

Real-time RT-PCR: Neue Ansätze zur exakten mRNA Quantifizierung

Michael Walter Pfaffl, Lehrstuhl für Physiologie, Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung & Umwelt, Freising

Die molekularen Technologien Genomics, Transcriptomics und Proteomics erobern immer mehr die klassischen Forschungsgebiete der Biowissenschaften. Die enorme Flut an gewonnenen Daten und Ergebnissen ist von überproportionalem Nutzen in der molekularen Diagnostik und Physiologie sowie die „Functional Genomics“. Immer neue ausgeklügelte Methoden und Anwendungen sind daher nötig um komplexe physiologische Vorgänge zu beschreiben. Da wir uns erst an Anfang dieser molekularen Ära befinden, ist es notwendig diese Techniken zu optimieren und komplett zu verstehen.

Eine dieser technisch ausgefeilten Methoden zur zuverlässigen und exakten Quantifizierung spezifischer mRNA, stellt die real-time RT-PCR dar. Dieser Artikel beschreibt im Wesentlichen die effizienz-korrigierte relative Quantifizierung, die Normalisierung der Expressionsergebnisse anhand eines nicht regulierten „Housekeeping Gens“, die Berechnung der real-time PCR Effizienz sowie die Verrechnung und statistische Auswertung der Expressionsergebnisse.

Alle beschriebenen Themenkomplexe können im Detail auf der korrespondierenden Internetseite (<http://www.wzw.tum.de/gene-quantification/>) in internationalen publizierten Originalarbeiten nachgeschlagen werden.

► Die Anwendung der Reversen Transkription (RT) mit folgender Polymerase Kettenreaktion (PCR) zum Nachweis einer spezifischen mRNA ist heute ein Routinewerkzeug in der Molekularbiologie. Der Nachweis kann qualitativ in klassischen RT-PCR Blockcyclern oder quantitativ mittels Echtzeit RT-PCR (real-time RT-PCR oder qRT-PCR) durchgeführt werden. Voraussetzung für einen zuverlässigen quantitativen Nachweis ist eine funktionierende mRNA Analytik, die exakte mRNA Quantifizierungsergebnisse bei ausreichender Genauigkeit und hoher Wiederholbarkeit liefert^[1]. Es stehen zwei generelle Quantifizierungsstrategien in der real-time RT-PCR zur Verfügung (Abb. 1):

Die **absolute Quantifizierung** wird anhand einer gegebenen Kalibrierkurve durchgeführt^[2], basierend auf einer Verdünnungsreihe von RT-PCR Produkten, Plasmid DNA, *in vitro* transkribierter RNA, synthetisierter DNA oder RNA Oligomere.

Bei der **relativen Quantifizierung** wird die Genexpression eines Zielgens auf ein weiteres nicht reguliertes „Housekeeping Gen“ (HKG) oder auf einen HKG Index, der sich aus mehreren HKG zusammensetzt, bezogen (Abb. 1). Man nennt diesen Vorgang auch

Normalisierung der Expressionsergebnisse^[3]. Die relative Quantifizierung lässt sich weiter optimieren, indem man die unterschiedlichen real-time PCR Effizienzen der untersuchten Faktoren mit berücksichtigt. Die effizienz-korrigierte relative Quantifizierung mittels real-time RT-PCR stellt bis dato die genaueste Form der mRNA Quantifizierung dar^[4].

Crossing Point

In der real-time PCR wird heute nicht mehr primär in DNA Produktmengen oder -konzentrationen gerechnet, sondern als Maß für die Quantifizierung der Startmenge werden die sog. Ct oder CP (= Crossing Point) Werte herangezogen. Sie entsprechen der Anzahl der PCR Zyklen die nötig sind um ein konstant definiertes Fluoreszenzniveau zu erreichen. Am CP befindet sich in allen Reaktionsgefäßen die gleiche Menge an neu synthetisierter DNA. Im Falle einer 100% Effizienz der PCR verdoppelt sich mit jedem Zyklus die DNA Produktmenge und analog dazu das Fluoreszenzsignal. Ein um eine Einheit geringerer CP entspricht der doppelten Menge an eingesetzten cDNA; respektive mRNA Startmenge.

