Journal of Microbiological Methods 166 (2019) 105745

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Estimating RNA numbers in single cells by RNA fluorescent tagging and flow cytometry

Mohamed N.M. Bahrudeen[a,1], Vatsala Chauhan[a,1], Cristina S.D. Palma[a], Samuel M.D. Oliveira[a,b], Vinodh K. Kandavalli[a], Andre S. Ribeiro[a,*]

[a] Laboratory of Biosystem Dynamics, BioMediTech, Faculty of Medicine and Health Technology, Tampere University, 33101 Tampere, Finland
[b] Department of Electrical and Computer Engineering, Center of Synthetic Biology, Boston University, Boston, USA

ARTICLE INFO

Keywords:
Flow cytometry
Time-lapse microscopy
MS2d-GFP RNA tagging
Single-cell RNA numbers

ABSTRACT

Estimating the statistics of single-cell RNA numbers has become a key source of information on gene expression dynamics. One of the most informative methods of in vivo single-RNA detection is MS2d-GFP tagging. So far, it requires microscopy and laborious semi-manual image analysis, which hampers the amount of collectable data. To overcome this limitation, we present a new methodology for quantifying the mean, standard deviation, and skewness of single-cell distributions of RNA numbers, from flow cytometry data on cells expressing RNA tagged with MS2d-GFP. The quantification method, based on scaling flow-cytometry data from microscopy single-cell data on integer-valued RNA numbers, is shown to readily produce precise, big data on in vivo single-cell distributions of RNA numbers and, thus, can assist in studies of transcription dynamics.

1. Introduction

Single-cell imaging and fluorescent proteins have become a key source of information on multiple processes in live cells, particularly gene expression (Kærn et al., 2005). Originally, they have been used for, e.g., quantifying cell-to-cell diversity in protein levels (Elowitz et al., 2002; Ozbudak et al., 2002; Pedraza and Van, 2005; Engl, 2018). Subsequent progresses in microscopy and in the engineering of synthetic fluorescent proteins have allowed observing in vivo individual proteins (Yu et al., 2006) and RNA molecules (Fusco et al., 2003; Golding et al., 2005; Trcek et al., 2012; Femino et al., 1998; Raj et al., 2008). This made possible, among other, the quantification of the effects and the identification of sources of transcriptional bursting (Golding et al., 2005; Yu et al., 2006; Chong et al., 2014).

While there are several methods to quantify RNA, such as RT-qPCR (Saiki et al., 1985)(Higuchi et al., 1993), microarrays (Bumgarner, 2013), RNA seq (Tang et al., 2009), and UMI-based single-cell RNA-seq (Kivioja et al., 2012; Islam et al., 2014), among other, only a few can visualize individual RNAs, such as RNA Fluorescence In Situ Hybridization (Singer and Ward, 1982), RNA aptamers (Bunka and Stockley, 2006), and MS2-GFP RNA tagging (Golding et al., 2005). The latter allows observing individual RNAs using a synthetic protein, MS2d-GFP, and a synthetic target RNA, coding for multiple binding sites for the MS2d capsid protein (Peabody, 1993). Due to the rapid and stable binding of multiple MS2d-GFP proteins to the several binding sites in a single RNA, time-lapse imaging detects individual RNAs as these are produced. This facilitates the identification of sources of intrinsic noise in RNA production (Golding et al., 2005), the dissection of rate-limiting steps in active transcription (Lloyd-Price et al., 2016; Kandavalli et al., 2016), and the quantification of propensities for threshold crossing in RNA numbers (Startceva et al., 2019), among other.

The quantification of RNA by MS2d-GFP tagging is not free from measurement noise. For example, in time-lapse confocal microscopy, it is not uncommon that tagged RNAs (Supplementary Fig. S1) intermittently disappear. In addition, the precision of the estimation of the number of tagged RNAs within a given 'RNA spot' decreases rapidly with the number of RNAs in the spot (Golding et al., 2005; Häkkinen et al., 2014). Further, it is laborious to collect data, since even when using tailored, state-of-the-art software for segmenting the microscopy images (e.g. (Martins et al., 2018)), it usually still requires manual corrections and, in the worst cases, the necessary information can be absent from the images (e.g. an existing spot might not be captured in the image, e.g. if not within a given z-plane).

One solution to these problems would be to complement the microscopy data on single-cell numbers of MS2d-GFP tagged RNAs with

* Corresponding author at: Arvo Ylpön katu 34, P.O.Box 100, 33014, Tampere University, Finland.
E-mail address: andre.sanchesribeiro@tuni.fi (A.S. Ribeiro).
[1] Equal first authors.

https://doi.org/10.1016/j.mimet.2019.105745
Received 14 June 2019; Received in revised form 11 October 2019; Accepted 15 October 2019
Available online 22 October 2019

flow cytometry data. This would allow to rapidly collect much larger amounts of data, and also reduce significantly the uncertainty in the data (e.g. cells that are not entirely imaged can be automatically removed from the dataset, by using combined information from various channels of the flow cytometer, and RNA spots would always be entirely present in each imaged cell). However, flow cytometry lacks spatial information, which so far has been used in the quantification of MS2d-GFP tagged RNAs (Golding et al., 2005; Häkkinen et al., 2014).

Recent approaches have successfully combined Fluorescence in situ hybridization (FISH) for RNA counting with flow cytometry (see e.g. (Arrigucci et al., 2017; Bushkin et al., 2015; Tiberi et al., 2018) for similar aims. However, achieving the same using the MS2d-GFP tagging technique is expected to be more complex because, unlike when using FISH, not only the MS2d-GFP tagged RNAs are fluorescent but also the cells' cytoplasm, due to the need for large numbers of free floating MS2d-GFP to readily detect newly formed RNAs.

To address this, and since MS2d-GFP tagged RNA have been shown to have constant fluorescence for a few hours following their formation (Tran et al., 2015; Lloyd-Price et al., 2016; Oliveira et al., 2016), we hypothesized that cells with tagged RNAs have, on average, higher fluorescence than otherwise (since the MS2d-GFP proteins attached to the RNA are 'immortalized'). As such, the total fluorescence of a cell should increase with the number of tagged RNAs that it accumulates. If so, it should be possible, from flow cytometry data on cells expressing MS2d-GFP tagged RNAs, to estimate the statistics of single-cell distributions of RNA numbers. Here we validate these hypotheses and show that flow cytometry data can be used to extract the mean, standard deviation, and skewness of single-cell distributions of RNA numbers that match those observed using microscopy.

2. Materials and methods

2.1. Chemicals

Measurements were performed in Luria-Bertani (LB) medium. The chemicals were: Tryptone and sodium chloride from Sigma Aldrich. Yeast extract was from Lab M (Topley House, Bury, Lancashire, UK). Antibiotics used are kanamycin and chloramphenicol, from Sigma-Aldrich. Inducers isopropyl β-D-1-thiogalactopyranoside (IPTG), anhydrotetracycline (aTc) and L-Arabinose (ara) were purchased from Sigma-Aldrich. For preparing microscopic gel pads we used agarose from Sigma-Aldrich.

2.2. Strains and plasmids

The *E. coli* strain used was DH5α-PRO, identical to DH5αZ1. Its genotype is deoR, endA1, gyrA96, hsdR17 (rK- mK+), recA1, relA1, supE44, thi-1, Δ(lacZYA-argF)U169, Φ80δlacZΔM15, F-, λ-, PN25/tetR, PlacIq/lacI and SpR. This strain produces the regulatory proteins required for tightly regulating the genetic constructs used (LacI, TetR and AraC).

The two genetic constructs used in this strain are: i) a multi-copy reporter plasmid responsible for producing MS2d-GFP proteins, controlled by the promoter $P_{LtetO-1}$, inducible by aTc; ii) A single-copy target F-plasmid is responsible for producing an RNA coding for mRFP1 up-stream of a 96 MS2 binding site array, controlled by the promoter $P_{Lac/ara-1}$, inducible by IPTG and L-Arabinose, ($P_{Lac/ara-1}$-mRFP1-96BS, Supplementary Fig. S2). We also used the *E. coli* DH5α-PRO strain carrying only the reporter plasmid. The plasmids were transferred into the host strain by standard molecular cloning techniques (Alberts et al., 2002).

