How to choose reference genes for gene expression analysis

QIAcuity digital PCR (dPCR) uses partition-based absolute quantification to reduce analytical variability and eliminate standard curves, providing more precise and reproducible measurements. However, both real-time PCR (qPCR) and dPCR gene expression experiments are influenced by variation in sample preparation, RNA quality, RNA input, and reverse transcription efficiency. Therefore, the use of reference genes is a fundamental requirement for qPCR and dPCR gene expression analyses (1, 2, 3).

Comparison of gene expression levels between samples is typically performed by normalization to reference genes whose expression is minimally affected by experimental conditions or sample types. Reference genes help account for technical variability, enabling more accurate and reliable measurement of expression differences arising from true biological effects. Despite their critical role in data quality, the importance of reference gene selection is often downplayed by the use of a single reference gene that is – without evidence – merely assumed to be sufficiently stable for the intended experiment.

To support proper reference gene selection and validation, preselected and wet-lab-tested QIAcuity dPCR RNA-Plex reference gene assays (cat.no. 250336, 250337) can be combined with genes of interest, providing a strong starting point for reliable normalization. The following five-step guide helps you select suitable reference gene candidates and confirm their stability under your experimental conditions, so you can incorporate the appropriate normalization into your gene expression analysis.

A simple five-step recipe

Here we present a step-by-step guide for reference gene selection and initial gene expression normalization.

1. Define the complete sample set and experimental conditions.
2. Build a data-driven reference gene candidate list.
3. Empirically evaluate candidate stability in representative samples.
4. Identify a suitable set of reference genes.
5. Normalize gene of interest quantities to the reference gene quantities.

Step 1: Define your sample set

The stability of a particular reference gene depends on the specific sample types and experimental conditions. The initial search should focus on candidate reference genes expected to be stably expressed under these conditions, while avoiding those involved in cellular processes likely to be affected by the experiment. To select the right reference gene, one should keep in mind that an expression level stable in liver may be unstable or not detectable in heart muscle. A gene stably expressed in healthy tissue may change in response to disease, treatment or over time.

Ask yourself:

- Which species are the samples from?
- What sample types are included?
- Are samples healthy or diseased?
- What treatments or time course define the experimental groups?

Best-practice tip: Choose reference genes for your experiment, not for biology in general.

Step 2: Build a shortlist of candidates

A list of candidate reference genes provides a starting point for evaluating potential options. The initial set of candidates is later refined down to the best-fitting reference genes for use in experimental testing.

Create a list of candidate reference genes using:

- RNA-seq datasets (4, 5)
- Transcriptomic atlases
- Published reference gene resources (6)
- Pre-qualified reference gene panels like our selection of candidate reference genes, which come with corresponding data on expected robustness and expression levels (Table 1).

A good candidate should:

- Show low variation across samples and conditions investigated
- Not be involved in pathways related to the genes of interest
- Not be co-regulated with other candidate reference genes
- Have a comparable expression level to the genes of interest

Many researchers default to selecting frequently used "traditional" reference genes such as GAPDH, ACTB, HPRT1, 18S, GUSB, B2M, TBP, SDHA or RPLP0 under the assumption that these genes are also suitable for their particular experiments. Yet many of these genes have become popular choices only because they are highly expressed across many cell types and perform essential cellular functions. However, a gene that is highly abundant and thereby universally expressed does not necessarily mean that it is stably expressed across tissues or conditions (7, 8, 9). Transcriptome datasets show that common 'housekeeping' genes often vary substantially between tissues (4, 5, 6).

A traditional reference gene should only be used if it is known to be stable in your specific sample set.

Best-practice tip: Use evidence, not habit, to select reference genes.

Step 3: Measure candidate reference gene stability

Empirical evaluation provides confidence that selected reference genes are stable within the samples and conditions being investigated. This step provides the data to narrow the initial candidate list down to the final selection of reference genes.

Measure

Measure the expression level of candidate reference genes (defined in Step 2) in a representative set of samples and conditions for your experiment (defined in Step 1).

Review

Compare expression levels of candidate reference genes in the tested samples. Check for reference genes that show large variations between samples or systematic differences between experimental groups or conditions. These issues could indicate instability of the candidate reference gene for the relevant experimental conditions. Also, check for genes that yield very low concentration measurements, which can introduce unwanted variability.

Rank

Rank candidate stability using appropriate tools such as geNorm (10), NormFinder (11), or BestKeeper (12). Candidates that receive higher stability scores from more than one tool can provide higher confidence in their selection. Because these tools use different assumptions, their rankings should not be interpreted in isolation. Interpret their rankings together with the requirements of your experimental design and the observed expression patterns within the tested representative samples.