Quantifizierungsstrategien in der real-time RT-PCR

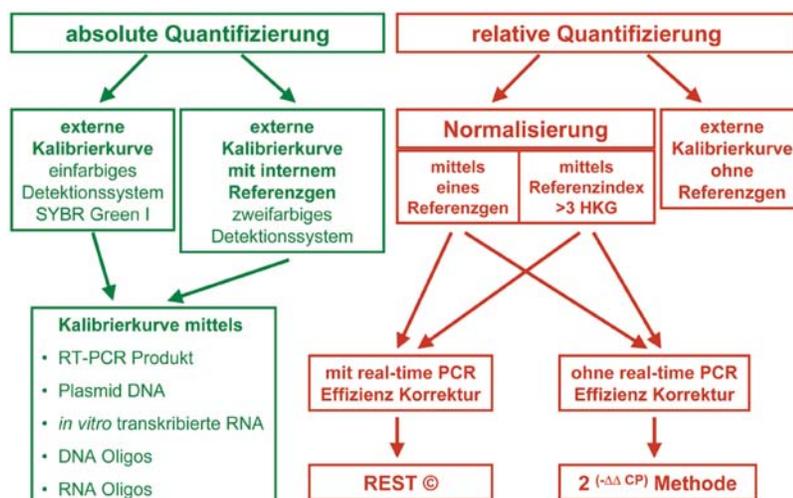


Abb. 1: Quantifizierungsstrategien in der quantitativen real-time RT-PCR.

Relative Quantifizierung

Bei der relativen Quantifizierung wird die Expression der Zielgene mit der eines nicht regulierten HKG normalisiert. Dabei werden nicht die absoluten Startkopienzahlen oder -konzentrationen bestimmt, sondern die Expression des zu untersuchenden Gens wird auf ein zweites, ubiquitär und homogen exprimiertes Gen bezogen. Die Vorteile der Normalisierung liegen in der Reduzierung der Varianz der Expressionsergebnisse, da Gewebe- und Matrixeffekte, unterschiedliche RNA Extraktionseffizienzen sowie Fehler bei der RT innerhalb einer experimentellen Probe gleichermaßen das Zielgen und das HKG betreffen. In den folgenden Berechnungen des Expressionsunterschiedes heben sich diese individuellen Probeneffekte wieder auf. GAPDH, ribosomale Untereinheiten (18S und 28S), Ubiquitin, Histon Untereinheiten und β -Aktin sind wohl die geläufigsten HKG die zur Normalisierung herangezogen werden. Die relative Expression des zu untersuchenden Gens in den behandelten experimentellen Proben wird auf ein Kontrollprobenmaterial bezogen. Die Berechnung des Expressionsunterschiedes (Ratio) kann über die sog. $\Delta\Delta$ CP Methode oder über genauere Effizienz korrigierte Modelle erfolgen. Dabei wird im ersten Schritt für jede untersuchte Probe der CP Wert des Referenzgens vom CP Wert des zu untersuchenden Gens subtrahiert (Δ CP = CP Zielgen – CP Referenzgen). Nach dieser Normierung wird vom Δ CP Wert der experimentell behandelten Proben der Δ CP Wert einer Kontrolle abgezogen ($\Delta\Delta$ CP); man kommt zum sog. „delta-delta CT“ Berechnungsmodell. Der relative Expressionsunterschied einer Probe zwischen der Behandlung und der Kontrolle (Ratio), normalisiert zum Referenzgen und bezogen auf eine Standardprobe, ergibt sich aus der arithmetischen Formel $2^{-\Delta\Delta$ CP}[5].

$$\begin{aligned} \Delta\text{CP} &= \text{CP Zielgen} - \text{CP Referenzgen} \\ \Delta\Delta\text{CP} &= \Delta\text{CP Behandlung} - \Delta\text{CP Kontrolle} \\ \text{Ratio} &= 2^{-\Delta\Delta\text{CP}} \end{aligned}$$