The high number of binding sites for MS2d and the high affinity of each site with MS2d proteins cause each target RNA, when tagged, to appear as a bright 'spot' (Fig. 1B and Supplementary section 1.2), soon after being transcribed (in < 1 min) and to exhibit constant fluorescence intensity for a long period of time (mean half-lives of ~140 min (Tran et al., 2015)). Finally, it has been shown that, in these cells, the protein expression level of the target gene is not affected by MS2d-GFP tagging and follows the RNA numbers (Startceva et al., 2019).

2.3. Growth media and induction of the reporter and target genes

From a glycerol stock (−80 °C), cells were streaked on a LB agar plate and incubated at 37 °C overnight. From this plate, a single colony was picked and inoculated into a fresh LB medium supplemented with appropriate antibiotics (35 μg/ml kanamycin and 34 μg/ml chloramphenicol) and grown overnight at 30 °C with aeration. From the overnight cultures, cells were diluted into fresh LB medium to an optical density (OD_{600}) of 0.03, and grown at 37 °C, 250 rpm. Once the cells reach the OD_{600} 0.3, aTc (100 ng/ml) was added to induce $P_{LtetO-1}$ for MS2d-GFP production. L-Arabinose (0.1%) was also added, at the same time, for pre-activation of the target promoter $P_{Lac/ara-1}$. After 50 min, IPTG was added (0, 6.25, 50, 100, 200, 300, 500, or 1000 μM) to activate the production of the RNA target for MS2d-GFP. Following 1 h, cells were observed to quantify RNA and proteins (microscopy or flow cytometry).

2.4. Spectrophotometry

Fluorescence intensities were measured by using a BioTek Synergy HTX Multi-Mode Microplate Reader with Gen5 software. From the overnight culture, cells were diluted to 1:1000 times in fresh LB medium and incubated at 37 °C with shaking, until an OD_{600} of 0.3. Afterward, cells were aliquoted into 96 well microplates, and allow them to grow while maintaining the same temperature and shaking. Following induction of the reporter and target genes (see Section 2.3), mean fluorescence intensities were recorded for 10 h at an interval of 10 min, using the excitation (485/20 nm) and emission (525/20 nm) filters. We performed 6 technical replicates for each condition. We found weak variability between replicates. Results are the averages and standard error of the means.

2.5. Microscopy and image analysis

A few μl of cells were sandwiched between the coverslip and an agarose gel pad (2%), and visualized by a 488 nm argon ion laser (Melles–Griot) and an emission filter (HQ514/30, Nikon), using a Nikon Eclipse (Ti-E, Nikon) inverted microscope with a 100× Apo TIRF (1.49 NA, oil) objective. Fluorescence images were acquired by C2+ (Nikon), a point scanning confocal microscope system. The laser shutter was open only during exposure time to minimize photo bleaching. Simultaneously with the confocal images, phase contrast images were also captured by a CCD camera (DS-Fi2, Nikon). All images were acquired with NIS-Elements software (Nikon). Microscopy images were analysed using the software 'CellAging' (Häkkinen et al., 2013). For details, see Supplementary Materials and Methods, Sections 1.1 and 1.2.

2.6. Flow cytometry and gating

Cells carrying the target and reporter genes were grown and induced as described in Section 2.3. For this, from 5 ml of the bacterial culture, cells were diluted 1:10000 into 1 ml PBS and vortexed for 10 s. In each measurement, 50,000 events were recorded using an ACEA NovoCyte Flow Cytometer (ACEA Biosciences Inc., San Diego, USA), equipped with a blue (488 nm) and a yellow laser (561 nm) for excitation. For detection of MS2d-GFP and RNA-MS2d-GFP, we used the fluorescein isothiocyanate (FITC) detection channel (530/30 nm filter) for emission, with a PMT voltage setting of 417. For detection of red fluorescence proteins (mRFP1), we used the PE-Texas Red fluorescence detection channel (615/20 nm) for emission, with a PMT voltage setting of 584. We set a flow rate of 14 μl/min and a core diameter of

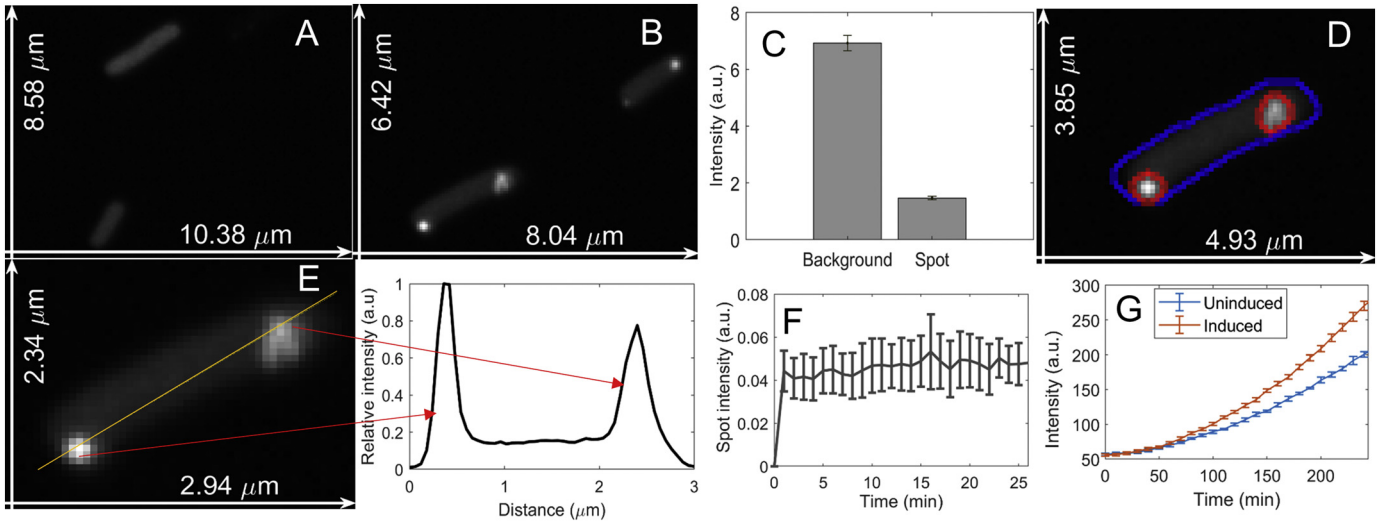

Fig. 1. A) Example microscopy image of cells carrying the reporter gene coding for MS2d-GFP, prior to the production of target RNAs. The cells are visible due to carrying a large amount of MS2d-GFP proteins; B) Example microscopy image of cells carrying the reporter gene coding for MS2d-GFP, after the production of target RNAs. The RNAs tagged with MS2d-GFP are visible as bright spots; (C) Mean total cell background fluorescence intensity and mean total fluorescence intensity of all RNA spots in individual cells (in arbitrary units), as measured by confocal microscopy (Methods, Section 2.5). Data from > 300 cells. The error bars are the standard error of mean. (D) Example image of a cell, along with the results of the segmentation of the cell border (blue line) and of the RNA spots within (red circles) using the tailored software 'SCIP' (Martins et al., 2018) (Methods, Section 2.5). (E) Left: example image of a cell along with a yellow line, manually introduced to obtain a fluorescence intensity profile using imageJ (Abramoff et al., 2004). Right: pixel intensity (in arbitrary units) along the yellow line shown on the left image. The peaks correspond to the regions where the two spots (tagged MS2d-GFP RNAs) are located. (F) Mean fluorescence intensity of individual tagged RNA molecules over time since first appearing. 10 tagged RNAs were tracked, all from cells with only one RNA. Also shown is the standard error of the mean (vertical bars). (G) Total fluorescence intensity (in arbitrary units) of cell populations over time, as measured by spectrophotometry, obtained from cells with target and reporter plasmids induced (brown line) and from cells with only the reporter plasmid induced (blue line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

7.7 μM. To avoid background signal from particles smaller than bacteria, the detection threshold was set to 5000 in FSC-H analysis. Data were extracted using the ACEA NovoExpress software.

