METHODOLOGY ARTICLE

Quantifying 35 transcripts in a single tube: model-based calibration of the GeXP multiplex RT-PCR assay

Pauline Marguardt^{1,2†}, Britta Werthmann^{1,3†}, Viktoria Rätzel^{1,4†}, Markus Haas¹ and Wolfgang Marwan^{1*}

Abstract

Background: Quantitative analysis of differential gene expression is of central importance in molecular life sciences. The Gene eXpression Profiling technology (GeXP) relies on multiplex RT-PCR and subsequent capillary electrophoretic separation of the amplification products and allows to quantify the transcripts of at least 35 genes with a single reaction and one dye.

Results: We provide a kinetic model of primer binding and PCR product formation as the rational basis for taking and evaluating calibration curves. The calibration procedure and the model predictions were validated with the help of a purposefully designed data processing workflow supported by easy-to-use Perl scripts for calibration, data evaluation, and guality control. We further demonstrate the robustness and linearity of guantification of individual transcripts at variable relative abundance of other co-amplified transcripts in a complex mixture of RNAs isolated from differentiating Physarum polycephalum plasmodial cells.

Conclusions: We conclude that GeXP analysis is a robust, sensitive, and useful method when the transcripts of tens to few hundred genes are to be precisely quantified in a high number of samples.

Keywords: Multiplex RT-PCR, DNA fragment analysis, Capillary electrophoresis, Cell differentiation, Marker genes, Physarum polycephalum

Background

The quantitative analysis of differential gene expression is of central importance in cell, developmental, and systems biology. Established experimental approaches have specific strengths and weaknesses and the method of choice depends on the required accuracy for quantitative detection, the available amount of RNA per sample, the number of different transcripts that are to be quantified in one sample, and the available budget [1, 2]. Transcript quantification by the next generation sequencing RNA-

* Correspondence: wolfgang.marwan@ovgu.de

Pauline Marguardt, Britta Werthmann and Viktoria Rätzel share first authorship

¹Magdeburg Centre for Systems Biology and Institute for Biology,

Otto-von-Guericke Universität, Magdeburg, Germany

BMC

¹The Beckman Coulter Capillary Electrophoresis product line including GeXP chemistry is now distributed and supported by AB SCIEX, Landwehrstr. 54, 64293 Darmstadt, Germany. (www.SCIEX. com).

Seg method, for example, allows to quantify the entire transcriptome with an accuracy depending on the number of reads analysed. However, the quantification of one sample is relatively expensive in terms of sequencing costs and requires sophisticated bioinformatic data processing [3, 4]. Real-time quantitative RT-PCR on the other hand is relatively inexpensive and easy to perform but only gives the transcript concentration for up to few genes per sample. A different, appealing, and useful approach is the GeXP multiplex RT-RCR developed by Beckman-Coulter.¹ This method relies on the





Full list of author information is available at the end of the article

[©] The Author(s), 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License. which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

simultaneous amplification of the cDNAs of a moderate number of genes in the same sample and the subsequent separation of the fluorescence-tagged amplification products by capillary electrophoresis. The number of transcripts that can be quantified in a single pot is limited by the number of fragments that can be thoroughly separated from each other. With the CEQ 8800[™] Genome Analysis System, an 8-capillary sequencer, at least 35 different transcripts can be readily quantified using a single fluorescent dye in a single tube [5], about ten times more than by one multiplex Real-time PCR assay [2]. Thus, GeXP can considerably save time, money, and RNA if transcripts of tens to hundreds of genes are to be quantified in a high number of samples.

The method is depicted in Fig. 1. Each transcript of interest is reverse transcribed with a gene-specific RT-primer (reverse primer) and the second strand of the cDNA is synthesized with a gene-specific forward primer. Both, the forward and the reverse primers carry additional sequence stretches at their 5'-ends that are identical for all gene-specific primer pairs used in the multiplex RT-PCR reaction. During PCR these sequences serve as specific binding sites for the two universal primers. The purpose of the universal primers is to amplify all cDNAs in the reaction mix independent of their length or sequence with equal efficiency during each cycle of the PCR. With the help of a spike-in RNA, an internal standard which is reverse transcribed together with the RNAs to be analysed, the amount of cDNA fragment amplified from each gene of interest can be normalized relative to the amount of fragment amplified from the spike-in RNA so that the relative transcript abundance for each gene becomes comparable between samples. By producing fragments of different length that are separated electrophoretically (Fig. 1), the relative abundances of the 35 chosen transcripts can be determined provided the system is calibrated appropriately. An appropriate calibration procedure and its validation is described in this paper. The normalized peak area for each amplification product which is determined during the electrophoretic separation can be fitted to the total RNA concentration by linear regression of the data points on a double logarithmic plot [6, 7]. Here, we demonstrate that this empirical calibration procedure complies with a kinetic model of cDNA synthesis and amplification. The kinetic model predicts for a set of different transcripts a series of curves of equal slope shifted in parallel to each other. We experimentally confirm the model predictions and the validity of the modelbased calibration procedure. We also provide so far missing necessary controls showing that the amplification efficiency of a given transcript is independent of the relative abundance of the transcripts of other genes that are co-amplified in a sample. Finally, we apply these methods to estimate the relative concentration of different transcripts of differentiation marker genes expressed in *Physarum polycephalum* plasmodial cells that are committed to sporulation [6] and show how the reproducibility of a high number of measurements can be computationally analysed. We also provide a set of Perl scripts as components of an integrated workflow for efficient data processing and evaluation, which is demonstrated in the context of the case study.

Results

Kinetic model

As the concentration of the transcripts of individual genes within a given sample of total RNA may vary by orders of magnitude relative to each other, the conditions of a GeXP experiment have to be adjusted such that the amplified fragments can be reliably detected side by side when separated on the capillary sequencer. This is achieved by empirically adjusting the concentration of the gene specific reverse primers so that at intermediate physiological concentration of each transcript the amount of cDNA amplified is in the same range for all transcripts that are to be detected in one reaction tube (Fig. 2) [5-7]. Hence the concentration of the primer for the reverse transcription of certain transcripts may become rate limiting for the amount of cDNA synthesized. In the following we describe a kinetic model of the RT-PCR reaction

 $RNA \xrightarrow{RT} cDNA^{1st} \xrightarrow{PCR} cDNA^{amp}$

as rational basis for evaluating the double logarithmically plotted calibration curves for the transcripts to be analysed. The mRNAs (RNA) are reverse transcribed to give a first-strand cDNA (cDNA^{1st}) which is then amplified with a universal primer set to obtain a PCR product (cDNA^{amp}) for each transcript which is electrophoretically separated on a capillary sequencer, as shown in Fig. 1).

Assuming that at maximum 2% of the total RNA isolated from a eukaryotic cell is mRNA, a 10 μ l RT reaction volume with the amount of total RNA as used in the experiments contains not more than 0.4 ng of mRNA. With approximately 17,000 genes expressed in a starving *P. polycephalum* plasmodium and an average length of a mature mRNA of approximately 1000 nucleotides [8], this is equivalent to an average concentration of 6.9×10^{-15} M for the mRNA of one gene corresponding to approximately 41,500 molecules per reaction volume. The smallest



RT primer concentration used in our RT reaction (Table S1) is 2.5×10^{-11} M, as compared to 2.7×10^{-8} M for the average and 5×10^{-8} M for the maximal concentration, respectively. For a specific mRNA of 6.9×10^{-12} M which is 1000 times more abundant than the average, even the smallest RT-primer concentration of 2.5×10^{-11} M would be in

excess to the concentration of the template in a RT reaction. Accordingly, the primer concentration for most if not all of the amplified fragments is expected to be in very large excess as compared to the template concentration and under these conditions the concentration of the reverse primer will not change significantly in the course of the RT reaction.



