Methods 50 (2010) 282–288

Contents lists available at ScienceDirect

Methods

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

Review Article

Single-cell gene expression profiling using reverse transcription quantitative real-time PCR

Anders Ståhlberg [a,b,*], Martin Bengtsson [b,c,1]

[a] Lundberg Laboratory for Cancer, Department of Pathology, Sahlgrenska Academy at University of Gothenburg, Gula Straket 8, 413 45 Gothenburg, Sweden
[b] TATAA Biocenter, Odinsgatan 28, 411 03 Gothenburg, Sweden
[c] Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK

ARTICLE INFO

Article history:
Accepted 7 January 2010
Available online 11 January 2010

Keywords:
Cell heterogeneity
Cell lysis
Gene expression
qPCR
Real-time PCR
Reverse transcription
RT-qPCR
Single cell
Single-cell biology
Single-cell gene expression

ABSTRACT

Even in an apparently homogeneous population of cells there are considerable differences between individual cells. A response to a stimulus of a cell population or tissue may be consistent and gradual while the single-cell response might be binary and apparently irregular. The origin of this variability may be preprogrammed or stochastic and a study of this phenomenon will require quantitative measurements of individual cells. Here, we describe a method to collect dispersed single cells either by glass capillaries or flow cytometry, followed by quantitative mRNA profiling using reverse transcription and real-time PCR. We present a single cell lysis protocol and optimized priming conditions for reverse transcription. The large cell-to-cell variability in single-cell gene expression measurements excludes it from standard data analysis. Correlation studies can be used to find common regulatory elements that are indistinguishable at the population level. Single-cell gene expression profiling has the potential to become common practice in many laboratories and a powerful research tool for deeper understanding of molecular mechanisms.

1. Introduction

Cells have a remarkable ability to cooperate and jointly construct complex structures such as tissues, organs and whole organisms. These constructions are normally accurately tuned and respond to stimuli with high precision. During development, cells differentiate to specialized cell types, each with particular functions in the environment they reside. In many aspects, individual cells exhibit a high degree of variability and responses to identical stimuli may be very different even in a seemingly homogeneous population [1–6].

Gene expression profiling is a pivotal research tool in molecular biology. By default, measurements are made on large pools of cells as this will increase reliability of the recordings. For tissues, different cell types are mixed uncontrollably and the measured gene expression profile has unknown contribution from different cell types. In addition, cell population measurement will not reveal how a particular transcript is distributed among the cells (Fig. 1A–B). Bulk measurements easily miss potentially important gene correlations (Fig. 1C–D) where single cell analysis would indicate coupled transcriptional regulations, which might be controlled by the same molecular mechanism (Fig. 1C–D) [7]. Observed heterogeneity may indicate the presence of specialized cell types or originate in the random nature of the transcription machinery [1–6].

A typical single cell contains ~1 pg mRNA, which is equivalent to a few hundred thousand molecules transcribed from about ten thousand genes [8]. The high sensitivity of reverse transcription quantitative real-time PCR (RT-qPCR) makes it possible to detect even a single molecule. RT-qPCR is also characterized by high reproducibility and wide dynamic range [9–11]. These properties make RT-qPCR suitable for single-cell gene expression profiling. Even if single-cell gene expression profiling using RT-qPCR has been successfully applied to several different applications [6,12–14], it has still not become common practice for laboratories. In this paper, we describe the workflow of single cell RT-qPCR including: cell collection, cell lysis, RT, qPCR and data analysis.

2. Description of method

Single cell RT-qPCR constitutes several sequential steps, outlined in Fig. 2. Our intention with this paper is to present the most common experimental approaches with suitable references and in

* Corresponding author. Address: Lundberg Laboratory for Cancer, Department of Pathology, Sahlgrenska Academy at University of Gothenburg, Gula Straket 8, 413 45 Gothenburg, Sweden. Fax: +46 31828733.
E-mail address: anders.stahlberg@neuro.gu.se (A. Ståhlberg).
[1] Present address: Nuevolution A/S, Ronnegade 8, 2100 Copenhagen, Denmark.

1046-2023/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2010.01.002

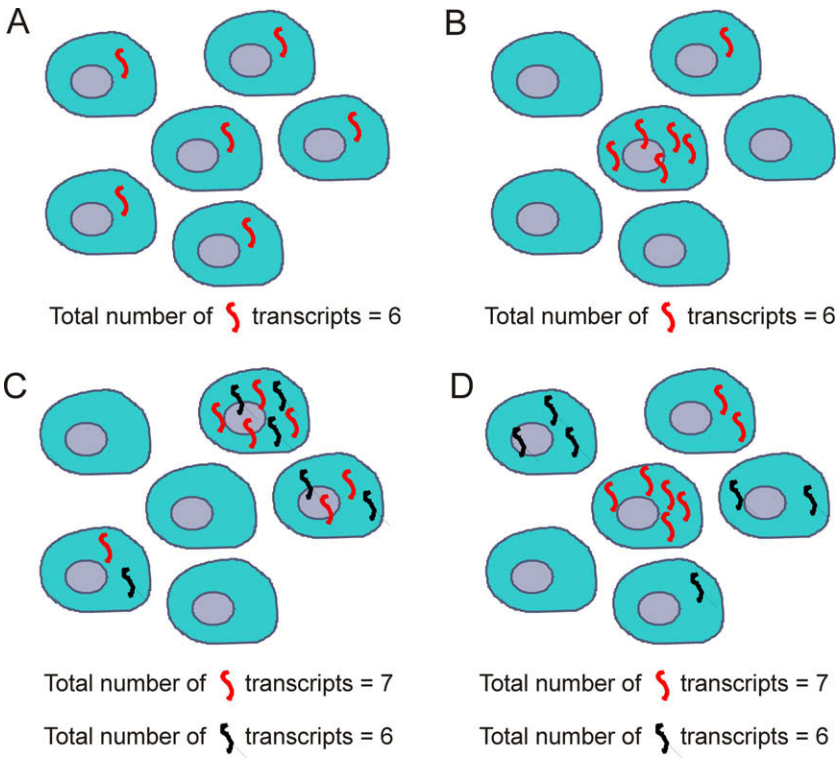

Fig. 1. Cell heterogeneity and correlated transcript levels. Single cell measurements can distinguish the cases in (A) and (C) from (B) and (D), respectively, while cell population measurements cannot.