We applied unsupervised gating (Razo-Mejia et al., 2018) to the flow cytometry data. We set the fraction of single-cell events whose data is used in the analysis (α) to 0.99, as it was sufficient to remove noncell events produced by debris, cell doublets, cell fragments, clump of cells, and other undesired events. Reducing α further did not change the results qualitatively. In addition, we removed events that did not exhibit fluorescence from free-floating MS2d-GFP by applying (manually) a minimum threshold (Supplementary Fig. S3, Right). Also, we removed < 0.01% of the events with highest FITC-H normalized by pulse width (F/W) values. Similarly, we removed < 0.01% of the events with highest R/W values. In all measurements by flow cytometry followed by data filtering, > 40,000 single-cells events were analysed per condition.

Finally, the total cell fluorescence differs with cell size (Supplementary Fig. S4, Right), while the concentration of MS2d-GFP does not (Supplementary Fig. S4, Left). To account for this, we normalized the FITC-H signal by the pulse Width (which differs with cell size (Cunningham, 1990; Traganos, 1984) denoted by F/W. Likewise, we also normalized the PETexasRedH signal by the pulse Width, denoted by R/W. For this reason, throughout the results section, we only refer to F/W and R/W when referring to flow cytometry data.

2.7. Mean, standard deviation, and skewness of single-cell distributions of RNA and protein numbers

We calculated the mean (M), Variance (Var), standard deviation (Sd), 3rd moment and skewness (S), of the distribution of single-cell RNA numbers (obtained from microscopy), and of the single-cell distributions of F/W, and R/W (obtained from flow cytometry), as shown in Table 1:

The standard error of M is calculated from $\frac{Sd(X)}{\sqrt{N}}$, where N is the sample size of X. Meanwhile, the standard error (SE) of Var, Sd, 3rd moment and S is estimated using a non-parametric bootstrap method (Carpenter and Bithell, 2000; DiCiccio and Efron, 1996), by performing 10^3 random resamples with replacement, to obtain the bootstrapped distributions of Var, Sd, 3rd moment and S.

Table 1
Mean (M), Variance (Var), standard deviation (Sd), 3rd moment and skewness (S), of a distribution of observed values of the sample items, X, where <.> stands for average.

Feature	M	Var	Sd	3rd moment	S
Definition	$\langle X \rangle$	$\langle (X - \langle X \rangle)^2 \rangle$	$\sqrt{\langle (X-\langle X \rangle)^2 \rangle}$	$\langle (X - \langle X \rangle)^3 \rangle$	$\frac{\langle (X - \langle X \rangle)^3 \rangle}{Sd^3}$

3. Results and conclusions

3.1. Time-course cell fluorescence in the presence and absence of RNA target for MS2d-GFP

We performed time-lapse microscopy measurements of *E. coli* cells carrying a gene coding for RNA target for MS2d-GFP, under the control of the Lac/Ara-1 promoter ($P_{Lac/ara-1}$). The cells also produce MS2d-GFP proteins from a multi-copy plasmid controlled by the $P_{LtetO-1}$ promoter (Materials and Methods, Section 2.2).

For RNAs target for MS2d-GFP to be readily detected, the cells need to contain multiple MS2d-GFP proteins (Golding et al., 2005). Due to this, their background is green fluorescent (Fig. 1A) and each target RNA appears as a bright spot in < 1 min after being produced (Tran et al., 2015) (Fig. 1B). In general, using these constructs and conditions, the cells produce from one to a few target RNAs during their lifetime (Häkkinen and Ribeiro, 2016).

In the absence of MS2d-GFP tagged RNAs, the total cell background fluorescence (i.e. the sum of the intensity of all pixels covering the cell

area) is nearly only due to free floating MS2d-GFPs (Supplementary Fig. S5). In addition, this total background fluorescence is higher than the fluorescence of single MS2d-GFP tagged RNA spots (Fig. 1C). Nevertheless, MS2d-GFP tagged RNAs are clearly visible to the Human eye (Fig. 1D) and detectable by image analysis (Martins et al., 2018), as the fluorescence intensity of pixels with a spot is much higher than in near-neighbour pixels (Fig. 1E). Thus, using spatial information, RNA spots can be segmented, e.g., by kernel density estimation with a Gaussian kernel (Häkkinen and Ribeiro, 2015). In addition, the variability in fluorescence intensity of pixels where spots are absence is much smaller than the difference in fluorescence intensity between pixels with and without a spot (Fig. 1E). Due to this, one can subtract the mean background fluorescence from a spot's total fluorescence to obtain a 'corrected' spot intensity, without risk of wrongly adding a 'false' RNA spot or removing a 'true' RNA spot. Unfortunately, these methods cannot be applied to flow cytometry data, as it only informs on total cell fluorescence.

Even though the spots' fluorescence is weaker than the total cell fluorescence, we hypothesized that the production of an RNA target for MS2d-GFP increases the total cell fluorescence, since the binding of MS2d-GFP to the target RNA will protect bound MS2d-GFP proteins from degradation or loss of fluorescence intensity (Tran et al., 2015). This is due to the weak disassociation rate constant of MS2d from the specific target RNA sequence (Dolgosheina et al., 2014), and the high stability and long lifetime of the fluorescence intensity of MS2d-GFP tagged RNAs. In particular, Fig. 1F shows that the RNA-MS2d-GFP complexes have a weak mean fluorescence decay rate of $\sim 8 \times 10^{-5}\,s^{-1}$, which correspond to long mean half-lives of ~140 min, in agreement with past reports (Tran et al., 2015; Golding and Cox, 2004; Golding et al., 2005). Consequently, following the production of an MS2d-GFP tagged RNA, as a cell produces more MS2d-GFP, a new equilibrium in the number of MS2d-GFP in the cytoplasm is expected to be reached, causing the total cell fluorescence to become higher.

To validate this hypothesis, we measured by spectrophotometry the cells' fluorescence over time, when and when not inducing the target gene with L-Arabinose and IPTG. Also, we measured cell grow rates. From Supplementary Fig. S6, the cell growth rate does not differ between the conditions. Meanwhile, from Fig. 1G, the activation of the target gene, as time progresses and tagged RNAs accumulate, causes a continuous increase in the mean cell fluorescence.

Next, we subjected cells with the target gene controlled by $P_{Lac/ara-1}$ (responsible for producing the RNA target for MS2d-GFP) to various IPTG concentrations (Methods). As a control, we performed the same measurements on the strain without the target gene (Methods). We measured by flow cytometry the single-cell fluorescence intensity relative to cell size, so as to account for differences in cell size (Cunningham, 1990; Traganos et al., 1984). In particular, we calculated the 'FITC-H' signal relative to the 'pulse Width', here onwards referred to as F/W (Methods, Section 2.6).

As a control, we further verified by microscopy that cells do not differ significantly in morphology, for different IPTG concentrations, by comparing their mean length along the major axis. We found no significant differences between conditions (Supplementary Fig. S7).

From Fig. 2, while both strains are subject to the inducers, only cells carrying the gene coding for the RNA target for MS2d-GFP show increased F/W for increasing IPTG, which is consistent with the increase in RNA numbers as measured by microscopy (Fig. 3A and D). It is also consistent with the results by spectrophotometry (note that, at 1 mM IPTG, the total cell fluorescence of cells of the strain carrying the target is also approximately 30% higher, as in Fig. 1G). Given this and all of the above, we conclude that the increase in F/W with increasing IPTG is solely due to the appearance of RNAs tagged with MS2d-GFP.