It has been shown that binding of a primer to its template during the reassociation (renaturation) of DNA follows the rate law of a second order reaction [9, 10]. This suggests that the rate limiting step for the reassociation is the nucleation reaction during which few bases of one strand bind to their fitting counterparts on the other strand. The subsequent zipping reaction must occur fast as compared to the nucleation event of base pairing because otherwise a first order instead of second order kinetics would be observed. This conclusion is plausible as the local concentration of fitting bases greatly increases if two strands are brought in spatial proximity and appropriate orientation as a consequence of the zipping reaction.

The concentration of the RT-primer for the transcript of each gene of interest has to be adjusted to attenuate the relative amount of DNA amplified from its RNA during the GeXP RT-PCR reaction while the concentration of the RTprimer (denoted as Primer in the following) is in large excess as compared to the reverse transcribed mRNA (see above). We accordingly assume that formation of the RNAprimer complex (RNA · Primer)

$RNA + Primer \rightarrow RNA \cdot Primer \xrightarrow{RT} cDNA^{1st}$

is the rate-limiting step of cDNA synthesis by reverse transcription (RT). Once priming of the RT reaction has occurred, the polymerization reaction of cDNA synthesis is equally fast for all 1st strand cDNA species (cDNA^{1st}) concurrently synthesized in the reaction mix.

Note that only a tiny fraction of the low molecular weight reactants is consumed during the RT reaction such that their concentrations do hardly change during the experiment. Based on the assumption of a ratelimiting bimolecular reaction of RT-primer binding to the RNA template, we expect an empirical rate law for the time-dependent decay $d[\text{RNA}_i]/dt$ of the concentration [RNA_i] of the transcript of each gene *i* according to

$$-\frac{d[\text{RNA}_i]}{dt} = k_{i,j} \cdot [\text{Primer}_{i,j}] \cdot [\text{RNA}_i]^n \tag{1}$$

where $k_{i,j}$ is the primer-specific rate constant for hybridization of the primer to its complementary RNA sequence, [Primer_{*i,j*}] is the concentration of primer *j* of given sequence and length, complementary to the corresponding sequence stretch of the mRNA of gene *i*, and *n* is the empirical order of the reaction. As the concentration of the primers in the RT reaction is in a sufficiently high excess over that of each gene-specific RNA (see estimation above), the primer concentrations can be approximated as remaining constant during the reaction. By defining the pseudo first order rate constant $k'_{i,j} = k_{i,j} \cdot [\text{Primer}_{i,j}]$ we obtain the empirical first order rate law

$$-\frac{d[\text{RNA}_i]}{dt} = k'_{i,j} \cdot [\text{RNA}_i]^n \tag{2}$$

where the value of $k'_{i,j}$ depends on the primer concentration that had been appropriately adjusted by the experimenter. Separation of the variables and integration over the reaction time *t* and the change in RNA_i concentration yields

$$\int_{[\text{RNA}_{i}]_{0}}^{[\text{RNA}_{i}]} \frac{1}{[\text{RNA}_{i}]^{n}} d[\text{RNA}_{i}] = -k_{i,j}^{'} \int_{t=0}^{t} dt.$$
(3)

In the simplest case of a first order reaction where n = 1, [RNA_{*i*}] decays exponentially during the reverse transcription reaction of length *t* starting from its initial concentration [RNA_{*i*}]₀ according to

$$[\operatorname{RNA}_{i}] = [\operatorname{RNA}_{i}]_{0} \cdot \exp\left(-k_{i,j}' \cdot t\right).$$
(4)

Alternatively, if the reaction order is different from one $(n \neq 1)$, the integration yields

$$\frac{[\text{RNA}_i]^{1-n} - [\text{RNA}_i]_0^{1-n}}{1-n} = -k'_{i,j} \cdot t$$
(5)

$$RNA_{i}]^{1-n} = k_{i,j}^{'} \cdot t(n-1) + [RNA_{i}]_{0}^{1-n}$$
(6)

and hence [RNA_i] decays according to

[

$$[\text{RNA}_{i}] = \sqrt[1-n]{k'_{i,j} \cdot t(n-1) + [\text{RNA}_{i}]_{0}^{1-n}}$$
(7)

In both cases, for n = 1 and for $n \neq 1$, the concentration $[cDNA_i^{1st}]$ of gene-specific first strand cDNA, synthesized during the entire RT reaction is

$$[cDNA_i^{1st}] = [RNA_i]_0 - [RNA_i]$$
(8)

The concentration of the gene *i*-specific cDNA fragment, $[cDNA_i^{amp}]$ which has been amplified by PCR from the first strand cDNA (cDNA_i^{1st})

$$cDNA_i^{1st} \xrightarrow{PCR} cDNA_i^{amp}$$

with the help of the universal primer pair (see Introduction) is expected to be

$$\left[cDNA_{i}^{amp} \right] = \alpha \cdot \left[cDNA_{i}^{1st} \right]$$
(9)

and according to eq. (8) dependent on the initial concentration of the gene *i*-specific transcript

$$\left[\mathrm{cDNA}_{i}^{amp}\right] = \alpha \cdot \left(\left[\mathrm{RNA}_{i}\right]_{0} - \left[\mathrm{RNA}_{i}\right]\right)$$
(10)

Overall amplification α of each 1st strand cDNA molecule after *c* cycles of PCR is expected to be $\alpha \leq 2^{c-1}$ because the first PCR cycle is required to synthesize the second strand on the cDNA, leaving *c* – 1 cycles for DNA amplification while a new second strand product can be synthesized from the first strand cDNAs during each cycle of PCR. Each newly synthesized second strand cDNA molecule then starts a new exponential cascade of amplification during the remaining cycles of PCR. With eq. (4) and (7), the concentration [cDNA_{*i*}^{*amp*}] of the gene *i*-specific cDNA fragment for *n* = 1 is

$$\begin{bmatrix} cDNA_{i}^{amp} \end{bmatrix} = \alpha \cdot [RNA_{i}]_{0} \\ \cdot \left(1 - \exp\left(-k_{i,j}^{'} \cdot t\right)\right)$$
(11)

and for $n \neq 1$

$$\left[\mathrm{cDNA}_{i}^{amp}\right] = \alpha \cdot \left(\left[\mathrm{RNA}_{i}\right]_{0} - \sqrt[1-n]{k'_{i,j} \cdot t(n-1) + \left[\mathrm{RNA}_{i}\right]_{0}^{1-n}}\right)$$
(12)

Calculating the concentration of cDNA [cDNA_i^{amp}] amplified from the gene *i*-specific mRNA as a function of the relative initial concentration $[RNA_i]_0$ in the RT reaction according to eq. (11) and (12) yields apparently linear relationships in a double logarithmic plot for empirical reaction orders of $n \le 1$ (Fig. S1). The concentration of any gene-specific mRNA present in the RTreaction is proportional to the concentration of total RNA [RNA_{tot}]₀ which is subjected to reverse transcription. Accordingly, a double logarithmic plot of the concentration of specific cDNAs amplified from the transcripts of different genes will yield a series of parallel shifted lines of equal slope and correspondingly different intercepts with the ordinate. The kinetic model with eq. (11) and (12) further predicts how the concentration of cDNA fragments depends on the initial concentration of the gene-specific reverse primer used for the RT reaction (Fig. S2).