detail describe the most appropriate experimental setup for analysis of cells in suspension. For successful single-cell gene expression profiling good laboratory practice is essential. All RNA work requires completely RNase free conditions. Furthermore, PCR contamination must be avoided, since even negative (zero values) are used in data analysis. Physical separation, i.e. different rooms, of pre-PCR, PCR and post-PCR is recommended.

3. Single cell collection and lysis

A majority of single-cell studies published deploy one of three methods to collect cells: Flow cytometry, glass capillaries, and laser capture. This review will describe the two former while laser capture and laser microdissection are described in detail elsewhere [15–17]. For some applications, the origin and surroundings of a cell is vital information and this excludes flow cytometry as collection method, leaving laser capture as best option. Where high numbers of cells are needed, and where collections of intact cells are of importance, flow cytometry is recommended. Glass capillaries, as we use them in this article, will also collect intact cells.

Single cell suspensions are prepared from tissue using mechanical separation, enzymatic treatment, and/or non-physiological buffers. The yield of functionally viable, dissociated cells will be dependent on several parameters such as concentrations, incubation times and temperatures. An informative website for different cell dissociation approaches is: www.tissuedissociation.com. As the expression of some genes may be altered by the cell treatment it is recommended that the expression of the genes of interest is quantified also in an untreated sample. For example, one part of the biological sample is saved for total RNA isolation and the remaining is used for cell dissociation and collection.

3.1. Single cell sorting using flow cytometry

Several flow cytometry instruments, such as FACSDiva and FACSVantage (both BD Biosciences, San Jose, CA, USA) can be used to sort out individual cells [18]. PCR plates (96-well) with lysis buffer should be prepared in advance. In addition to standard flow cytometry calibration, the instrument needs to be carefully calibrated to deposit single cells in the center of each collection tube. This can easily be tested by sorting ~50 beads/cells on the plastic

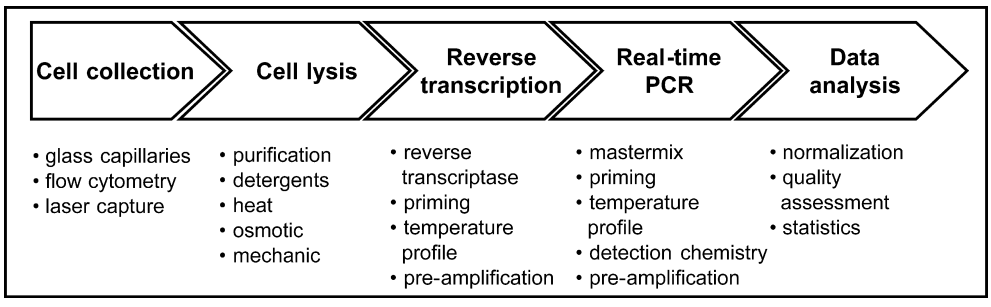

Fig. 2. Overview of single-cell gene expression profiling using RT-qPCR.

film covering the 96-well PCR plate. It can also be worthwhile to check the calibration every second plate, because the sorting arm may be displaced over time. A small volume of lysis buffer will increase the risk of a cell sticking to the wall of the tube, while too large volume will likely interfere with downstream reactions. We have found that ~5 μl lysis buffer is a suitable volume to work with using flow cytometry.

For practical reason, a significant number of cells are needed for calibration, and thousands of cells are wasted by the flow cytometry instrument. For most cell types it is recommended to keep the cells on ice before sorting. If specific antibodies are used to sort out subpopulations, cells need to be kept in medium that do not interfere with the fluorophores used (PBS with ~2.5% serum is usually compatible with both the fluorescence-activated cell sorting procedure and keeping the cells viable). If no fluorophores are used, cells should be kept on cell type-specific medium to avoid altered gene expression. Viable cells can be enriched by flow cytometry by adding 7-aminoactinomycin D (Sigma–Aldrich) before sorting. As a positive control 1–2 wells may be used to collect ~30 cells. This number of cells will level out most single cell variability, and if these controls result in negative expression data, reliable single cell data for that particular gene will most likely be hard to obtain. After finished sorting, each 96-well PCR plate should be sealed and put on dry ice and stored at −80 °C until RT. As a control for contamination, some liquid without cells can be collected from the flow cytometry instrument to eliminate the chance that mRNAs from lysed cells have contaminated the collected cells.

3.2. Single cell collection using glass capillaries

This collection method requires a phase contrast microscope, preferably inverted, mounted with micromanipulators [6,19]. Shortly prior cell collection, the dissociated cells should be re-plated on Petri dishes, allowing cell adhesion. Usually, several dishes can be prepared and maintained in optimal conditions until single cell collection. Depending on the use of fluorophores, dissociated cells should be kept on cell specific medium or medium which do not interfere with the fluorophores used. Cell collection is simplified if the cells have attached slightly to the dish and are physically separated from each other. Cells may be washed before collection to remove dead cells and loose debris that could contaminate the pipette.