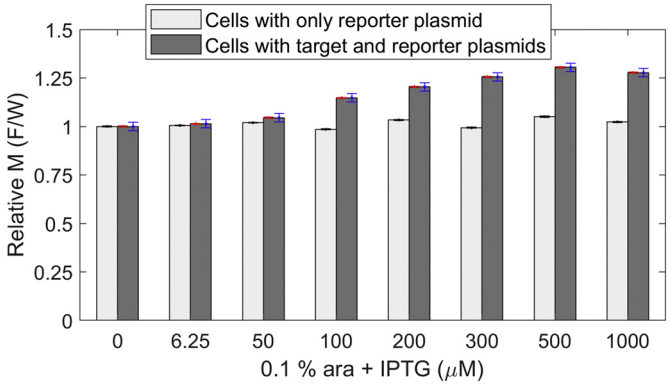

Fig. 2. (Light grey bars) Mean F/W values of the strain carrying only the multi-copy plasmid carrying the reporter gene, at various IPTG concentrations (x-axis), relative to its mean F/W value at the 0 μM IPTG condition. Its black error bars are the standard error of mean, estimated from the cells in each condition (Methods, section 2.7), relative to its mean F/W value at the 0 μM IPTG condition. (Dark grey bars) Mean F/W values of the strain with both the single-copy F-plasmid with the target gene and the multi-copy plasmid with the reporter gene at various IPTG concentrations (x-axis), relative to its mean F/W value at the 0 μM IPTG condition. The red error bars are the standard error of mean, estimated from the cells in each condition (Methods, section 2.7), relative to its mean F/W value at the 0 μM IPTG condition. The blue error bars result from the standard error of mean, relative to its mean F/W value at 0 μM IPTG condition, after adding the empirical variability between all measurements using cells with only the reporter gene. This estimation is explained in Supplementary Methods, section 1.6. In all conditions, cells were given 0.1% of L-Arabinose (Methods, Section 2.3).

3.2. *Relationship between the statistics of single-cell RNA numbers obtained by confocal microscopy and single-cell F/W obtained by flow cytometry*

We next measured by microscopy and image analysis (Methods, section 2.5) the RNA numbers produced by our gene of interest, under the control of $P_{Lac/ara-1}$, for different concentrations of IPTG. Fig. 3A-3C show the mean, standard deviation, and skewness of the single-cell distribution of these numbers (Methods, Section 2.7) as a function of IPTG, respectively.

Next, we extracted the same three statistics of the single-cell distribution of F/W values obtained in the same conditions by flow cytometry. Results are shown in Fig. 3D-3F. Supplementary Fig. S8 (Left) shows the probability density functions of the single-cell F/W values, for each condition.

Given this, we investigated the relationship between the statistics for F/W and the statistics for RNA numbers per cell. Results in Supplementary section 1.5, show that there is a linear fit between the two Means, the two Variances and, the two third moments, respectively.

Given these linear relationships, to evaluate whether the moments of single cell RNA numbers from microscopy and single cell F/W values of flow cytometry are correlated, we plotted the results of Mean (M), Variance (Var) and the third moment obtained by microscopy against the results of M, Var and the third moment obtained by flow cytometry in scatter plots (Fig. 4A-C). Next, we did a linear fit to the data, which was performed using the linear regression fitting method explained in Supplementary Methods, section 1.4. The adjusted R^2 values and corresponding *p*-values of the linear fit are shown in Supplementary Table S1. We find that Mean, Var and the third moment are well fitted by a line (in Fig. 4).

Hence, we conclude that there is a good linear fit between the Mean, Var and the 3rd moment of the single-cell distributions of RNA numbers obtained by microscopy, and the Mean, Var, and the 3rd moment of the single-cell distributions of F/W values obtained by flow cytometry, respectively.

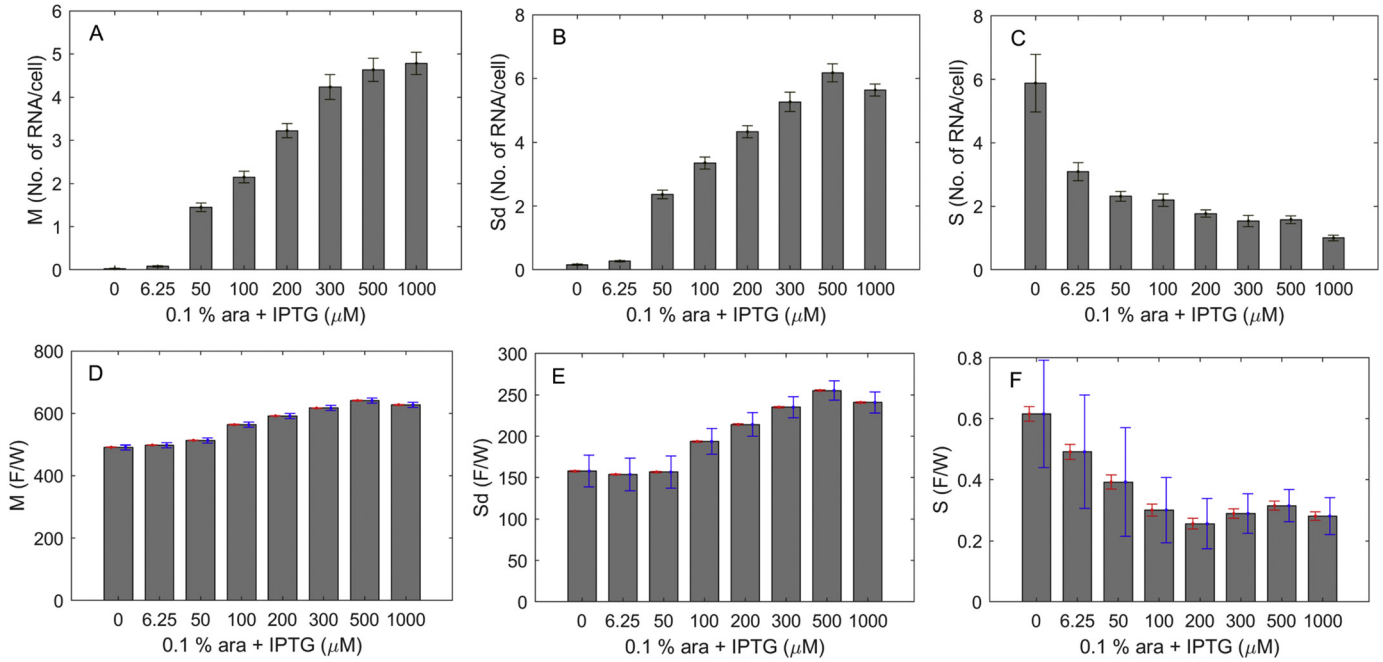

Fig. 3. (A) Mean, M, (B) Standard deviation, Sd, and (C) Skewness, S, of single-cell distributions of integer-valued RNA numbers obtained by microscopy, as a function of IPTG concentration (x-axis). The standard error of M, Sd and S of RNA numbers was estimated as described in Methods, section 2.7. (D) Mean, (E) Standard deviation, and (F) Skewness of the single-cell distribution of F/W values obtained by flow cytometry. The red error bars are standard errors of the statistics (Methods, Section 2.7). The blue error bars are the standard error of the statistics after adding variability estimated from eight technical replicates of cells carrying only the reporter gene (Supplementary Methods, Section 1.6). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These results are, as expected, dependent on degree of background noise, produced by MS2d-GFP (random motion and measurement error generate spatial heterogeneity). This noise could differ, e.g., in different environments or if different plasmids were used to produce MS2d-GFP. We thus tested the effects of increased background noise on our estimation of M, Var, and 3rd moment in cells observed by Flow Cytometry. For this, we modelled increasing background noise by adding increasingly higher Gaussian noise to the total cell fluorescence (F/W) obtained by Flow Cytometry. These added noises are shown in Supplementary Fig. 10A.

The consequences of adding the increasingly higher noise on the mean, variance, and the 3rd moment of the single-cell fluorescence distributions, as measured by Flow Cytometry, are shown, respectively, in Supplementary Fig. 10B, 10C, 10D.

Visibly, from Supplementary Fig. 10B, the addition of Gaussian noise to the single-cell F/W distribution at different IPTG concentrations (Noise corrupted F/W distribution), does not perturb the mean. In particular, even though the Gaussian noise was gradually increased from $\sigma = 0$ to 400, the best fitting lines between the mean of the noise corrupted F/W distribution and mean RNA numbers per cell obtained from microscopy are all identical to the best fitting line when using the original F/W distribution (Supplementary Fig. S10B).