Experimental validation of the calibration procedure

The kinetic model of the GeXP RT-PCR predicts that for a set of transcripts one obtains a series of parallel shifted lines when the concentration of amplified gene-specific cDNA fragment is double logarithmically plotted versus the total RNA concentration used in the assay. For experimental validation of this prediction, we used *Physarum polycephalum* total RNA isolated from plasmodial cells as a source of typical eukaryotic mRNAs [8].

Starving plasmodial cells can be induced to sporulation by a brief pulse of far-red light which is sensed by a phytochrome photoreceptor [8, 11, 12]. The cellular commitment to sporulation is associated to the differential expression of developmentally regulated genes [13]. Based on transcriptome data [13, 14] we have previously designed gene specific primers for the quantification of the transcripts of a set of 35 genes with a single GeXP RT-PCR reaction. The concentrations of the genespecific reverse primers were attenuated (adjusted) so that amplification of each of the 35 cDNAs yields a sufficient amount of each fragment to be resolved side-byside during electrophoretic separation (Fig. 2; Table S1) [6]. The set comprised genes that are up- or downregulated at different strengths as well as genes with constant or almost constant expression level that can serve as a reference.

Fragment quantification by capillary electrophoresis

For the initial validation of the performance of the electrophoretic quantification of cDNA fragments, we used two samples of RNA, one prepared from uninduced plasmodial cells (controls) and one from far-red light induced plasmodial cells harvested at 6 h after the far-red pulse. The induced cells were committed to sporulation but did not yet show any morphogenetic alterations. In the cells of the two groups, a number of genes are upor down-regulated [6]. A known amount of spike-in RNA was added to the two samples of total RNA isolated from the cells of the two groups and the cDNAs of the 35 transcripts were amplified by GeXP RT-PCR. Subsequently, a DNA size standard was added to each of the two samples and the fluorescence tag-labelled fragments were separated electrophoretically on the CEQ 8800[™] 8-capillary sequencer (see Fig. 1c). The fragments were quantified through their peak area and the amount of each gene-specific fragment was normalized relative to the amount of fragment amplified from the spike-in RNA with the help of the Perl scripts (Supplementary Information) to give the normalized peak area (NPA).

In order to test the performance of the electrophoretic separation process and to validate the fidelity of the peak quantification software, we prepared a series of mixtures, each containing three pairs of fluorescently labelled DNA fragments of different relative concentration, mimicking different concentrations of cDNAs as obtained by RT-PCR. Each mixture was separated on the CEQ 8800™ capillary sequencer and the normalized peak area (NPA) determined for three chosen pairs of fragments that were amplified from an up-regulated and one downregulated transcript, respectively. For the three pairs of DNA fragments, the ratio of the normalized peak areas for the corresponding two fragments was plotted against the relative concentration of the two samples in the mixture. In the three pairs of fragments the normalized peak area for each fragment was proportional to its relative amount in the mixture (Fig. S3), confirming that the peak detection algorithm of the software delivered with CEQ 8800[™] capillary sequencer worked correctly.

Calibration curves for the 35 transcripts

In order to obtain calibration curves for each of the 35 transcripts, we mixed two RNA samples (from lightinduced and control cells) in a 1:1 ratio in order to obtain a sample that contained even the differentially regulated transcripts at intermediate concentration (Fig. 2). This RNA mixture was serially diluted and, after addition of spike-in RNA to each sample of the dilution series, the cDNA fragments as amplified by the GeXP multiplex RT-PCR were separated electrophoretically. The normalized peak area determined for each cDNA fragment was plotted versus the concentration of total RNA. The data points obtained for all fragments were fitted to straight lines in the double logarithmic plot. The slopes as obtained by the 35 fits were averaged and the best-fitting intercept for each of the 35 curves was calculated. Accordingly, this procedure gave a series of fitted curves that only differed in being shifted with

respect to the abscissa (Fig. 3). As predicted by the kinetic model, for none of the genes any systematic deviation of the data points from the fitted straight line was seen (Fig. 3). Hence, all fragments were amplified with the same efficiency by the universal primers. The parallel shift relative to the abscissa, according to the kinetic model, is due to the different absolute concentrations of the mRNAs and to the different amounts of cDNA produced during the RT reaction at the particular concentration of the RT-primer. Noteworthy, the slopes of the curves clearly were smaller than 1 as observed in most experiments performed in our lab. A slope of 1 would be expected if the cDNA synthesis occurred with first order kinetics as assumed in eq. (11). The kinetic model therefore suggests that the RT reaction occurs with a reaction order of $n \le 1$ (eq. (12)). The value of nobserved in each particular experiment might slightly vary depending on the experimental conditions.

Efficiency of concurrent amplification of different transcripts in one sample

Since cDNA fragments must be of different lengths in order to get separated during capillary electrophoresis and the efficiency of amplification might depend on the fragment length, we wished to confirm that the amount of cDNA amplified for any given transcript is independent of other fragments that are concurrently amplified in the same sample. Therefore, we mixed the RNA sample from the control cells with the sample from the lightinduced cells taking different relative amounts and estimated for each gene how the normalized peak area changed with the relative amount of the RNA from light-induced cells or from the dark controls (Fig. 4). For each fragment, the normalized peak area and hence the amount of amplified cDNA was a linear function of the mixing ratio with a slope depending on whether genes were up-regulated or down-regulated. Curves for genes that are not regulated in response to the photo-stimulus fitted to a straight line that was parallel to the abscissa (Fig. 4). Hence, the amount of amplified DNA for each gene was found to be proportional to the concentration of the corresponding mRNA in the sample. Likewise, amplification of the cDNAs of those genes that are not or rarely differentially regulated (*dspA*, *spiA*, *tspA*, etc.) confirms that the amplification efficiency indeed does not depend on the concentration of other (differentially regulated) mRNAs that are concurrently amplified in the sample.

An experimental result consistent with the computational predictions of the kinetic model of the GeXP RT-PCR (eq. (12)) was obtained by changing the concentration of one primer while the concentration of all other primers was kept constant. Until saturation, the normalized peak area increased with the RT-primer concentration, indicating that the reverse primer was rate-limiting even though its concentration was high relative to the concentrations of the other RT-primers in the mix (Fig. S2).

Taken together, these experimental results confirm that all fragments are amplified with the same efficiency during PCR with the universal primer set regardless of their sequence or the length of the amplification products. Amplification of individual fragments is not influenced by other fragments that are concurrently amplified in the sample. This confirms (at least for the primer set chosen and for the experimental conditions applied) that the relative changes in the mRNA concentration of any fragment robustly translate into relative changes in the amount of detected amplified cDNA during a GeXP multiplex RT-PCR experiment. Certainly, equal amplification efficiency for all cDNAs quantified within one essay should be verified for any new set of primers and any new set of transcripts to be analysed. The Perl scripts provided as supplementary material (see below) facilitate this task.

Quantification of differentiation marker genes in

plasmodial cells before and after induction of sporulation By quantifying the light-induced expression of differentiation marker genes in *Physarum polycephalum* plasmodial cells, we present a workflow including the application of Perl scripts for calibration and data evaluation.