Dieses Berechnungsschema setzt eine Verdoppelung der DNA Menge in jedem Zyklus voraus. Man geht von einer optimalen real-time PCR Effizienz in allen Proben aus, was natürlich nicht der Praxis entspricht. Die „wahre real-time PCR Effizienz“ dürfte sich unter optimierten Reaktionsbedingungen im Bereich von 1,7 bis 1,9 bewegen und weist Schwankungsbreiten von 1,5 bis über 2,0 auf. Geringste Schwankungen in der Effizienzen von Zielgen zu Referenzgen führen zu enormen Unterschieden in den Expressionsunterschieden[6]. Deswegen sind Berechnungsmodelle entwickelt worden die der unterschiedlichen Effizienzen in den ex-

perimentellen Proben Rechnung trägt. Ausgehend vom oberen „delta-delta CT“ Modell wurde das effizienz-korrigierte relative Quantifizierungsmodell entwickelt[4].

$$\text{Ratio} = \frac{(E_{\text{Zielgen}})^{\Delta\text{CP}_{\text{Zielgen}}(\text{Kontrolle} - \text{Behandlung})}}{(E_{\text{Referenzgen}})^{\Delta\text{CP}_{\text{Referenzgen}}(\text{Kontrolle} - \text{Behandlung})}}$$

Das Berechnungsmodell setzt sich aus der Berechnung des Expressionsunterschiedes zwischen Behandlung und Kontrolle im Zielgen im Zähler und aus der Berechnung des Expressionsunterschiedes des Referenzgens im Nenner zusammen. Idealerweise ist das Referenzgen nicht reguliert und sowohl in der Behandlung als auch in der Kontrolle sind die CPs identisch, d.h. die CPs heben sich in der Berechnungsformel auf. Der Nenner wird gleich eins, das Referenzgen fällt heraus und die berechnete Ratio ist nur von den Expressionsunterschieden des Zielgens abhängig.

$$\text{Ratio} = \frac{(E_{\text{Zielgen}})^{\Delta\text{CP}_{\text{Zielgen}}(\text{Kontrolle} - \text{Behandlung})}}{(E_{\text{Referenzgen}})^0}$$

$$\text{Ratio} = (E_{\text{Zielgen}})^{\Delta\text{CP}_{\text{Zielgen}}(\text{Kontrolle} - \text{Behandlung})}$$

Berechnungssoftware – REST

Hat man nun, wie üblich in einem wissenschaftlichen Experiment, mehrere Probanden oder Wiederholungen die behandelt sind und weitere die als Kontrolle dienen, so kommt mit dem oben beschriebenen relativen Modell (n = 1) nicht mehr weiter. Man muss die Gruppe der Behandelten als auch die Gruppe der Kontrollen zusammenfassen. Dies führte dazu eine Software zu schreiben die dem Problem Rechnung trägt und die Expressionsunterschiede von zwei unterschiedlich großen Gruppen berechnet als auch statistisch ausgewertet (REST = Relative Expression Software Tool)[7]. Diese Software fasst die Gruppen als Mittelwerte zusammen und berechnet daraus die mittleren Expressionsunterschiede (R) der Gruppen, normalisiert über ein bestimmtes HKG, sowie die Varianzen der einzelnen Expressionen. Die Ergebnisse werden statistisch getestet und die Expressionsunterschiede werden in einem Ausgabefenster mit ihren Signifikanzlevels ausgegeben. Die Statistik basiert auf einem sehr robusten und von einer Normalverteilung unabhängigen Randomisierungstest, bei dem beliebig viele tausend Randomisierungen und Wiederholungen durchgeführt werden können[7].

$$R = \frac{(E_{\text{Zielgen}})^{\Delta\text{CP}_{\text{Zielgen}}(\text{MW Kontrolle} - \text{MW Behandlung})}}{(E_{\text{Referenzgen}})^{\Delta\text{CP}_{\text{Referenzgen}}(\text{MW Kontrolle} - \text{MW Behandlung})}}$$