Meanwhile, we expect increasing variance in the noise-corrupted F/W distributions. However, the best fitting line between variance of the noise corrupted F/W distributions and variance of the single-cell RNA numbers distribution shows that only the intercept changes, not the slope (Supplementary Fig. S10C). As such, one can reliably quantify the variance of RNA numbers from the variance of noise corrupted F/W distributions.

Finally, the third moment of noise is not expected to change with increasing Gaussian noise. This can be seen at low noise levels (0 to 200), as the best fitting line between the third moment of noise corrupted F/W distributions and the third moment of the distribution of RNA numbers per cell is almost the same. However, at higher noise levels (300 and 400), the best fitting line shifted slightly (Supplementary Fig. S10D). This may be because the standard

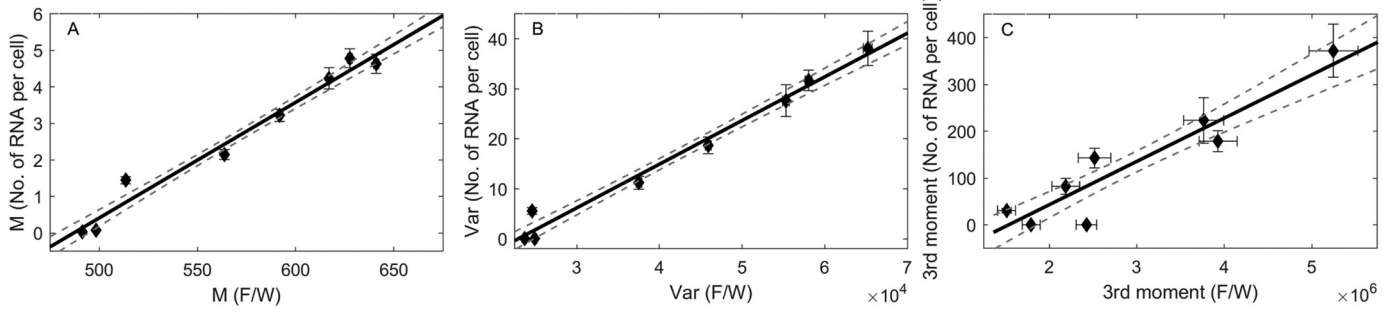

Fig. 4. Scatter plots between (A) Mean (M), (B) Variance (Var), (C) 3rd Moment of the single-cell distributions of F/W values obtained by flow cytometry against M, Var and 3rd Moment of the single-cell distributions of RNA numbers in individual cells obtained by Microscopy for various induction strengths (0, 6.25, 50, 100, 200, 300, 500, and 1000 μM IPTG). The error bars of the points on x and y directions are standard errors estimated as in Methods, section 2.7. In each plot, we obtained the best linear fit (black straight line) as described in Supplementary Methods, section 1.4. The dotted lines are the standard error of the fitted line.

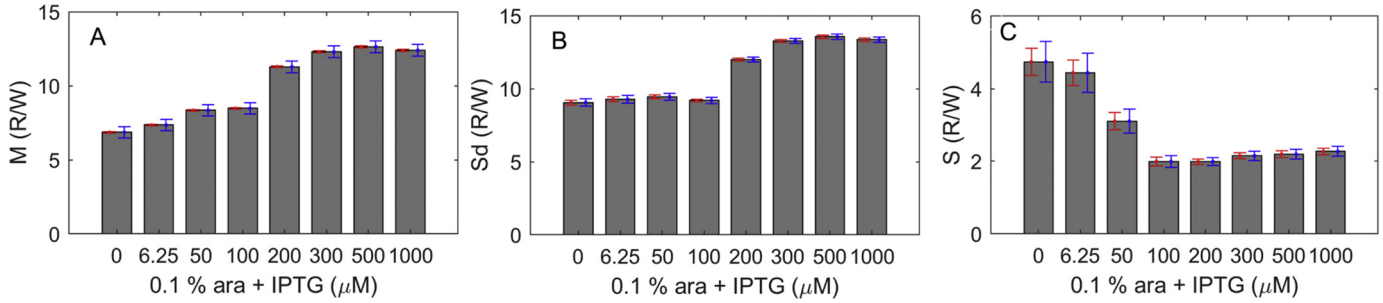

Fig. 5. (A) Mean, M, (B) Standard deviation, Sd, and (C) Skewness, S, of single-cell distributions of R/W values, when subject to various IPTG concentrations. The red error bars are standard errors (Methods, section 2.7). The blue error bars are the standard error of the statistics after adding empirical variability estimated from cells carrying only reporter gene (Supplementary section 1.6 and Supplementary Fig. S12). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

deviation, at the lower induction levels, is smaller than the added noise. In that regime, the parameters of the best fitted line start being sensitive to Gaussian noise, which increases the error of the estimation of the third moment.

3.3. Validation of the quantification of MS2d-GFP tagged RNAs from flow cytometry data

If the signal detected by Flow cytometry is produced by MS2d-GFP tagged RNAs, one should be able to detect the corresponding proteins produced from these RNAs (in particular, mRFP1 red fluorescent proteins, see Methods). To test this, from the same flow cytometry measurements as above, we also extracted the single-cell distribution of PETexasRed-H and normalized these signals by the Pulse Width (denoted as R/W). From the single-cell R/W distribution, we obtained its mean, standard deviation and skewness for each induction strength (Fig. 5A-C). Supplementary Fig. S8 (right) shows the probability density function of R/W for each induction strength.

To assess if the protein statistics (Fig. 5A-C) follows the RNA statistics (Fig. 3D-F), we plotted the values of each statistic in scatter plots (Fig. 6A-C) and fitted with a linear fit. The adjusted R^2 values and corresponding *p*-values of the linear fit are shown in Supplementary Table S2. From the Figures and Table, all three statistics are well fitted by a line. Given the adjusted R^2 values and p-values, we conclude that there is a strong linear fit between those statistics of the single-cell distribution of FITC-H normalized by Pulse width and PETexasRed-H normalized by Pulse width obtained by flow cytometry, respectively. These results confirm that the statistics of distribution of F/W values in Fig. 3D-F should be the result of single-cell distribution of MS2d-GFP tagged RNAs.

3.4. Estimation of mean, standard deviation and skewness of the single-cell distribution of RNA numbers from single-cell F/W values

From the above, it should be possible to estimate the statistics of the distribution of single-cell RNA numbers from the single-cell distribution of F/W values. In particular, it should be possible to 'map' the flow cytometry data to the microscopy data. E.g. one could calibrate two, or more, data points (conditions) of the flow-cytometry data to the corresponding points (conditions) of the microscopy data. Then, we could estimate the RNA numbers statistics of the remaining F/W data points by linear interpolation and/or extrapolation. From this, we can obtain an absolute RNA count scale for estimating the mean, standard deviation, and skewness of the single-cell distribution of RNA numbers from flow-cytometry data.

We start by calibrating the difference between a pair of conditions from flow-cytometry data (e.g. 0 and 1000 μM IPTG) to the difference between the corresponding pair of conditions from microscopy data. This process has to be done independently for the mean, variance, and third moment, but one can use any pair of conditions for each of the moments.

Here we use the data in Fig. 4A-C to obtain the necessary pairs of data points. For this, we started by testing all possible combinations of pairs of data points (Fig. 4A-C contain 8 data points each, and thus, there are 28 possible pairs of data points). Out of these, there are several pairs that provide calibration lines that are consistent between them, and thus can be used to obtain reliable results.

In order to find the largest group of calibration lines that are consistent between, we plotted the y-intercepts against the slopes of the 28 calibration lines. Then we calculated the location of a 'Median point' in that graph whose x-coordinate is the median of the slopes and the y-coordinate is the median of the y-intercepts of the calibration lines (Fig. S9A-S9C).