Stimulation of cells, sample preparation, and biological replicates

Plasmodial cells were starved for six to seven days to gain sporulation competence. Sporulation-competent plasmodial cells, each spreading the surface of a Petri dish were stimulated with a short pulse of far-red light for the induction of sporulation. Before and at different time points after application of the stimulus, samples of each plasmodial cell were harvested and frozen in liquid nitrogen for the subsequent isolation of RNA. The remainder of the plasmodium was further incubated overnight to verify whether the developmental decision to sporulation had occurred [15]. This approach gave a collection of time series where each time series consisted of samples taken from the same individual plasmodial cell. In a control experiment, light stimulation was omitted while samples of individual plasmodial cells were taken correspondingly. From each frozen sample, total RNA was isolated and the expression pattern of the set of 35 genes (Table S1) were quantified as described above.

Data processing workflow and Perl scripts

The electrophoretic separation of the GeXP multiplex RT-PCR samples was performed in the CEQ 8800™



induced and control plasmodial cells (see Fig. 2) was serially diluted and the 35 cDNA fragments quantified after capillary electrophoretic separation. Each RNA concentration was analysed twice and each of the obtained data points, normalized peak area (NPA) versus concentration of total RNA, was displayed in a double logarithmic plot (logarithm to the base 10). All data points were fitted to a straight line of the same slope as described in Materials and Methods. The R² value displayed in each panel gives the correlation coefficient of the first linear regression fit (see Methods for details). The red line in each panel represents the straight line of equal slope finally fitted to the data for each of the 35 transcripts (second linear regression fit; see Methods for details). For names of genes see Table S1. The figure was prepared with R, version 3.5.2



Genome Analysis System run via the so-called Control Center, a software package for operating the system and for storing the data. All further data processing steps were performed by Perl scripts as described in the Supplementary Information (see also Figs. S4 and S5).

As a standard quality control for the electrophoretic separation, the electropherogram and the detected fragments were graphically displayed and visually inspected calling the respective scripts (Fig. S4). All further data processing steps were executed using the .ceqfrag files that list peak height, peak area and length (in nucleotides) of each fragment. The Perl scripts *gep_makestd* and *gep_calcamount* perform a batch evaluation of these files to generate the standard curve and to calculate the relative transcript abundances from the measured data, respectively. This is achieved by taking the fitted mean slope and the corresponding intercept of the standard curve for each transcript into account. The final output is a table in the form of a file with tab-separated values that contains for each quantified fragment the fragment ID number, the length (nt's), peak area, normalized peak area (NPA), the relative amount of cDNA, and the detection limit of each fragment during the respective

