study of mouse tissues and cell

types[40]. We constructed a bed file with enhancers, where each enhancer had an entry with values (center, center + 1bp) in field 2 and 3. The bed file was converted from mm9 to mm10 with liftOver (linux executable downloaded from UCSC Genome Browser and appropriate Chain file). We intersected SNPs with enhancer regions using bedtools[44].

For each gene, we considered SNPs located within the immediate 100 bp upstream and downstream from peak of the enhancer region. Which genes each enhancer is determined to interact with is described in the above section. The particular commands used can be found at the bottom of this document. We calculated the number of SNPs in enhancer regions for each gene and sorted the values after their p-values of significantly different burst kinetics between the alleles. A rolling median of 50 genes was applied to the SNP counts. For testing the significant different between SNP counts between significant and non-significant genes the t-test was used. The result for burst frequency is presented in Figure 4c and for burst size presented in Extended Data Figure 5g.

Analysis of Sox2 kinetics in a monoallelic Sox2-SE^{distal} deletion ES cell line. The WT and 2R (monoallelic CAST mutant) cells[24] were prepared and sequenced as described above. The kinetic inference was performed as described above and confidence intervals were obtained for the burst frequency and size of Sox2 for the 129 and CAST allele in both conditions respectively. The hypothesis testing algorithm was then applied to the deleted condition for burst frequency and size. To investigate how well the inference can detect a perturbation to the kinetics we simulated (100 times) a 90% drop in expression (the known effect of losing the Sox2 enhancer in question; we also see roughly such a reduction on mean expression) if that would be only resulting from a loss of burst size or frequency. The two landscapes in parameter space for the two scenarios are shown in green and red, respectively, in Figure 4e.

Computation of RNA molecules from single-molecule RNA FISH images.

The image files were analyzed using rajlabimagetools in the MATLAB environment. We used the standard pipeline in rajlabimagetools to stack Z images and to perform background correction. All cells were manually segmented and background corrected pixel intensities for specific cells exported as data tables. The quantification of RNA dots from these images were performed in Python using the scikit-image package[45] and function *blob_log* (with paramers set to: min_sigma=1, max_sigma=40, num_sigma=5, threshold=1/10 the threshold for the channel identified in rajlabimagetools). Visual inspection validated that counts corresponded to robust spots and that spots covering several pixels were not frequently overcounted. Furthermore, three counts were removed for each cell to correct for false positive signals of multiple confounding sources. Parameters estimation was performed on the RNA molecule distributions over cells (as for scRNA-seq data) and cell-type differences was investigated with the profile likelihood method. The result is presented in Figure 3e. Examples of images obtained are shown in Extended Data Figure 6. The distribution of counts for each genes measured for each cell type is presented in Extended Data Figure 7a-d. Comparisons of kinetic parameters obtained my smFISH and scRNA-seq for these genes are presented in Extended Data Figure 7e-f. Confidence intervals for each gene is presented in Extended Data Figure 7g-j.

Robustness of inference to cell numbers and technical noise in single-cell RNA-seq. We first simulated cellular snapshots to determine the robustness of the inference procedure to the number of cellular observations. For a number of parameters in the parameter space corresponding to specific genes we generated data sets (50 times for each number of cells) with 25, 50, 100, 150, 300 and 500 single-cell observations. We inferred the parameters using maximum likelihood and presented the distributions of inferred parameters as standard boxplots in Extended Data Figure 2.

Next, we determined the robustness of the inference to stochastic losses of RNA molecules mimicking the incomplete sampling of RNAs in single-cell RNA-sequencing protocols[46].

We first simulated 100 cellular observations and then subjected each observation to a stochastic loss by removing each RNA molecules with the probability (1-sensitivity). We iterated this procedure 50 times and noted the inferred burst frequency and size. The results for burst frequency were shown directly in standard boxplots in Extended Data Figure 2 whereas we scaled the inferred burst sizes by dividing with the sensitivity (as the burst size will linearly scale with the sensitivity parameter used). These scaled burst size estimates were shown as standard boxplots in Extended Data Figure 2. Note that the inference procedure is robust to smaller cell numbers, even though increasing cell numbers reduces inference variation. Importantly, the inference was robust to low sensitivities (at and beyond the sensitivity of Smart-seq2 estimated to 35%) when applying technical noise to simulated data to mimic the incomplete sampling of RNAs during single-cell RNA-seq library construction.