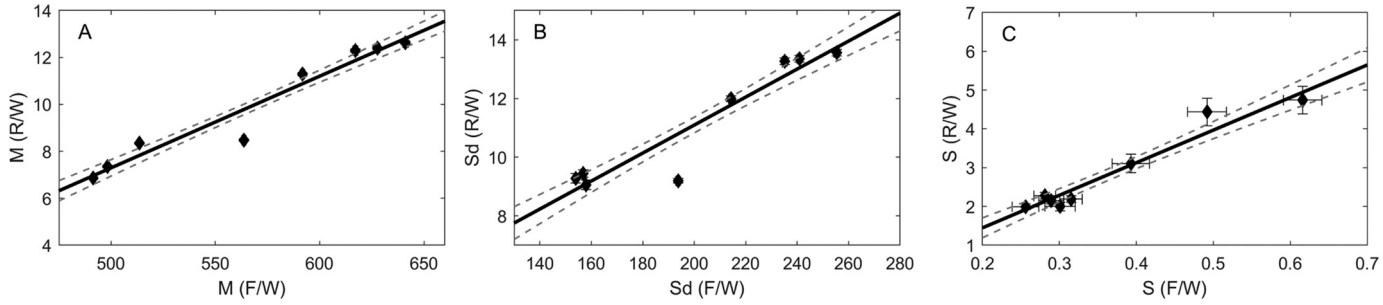

Fig. 6. Scatter plots between (A) Mean (M), (B) Standard deviation (Sd), (C) Skewness (S) of the single-cell distributions of F/W values against M, Sd and S of the single-cell distributions of R/W values for various induction strengths, differing in IPTG concentration (0, 6.25, 50, 100, 200, 300, 500, and 1000 μM IPTG). The error bars of the points are the standard errors (red error bars as in Fig. 3D-F and Fig. 5A-C). In each plot, we obtained the best linear fit (black straight line) as described in Supplementary Methods, section 1.4. The dotted lines are the standard error of the fitted line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

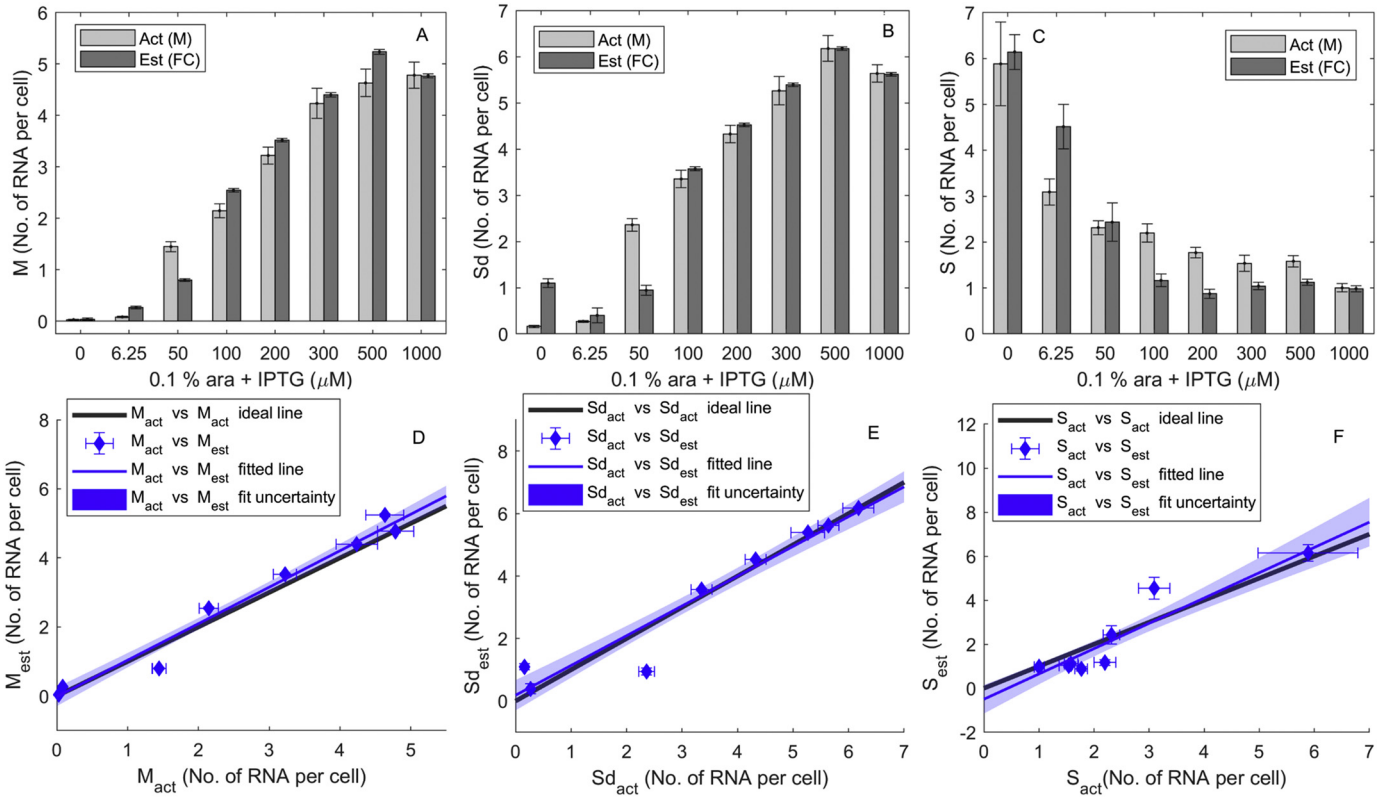

Fig. 7. (A) Mean single-cell RNA numbers estimated from Flow cytometry data, using microscopy data (Mean RNA numbers per cell) in the minimum (0 μM IPTG) and maximum induction (1000 μM IPTG) conditions for the calibration. (B) Standard deviation of single-cell RNA numbers estimated from Flow cytometry data, using the microscopy data (Variance of RNA numbers per cell) in 6.25 μM IPTG and maximum induction conditions (1000 μM IPTG) for the calibration. (C) Skewness of single-cell RNA numbers estimated from Flow cytometry, using the microscopy data (3rd moment of RNA numbers per cell) in 50 and 1000 μM IPTG for the calibration. Light grey bars are the actual values obtained from microscopy data and dark grey bars are the estimated values from flow cytometry data (F/W). (D) Scatter plot between estimated and actual mean values of single-cell RNA numbers. The blue points along with their standard error bars are the estimated mean of single-cell RNA numbers (M_{est}), plotted against the corresponding actual values (M_{act}). Also shown is the best linear fit to the blue points (blue line) along with the uncertainty of the fit (blue area). Finally, it is shown the 'ideal' linear fit (black line). The black line crosses 0 at the y-axis and has an inclination of 1, which would correspond to the estimated values being identical to the actual values. (E) Scatter plot between estimated and actual standard deviations of single-cell RNA numbers. The blue points along with their standard error bars are the estimated standard deviations of singe-cell RNA numbers (Sd_{est}), plotted against the corresponding actual values (Sd_{act}). Also shown is the best linear fit to the blue points (blue line) along with the uncertainty of the fit (blue area). Finally, it is shown the 'ideal' linear fit (black line). The black line crosses 0 at the y-axis and has an inclination of 1, which would correspond to the estimated values being identical to the actual values. (F) Scatter plot between estimated and actual skewness of single-cell RNA numbers. The blue points along with their standard error bars are the estimated skewness of single-cell RNA numbers (S_{est}), plotted against the corresponding actual values (S_{act}). Also shown is the best linear fit to the blue points (blue line) along with the uncertainty of the fit (blue area). Finally, it is shown the 'ideal' linear fit (black line). The black line crosses 0 at the y-axis and has an inclination of 1, which would correspond to the estimated values being identical to the actual values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Next, we found that using the 33% points with smaller Euclidean distance to the Median point, one obtains consistent calibration lines for the mean, variance, and third moment. These lines are shown, respectively, in Supplementary Fig. S9D-S9F. As expected, these set of consistent lines correspond to using pairs of data point that differ significantly between them in the Fig. 4A-4C (e.g. the pair of conditions 0 and 1000 μM IPTG).