6d, 5 min	FR 6d, 15 min FR 7d, 20 min Ff	7d, 30 min FR	6d, 5 min FR	6d, 15 min FR 7d,	20 min FR 7d, 30 min FR
psgA Time (h) P1 P2 P3 0 100 100 100 2 56 39 85 6 10 27 68 8 4 6 18 11 14 12 7	P4 P5 P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 P16 P1 M05	P18 P19 P20 P21 P23 P24 P25 P26 0100 1000	Time (h) P1 P2 P3 P4 P5 P6 0 100	66 P7 P8 P9 P10 P11 P12 P13 P14 P15 C05 #83 1050 100	P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P26 P27 P20 P21 P21 P20 P21 P20 P21 P20 P21 P21 P20 P21 P21 <thp21< th=""> <thp21< th=""> <thp21< th=""></thp21<></thp21<></thp21<>
pldC Time (h) P1 P2 P3 0 1000 38 100 2 19 58 75 6 2 27 4 8 1 100 1 11 2 49 2	P5 P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 P16 P16 P100 28 47 100 100 100 26 100 100 27 100	P 113 P19 P20 P21 P22 P23 P24 P25 P26 0 000 1000 1000 1000 1000 1000 1000 1000 1000 1000 1100 1000 1100 111 1	Time (h) P1 P2 P3 P4 P5 P0 0100 38 100 130 38 2 13 65 82 27 100 6 2 28 2 15 13 8 1100 0 5 17 11 1 49 2 2 10	66 77 78 79 710 711 712 713 714 715 000 1500 1500 1500 1500 1500 1500 1500 1500 1500 1500 1500 1500 1500 1500 150 <td< th=""><th>P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P25 100 1000 11 1000 11 1000 11 1000 11 1000 11 16 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 <t< th=""></t<></th></td<>	P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P25 100 1000 11 1000 11 1000 11 1000 11 1000 11 16 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 11 1000 <t< th=""></t<>
pumA Time (h) P1 P2 P3 0 100 8 100 2 155 12 77 6 32 36 45 8 29 169 20 11 19 100 45	P4 P5 P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 P16 P1 1000 1000 104 1000 1000 106 106 1000 100 100 100 100	P18 P19 P20 P21 P22 P23 P24 P25 P26 0.000 0.000 1000 2.01 1000 2.5 2.01 199 0.100 0.000 1000 2.61 1000 2.5 2.61 1999 0.100 1000 2.68 9.89 9.90 2.62 2.63 3.01 0.102 2.61 5.85 9.00 1000 4.55 2.8 4.84 3.01 2 3.65 3.90 1000 3.01 168 190 1601 1 1.45 3.85 3.97 7.00 3.81 168 190 1601	Time (h) P1 P2 P3 P4 P5 P6 0 100 9 100 100 100 10 2 165 13 60 23 28 6 38 40 49 27 36 8 30 71 14 27 10 11 25 100 47 6 21 1	6 P7 P8 P9 P10 P11 P12 P13 P14 P15 000 1000	P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P26 1000 1000 1000 1000 1001 101 25 1000 17 15 16 5 776 1055 1000 15 75 0 1000 109 19 2 8 100 101 48 12 15 55 12 105 00 75 2 12 0 47 38 4 36 14 30 28 46 80 87 32 166 186 186 186 145 4 8 15 14 55 14 00 0 4 48 1000 100 48
anxA Time (h) P1 P2 P3 0 100 9 100 2 19 3 86 6 4 3 49 8 2 51 50	P4 P5 P6 P9 P1 P11 P13	7 P18 P19 P20 P21 P22 P23 P24 P25 P26 0 58 100 100 9 100 100 100 56 100 1 000 25 39 38 20 35 10 56 100 7 19 28 38 9 13 86 100 47 1 28 14 16 100 7 3 3 30 38 33	Time (h) P1 P2 P3 P4 P5 P6 0 100 111 100 100 100 12 2 13 5 180 35 25 6 4 4 39 38 46 8 2 51 41 5 10	P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 100 32 1000 1000 35 1000 1000 100 1000 1000 100 1000 1	P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P26 0 000 000 000 000 100 100 100 100 100 1
11 7 100 S1 meiß Time (h) P1 P2 P3 0 1000 18 98 2 16 13 47 6 12 23 100 6 14 100 68	1 12 5 5 6 2 2 1000 1 58 1000 2 5 0 P4 P5 P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 P16 P1 100 100 100 100 100 100 30 30 100 100 10	3 10 9 2 39 11 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2	11 5 100 41 0 13 Time (h) P1 P2 P3 P4 P5 P6 0 100 18 100 100 100 100 100 1 2 23 10 54 19 23 6 16 18 95 20 13 8 7000 65 20 13	1 8 6 30 1000 12 49 1000 5 0 26 P7 P8 P9 P10 P11 P12 P13 P14 P15 100 1000 100 176 35 1000 1000 100 100 100 95 60 91 178 19 49 28 88 96 177 24 57 49 1000 27 58 97 15 175 2	D 6 4 10 5 1 50 19 4 27 12 7 P16 P17 P18 P19 P20 P21 P22 P23 P24 P25 P26 1 187 86 300 100 100 100 100 100 100 100 99 193 100 187 800 68 47 8 38 8 8 9 2 1 2 9 91 100 59 180 006 68 47 8 38 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
ribB Time (h) P1 P2 P3 0 95 100 96 2 95 85 95 6 90 85 100	1 1	1 1	1 9 9 9 94 5 10 1 Time (h) P1 P2 P3 P4 P5 P6 0 92 100 100 89 96 1 2 94 88 92 96 95 1 6 96 84 98 92 100 1	1 24 12 97 100 48 50 22 17 21 15 24 12 197 100 48 50 22 17 21 16 97 P8 P9 P10 P11 P12 P13 P14 P15 98 172 91 100 98 100 90 94 99 95 98 170 92 92 92 80 90 99 99 99 99 99 99 99 99 99 99 99 99 90 99 90	Im Im<
8 100 98 97 11 96 98 97 gapA Time(h) P1 P2 P3 0 9 2 5 2 5 1 5 6 10 10 10	192 194 191 72 86 88 96 66 184 190 77 199 192 192 93 87 196 88 81 190 177 199 192 192 93 87 196 188 81 190 177 199 192 193 196 196 196 88 81 190 177 14 195 196 190 110 11 12 13 14 15 16 15 1 1 16 16 1 16 15 1 1 16 18 1 16 18 1 16 18 1 16 18 1 16 18 1 16 18 16 18 1 16 18 1 16 18 1 16 18 1 16 18 1 16 18 1	1 779 88 97 73 94 100 100 91 95 2 86 98 91 84 91 93 96 89 100 7 918 919 920 921 922 923 924 925 926 6 14 37 38 8 36 60 0 18 6 8 40 38 26 14 30 37 38 13 06 60 0 18 6 8 40 97 38 26 14 80 77 24 19 13	8 100 99 88 90 98 1 11 94 92 87 100 96 1 Time (h) P1 P2 P3 P4 P5 P0 0 7 2 5 21 16 2 7 1 4 41 11 1	52 83 79 91 100 0 89 100 197 193 56 89 88 87 99 92 92 100 197 193 56 89 88 87 99 92 92 100 197 193 57 7 38 101 6 7 29 17 9 13 26 8 9 101 13 22 13 14 P15 13 26 8 9 10 10 36 22 18 178 17 9 13 26 9 11 36 22 18 178 13 13 26 19 13 26 10<	170 92 175 100 98 96 98 95 95 100 93 198 100 198 95 100 173 96 100 173 96 100 175 96 96 97 95 100 175 96 100 175 96 96 97 95 100 175 96 96 97 95 100 175 96 96 96 97 95 100 175 97 9
pldA Time (h) P1 P2 P3 0 10 10 10 10 10 10 10 10 10 10 10 10 10	All All <th>100 45 20 21 101 102 102 103</th> <th>6 24 13 3 20 45 18 176 52 11 100 100 100 100 50 16 Time (h) P1 P2 P3 P4 P5 P6 0 6 6 91 1 1 2 20 4 192 1 4</th> <th>13 4 4 7 4 21 4 7 4 1 4 1 4 1 4 1 4 4 4 4 4 4 4 4 4</th> <th>3 3 3 3 4 3</th>	100 45 20 21 101 102 102 103	6 24 13 3 20 45 18 176 52 11 100 100 100 100 50 16 Time (h) P1 P2 P3 P4 P5 P6 0 6 6 91 1 1 2 20 4 192 1 4	13 4 4 7 4 21 4 7 4 1 4 1 4 1 4 1 4 4 4 4 4 4 4 4 4	3 3 3 3 4 3
6 51 93 74 8 49 100 48 11 100 85 100 ligA Time (h) P1 P2 P3 0 5 3 8	381000 581 581 591 100<	0 18 19 19 18 53 45 100 100 10 18 48 7 100 18 19 100 90 89 99 19 10 1 16 100 100 172 100 94 78 100 100 7 118 119 720 721 722 723 724 725 726 8 27 7 8 112 3 10 1 1 4	6 41 80 80 40 100 14 8 38 100 52 100 58 1 11 100 98 100 95 86 1 Time (h) P1 P2 P3 P4 P5 P6 0 7 4 4 4 27 31	1000 56 100 22 28 30 41 56 12 21 13 100 13 50 58 38 300 32 44 14 100 </th <th>175 100 48 28 36 50 38 100 58 27 44 189 123 100 38 36 107 107 50 56 100 58 127 48 50 50 50 50 50 50 50 50 50 50 50 50 100</th>	175 100 48 28 36 50 38 100 58 27 44 189 123 100 38 36 107 107 50 56 100 58 127 48 50 50 50 50 50 50 50 50 50 50 50 50 100
2 5 12 10 6 18 67 9 8 34 100 51 11 100 93 100 pwiA Time (h) P1 P2 P3	1007 28 18 9 101 14 6 12 25 7 78 15 1000 36 41 84 34 43 34 43 34	1 20 16 14 17 53 63 14 2 16 0 28 82 13 14 43 600 78 51 14 0 28 82 14 43 55 85 87 51 12 10 24 99 46 45 55 85 87 51 21 100 100 100 100 100 100 100 100 10 148 920 721 722 723 724 725 726	2 6 11 10 1956 24 2 6 22 70 8 94 27 3 8 41 100 96 100 100 10 11 100 96 100 100 100 10	16 100 111 11 8 7 113 30 9 10 13 59 17 46 21 42 30 52 36 52 34 2 19 74 38 58 58 58 59 57 66 77 66 100 100 100 100 100 100 100 100 100 100 100	1 22 11 21 21 81 4 51 46 46 7 7 1 14 8 100 30 7 1 9 1 28 100 12 40 9 3 84 100 50 7 9 1 29 38 24 10 10 100 100 100 100 100 100 100 100
0 6 2 2 2 2 1 1 12 6 6 21 29 8 13 82 85 11 100 100 100	21 3 0 2 51 1 2 51 1 2 22 50 1501 1 3 38 38 10 38 151 1 2 22 50 1501 1 38 39 10 28 14 47 28 30 26 16 163 1502 1 38 39 10 28 14 47 28 30 26 16 163 16 36 16	7 4 5 5 13 3 19 0 1 2 4 99 11 3 19 6 4 24 6 9 172 36 12 20 11 34 5 5 8 16 3100 169 1 85 7 48 72 30 41 0 47 160 100 100 100 100 100 100 100	0 7/ 1 22 7/ 3 2 0 1 5 1 1 6 5 19 26 182 17 8 13 68 7/8 180 37 11 100 100 100 100 10	191 1.6 1.2 5 5 2 2 1 4 3 0 5 0 4 5 1 7 1 7 89 11 26 16 46 32 32 40 1 21 30 25 56 56 10 46 33 100 100 100 100 100 100 100 100 33 100 100 100 100	0 9 2 5 4 11 3 15 3 1 2 1 9 9 1 1 3 19 3 15 10 1 75 13 3 24 9 3 24 51 10 1 75 14 7 9 3 24 51 10 1 75 84 9 3 10 7 35 36 7 17 1 70 84 98 100 7 4 36 36 29 45 35 28 45 3 36
7gSA Time (h) (P1 (P2 (P3 (A 2 10 (2))) 6 (2) (100 (13)) 8 (3) (100 (2)) 11 (100 (22) (100)	94 95 96 97 96 99 10 911 911 911 911 911 911 91 91 911 911 911 91 91 911 911 911 911 91 91 911 <td< th=""><th>P18 P19 P20 P21 P23 P24 P25 P26 0 31 2.21 2 3 2 2 2 4 0 2 4 1 2 4 1 2 4 1 2 4 1 2 4 1 2 4 1 2 4 1</th><th>Immet(h) P1 P2 P3 P4 P5 P6 0 1 45 81 2 61 2 1 46 50 0 3 6 2 1000 15 56 14 : 8 3 29 18 56 14 : : 11 100 20 100 100 100 100 100</th><th>76 77 78 79 710 71 72 73 74 71 7 71<</th><th>P16 P17 P17</th></td<>	P18 P19 P20 P21 P23 P24 P25 P26 0 31 2.21 2 3 2 2 2 4 0 2 4 1 2 4 1 2 4 1 2 4 1 2 4 1 2 4 1 2 4 1	Immet(h) P1 P2 P3 P4 P5 P6 0 1 45 81 2 61 2 1 46 50 0 3 6 2 1000 15 56 14 : 8 3 29 18 56 14 : : 11 100 20 100 100 100 100 100	76 77 78 79 710 71 72 73 74 71 7 71<	P16 P17
6d, 5 min	FR 6d, 15 min FR 7d, 20 min FF	7d, 30 min FR	6d, 5 min FR	6d, 15 min FR 7d,	20 min FR 7d, 30 min FR