Borosilicate glass capillaries (Hilgenberg, GmbH) with outer diameter of 1.5–1.6 mm and wall thickness of 0.16 mm can be pulled to pipettes using a patch-clamp pipette puller (Heka PIP5). The diameter of the tip should be ~10 μm, but the width is adjusted to cell size. This is substantially wider than standard patch-clamp pipettes and large enough to allow passage of an intact cell ensuring collection of all mRNA. An alternative method is to penetrate the cell membrane with a patch-clamp pipette (diam. ~0.5 μm) and only collect the cytoplasm [20,21]. This has the advantage of allowing mRNA harvest from intact tissue, and in conjunction with patch-clamp recording. However, it is technically very challenging which limits the number of samples that can be collected within reasonable time limits. In addition, normalization of expression levels poses a particular difficulty when an unknown fraction of the cell's mRNA is harvested.

The glass pipette is mounted on a hydraulic micromanipulator (Narishige or equivalent) on the inverted microscope and connected by a flexible plastic tubing to the mouth of the researcher. If the pipette size is right it is readily feasible to control the pressure inside the pipette and collect individual cells with minimum volume of extracellular solution ($\ll$0.1 μl). Pipettes should be emptied in standard PCR tubes containing 2 μl lysis solution. This small volume allows high concentration of detergents needed for cell lysis and RNase inhibition. To simplify the process of emptying the glass capillary, a stand with tube holder and an adjustable capillary mount is recommended (Fig. 3). Careful breaking of the glass tip to the bottom of the tube facilitates complete emptying of the capillary. We have not detected any inhibition of RT or difference in transcript number for cells with or without small traces of borosilicate glass (data not shown). PCR tubes containing single cells should be stored on ice during collection and kept in −80 °C until RT. As a control for contamination and inhibition of downstream reactions, ~1 μl of the buffer or medium surrounding the cells should be collected and analyzed together with the single-cell samples.

3.3. Cell lysis

The amount of protein, salts and debris in the single-cell samples is very small, allowing us to omit mRNA purification. We have

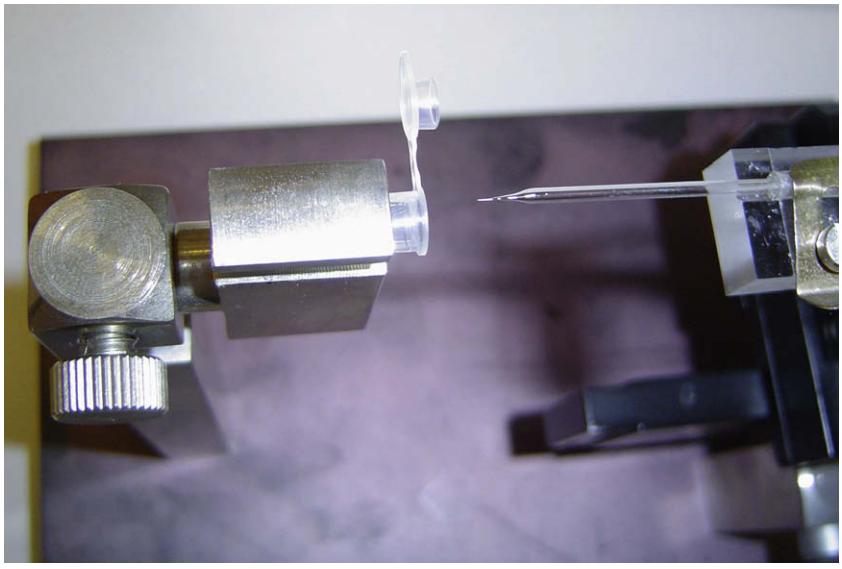

Fig. 3. Glass capillary emptying. The tip of the glass capillary contains a single cell that is emptied into the PCR tube containing 2 μl lysis buffer. To speed up and simplify the procedure, our local workshop constructed a stand for the tube and a moveable capillary holder allowing precise emptying.

pooled up to ten cells, each collected individually, and not observed any inhibitor effects on reverse transcription [19]. However, for efficient and complete RT we need to ascertain that the cell is lysed and that all available mRNA is accessible to the enzyme. Cell lysis can be performed with heat, mechanic forces, osmotic pressure, enzymatic treatment and detergents. In addition to efficient cell lysis, RNase activity needs to be blocked. RNase activity can be inhibited by RNase inhibitors, such as RNaseOUT (20 U, Invitrogen), and chaotropic salts. For mammalian cells, especially mouse β-cells, we have found that the use of guanidine thiocyanate is efficient to lyse cells. In addition, low concentrations (~50 mM) of guanidine thiocyanate is compatible with RT, for some genes it may even enhance the RT efficiency and act as an RNase inhibitor [19]. The maximum concentration of guanidine thiocyanate that may be used in the lysis buffer depends on the dilution factor between lysis and reverse transcription. For mouse β-cells, 2 μl lysis buffer containing 250 mM guanidine thiocyanate is sufficient for efficient lysis and also compatible with 10 μl RT reactions. RNA/DNA carrier may be added to the lysis buffer, but we have not found any positive effects from the use of carrier in our experimental setup. We have also evaluated heat incubation during lysis (80 °C, 5 min) but it does not significantly affect mRNA recovery (data not shown). Commercial single cell lysis buffers, such as CellsDirect (Invitrogen) and CelluLyser (TATAA Biocenter), may also be used without additional extraction or purification.

3.4. RNA spike

As a control for the entire workflow from single cell collection to data analysis, an RNA spike may be added in the lysis buffer.