Best-practice tip: Trust, but verify.

Step 4: Select multiple reference genes when feasible

The accuracy and validity of the normalization depend on the right choice of reference genes.

Guidelines recommend using multiple reference genes for normalization. If a single reference gene is used, changes in its expression can be easily mistaken for changes in the expression of the gene of interest. Using multiple validated reference genes helps to reduce this bias and improves the accuracy of relative expression estimates.

When selecting a set of reference genes to use, remember the following considerations:

- Including multiple reference genes generally increases experimental robustness by compensating for expression variations from an individual reference.
- More is not always better. The goal is to identify a sufficient set of stable reference genes. Including additional candidate references that have lower stability rankings may compromise accuracy or precision. The reference gene selection tools mentioned in Step 3 can help not only rank candidates but also define a sufficient set of reference genes to use.

Best-practice rule: One reference gene is an assumption; multiple reference genes give robustness.

Step 5: Normalize the expression data

The preceding steps focus on selecting reference genes for a gene expression experiment. This step discusses the use of the previously selected reference genes when performing data analysis.

The following steps summarize a normalization approach that is analogous to fold-change calculations typically performed in gene expression analysis with real-time PCR (10).

1) Analysis starts from the target quantities (i.e., target concentrations) obtained by digital PCR for the selected reference genes and genes of interest.
2) For each sample, calculate the geometric mean of the concentrations for the set of reference genes. The geometric mean is used rather than the arithmetic mean to minimize the impact of fluctuations for reference genes with outlying higher or lower expression levels compared to the other references (10).
3) For each gene of interest in the sample, calculate a normalized expression value. This value is obtained by the ratio of the gene of interest to the geometric mean of the reference genes. These normalized values represent relative measures of the gene of interest and help account for upstream technical variations that may not be captured when comparing samples using raw target concentration values.
4) Experiments often compare target abundances between different samples or relative to a control sample (e.g., treated vs untreated or different timepoints).
   The corresponding fold differences between samples are obtained from the ratio of the normalized expression values for a target gene.

Depending on experimental requirements, it can be helpful to report the normalized results in different ways.
This includes presenting results as percentages, ratios, fractional abundances or normalized copies (2), which requires adopting different normalization strategies.

Think of the reference genes like a tare on a scale

Reference genes establish a consistent baseline for comparing target abundances between different samples. They help distinguish changes in the target abundance levels that are due to biological effects rather than technical variations related to sample preparation and upstream manipulations.

Best-practice tip: Normalize consistently.

Quick Checklist

Before starting your experiment, ask:

- ☐ Did I consider all relevant samples and conditions for my experiment?
- ☐ Did I evaluate candidate reference genes using available expression data?
- ☐ Did I avoid reference genes linked to the biology of interest?
- ☐ Are my reference genes stably expressed in my samples and conditions?
- ☐ Am I using a sufficient number of reference genes?
- ☐ Are my reference genes expressed at appropriate levels relative to my genes of interest?
- ☐ Am I using the geometric mean of my reference genes for normalization?

If the answer is "yes" to these questions, the reference gene strategy will provide a solid foundation for your future dPCR gene expression analysis.

Table 1. Reference genes available in the QIAcuity dPCR RNA-Plex Assay portfolio.