Fig. 5 Comparative display of single cell gene expression time series in response to a sporulation-inducing far-red light stimulus. Individual plasmodial cells (P1 to P26) starved for 6 or 7 days to gain sporulation competence were exposed to a far-red light pulse for 5 to 35 min. Before and at four time points after the light pulse, samples were taken from each plasmodium and the gene expression pattern analysed to give a time series for each individual cell. The relative changes of expression values were normalized to the maximal value within the time series and displayed side by side for each plasmodium and gene. The gene expression pattern in each of the RNA samples was analysed twice with independent RT-PCR reactions and based on separate calibration curves. The first measurement and the technical replicate are displayed in the left and in the right panels, respectively. The figure was prepared with Microsoft Excel for Mac, version 14.7.7

separation run. Two values, relative amount of cDNA and detection limit, are calculated based on the calibration curves. We routinely made a new standard curve for each kit in order to account for possible variations in the performance of the kit components. Details on data processing and on the determination of the transcriptspecific detection limits are given in the legend to Figs. S4 and S5. Samples of all file types used are available together with the Perl scripts (for details see Supplementary Information). With the help of the two Perl scripts, the data processing for the generation of the calibration curves and for calculating the amounts of cDNA for each fragment and sample becomes fully automatic due to the batch processing of the files. We finally store all values for calibration curves as well as for all other measurements in an SQL database for documentation and further data processing and evaluation.

Light-induced gene expression and technical replicates

For visual comparison of the light-induced differential regulation of the differentiation marker genes in individual plasmodial cells, the expression values of the 35 genes were comparatively displayed for each sample (Fig. 5; for the complete data set see Fig. S6). Each data point of a time series originates from material taken from the same plasmodial cell, allowing to compare time series between individual cells. Strongly up- or downregulated genes are clearly evident while the expression of *ribB* was almost constant. Accordingly, *ribB* might be used as a reference gene to normalize the expression values in each sample in order to correct for pipetting errors or inaccuracies in the quantification of the total RNA samples analysed. Nevertheless, we displayed all expression values as they were obtained, i.e. without normalization. However, the five expression values for each gene in a time series taken from a given cell were then normalized to the maximal value to reveal the percental changes as a function of time in order to highlight the relative changes that occurred as a function of time.

In order to estimate how reproducible the measurements are, the RNA samples were measured again in an independent GeXP RT-PCR experiment which was evaluated based on a fresh calibration curve. The results of the technical replicate were displayed side by side with the first measurement (Fig. 5; Fig. S6) and revealed that the time series gave similar results. For a quantitative evaluation, the reproducibility of the measurements was analysed with the help of a frequency distribution. For each single expression value obtained in the first and in the second measurement (the technical replicate) for any given sample and gene, the ratio between the two values was calculated and the absolute frequency of the log2 (ratio) displayed in the form of a histogram. For the experiment performed with light-induced plasmodia (Fig. 6a) the variation was considerably higher than for a second experiment performed with plasmodia that had not been exposed to any light (dark controls; Fig. 6b). For the light-induced plasmodial cells, 77% of the values varied by a factor of two or less and 9% of the values by a factor of four or more. One reason for the high variation observed in the light-induced plasmodia might be that genes are strongly up- or down-regulated. Two measurements of a transcript, the concentration of which is close to the detection limit, may be inaccurate and therefore give a relatively high value for the ratio between first and second measurement. In contrast, the values measured in the dark controls were highly reproducible (Fig. 6b).

Discussion

We have designed and experimentally validated a calibration procedure for Gene eXpression Profiling (GeXP), a quantitative multiplex RT-PCR method for the simultaneous quantification of tens of transcripts. Our calibration procedure is rationally based on a kinetic model of primer binding, reverse transcription, and amplification. We have also tested an essential underlying assumption of GeXP that amplification of cDNAs synthesized from an individual transcript occurs independently of the relative abundances of the transcripts that are co-amplified in the sample. Finally, we provide an easy-to-use pipeline of Perl scripts for calibration, data evaluation, and quality control of GeXP. Our method has been successfully applied in several recent studies [6, 15–20].

Methods

Growth and preparation of sporulation-competent plasmodia

Plasmodia of the apogamic strain WT31 [21] which is wild-type with respect to the photocontrol of sporulation were grown for four days at 24 °C in 1.5 L of growth medium [22] in a 5 L fermenter (Minifors, Infors HT, Bottmingen, Switzerland) inoculated with 2% of a 3.5 days old shaken culture, supplied with 0.1 L of air per minute, and stirred at 250 rpm with a marine propeller. Microplasmodia were harvested and applied to starvation agar plates (9 cm diameter) as described [21]. Plates were incubated at 22 °C in complete darkness for 6 days if not stated otherwise. During this time period, one multinucleate macroplasmodium develops on each plate and, by starvation, becomes competent for sporulation. The incubation temperature is critical in order to avoid unwanted spontaneous sporulation. For the induction of sporulation, competent plasmodia were exposed to a farred light pulse ($\lambda \ge 700$ nm, 13 W/m²). The far-red light was generated by Concentra Weißlicht lamps (Osram, Munich, Germany) which was passed through an Orange 478 combined to a Blue 627 plexiglass filter



(Röhm, Darmstadt, Germany). After light exposure, plasmodia were returned to the dark and incubated at 22 °C until the samples were taken. At five subsequent time points, a portion of each plasmodium was harvested (approximately one sixth of its original biomass for each sample), shock-frozen in liquid nitrogen and stored at -80 °C until RNA isolation. After taking the samples, the remainder of each plasmodium was maintained overnight in the dark to reveal the developmental decision (i.e. whether or not sporulation had occurred). Control plasmodia were treated identically except that the light pulse was omitted (dark controls). All manipulations were done under sterile conditions and under dim green safe light as previously described [21].