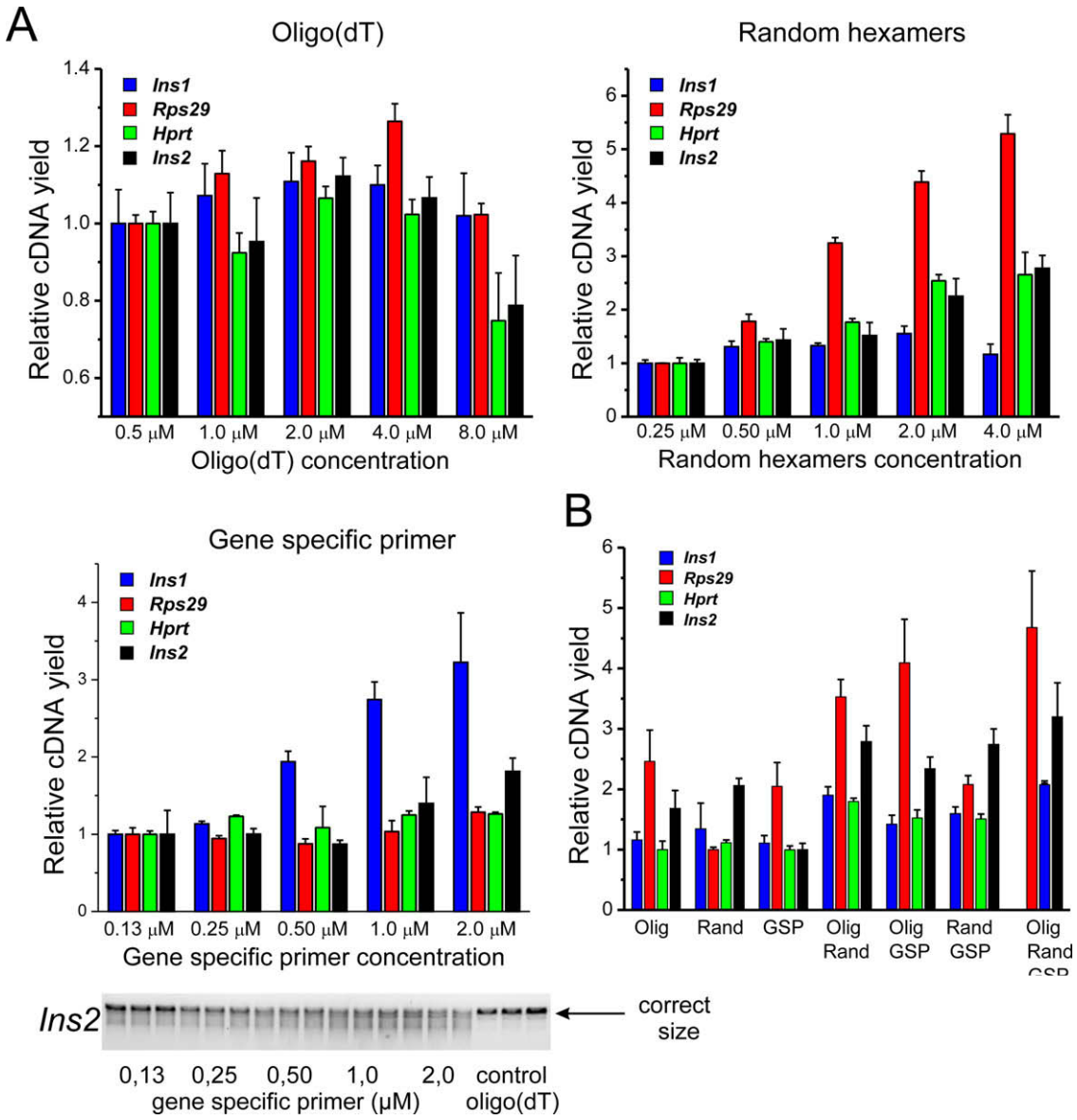

Fig. 4. RT priming. Identical amounts of purified total RNA was used as starting material and four genes were measured (*Ins1*, *Ins2*, *Rps29* and *Hprt*). (A) Determination of optimal RT primer concentration using oligo(dT), random hexamers and gene specific primer. The gene specific primer was identical to reverse qPCR primer. Agarose gel electrophoresis analysis for *Ins2* revealed that increasing concentration of gene specific primer resulted in formation of erroneous PCR products. Melting curve analysis could not distinguish between correct and erroneous PCR products. Unspecific PCR products were observed for *Ins2*, *Hprt* and *Rps29* using gene-specific primers. (B) Relative RT reaction yields are shown for various primer combinations. 2.0 μM oligo(dT), 2.0 μM random hexamers and 0.25 μM GSP was used for all RT primer combinations. Agarose gel inspection showed formation of erroneous PCR products for *Hprt* and *Ins2* using gene specific primer. The *Ins1* assay was specific for all conditions while *Rps29* occasionally formed erroneous PCR products using gene specific primer. Any combination of priming methods was invariably superior or equal to the single best priming method used. Relative cDNA yield was arbitrarily set to a value of one for the gene with lowest expression in the group of genes with the lowest RT primer concentration tested in (A) and in left group (Olig) in (B) [22]. Values are means ± SEM for 3 separate experiments. *Abbreviations used:* Olig, oligo(dT); Rand, random hexamer; GSP, gene specific primer.