Goodness-of-fit test. In order to assess the goodness of fit for the inferred parameters, we used a Monte-Carlo based goodness of fit measure. For each gene, we simulated a set of observations from the inferred parameters (n = number of observations for the gene) and calculated the chi-square score by $\frac{(Expected - Observed)^2}{Expected}$ for a binned gene expression distribution. This was repeated 1000 times to derive a Monte-Carlo null distribution. The inferred parameters were considered to have a good fit if the chi-square score is smaller than at least 5% of the simulated cases. We used the Monte Carlo approach as the number of observations were too low for many expressed genes to have sufficient observations for a standard Chi-square goodness-of-fit test. The result is presented in Extended Data Figure 1c.

Comparision between using rpkm values and UMI counts.

The Aikaike information criteria (AIC) for the model fit was calculated by

$$AIC = 2k - 2\ln(\hat{L})$$

Where $\hat{L}$ is the maximum value of the likelihood function of the model and $k = 3$.

The AIC was calculated for rpkm and UMI values for the same gene and compared with a scatter plot. The result is presented in Extended Data Figure 1d. The kinetic parameters obtained were compared with scatter plots and the results are presented in Extended Data Figure 1e and 1f.

Independence of alleles

The fraction of cells with biallelic expression of a gene was compared to the fraction of silent cells to assess whether alleles transcribe genes independently. See details in Deng et. al 2014[12]. The result is presented in Extended Data Figure 1i.

Investigating the effect of gene length on kinetic parameters.

We binned genes according to gene length (n=30) and investigated the correlation between gene length and burst frequency and size. These results are presented in Extended Data Figure 3a and 3b.

Effect of transcript length on burst size

In order to investigate the effect of transcript length on burst size compared to gene length we binned genes according to gene length (n=20). For each bin, we ranked the genes by the mRNA transcript lengths. For each gene in each bin, we calculated the difference in their burst size from the median burst size of that bin. The differences from all genes with rank 1 (from all initial bins by gene length) are shown in first boxplot, the results from genes ranked second in second boxplot, and so forth, in Extended Data Figure 3c. If mRNA transcript length would confound burst size estimates (e.g. due to Reverse transcription drop-offs during RT) one would expect to see biases for

shorter versus longer mRNA transcripts (when compared for similar gene lengths). We tried different bin sizes without observing strong mRNA transcript effects (in addition to the strong gene length effect).

Power calculation. We investigated the power of detecting 4-fold changes in kinetics for both burst frequency and size for the same parameters used for the sensitivity analysis. Sets of observations with 25, 50, 100, 250, 500, and 1000 cells were simulated (100 times each) with a 4-fold burst frequency change. A similar simulation was performed with a 4-fold burst size change (while keeping the burst frequency fixed). The hypothesis testing algorithm described above was then applied and the p value was noted. The fraction of instances with p-value below 0.05 was calculated for each simulated set of observations and presented in Extended Data Figure 4.

Cell cycle analysis. The cells were classified according to known gene markers for the different stages of the cell cycle as described in Reinius et. al (2016)[13]. Shortly, the 50 most variable known cell cycle genes were identified and used for principal component analysis whereby three phases could be separated (G1, S and G2M). After this classification the counts for the cells were split up according to their stage in the cell cycle and the parameter inference was done separately. Each pair of groups were tested for differential kinetics. The gene-sets which were determined to be significant after FDR-correction (0.05%) were analyzed for GO-term enrichment using PantherDB. The result is presented in Extended Data Figure 10.

Conservation of bursting kinetics between human-mouse orthologs in fibroblasts. Single-cell RNA-seq data for 163 individual human fibroblasts corresponding to UCF1014[47] were downloaded and processed (EGAS00001001009). We phased the haplotypes using scphaser[48], see Edsgard et al. for details. We used the haplotype denoted hapA in the comparison to the C57 allele in the mouse

fibroblasts. Note that the phasing method is not informative with regards to the whether the haplotype is maternal or paternal. Next, we inferred transcriptional kinetics for 2,484 genes out of the 3,111 genes with two or more SNPs. We further only considered one-to-one orthologs (using Ensembl) between human and mouse. To determine whether transcriptional kinetics was conserved beyond what can be accounted for by a conservation of mean expression level, we devised a specific test. The conservation of each one-to-one ortholog was assessed by considering the kinetics of the 50 most similar (in terms of mean expression) genes in both species, respectively. We considered the kinetics to be consistent if the burst frequency or size are on the same side of the median for the kinetics of the 50 genes. Since we could observe that the two measurements of consistency (burst frequency and size) are informative of each other, we decided to only consider the consistency of burst frequency in the resulting significance test. As a control, we selected a random gene out of the top 50 genes in the two species respectively and checked their consistency, which is expected to be half of the time. Real one-to-one orthologs had a significantly increased consistency demonstrating a conservation of the positioning of genes in the parameter space drawn by burst size and frequency. The result is shown in Extended Data Figure 9.