Preparation of total RNA

For extraction of total RNA, approximately 15 mg of fresh weight material which had been taken from an individual plasmodium was used. Samples were dissolved in 1.5 ml PeqGOLD TriFastTM solution (Nr 30-2020; PeqLab, Erlangen) by disrupting the cell material in the presence of glass beads $\phi = 0.5$ mm using a Precellys 24-Dual homogenizer, (Peqlab, Erlangen, Germany) for 10s at room temperature. The homogenate was transferred to 2 ml PeqGOLD PhaseTrap tubes (PeqLab; Nr 30-0150A-01). After addition of 300 µl CHCl₃, vigorous shaking and centrifugation of the PhaseTrap tubes, the aqueous phase containing the RNA was transferred to a fresh Eppendorf cap following the manufacturer's instructions for the entire procedure. After addition of 0.9 volumes of a 4:1 mixture of ethanol and 1 M acetic acid, the RNA was precipitated at - 20 °C for at least 24 h. The precipitate was spun down for 15 min at 18,000 x g and 4 °C, the pellet washed once with 400 µl ice-cold 3 M sodium acetate and twice with ice-cold 70% ethanol followed by centrifugation for 10 min at 18,000 x g, 4 °C after each washing step. Finally, the RNA was dissolved in RNAse-free water. For removal of contaminating genomic DNA, 90 µl of each RNA sample were mixed with 10 µl rDNase buffer and 1 µl rDNase (rDNase Set Macherey-Nagel: 740963) and incubated for 12 min at 37 °C. The total RNA was finally purified using the RNA Clean-up Kit (Macherey-Nagel: 740948.250). Approximately 300 ng aliquots checked for integrity of the 18 S and 28 S rRNA by agarose gel electrophoresis. For subsequent steps, samples were standardized to 20 ng/µl of total RNA.

Multiplex gene expression profiling (GeXP)

Multiplex RT-PCR reactions were performed to simultaneously amplify 35 transcripts and the amplification products were quantified through separation on a Beckman Coulter 8-capillary sequencer (CEQ 8800). The GenomeLab GeXP Start Kit (Beckman Coulter; A85017) along with the ThermoStart Taq Polymerase Kit (Beckman Coulter; A25395) was used according to the GeXP Chemistry Protocol (The Beckman Coulter Capillary Electrophoresis product line including GeXP chemistry is now distributed and supported by AB SCIEX, Landwehrstr. 54, 64293 Darmstadt, Germany (www.SCIEX.com)). cDNA synthesis and polymerase chain reaction were performed in half of the recommended volume leaving the recommended concentrations unchanged. Twenty ng of total RNA was used for each reverse transcription (see below). Controls without template were included to check the kit components for contaminating nucleic acids and controls without reverse transcriptase were used to check for contaminating DNA in the RNA samples.

Primers used

Forward and reverse gene-specific primers for a set of 35 genes were originally desinged by [6] to generate amplified DNA fragments with similar GC contents and melting temperature of different lengths ranging between 114 and 357 bp (Table S1). Each primer contained a universal priming sequence at the 5'- end and a gene-specific sequence at the 3'-end. The concentration of each gene-specific reverse primer in the pre-mixed reverse primer plex was experimentally adjusted to give a fluorescence signal for each fragment that was in the linear sensitivity range of the measurement system. The attenuated concentration of each reverse primer was adjusted to a value between 0.5 to $0.00025 \,\mu$ M (Table S1).

cDNA synthesis

The RT reaction was performed in 96-well plates in a Biometra (T Professional) thermal cycler with 20 ng of total RNA (except for the standard curves), 1 x RT Buffer, 1 μ L of reverse primer mix, 1 μ L KAN^r RNA and 0.5 μ L reverse transcriptase (20 U/ μ l) in a total volume of 10 μ L with the following temperature program: 1 min at 48 °C, 5 min at 37 °C, 60 min at 42 °C, 5 min at 95 °C, hold at 4 °C. The plate was briefly centrifuged to recollect the sample volume at the bottom of the wells.

Polymerase chain reaction

4.7 μ L of the 10 μ L RT reaction were used as template for the PCR. The PCR was performed with 0.35 μ L thermo-start DNA polymerase, 5 mM MgCl₂, 1 x PCR buffer and pre-mixed forward primer plex (20 nM final concentration) in a volume of 10 μ L under the conditions: 10 min at 95 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 55 °C, and 1 min at 68 °C; hold at 4 °C

Fragment analysis

Amplification products were separated and quantified according to the manufacturer's instructions. Briefly, the PCR products were pre-diluted 20-fold with 10 mM Tris-HCl buffer pH 8.0. One μ L of the diluted sample was separated with 39 μ L Sample Loading Solution (Beckman Coulter) and 0.4 μ L DNA size standard-400 (Beckman Coulter) on the GenomeLab Separation Capillary Array of the CEQ8800 (Beckman Coulter) using the Frag-3 separation program (denaturation for 120 s at 90 °C; sample injection for 30 s at 2.0 kV; separation of

Calibration curves

A calibration curve for each transcript *i* was established to calculate its relative concentration in the analysed samples from the corresponding normalized fluorescence signal values integrated over the peak areas of the respective amplified DNA fragments as separated on the capillary sequencer. To obtain the calibration curve, total RNA samples from dark- and light-treated plasmodia were mixed as described in the Results section to give an appropriate analytical standard. The mix was repeatedly diluted, two-fold in each step, to obtain a set of concentrations of total RNA ranging from 100 ng to 0.78 ng per reaction volume. Each sample was measured twice with two independent RT-PCR-reactions. The relative fluorescence signal F_i (which is identical to the normalized peak area NPA) as measured for each DNA fragment amplified from the mRNA of gene *i* was obtained by normalizing its peak area relative to the peak area of the KAN^r control RNA which served as internal standard. The logarithm of the normalized mean values was plotted against the logarithm of the relative total RNA concentration as obtained in the dilution series [7]. By taking the logarithm for the base 10, a straight line in the form of log $F_i = a_i \cdot \log[\text{RNA}]_{tot} + b_i$ was computed by linear regression for the n data points of each fragment amplified from the mRNA of gene i (1st linear regression) according to

$$a_{i} = \frac{\sum_{m=1}^{n} (\log[\text{RNA}]^{m}_{tot} - \overline{\log[\text{RNA}]_{tot}}) \cdot (\log F_{m,i} - \overline{\log F_{i}})}{\sum_{m=1}^{n} (\log[\text{RNA}]^{m}_{tot} - \overline{\log[\text{RNA}]_{tot}})^{2}}$$

The slopes a_i obtained for all gene-specific fragments in the 1st linear regression were averaged and the data points for each specific fragment amplified from the mRNA of gene *i* were again fitted to a straight line by a 2nd linear regression that computes the intercept with the ordinate $b_{i\overline{a}}$ which is obtained for the average slope $\overline{a} = \sum_{i=1}^{35} (a_i)$ according to

 $b_{i\overline{a}} = \overline{\log F_i} - \overline{a} \cdot \overline{\log[\text{RNA}]_{tot}}.$

Fitting parameters for the 1st and for the 2nd linear regression were stored for documentation and quality control purposes. Data processing was done with the Perl scripts provided in conjunction with this paper (for details see legends to Figs. S4 and S5). The data points of all 35 fragments fitted the calibration curves without any systematic deviation from the straight lines. The relative concentration [RNA]_i^{rel} for each transcript of gene *i* was calculated on the basis of its specific calibration curve as [RNA]_i^{rel} = 10^A with $A = \frac{\log F_i - b_i}{\overline{\sigma}}$.