However, the use of RNA spike will not reveal any information about cell quality, lysis efficiency, mRNA accessibility and mRNA quality. Setting up single-cell gene expression profiling for the first time or changing the experimental setup, RNA spikes may be useful to include as a control for RNA stability, RT and qPCR. From our experience, data from the RNA spike always results in almost identical Cq-values if all single cells are run in parallel. RNA spikes can either be commercial or generated by in vitro transcription [19]. The sequence of the RNA spike should be unique compared to the transcriptome of the single cells analyzed. Alternatively, the RNA spike should be used in high excess compared to the endogenous expression.

3.5. Reverse transcription

High RT efficiency is needed to ensure that low transcript numbers can be detected and quantified by qPCR. There are several different reverse transcriptases and priming methods available. We have previously shown that the variability in RT efficiency is highly gene dependent [22,23]. Therefore, we prefer to use a blend of oligo(dT) and random hexamers to prime the reverse transcription reaction (Fig. 4). Gene-specific primers can bind unspecifically to the mRNA [22] leading to cDNA similar to that of random hexamer priming, but with the gene specific primer in the 3′ end. If the gene specific primer is identical to the reverse PCR primer, the specificity will be reduced in the PCR and the number of products unrelated to the intended target will increase. Use of gene specific primer identical to the reverse primer should therefore be carefully optimized before use or avoided completely. It may also be worthwhile to test different reverse transcriptases, as the efficiency varies greatly and gene-dependently [23].

For many genes we have seen that the use of SuperScript III (Invitrogen) gives sufficient cDNA yield. We frequently use the following reverse transcription protocol: A mix of a single cell in 2 μl lysis buffer (containing 50 mM guanidine thiocyanate, Sigma–Aldrich), 0.5 mM dNTP (Sigma–Aldrich), 2.0 μM oligo (dT_{15}, Eurofins MWG Operon), 2.0 μM random hexamers (Eurofins MWG Operon) is incubated at 65 °C for 5 min in a total volume 6.5 μl. dNTP, oligo(dT_{15}) and random hexamers can be added directly to the frozen cells without thawing the samples. Beware of batch differences of oligo(dT) and random hexamers; check new batches before using them. Then cool the samples on ice or cooling block to <25 °C. Add 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 5 mM dithiothreitol, 20 U RNaseOUT and 50 U SuperScript III (all Invitrogen) to total volume of 10 μl. The final concentration of guanidine thiocyanate in RT should be $\leqslant$50 mM. The following temperature profiles may be used for RT: 25 °C for 5 min, 50 °C for 60 min, 55 °C for 15 min and 70 °C for 15 min. Only minor differences have been observed using other incubation times and temperatures as long the reverse transcriptases have not been inactivated by heat. To increase the amount of mRNA, pre-amplification methods may be applied [24,25].

The quantitative contribution from genomic DNA to the expression signal is generally insignificant, as it would only increase the copy number by two for most genes. Nevertheless, PCR primers should whenever possible be designed to cross introns. DNase treatment is another option, but difficult to evaluate in single-cell samples. For +/− measurements however, both these precautions are required.

3.6. Quantitative real-time PCR

The number of genes that can be analyzed from a single cell is limited by the number of transcripts of the studied genes. We have shown that >∼20 target molecules per PCR are needed for accurate quantification [19]. Furthermore, we have shown that highest reproducibility for measurements of limited samples is achieved when the dilution between RT and qPCR is kept at a minimum, even at the expense of qPCR-replicates [19]. For example, the accuracy in qPCR will be higher if all cDNA is quantified in a single qPCR reaction instead of splitting the cDNA sample in two and analyze the duplicate. However, loading too much of the RT reaction may inhibit the qPCR, in part due to inhibitory effects of reverse transcriptase enzyme on PCR [19,21,26–28]. The inhibition depends on the amount of reverse transcriptase and Taq polymerase used in the two reactions. Using SuperScript III (Invitrogen) and JumpStart (Sigma–Aldrich) the ratio should be <2 (U/U). This ratio may be different using other enzymes. Practically, about 5–10 genes can successfully be analyzed from the same cell, a higher number would require pre-amplification of cDNA.

Pre-amplification of single-cell cDNA allows virtually limitless number of measured genes, and has been used on single cells in concert with microarrays [29]. In a pioneering work by Brady and Iscove [30], it was shown that polyA-cDNA can be globally amplified, followed by gene-specific PCR on the resulting pool. Alternatively, a multiplex PCR on the single-cell cDNA, using low concentrations of all gene-specific primers, results in a population of PCR products that are quantified by qPCR. This way, up to 40 [31] and 100s (TaqMan® PreAmp Master Mix, p/n 4391128, Applied Biosystems) of transcripts have been analyzed in a single cell. However, pre-amplification introduces one extra source of potential bias which needs to be verified. For more information on pre-amplification methods, the reader is referred elsewhere [24,25].

Normal qPCR guidelines are used to design working assays for single-cell gene expression profiling [9,32]. Assays should be gene specific, have high sensitivity, dynamic range and reproducibility [9,32]. For most real-time PCR instruments, typical quantification cycle (Cq) values are between 30 and 40 analyzing single cells. Many single cells will and should not result in any detectable PCR product (no Cq-value). Using melting curve analysis, non-specific PCR products may be identified. To eliminate detection and quantification bias, assays should not give rise to any primer-dimers, at least not before Cq = 40. From our experience is the assay specificity, i.e. lack of erroneous PCR products such as primer-dimers, one of the most important parameter when optimizing single cell assays. If non-specific PCR fragments are formed in addition to the correct PCR product, data should only be used qualitatively. Probe-based qPCR assays can be used but they will not reveal important information about interfering PCR products that are formed. Another alternative is to use one-step RT-qPCR assays. Running RT and qPCR in the same reaction tube eliminates one handling step and consequently reduces the risk of contamination. The main drawback with one-step RT-qPCR is that RT and qPCR most often have different optimal reaction conditions, reducing the sensitivity, efficiency and reproducibility of respective reactions. All assays should be validated by gel electrophoresis, even if probes are used.