Real-time PCR Effizienzberechnung

Die PCR Effizienz eines bestimmten Faktors lässt sich durch verschiedenste Methoden bestimmen. Am häufigsten wird eine Standardkurve aus unterschiedlichen Verdünnungsstufen erstellt^[3], aus dessen Steigung die real-time PCR Effizienz bestimmt werden kann (Abb. 2a). Hat man genügend cDNA (respektive mRNA) Startmatrize zur Verfügung, so verwendet man dafür eine probenindivi-

duelle cDNA Verdünnungsreihe. Falls nicht wird die Verdünnungsreihe aus einer repräsentativ Mischprobe, die sich aus allen verwendeten experimentellen cDNA Proben eines Gewebes zusammensetzt, erstellt. Die eingesetzte Menge an cDNA (mRNA) wird in einer logarithmischen Funktion gegen die Zykluszahl dem CP dargestellt. Die Effizienz berechnet sich nach der Formel: $E = 10^{-1/\text{Steigung}}$ ^[3]. Diese Methode ist sehr arbeitsaufwendig und weist in der Mischprobe

stimmung erhebliche Mängel auf, weil sie nicht auf probenindividuelle Unterschiede eingeht. Im Übrigen überschätzt sie die eigentliche real-time PCR Effizienz. Nicht selten errechnen sich Effizienzen über 2,0 bis 2,2, was theoretisch nicht plausibel ist.

Eine zweite Möglichkeit bietet die Berechnung der Effizienz aus dem absoluten Fluoreszenzanstieg (Abb. 2b) in der exponentiellen Phase^[8,9]. Die exponentielle Phase wird in einer linearen Regression probenindividuell zusammengefasst. Die Steigung der Regression entspricht der PCR Effizienz. Diese Methode unterliegt großen Schwankungen, da der Wissenschaftler selbst entscheidet welche Punkte er in die Effizienz Berechnung mit einbezieht und welche nicht. Bei dieser Methode liegen die Effizienzen von 1,3 bis 1,6 und unterschätzt in unseren Augen die wahre Effizienz.

Eine weitere Methode ist die Berechnung anhand mathematischer Algorithmen innerhalb einer einzelnen experimentellen Probe. Dies hat den Vorteil, dass man für jede Probe eine individuelle PCR Effizienz erhält und diese in der Effizienz-korrigierten relativen Quantifizierung berücksichtigen kann. Dabei werden nicht nur einzelne arbiträr ausgesuchte Zyklen zur Berechnung herangezogen, sondern die Fluoreszenz Kurve, die alle 40 oder mehr Zyklen berücksichtigt (Abb. 2c). Der Fluoreszenzverlauf wird an ein 4-parametrisches sigmoidales oder logistisches Modell mit sehr hoher Korrelation ($r = 0.99$) angepasst und mittels der Variablen **b** kann die wahre Effizienz geschätzt werden^[10]. Bei dieser Methode erfasst man die Effizienz nur im Wendepunkt und die Variable **b** gibt nicht die eigentliche real-time PCR Effizienz an. Man muss sie noch umrechnen und erhält dabei ähnliche niedrige Effizienzen von 1,3 bis 1,7 und unterschätzt wiederum die wahre Effizienz.

Die innovativste Methode der Effizienzberechnung basiert auf einer Berechnung der Effizienz anhand eines exponentiellen Modells, was der Kinetik der PCR am besten entspricht (Abb. 3)^[11]. Das gesamte Berechnungsschema setzt sich aus mehreren Schritten zusammen: Als erstes wird die Hintergrundfluoreszenz anhand einer linearen Regression bestimmt. Mittels eines „standardized residual“ Algorithmus werden die Fluoreszenzpunkte gesucht, die nicht mehr auf dem leicht steigenden linearen Trend der Hintergrundfluoreszenz liegen (Abb. 3 Inlay). Dies sind die ersten CP, die sich anhand eines neuen Trend, einer exponentiellen PCR Funktion, verhalten. Nun sucht man die CP die nicht mehr dem exponentiellen Verhalten der PCR Reaktion entsprechen. Dafür werden alle CP in ein logistisches Modell mit einbezogen und das zweite Maximum (second Derivative Maxi-