Common gene name	Expression group	Human averaged NGS expression (nTPM)	Tissue stability (%)	Human GeneGlobe Assay ID	Mouse GeneGlobe Assay ID	Rat GeneGlobe Assay ID	Genes influencing the reference gene expression (IPA - QDI)	Other substances / conditions influencing reference gene expression (IPA - QDI)
RPS29	high	1915.170781	78.01	DRH0118164	DRM0114305	DRR0112408	MYC, CTNNB1, EGF, GATAD2B, LARP1	sirolimus, Immunoglobulin, GABA, severe acute respiratory syndrome coronavirus 2, dihydrotestosterone
RPL32	high	1371.794805	78.70	DRH0118161	DRM0114302	DRR0112405	MTOR, MLXIPL, POLG, MYC	lipopolysaccharide, LH, insulin, Influenza A virus (A/Japan/305/57(H2N2)), sirolimus, heavy metal
RPL18	high	1093.448645	74.60	DRH0118159	DRM0114300	DRR0079733	YAP1, NFE2L2, CTNNB1, MYC, MYCL, SYVN1	1,2-dithiole-3-thione, sirolimus, torin1, Hbx
DDX17	medium high	168.367290	82.68	DRH0118127	DRM0114268	DRR0112375	MTDH, YAP1, TP53, THRAP3	lipopolysaccharide, beta-estradiol, MYOPARR, Hbx, rebemadlin, EAhy926 cells
TRAPPC1	medium high	108.250971	87.88	DRH0118174	DRM0114314	DRR0112418	MRTFA, KLF3, AKT1, MRTFB	EWSR1-FLI1, levodopa
EIF3I	medium high	106.148211	79.92	DRH0118131	DRM0114272	DRR0112378	MYC, TGFBR2, AGT, KRAS, SUMO2, RUNX2, E2F1	tretinoin, MLN7243, CD 437
HNRNPH3	medium	65.172400	88.18	DRH0118136	DRM0114277	DRR0112383	IL4, HNRNPA1, CD79A, STAT6	methylselenic acid, B lymphocytes, Ebna-LP, hydroxyurea, Ebna2
SNRPA	medium	44.819499	83.21	DRH0118170	DRM0114311	DRR0112414	DAP3, IL4, PRMT7, IRF5, NFE2L2	lipopolysaccharide, pristane, Influenza A virus (A/Bangkok/RX73(H3N2)), 1,2-dithiole-3-thione, imiquimod
SNW1	medium	35.090660	80.62	DRH0118171	DRM0114312	DRR0112415	TNF, MYC, MITF, PLN, PPP2R1A, CIP2A	Jurkat E6.1 cells, Major late, Pc 3 cells, NOMO

Common gene name	Expression group	Human averaged NGS expression (nTPM)	Tissue stability (%)	Human GeneGlobe Assay ID	Mouse GeneGlobe Assay ID	Rat GeneGlobe Assay ID	Genes influencing the reference gene expression (IPA - QDI)	Other substances / conditions influencing reference gene expression (IPA - QDI)
PAF1	medium	34.464959	84.70	DRH0118146	DRM0114287	DRR0112391	TP53, SUMO1, WNT3A, TARDBP, Cdkn2a, PAF1	CIL56, acute myeloid leukemia cell lines, GNE-987, verteporfin
DDX23	medium	29.866645	85.10	DRH0118128	DRM0114269	DRR0112376	DAP3, FOXA2, NPY	GABA, ACH-2 cells, THP-1 cells, MLN7243, Jurkat E6.1 cells, Major late, Ie86
MAEA	medium	27.112640	84.91	DRH0118141	DRM0114282	DRR0112386	TNF, NR3C1, CSF2	
PUM1	medium	26.022273	84.15	DRH0118153	DRM0114294	DRR0112398	CEBPA, CALCA, PTH, GCG, DICER1, TGFBR1	desmopressin, P-TEFb, Ebna3c, interferon beta-1a
PPIE	medium	22.392089	87.76	DRH0118151	DRM0114292	DRR0112396	DAP3, C9orf72	Major late, NOMO, WAKMAR1, Wy 40905
MTREX	medium	18.763062	84.53	DRH0118144	DRM0114285	DRR0112389	MYC, ESR2, NRDE2, MTREX	hydroxyurea, GABA
RNF20	medium	17.647555	85.60	DRH0118156	DRM0114297	DRR0112401	IRF2, RPA1, PPME1, SMURF2, SPIB	rebemadlin, let-7 (includes others), palmitic acid, etoposide, fingolimod
TAF11	medium	13.835139	85.05	DRH0118172	DRM0114313	DRR0112416	TRIM24	growth hormone, tretinoin, cyclosporin A, alitretinoin, CB 1093, triiodothyronine, reverse, CD3 (complex), NEF
RNF216	medium	12.158566	82.63	DRH0118157	DRM0114298	DRR0112402		BAY 41-2272, epigallocatechin-gallate, Sendai virus, NS 2028, vesicular stomatitis virus, oxyhemoglobin
PDCD7	medium	11.890181	88.05	DRH0118148	DRM0114289	DRR0112393	NR3C1	arsenic trioxide, imatinib, plasma cells, B lymphocytes, T-2 toxin, olaparib, enzalutamide
SLC35E1	low	8.098024	78.70	DRH0118169	DRM0114310	DRR0112413		heavy metal
ARFGEF2	low	7.563072	63.66	DRH0118122	DRM0114265	DRR0112370	FLNA, CALCA, GCG, PTH, OSM, ERG	nocodazole, guanosine 5'-O-(3-thiotriphosphate), desmopressin, brefeldin A
SCFD2	low	5.435229	87.91	DRH0118166	DRM0114307	DRR0112410		
GPATCH2	low	4.407855	87.27	DRH0118133	DRM0114274	DRR0112380	FBXO44	
KIN	low	2.229007	87.69	DRH0118140	DRM0114281	DRR0112385	TP53, MYC, ATF2, XPC, KIN, XPA	