Next, using these calibration lines (see Supplementary Section 1.7), we estimated the mean, standard deviation and skewness (along with their standard errors) of the single-cell distribution of RNA numbers from the distribution of F/W values. Fig. 7A shows the estimated mean of the single-cell distribution of RNA numbers from flow cytometry data, for each condition, using the calibration line obtained by using the pair of conditions 0 μM IPTG and 1000 μM IPTG. Fig. 7B shows the estimated standard deviation of the single-cell distribution of RNA numbers from flow cytometry data, for each induction level, using the calibration line obtained using the pair of conditions 6.25 μM IPTG and 1000 μM IPTG. Fig. 7C shows the estimated skewness of the single-cell distribution of RNA numbers from flow cytometry data, for each induction level, using the calibration line obtained using the pair of conditions 50 μM IPTG and 1000 μM IPTG.

To evaluate the accuracy of the estimated mean, standard deviation, and skewness from flow cytometry data, we plotted them against the corresponding actual values, obtained by microscopy (Fig. 7D-F). If the estimations are accurate, one expects the best-fitting line to these points (black lines in Fig. 7D-F) to exhibit a 45-degree inclination and to intercept the y-axis at zero. To test this, we plotted also the 'ideal line' (black lines in Figs. 7D-F). Next, we compared by analysis of covariance (McDonald, 2009) whether the best fitting line and the ideal line could be distinguished in slope and intercept, in a statistical sense. Results of these tests for the mean, standard deviation, and skewness (Supplementary Table S3) show that the best fitting line cannot be distinguished from the ideal line, from which we conclude that the estimations are accurate.

Given the above, we conclude that collecting data using microscopy from two conditions differing in RNA numbers, allows accurate estimations of the mean, standard deviation, and skewness of single cell distributions of RNA numbers from the distribution of total cell fluorescence measured byflow cytometry in multiple conditions differing in induction strength, using MS2d-GFP tagging of RNA.

Next, as in Section 3.2, we tested for the robustness of these estimations by adding increasingly higher Gaussian noise to the empirical F/W values. Results of these estimations using the noise corrupted F/W values are shown in Supplementary Fig. S11. Visibly, the estimations of the mean and standard deviation are not significantly affected. Meanwhile, the estimations of skewness are only affected for the 3 lowest induction conditions, similar to the results in Section 3.2, for similar reasons.

It is noted that the added Gaussian noise is much above what we expect to observe in real data collected from cells with the MS2d-GFP technology. I.e., the highest artificially added noise is much higher than the observed noise at the lowest induction conditions. Specifically, e.g., in the case at 6.25 μM IPTG induction, the observed standard deviation is ~155, while we added up to $\sigma = 400$ artificial Gaussian noise.

4. Discussion

Presently, FISH and MS2d-GFP RNA tagging are two of the preferred technologies for visualizing and quantifying RNA numbers in individual cells (Raj and van Oudenaarden, 2009). While the latter is likely more intrusive, it has some advantages, such as allowing to track the dynamics of RNA production in live cells, which has been used to dissect the underlying kinetic steps of transcription initiation, not possible otherwise (Lloyd-Price et al., 2016). So far, the use of MS2d-GFP RNA tagging has required microscopy and subsequent image analysis, which heavily limits the amount of data that can be produced. Further, image analysis introduces many errors (even with manual corrections). The ability to extract information using this technique from flow cytometry would overcome both limitations.

We have shown that it is possible to perform flow cytometry of cells expressing MS2d-GFP and RNA targets for MS2d-GFP and accurately estimate the mean, standard deviation, and skewness of the single-cell distribution of RNA numbers. Importantly, we have shown that the results cannot be distinguished, in a statistical sense, from those obtained by microscopy followed by manually corrected image analysis. Also, we have shown (Fig. 4) that the estimations of integer valued RNA numbers in individual cells are highly correlated with single-cell fluorescent protein levels, which is strong evidence of the accuracy of the estimations.

Interestingly, the estimations of mean single-cell RNA numbers from flow-cytometry data only exhibit significant discrepancy with the microscopy data when RNA production is weaker (Fig. 7). Past studies using microscopy and image analysis of cells with MS2d-GFP tagged RNAs (Häkkinen et al., 2014; Häkkinen and Ribeiro, 2015, 2016) suggest that these discrepancies arise mostly from errors in the microscopy data, which is based on ~500 cells per condition. In comparison, flow-cytometry data is based on ~40,000 cells (each of which randomly collected from a well-stirred medium). Consequently, the microscopy data is more prone to errors due to small sample size, particularly in weak expression conditions, where it is harder to select images of cells that are good representatives of the population. Nevertheless, regarding the estimations from flow-cytometry data, it is worth noting the unexpected value for the standard deviation at 0 μM IPTG (Fig. 7B), likely due to random biological variability.

In general, our results indicate that estimations of the statistics of single-cell RNA numbers can largely be performed from flow cytometry data and then be complemented by microscopy measurements (with scaling only requiring population images in two conditions differing in mean RNA numbers per cell). The large number of cells that can be observed by flow-cytometry promises precise estimations these statistics. Namely, we note that the estimations of single-cell RNA number statistics performed here are accurate not only in what concerns mean and standard deviation, but also skewness, which is in itself evidence of the accuracy of the estimations. Relevantly, this ensures that this technique can be used to estimate the propensity of a specific transcription kinetics to overcome thresholds in RNA and protein numbers (which is of significance in the context of small genetic circuits, among other). Finally, it is worth noting that, in principle, the method is readily applicable to cells with fluorescently tagged RNA using FISH technology. In this case, the methodology is expected to contribute in decreasing the effects of noise due to auto fluorescence from natural cellular components.

Overall, we expect the methodology proposed here to be useful in studies of in vivo transcription at single-molecule level, by adding more reliability to the conclusions, as these will be based on larger number of cells (by 2 to 3 orders of magnitude when compared to when collecting data by microscopy and image analysis). Also, much more conditions can be tested, due to the incomparably faster speed by which results can be obtained, compared to when using microscopy and image analysis.

Funding

Work supported by Tampere University Graduate Program (Finland) [to M.M.B. and V.C.]; Finnish Academy of Science and Letters [to C.P.]; Pirkanmaa Regional Fund [to V.K.K.]; Academy of Finland [295027 to A.S.R.]; and Jane and Aatos Erkko Foundation [610536 to A.S.R.]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mimet.2019.105745.

References

Abramoff, M.D., Magalhaes, P.J., Ram, S.J., 2004. Image Processing with ImageJ. Biophotonics International 11, 36–42.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Molecular Biology of the Cell. Garland Science, New York.

Arrigucci, R., Bushkin, Y., Radford, F., Lakehal, K., Vir, P., Pine, R., Martin, D., Sugarman, J., Zhao, Y., Yap, G.S., Lardizabal, A.A., Tyagi, S., Gennaro, M.L., 2017. FISH-flow, a protocol for the concurrent detection of mRNA and protein in single cells using fluorescence in situ hybridization and flow cytometry. Nat. Protoc. 12, 1245–1260. https://doi.org/10.1038/nprot.2017.039.

Bumgarner, R., 2013. DNA microarrays: types, applications, and their future. Curr. Protoc. Mol. Biol. Chapter 22, Unit 22, 1. https://doi.org/10.1002/0471142727.mb2201s101.

Bunka, D.H.J., Stockley, P.G., 2006. Aptamers come of age – at last. Nat. Rev. Microbiol. 4, 588–596. https://doi.org/10.1038/nrmicro1458.

Bushkin, Y., Radford, F., Pine, R., Lardizabal, A., Mangura, B.T., Gennaro, M.L., Tyagi, S., 2015. Profiling T cell activation using single molecule-FISH and flow cytometry. J. Immunol. 194, 836–841. https://doi.org/10.4049/jimmunol.1401515.

Carpenter, J., Bithell, J., 2000. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat. Med. 19, 1141–1164. https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F.

Chong, S., Chen, C., Ge, H., Xie, X.S., 2014. Mechanism of transcriptional bursting in bacteria. Cell 158, 314–326. https://doi.org/10.1016/j.cell.2014.05.038.

Cunningham, A., 1990. Fluorescence pulse shape as a morphological indicator in the analysis of colonial microalgae by flow cytometry. J. Microbiol. Methods 11, 27–36. https://doi.org/10.1016/0167-7012(90)90044-7.