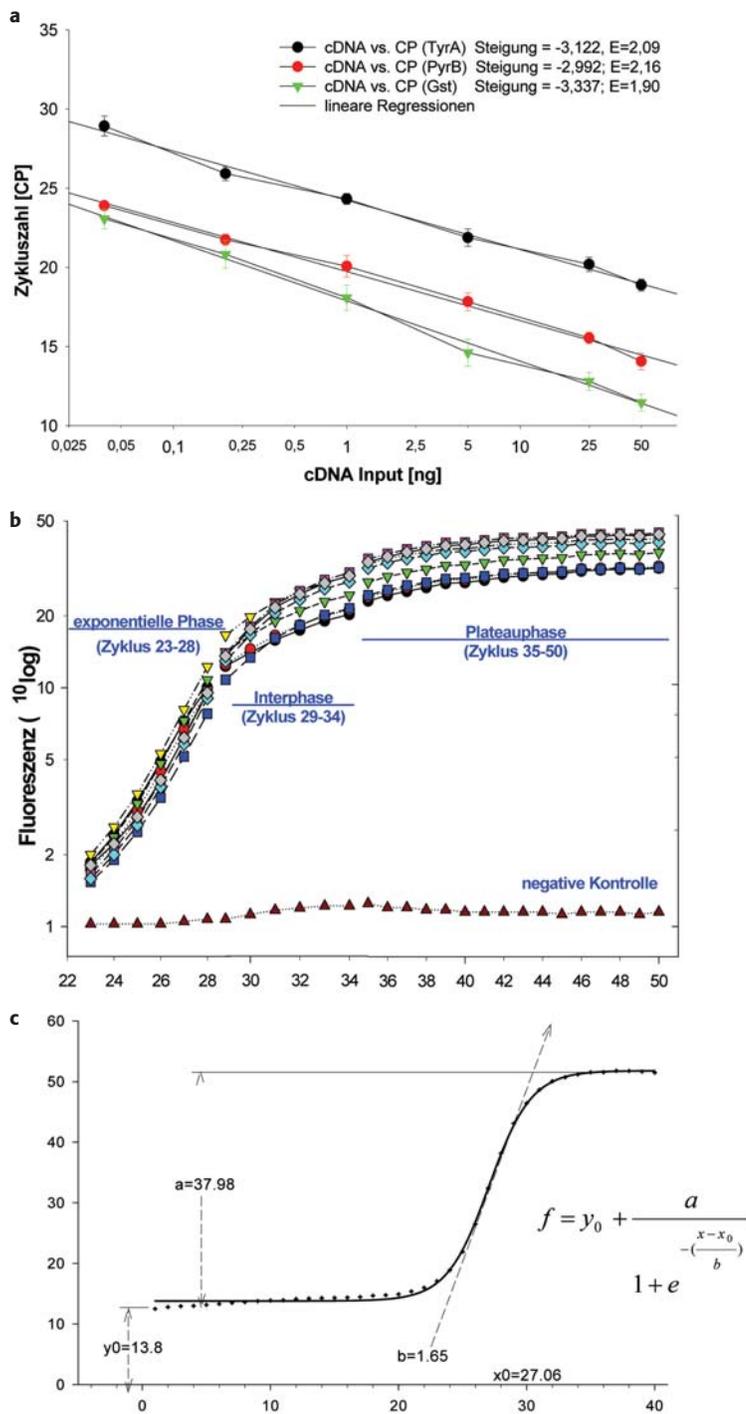


Abb. 2: Berechnungsmodelle für die real-time PCR Effizienz. 2a: Berechnung anhand einer Verdünnungsreihe; 2b: Berechnung anhand des absoluten Fluoreszenzanstiegs; 2c: Berechnung anhand mathematischer Algorithmen (y_0 = Hintergrundfluoreszenz; a = maximale Fluoreszenz - y_0 ; x_0 = erstes Maximum = Wendepunkt der Funktion; b = Steigung der Gerade in x_0).