The following qPCR protocol has been proven to work well: 10 μl reactions containing 10 mM Tris (pH 8.3), 50 mM KCl, 3 mM $MgCl_2$, 0.3 mM dNTP, 1 U JumpStart Taq polymerase (all Sigma–Aldrich), 0.5 × SYBR Green I (Invitrogen), 200–400 nM of each primer (Eurofins MWG Operon) and 1–4 μl cDNA (cDNA should be diluted to avoid qPCR inhibition from the RT reaction). Premade qPCR mixes can also be used. However, assay performance usually differs somewhat between qPCR mixes and should not be changed between runs for samples to be compared. Primers should be designed according to standard procedure [9,11]. Primer3 (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi.) is one freely available primer design software that may be used. For samples with low RNA quality short amplicon lengths are recommended [33,34]. The following temperature profile may be ap-

plied for most qPCR instruments: 95 °C for 3 min followed by 50 cycles of amplification (95 °C for 20 s, 60 °C for 20 s, and 72 °C for 20 s). The predenaturation step is to activate the antibody-inactivated Taq DNA polymerase and the cycling program is for primers designed with melting temperatures of 60 °C using Primer3 and amplicon lengths <250 bp.

3.7. Data analysis

Single-cell gene expression data may either be determined as relative quantities or as absolute mRNA/cDNA numbers. The efficiency of reverse transcription is gene dependent and <100% [22,23]. Hence, the number of cDNA molecules is a lower-limit estimate of the number of mRNA molecules present in the cell. The RT efficiency can be estimated experimentally by using known amounts of synthetic RNA mimicking the properties of the native mRNA. However, producing full length mRNAs is challenging and it is common to use a truncated RNA molecule, which may not be reverse transcribed with the same efficiency as the native mRNA. Alternatively, the RT-qPCR variability can be used to determine the reverse transcription efficiency using mathematical modeling [19]. The number of cDNA molecules can be determined using standards with known concentrations as described [9–11]. If double-stranded DNA standards are used, one cycle should be subtracted from each acquired Cq-value to compensate for the fact that cDNA is single stranded [22].

Fig. 5A shows representative Cq-values of two genes with medium (*Chgb*) and high (*Ins2*) expression in 37 individual mouse β-cells [19]. Fig. 5B shows the distribution of *Ins2*-expressing cells in linear-scale and Fig. 5C shows the distribution in $\log_{10}$-scale. It is a highly skewed distribution where ~40% of all transcripts originate from only three cells, resulting in a better fit for lognormal distribution than for normal distribution ($p = 0.63$ and $p < 0.01$, respectively, with a Shapiro–Wilk normality test where a high value corresponds to a good fit). In lognormal distributions, the typical cell is best characterized by the geometric mean. Extrapolation of cell population data to single cell data by dividing the total expression by the number of cells, i.e. arithmetic mean, will overestimate the transcript number for a typical single cell in a population.

Fig. 6 shows how the mRNA copy number of three hypothetical genes may vary over time. Genes 1 and 3 have correlated expression, while gene 2 is uncorrelated with genes 1 and 3. Gene expression in mammalian cells occurs in bursts [4,5,35]. The total amount of a specific mRNA at a given time point in a cell is determined by the frequency and duration of the bursts and the mRNA degradation. The observed distributions with lognormal features are consistent with the burst theory [5]. The nature and consequences of the stochastic behavior of gene expression have recently gained new insights but still this is a relatively new field of research [4].

Maybe the most critical step in single-cell gene expression profiling using RT-qPCR is to ensure that expression data between single cells are comparable. How can we compensate for differences in accessible mRNA content among individual cells, and how do we separate between experimental and biological variability in the mRNA pool of individual cells? Today, single cell data are reported as absolute mRNA/cDNA copies or as relative quantities without any normalization except to the number of cells (one). The use of RNA spike will only compensate for technical errors downstream of cell lysis. The expression of all genes changes over time, consequently the use of reference genes is not valid to normalize single cell data (Fig. 6). In addition, most actively expressed genes will at some time point be expressed below the detection limit of the RT-qPCR assay due to the stochastic behavior of gene expression. A strong correlation between the geometric mean and the frequency of cells with observed transcripts of a specific gene is usually observed. For example, single β-cells only occasionally lack *Ins2* (high expressed) expression, while *Chgb* (medium ex-