DiCiccio, T.J., Efron, B., 1996. Bootstrap confidence intervals. Stat. Sci. 11, 189–228. https://doi.org/10.1214/ss/1032280214.

Dolgosheina, E.V., Jeng, S.C.Y., Panchapakesan, S.S.S., Cojocaru, R., Chen, P.S.K., Wilson, P.D., Hawkins, N., Wiggins, P.A., Unrau, P.J., 2014. RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. ACS Chem. Biol. 9, 2412–2420. https://doi.org/10.1021/cb500499x.

Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. Science 297, 1183–1186. https://doi.org/10.1126/science.1070919.

Engl, C., 2018. Noise in bacterial gene expression. Biochem. Soc. Trans. 47, 209–217. https://doi.org/10.1042/bst20180500.

Femino, A.M., Fay, F.S., Fogarty, K., Singer, R.H., 1998. Visualization of single RNA transcripts in situ. Science 280, 585–590. https://doi.org/10.1126/science.280.5363.585.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H., Bertrand, E., 2003. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. Curr. Biol. 13, 161–167. https://doi.org/10.1016/S0960-

9822(02)01436-7.
Golding, I., Cox, E.C., 2004. RNA dynamics in live Escherichia coli cells. Proc. Natl. Acad. Sci. 101, 11310–11315. https://doi.org/10.1073/pnas.0404443101.
Golding, I., Paulsson, J., Zawilski, S.M., Cox, E.C., 2005. Real-time kinetics of gene activity in individual bacteria. Cell 123, 1025–1036. https://doi.org/10.1016/j.cell.2005.09.031.
Häkkinen, A., Ribeiro, A.S., 2015. Estimation of GFP-tagged RNA numbers from temporal fluorescence intensity data. Bioinformatics 31, 69–75. https://doi.org/10.1093/bioinformatics/btu592.
Häkkinen, A., Ribeiro, A.S., 2016. Characterizing rate limiting steps in transcription from RNA production times in live cells. Bioinformatics 32, 1346–1352. https://doi.org/10.1093/bioinformatics/btv744.
Häkkinen, A., Muthukrishnan, A.B., Mora, A., Fonseca, J.M., Ribeiro, A.S., 2013. CellAging: a tool to study segregation and partitioning in division in cell lineages of Escherichia coli. Bioinformatics 29, 1708–1709. https://doi.org/10.1093/bioinformatics/btt194.
Häkkinen, A., Kandhavelu, M., Garasto, S., Ribeiro, A.S., 2014. Estimation of fluorescence-tagged RNA numbers from spot intensities. Bioinformatics 30, 1146–1153. https://doi.org/10.1093/bioinformatics/btt766.
Higuchi, R., Fockler, C., Dollinger, G., Watson, R., 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Nat. Biotechnol. 11, 1026–1030. https://doi.org/10.1038/nbt0993-1026.
Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., Lönnerberg, P., Linnarsson, S., 2014. Quantitative single-cell RNA-seq with unique molecular identifiers. Nat. Methods 11, 163–166. https://doi.org/10.1038/nmeth.2772.
Kærn, M., Elston, T.C., Blake, W.J., Collins, J.J., 2005. Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. 6, 451–464. https://doi.org/10.1038/nrg1615.
Kandavalli, V.K., Tran, H., Ribeiro, A.S., 2016. Effects of σ factor competition are promoter initiation kinetics dependent. Biochim. Biophys. Acta - Gene Regul. Mech. 1859, 1281–1288. https://doi.org/10.1016/j.bbagrm.2016.07.011.
Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., Taipale, J., 2012. Counting absolute numbers of molecules using unique molecular identifiers. Nat. Methods 9, 72–74. https://doi.org/10.1038/nmeth.1778.
Lloyd-Price, J., Startceva, S., Kandavalli, V., Chandraseelan, J.G., Goncalves, N., Oliveira, S.M.D., Häkkinen, A., Ribeiro, A.S., 2016. Dissecting the stochastic transcription initiation process in live Escherichia coli. DNA Res. 23, 203–214. https://doi.org/10.1093/dnares/dsw009.
Martins, L., Neeli-Venkata, R., Oliveira, S.M.D., Häkkinen, A., Ribeiro, A.S., Fonseca, J.M., 2018. SCIP: a single-cell image processor toolbox. Bioinformatics 34, 4318–4320. https://doi.org/10.1093/bioinformatics/bty505.
McDonald, J.H., 2009. Handbook of Biological Statistics, 2nd ed. Sparky House Publishing, Baltimore, MD.
Oliveira, S.M.D., Häkkinen, A., Lloyd-Price, J., Tran, H., Kandavalli, V., Ribeiro, A.S., 2016. Temperature-dependent model of multi-step transcription initiation in Escherichia coli based on live single-cell measurements. PLoS Comput. Biol. 12, 1–18. https://doi.org/10.1371/journal.pcbi.1005174.
Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., Van Oudenaarden, A., 2002. Regulation of noise in the expression of a single gene. Nat. Genet. 31, 69–73. https://doi.org/10.1038/ng869.
Peabody, D.S., 1993. The RNA binding site of bacteriophage MS2 coat protein. EMBO J. 12, 595–600.
Pedraza, J.M., Van, O., 2005. A. Noise propagation in genetic networks. Science 307, 1965–1969. https://doi.org/10.1126/science.1109090.
Raj, A., van Oudenaarden, A., 2009. Single-molecule approaches to stochastic gene expression. Annu. Rev. Biophys. 38, 255–270. https://doi.org/10.1146/annurev.biophys.37.032807.125928.
Raj, A., Van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., Tyagi, S., 2008. Imaging individual mRNA molecules using sets of singly labeled probes. Nat. Methods 5, 877–879. https://doi.org/10.1038/nmeth.1253.
Razo-Mejia, M., Barnes, S.L., Belliveau, N.M., Chure, G., Einav, T., Lewis, M., Phillips, R., 2018. Tuning transcriptional regulation through signaling: a predictive theory of allosteric induction. Cell Syst. 6, 456–469.e10. https://doi.org/10.1016/j.cels.2018.02.004.
Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230, 1350–1354. https://doi.org/10.1126/science.2999980.
Singer, R.H., Ward, D.C., 1982. Actin gene expression visualized in chicken muscle tissue culture by using in situ hybridization with a biotinated nucleotide analog. Proc. Natl. Acad. Sci. U. S. A. 79, 7331–7335. https://doi.org/10.1073/pnas.79.23.7331.
Startceva, S., Kandavalli, V.K., Visa, A., Ribeiro, A.S., 2019. Regulation of asymmetries in the kinetics and protein numbers of bacterial gene expression. Biochim. Biophys. Acta - Gene Regul. Mech. 1862, 119–128. https://doi.org/10.1016/j.bbagrm.2018.12.005.
Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., Lao, K., Surani, M.A., 2009. mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods 6, 377–382. https://doi.org/10.1038/nmeth.1315.
Tiberi, S., Walsh, M., Cavallaro, M., Hebenstreit, D., Finkenstädt, B., 2018. Bayesian inference on stochastic gene transcription from flow cytometry data. Bioinformatics 34, 647–655. https://doi.org/10.1093/bioinformatics/bty568.
Traganos, F., 1984. Flow cytometry: principles and applications. I. Cancer investigations. Cancer Investig. 2, 149–163.
Tran, H., Oliveira, S.M.D., Goncalves, N., Ribeiro, A.S., 2015. Kinetics of the cellular intake of a gene expression inducer at high concentrations. Mol. BioSyst. 11, 2579–2587. https://doi.org/10.1039/c5mb00244c.
Trcek, T., Chao, J.A., Larson, D.R., Park, H.Y., Zenklusen, D., Shenoy, S.M., Singer, R.H., 2012. Single-mRNA counting using fluorescent in situ hybridization in budding yeast. Nat. Protoc. 7, 408–419. https://doi.org/10.1038/nprot.2011.451.
Yu, J., Xiao, J., Ren, X., Lao, K., Xie, X.S., 2006. Probing gene expression in live cells, one protein molecule at a time. Science 311, 1600–1603. https://doi.org/10.1126/science.1119623.