Abbreviations

Bp: Base pair; cDNA: Complementary DNA; GeXP: Gene eXpression Profiling; KAN^r: Kanamycin resistance gene; L: Liter; PCR: Polymerase Chain Reaction; rpm: Rotations per minute; RT-PCR: Reverse Transcription Polymerase Chain Reaction

Supplementary Information

The online version contains supplementary material available at https://doi. org/10.1186/s12896-021-00689-4.

Additional file 1.

Additional file 2 Table S1. Sequences and concentrations of primers used for GeXP-RT-PCR.

Additional file 3 Fig. S6. Comparative display of single cell gene expression time series. The figure legend is provided as part of Additional file 1.

Acknowledgments

We thank PD Dr. Marcus Hauser for discussion and valuable suggestions, Dr. Xenia Hoffmann for help with the literature search and Drs. Manfred Souquet and Marcus Hauser for critical reading of the manuscript.

Authors' contributions

All authors contributed to the development of the method. MH and WM conceived of the study and participated in its design and coordination. BW, VR and PM performed measurements. MH designed the data processing workflow and wrote the Perl scripts. WM developed the kinetic model. PM revised the model and performed model calculations. PM and BW prepared some of the Figs. WM wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was partly funded by the Deutsche Forschungsgemeinschaft (DFG; MA 1516/10–1). BW and PM were financially supported by the International Max Planck Research School through the Research Center Dynamic Systems, Magdeburg, funded by the state of Saxony Anhalt. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The bioinformatics pipeline used for data evaluation is available from https://github.com/markushaas/genexpro4ceq8800

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have not competing interests. EKF-diagnostic GmbH and Novartis were not involved in this study and did not provide any financial support.

Author details

¹Magdeburg Centre for Systems Biology and Institute for Biology, Otto-von-Guericke Universität, Magdeburg, Germany. ²Present address: Institut für Medizinische Mikrobiologie und Krankenhaushygiene, Otto-von-Guericke Universität, Magdeburg, Germany. ³Present address: EKF-diagnostic GmbH, Ebendorfer Chaussee 3, 39179 Barleben, Germany. ⁴Present address: Novartis Technical Operation – Solides, Barleben, Otto-von-Guericke Allee 1, 39179 Barleben, Germany.

Received: 14 October 2020 Accepted: 23 March 2021 Published online: 14 April 2021

References

- Roth CM. Quantifying gene expression. Curr Issues Mol Biol. 2002;4(3):93– 100.
- VanGuilder H, Vrana K, Freeman W. Twenty-five years of quantitative PCR for gene expression analysis. BioTechniques. 2008;44(Supplement):619–26.
- Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010;11(10):R106. https://doi.org/10.1186/gb-2010-11-10-r106.
- 4. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:31–21.
- Hayashi E, et al. Multiplexed, quantitative gene expression analysis for lettuce seed germination on GenomeLabTM GeXP genetic analysis system, Beckman Coulter Application Information A-10295A; https://ls.beckma ncoulter.co.jp/files/ appli_note/A_10295A.pdf. Accessed 8 Apr 2021.
- Hoffmann X-K, Tesmer J, Souquet M, Marwan W. Futile attempts to differentiate provide molecular evidence for individual differences within a population of cells during cellular reprogramming. FEMS Microbiol Lett. 2012;329(1):78–86. https://doi.org/10.1111/j.1574-6968.2012.02506.x.
- Rai AJ, Kamath RM, Gerald W, Fleisher M. Analytical validation of the GeXP analyzer and design of a workflow for cancer-biomarker discovery using multiplexed gene-expression profiling. Anal Bioanal Chem. 2009;393(5): 1505–11. https://doi.org/10.1007/s00216-008-2436-7.
- Schaap P, et al. The *Physarum polycephalum* genome reveals extensive use of prokaryotic two-component and metazoan-type tyrosine kinase signaling. Genome Biol Evol. 2016;8:109–25.
- Marmur J, Rownd R, Schildkraut CL. Denaturation and renaturation of deoxyribonucleic acid. In: Davidson JN, Cohn WE, editors. Progress in Nucleic Acid Research. New York: Academic Press, Inc.; 1963.
- Wetmur JG, Davidson N. Kinetics of renaturation of DNA. J Mol Biol. 1968; 31(3):349–70. https://doi.org/10.1016/0022-2836(68)90414-2.
- Lamparter T, Marwan W. Spectroscopic detection of a phytochrome-like photoreceptor in the myxomycete *Physarum polycephalum* and the kinetic mechanism for the photocontrol of sporulation by P_{fr}. Photochem Photobiol. 2001;73(6):697–702. https://doi.org/10.1562/0031-8655(2001)073< 0697:SDOAPL>2.0.CO;2.
- Starostzik C, Marwan W. A photoreceptor with characteristics of phytochrome triggers sporulation in the true slime mould *Physarum polycephalum*. FEBS Lett. 1995;370(1-2):146–8. https://doi.org/10.1016/0014-5793(95)00820-Y.
- Barrantes I, Glöckner G, Meyer S, Marwan W. Transcriptomic changes arising during light-induced sporulation in *Physarum polycephalum*. BMC Genomics. 2010;11(1):115. https://doi.org/10.1186/1471-2164-11-115.
- Barrantes I, Leipzig J, Marwan W. A next-generation sequencing approach to study the transcriptomic changes during the differentiation of *Physarum* at the single-cell level. Gene Regulation Systems Biol. 2012;6:127–37. https:// doi.org/10.4137/GRSB.S10224.
- Rätzel V, Marwan W. Gene expression kinetics in individual plasmodial cells reveal alternative programs of differential regulation during commitment and differentiation. Develop Growth Differ. 2015;57(5):408–20. https://doi. org/10.1111/dgd.12220.
- Rätzel V, Ebeling B, Hoffmann XK, Tesmer J, Marwan W. *Physarum polycephalum* mutants in the photocontrol of sporulation display altered patterns in the correlated expression of developmentally regulated genes. Develop Growth Differ. 2013;55(2):247–59. https://doi.org/10.1111/dgd.12029.
- Walter P, Hoffmann X-K, Ebeling B, Haas M, Marwan W. Switch-like reprogramming of gene expression after fusion of multinucleate plasmodial cells of two *Physarum polycephalum* sporulation mutants. Biochem Biophys Res Commun. 2013;435(1):88–93. https://doi.org/10.1016/j.bbrc.2013.04.043.
- Werthmann B, Marwan W. Developmental switching in *Physarum* polycephalum: Petri net analysis of single cell trajectories of gene expression

indicates responsiveness and genetic plasticity of the Waddington quasipotential landscape. J Phys D Appl Phys. 2017;50(46):464003 (464018pp). https://doi.org/10.1088/1361-6463/aa8e2b.

- Rätzel V, Werthmann B, Haas M, Strube J, Marwan W. Disentangling a complex response in cell reprogramming and probing the Waddington landscape by automatic construction of Petri nets. BioSystems. 2020;189: 104092 (104018pp). https://doi.org/10.1016/j.biosystems.2019.104092.
- Pretschner A, Pabel S, Haas M, Heiner M, Marwan W. Regulatory dynamics of cell differentiation revealed by true time series from multinucleate single cells. Front Genet. 2021;11:612256. https://doi.org/10.3389/fgene.2020.612256.
- Starostzik C, Marwan W. Kinetic analysis of a signal transduction pathway by time-resolved somatic complementation of mutants. J Exp Biol. 1998;201(Pt 13):1991–9.
- 22. Daniel JW, Baldwin HH. Methods for culture of plasmodial myxomycetes. Methods Cell Physiol. 1964;1:9–41.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- · thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