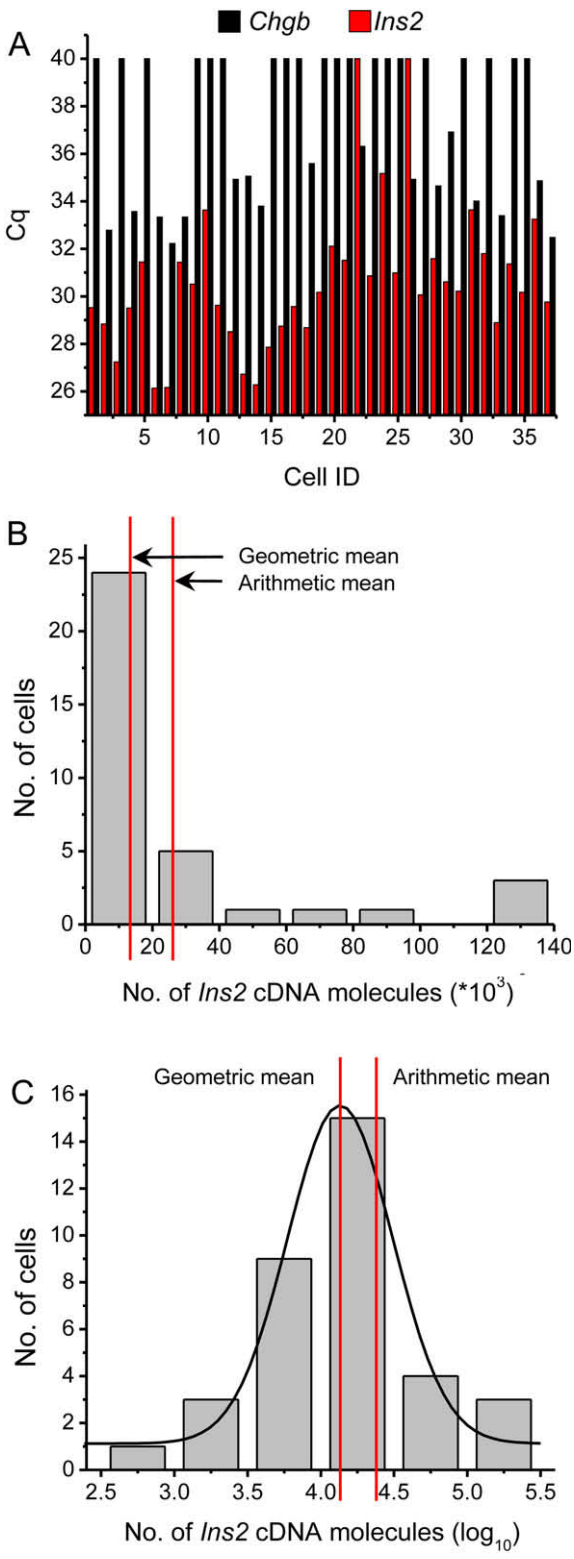

Fig. 5. Single cell gene expression data. (A) Cq-values of *Chgb* and *Ins2* from 37 single cells. Cells not expressing, respectively, gene was arbitrarily given a value of 40. (B) Distribution of *Ins2* expression shown in linear-scale. Cq-values of *Ins2* were converted to number of cDNA molecules using standard curves of known concentrations. (C) Distribution of *Ins2* expression shown in $\log_{10}$-scale. The geometric and arithmetic mean are shown in both linear- and $\log_{10}$-scale. Data have been published before [19].

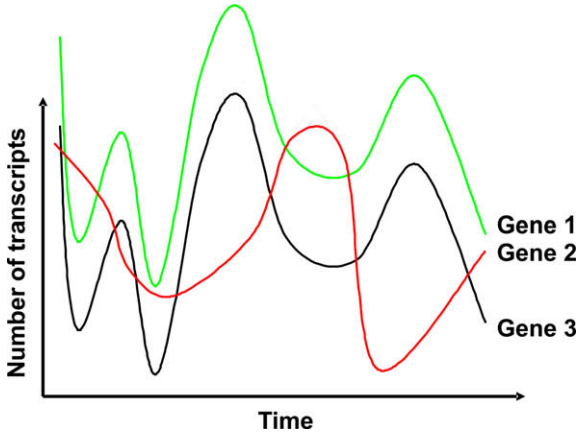

Fig. 6. mRNA expression over time. The transcription process is known to occur in bursts. The total amount of a specific mRNA is determined by the frequency and duration of transcriptional bursts and mRNA degradation. The green (Gene 1) and black (Gene 3) curves illustrate correlated mRNAs, while the red curve (Gene 3) represents an uncorrelated mRNA (for illustration only).

pressed) is only detectable in about 50% of the cells. (Fig. 5A) [19]. Another confounding effect when analyzing single cell data is that compared cells usually are not exactly in the same phase in the cell cycle. In many experimentally setups, cell cycle synchronization is impossible or even unwanted. Exactly how the cell cycle affects gene expression is not known. One interesting approach to test for normalizing single cell data would be to inject a specific mRNA molecule with known concentration into the cell using glass capillaries, followed by standard cell collection and measurement.

The fact that a minority of cells constitute the major part of the total number of transcripts for a given gene in a cell population results in that observed correlations at cell population level actually may occur in different cells (Fig. 1C–D). In a previous study [6], we found that *Ins1*, *Ins2* and *Actb* were all up-regulated in β-cells at cell population level when the glucose concentration was increased. Interestingly, single-cell correlations between the genes revealed that *Ins1* and *Ins2* expression increased in the same cells (high correlation), while *Actb* was up-regulated in other cells; i.e. *Actb* showed no single-cell correlation with *Ins1*/*Ins2*. The positive correlation between *Ins1* and *Ins2* and that they share almost identical promoter regions [36] suggest that they have synchronized transcriptional bursts (Fig. 6) [3,4,7]. Consequently, correlation at single cell level can be used as a tool to identify genes that truly share regulatory mechanisms, not just being affected by the same stimulus.

4. Concluding remarks

Single-cell gene expression profiling is a powerful tool to better understand molecular mechanisms. However, single-cell gene expression profiling is rarely applied, in part due to a lack of understanding of single-cell biology. The stochasticity of gene expression makes it difficult to separate technical variation from biologically relevant cell-to-cell variation. For example, some cells in a larger population are expected to be expressed below the detection limit, even for well-optimized RT-qPCR assays. Many laboratories have access to equipment for single cell collection and real-time PCR instruments. Due to the very few transcripts being measured, each step in the process from cell preparation to data analysis need to be carefully optimized and obviously follow good laboratory practice, but each step is otherwise identical as for cell population measurement. When single cell experiments are applied for the first time, it may be worthwhile to analyze a few highly expressed genes to get started, even if the genes are not of biological interest. One major advantage using RT-qPCR for single-cell gene expression profiling is that RT-qPCR for cell population measurement is standard procedure for many laboratories, which limits the cost, time and effort to perform the experiments. Analyzing individual cells opens up new doors for molecular biologists and the research field of single-cell biology is growing rapidly.