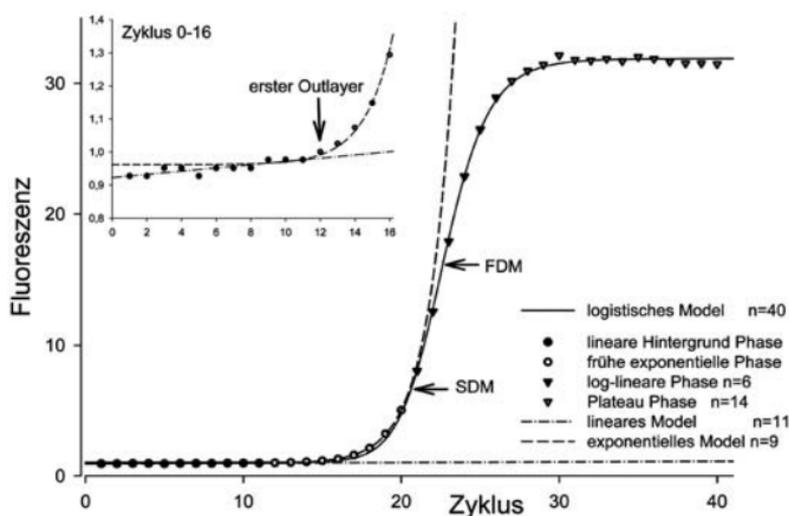


Abb. 3: Standardisierte Berechnung der real-time PCR Effizienz in einem einzigen Reaktionsansatz anhand eines exponentiellen Berechnungsmodell.

mum = SDM) berechnet. Ab dem SDM nimmt die Fluoreszenz nicht mehr im exponentiellen Trend zu. Der letzte CP vor dem SDM ist auch der letzte Punkt mit exponentiellem Fluoreszenzzuwachs und bildet den letzten Datenpunkt im exponentiellen PCR Modell. Im Allgemeinen werden 6-10 CP zur Berechnung der Effizienz herangezogen. Das vorgestellte Modell zeigt Effizienzen von 1,65 bis 1,90 und im Vergleich mit anderen Effizienz Berechnungsmodellen die geringste Varianz und höchste Genauigkeit^[11].

In Zukunft soll eine Software erarbeitet werden, die die Effizienz innerhalb einer einzelnen experimentellen Probe aus den Fluoreszenzrohdaten der real-time Plattform berechnet und dann in eine neue effizienz-korrigierte Quantifizierungs- und Statistiksoftware mit einbezieht.

Literatur

[1] **Bustin, S.A.** (2000): Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol.*, 25: 169-193

[2] **Pfaffl, M.W., and Hageleit, M.**, (2001): Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnology Letters*, 23: 275-282

[3] **Rasmussen, R.** (2001): Quantification on the LightCycler. In: Meuer, S, Wittwer, C, and Nakagawara, K, eds. *Rapid Cycle Real-time PCR, Methods and Applications* Springer Press, Heidelberg; ISBN 3-540-66736-9, 21-34

[4] **Pfaffl, M.W.** (2001): A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, 29(9): e45

[5] **Livak, K.J., and Schmittgen, T.D.** (2001): Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ Method. *Methods.*, 25(4): 402-408

[6] **LightCycler Relative Quantification Software** (2007): Version 1.0, Roche Molecular Biochemicals

[7] **Pfaffl, M.W., Horgan, G.W., and Dempfle, L.** (2002): Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.*, 30(9): e36

[8] **Peccoud, J., and Jacob, C.** (1996): Theoretical uncertainty of measurements using quantitative polymerase chain reaction. *Biophys J.*, 71(1): 101-108

[9] **Pfaffl, M.W.** (2007): Development and validation of an externally standardised quantitative Insulin like growth factor-1 (IGF-1) RT-PCR using LightCycler SYBR Green I technology. In: Meuer, S., Wittwer, C., and Nakagawara, K., eds.: *Rapid Cycle Real-time PCR, Methods and Applications* Springer Press, Heidelberg, ISBN 3-540-66736-9, 281-191

[10] **Tichopad, A., Dzidic, A., and Pfaffl, M.W.** (2003): Improving quantitative real-time RT-PCR reproducibility by boosting primer-linked amplification efficiency. *Biotechnology Letters* 24: 2053-2056

[11] **Tichopad, A., Dilger, M., Schwarz, G., and Pfaffl, M.W.** (2003): Real-time PCR efficiency standardized estimation from single reaction setup. *Nucleic Acids Res.*, 31(20): e122

Korrespondenzadresse:

PD Dr. Michael Walter Pfaffl
Lehrstuhl für Physiologie
Wissenschaftszentrum
Weihenstephan für Ernährung,
Landnutzung & Umwelt
Weihenstephaner Berg 3
D-85354 Freising
Tel.: 08161-713511
Fax: 08161-714204
michael.pfaffl@wzw.tum.de
www.wzw.tum.de/gene-
quantification/