Code for creating enhancer magnitudes between cell types.

wget ftp://ftp.ncbi.nlm.nih.gov/geo/samples/GSM723nnn/GSM723008/suppl/GSM723008_RenLab-Input-MEF-DM266.bam

wget ftp://ftp.ncbi.nlm.nih.gov/geo/samples/GSM851nnn/GSM851277/suppl/GSM851277_RenLab-H3K27ac-MEF-DM708.bam

../bin/macs14 -t GSM851277_RenLab-H3K27ac-MEF-DM708.bam -c GSM723008_RenLab-Input-MEF-DM266.bam -f BAM -g mm -n MEF

wget ftp://ftp.ncbi.nlm.nih.gov/geo/samples/GSM723nnn/GSM723020/suppl/GSM723020_RenLab-Input-mESC-DM162.bam

wget ftp://ftp.ncbi.nlm.nih.gov/geo/samples/GSM851nnn/GSM851278/suppl/GSM851278_RenLab-H3K27ac-mESC-DM754.bam

../bin/macs14 -t GSM851278_RenLab-H3K27ac-mESC-DM754.bam -c GSM723020_RenLab-Input-mESC-DM162.bam -f BAM -g mm -n mESC

Code for SNP density calculation in enhancer regions

```
while read p; do grep $p Ren_supplementary_table7.csv; done < all_genes_per.txt | awk '{printf("%s\t%s\t%s\t%s\n", $1, $2, $2+1, $4)}' - > every_enhancer_site.bed

./liftOver every_enhancer_site.bed mm9ToMm10.over.chain.gz every_enhancer_site_mm10.bed every_enhancer_site_mm10_unlifted.bed

bedtools sort -i every_enhancer_site_mm10.bed > every_enhancer_site_mm10_sorted.bed

bedtools window -b every_enhancer_site_mm10_sorted.bed -a CAST_EiJ.mgp.v5.snps.dbSNP142_PASS_only_renamed.vcf.gz -w 100 | awk '{printf("%s\t%s\t%s\t%s\n", $1, $2, $13, $14)}' - > every_snp_enhancers_per.bed
```

Supplemental Methods References

29. Selvaraj, S., R Dixon, J., Bansal, V. & Ren, B. Whole-genome haplotype reconstruction using proximity-ligation and shotgun sequencing. *Nat. Biotechnol.* **31,** 1111–1118 (2013).
30. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**, 171-181 (2014).
31. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10,** 1096–1098 (2013).
32. Picelli, S. *et al.* Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24,** 2033–2040 (2014).
33. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
34. Ramsköld, D., Wang, E. T., Burge, C. B. & Sandberg, R. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Comput Biol* **5,** e1000598 (2009).
35. Storvall, H., Ramsköld, D. & Sandberg, R. Efficient and comprehensive representation of uniqueness for next-generation sequencing by minimum unique length analyses. *PLoS ONE* **8,** e53822 (2013).
36. Vu, T. N. *et al.* Beta-Poisson model for single-cell RNA-seq data analyses. *Bioinformatics* **32,** 2128–2135 (2016).
37. Pawitan, Y. *In All Likelihood*. (Oxford University Press, 2013).
38. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nat. Methods* **14,** 1198–1204 (2017).
39. Dreos, R., Ambrosini, G., Groux, R., Cavin Périer, R. & Bucher, P. The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. *Nucleic Acids Res.* **45,** D51–D55 (2017).
40. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488,** 116–120 (2012).
41. Chen, G., *et al.* Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. *Genome Research* 26: 1342-1354.
42. Shao, Z., Zhang, Y., Yuan, G.-C., Orkin, S. H. & Waxman, D. J. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol.* **13,** R16 (2012).
43. Keane, T. M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477,** 289–294 (2011).

44. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
45. van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2,** e453 (2014).
46. Ramsköld, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30,** 777–782 (2012).
47. Borel, C. *et al.* Biased allelic expression in human primary fibroblast single cells. *Am. J. Hum. Genet.* **96,** 70–80 (2015).
48. Edsgärd, D., Reinius, B. & Sandberg, R. scphaser: haplotype inference using single-cell RNA-seq data. *Bioinformatics* **32,** 3038–3040 (2016).