Acknowledgments

This work was partly supported from the Swedish Society for Medical Research, Socialstyrelsens Foundation, Sigurd and Elsa Goljes Minne, Foundation Blanceflor Boncompagni-Ludovisi née Bildt and Wilhelm and Martina Lundgren's Research Foundation.

References

[1] M.B. Elowitz, A.J. Levine, E.D. Siggia, P.S. Swain, Science 297 (2002) 1183–1186.
[2] W.J. Blake, M. Kaern, C.R. Cantor, J.J. Collins, Nature 422 (2003) 633–637.
[3] J.M. Raser, E.K. O'Shea, Science 309 (2005) 2010–2013.
[4] A. Raj, A. van Oudenaarden, Cell 135 (2008) 216–226.
[5] A. Raj, C.S. Peskin, D. Tranchina, D.Y. Vargas, S. Tyagi, PLoS Biol. 4 (2006) e309.
[6] M. Bengtsson, A. Ståhlberg, P. Rorsman, M. Kubista, Genome Res. 15 (2005) 1388–1392.
[7] D. Volfson, J. Marciniak, W.J. Blake, N. Ostroff, L.S. Tsimring, J. Hasty, Nature 439 (2006) 861–864.
[8] E.S. Kawasaki, Ann. NY Acad. Sci. (2004) 92–100.
[9] T. Nolan, R.E. Hands, S.A. Bustin, Nat. Protoc. 1 (2006) 1559–1582.
[10] M.W. Pfaffl, Nucleic Acids Res. 29 (2001) 2002–2007.
[11] S.A. Bustin, J. Mol. Endocrinol. 29 (2002) 23–29.
[12] L. Warren, D. Bryder, I.L. Weissman, S.R. Quake, Proc. Natl. Acad. Sci. USA 103 (2006) 17807–17812.
[13] J. Gründemann, F. Schlaudraff, O. Haeckel, B. Liss, Nucleic Acids Res. 36 (2008) e38.
[14] A. Diercks, H. Kostner, A. Ozinsky, PLoS One 4 (2009) e6326.
[15] V. Espina, J.D. Wulfkuhle, V.S. Calvert, A. van Meter, W. Zhou, G. Coukos, D.H. Geho, E.F. Petricoin III, L.A. Liotta, Nat. Protoc. 1 (2006) 586–603.
[16] K.M. Keays, G.P. Owens, A.M. Ritchie, D.H. Gilden, M.P. Burgoon, J. Immunol. Methods 302 (2005) 90–98.
[17] G.I. Murray, Acta Histochem. 109 (2007) 171–176.
[18] A. Ståhlberg, M. Bengtsson, M. Hemberg, H. Semb, Clin. Chem. 55 (2009) 2162–2170.
[19] M. Bengtsson, M. Hemberg, P. Rorsman, A. Ståhlberg, BMC Mol. Biol. 9 (2008) 63.
[20] B. Liss, O. Franz, S. Sewing, R. Bruns, H. Neuhoff, J. Roeper, EMBO J. 20 (2001) 5715–5724.
[21] B. Liss, Nucleic Acids Res. 30 (2002) e89.
[22] A. Ståhlberg, J. Håkansson, X. Xian, H. Semb, M. Kubista, Clin. Chem. 50 (2004) 509–515.
[23] A. Ståhlberg, M. Kubista, M.W. Pfaffl, Clin. Chem. 50 (2004) 1678–1680.
[24] A. Dafforn, P. Chen, G. Deng, M. Herrler, D. Iglehart, S. Koritala, S. Lato, S. Pillarisetty, R. Purohit, M. Wang, S. Wang, N. Kurn, Biotechnology 37 (2004) 3–7.
[25] E. Wang, L.D. Miller, G.A. Ohnmacht, E.T. Liu, F.M. Marincola, Nat. Biotech. 18 (2000) 457–459.
[26] L.N. Sellner, R.J. Coelen, J.S. Mackenzie, Nucleic Acids Res. 20 (1992) 1487–1490.
[27] D.P. Chandler, C.A. Wagnon, H. Bolton Jr., Appl. Environ. Microbiol. 64 (1998) 669–677.
[28] O. Suslov, D.A. Steindler, Nucleic Acids Res. 33 (2005) e181.
[29] N.N. Iscove, M. Barbara, M. Gu, M. Gibson, C. Modi, N. Winegarden, Nat. Biotech. 20 (2002) 940–943.
[30] G. Brady, N.N. Iscove, Methods Enzymol. 225 (1993) 611–623.
[31] A. Peixoto, M. Monteiro, B. Rocha, H. Veiga-Fernandes, Genome Res. 14 (2004) 1938–1947.
[32] S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, Clin. Chem. 55 (2009) 611–622.
[33] S. Fleige, M.W. Pfaffl, Mol. Aspects Med. 27 (2006) 26–39.
[34] S. Fleige, V. Walf, S. Huch, C. Prgomet, J. Sehm, M.W. Pfaffl, Biotechnol. Lett. 28 (2006) 601–613.
[35] J.R. Chubb, T. Trcek, S.M. Shenoy, R.H. Singer, Curr. Biol. 16 (2006) 1018–1025.
[36] S.J. Giddings, L.R. Carnaghi, J. Biol. Chem. 263 (1988) 3845–3848.