Genes are categorized by average expression level in human tissues (high: >1,000 nTPM; medium-high: 100-1,000 nTPM; medium: 10-100 nTPM; low: 1-10 nTPM). Average expression values are based on log-normalized GTEx RNA-seq data (nTPM) across 51 human tissues. Tissue stability (%) reflects the consistency of gene expression across tissues, with higher values indicating lower expression variability. Human, mouse, and rat GeneGlobe Assay IDs are provided for each reference gene. Known gene regulators and other biological factors influencing gene expression were identified using Ingenuity Pathway Analysis (IPA, QIAGEN Digital Insights).

References

1. Huggett JF, Foy CA, Benes V, et al. The digital MIQE guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments. *Clin Chem.* 2013;59(6):892-902. doi:10.1373/clinchem.2013.206375
2. dMIQE Group, Huggett JF. The Digital MIQE Guidelines Update: Minimum Information for Publication of Quantitative Digital PCR Experiments for 2020. *Clin Chem.* 2020;66(8):1012-1029. doi:10.1093/clinchem/hvaa125
3. Bustin SA, Ruijter JM, van den Hoff MJB, et al. MIQE 2.0: Revision of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments Guidelines. *Clin Chem.* 2025;71(6):634-651. doi:10.1093/clinchem/hvaf043
4. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015;348(6235):648-660. doi:10.1126/science.1262110
5. Söllner JF, Leparc G, Hildebrandt T, et al. An RNA-Seq atlas of gene expression in mouse and rat normal tissues. *Sci Data.* 2017;4:170185. Published 2017 Dec 12. doi:10.1038/sdata.2017.185
6. Hounkpe BW, Chenou F, de Lima F, De Paula EV. HRT Atlas v1.0 database: redefining human and mouse housekeeping genes and candidate reference transcripts by mining massive RNA-seq datasets. *Nucleic Acids Res.* 2021;49(D1):D947-D955. doi:10.1093/nar/gkaa609
7. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet.* 2013;29(10):569-574. doi:10.1016/j.tig.2013.05.010
8. de Jonge HJ, Fehrmann RS, de Bont ES, et al. Evidence based selection of housekeeping genes. *PLoS One.* 2007;2(9):e898. Published 2007 Sep 19. doi:10.1371/journal.pone.0000898
9. Krzystek-Korpacka M, Diakowska D, Bania J, Gamian A. Expression stability of common housekeeping genes is differently affected by bowel inflammation and cancer: implications for finding suitable normalizers for inflammatory bowel disease studies. *Inflamm Bowel Dis.* 2014;20(7):1147-1156. doi:10.1097/MIB.0000000000000067
10. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002;3(7):RESEARCH0034. doi:10.1186/gb-2002-3-7-research0034
11. Andersen CL, Jensen JL, Ørntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 2004;64(15):5245-5250. doi:10.1158/0008-5472.CAN-04-0496
12. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett.* 2004;26(6):509-515. doi:10.1023/b:bile.0000019559.84305.4

Ordering Information

Product	Contents	Cat. no.
QIAcuity One, Four or Eight Platform System	Digital PCR instrument for detecting up to 5 fluorescent dyes or up to 8 fluorescent dyes operating under software version 3.1 or higher, enabling up to 12-plex analysis	911021, 911042, 911052
dPCR RNA-Plex Assays	One tube with lyophilized assay, configurable fluorophore, 250 to 750 reactions (12 µL reaction in Nanoplate 8.5K)	250336, 250337
QIAcuity OneStep High Multiplex PCR Kit	4x OneStep High Multiplex PCR Master Mix, 100x OneStep High Multiplex RT Mix, Enhancer GC, RNAse-free water, QN Internal Control RNA; for 100 (or 500) reactions in Nanoplate 26K and 333 (or 1666) reactions in Nanoplate 8.5K	250135, 250136
QIAcuity High Multiplex PCR Kit	1 mL 4x concentrated QIAcuity High Multiplex Probe PCR Master Mix, 2x1.9 mL RNAse-free Water, QN Internal Control DNA	250135, 250136

For up-to-date licensing information and product-specific disclaimers, see the respective QIAGEN kit handbook or user manual. QIAGEN kit handbooks and user manuals are available at www.qiagen.com or can be requested from QIAGEN Technical Services or your local distributor.

Find out more about the QIAcuity Digital PCR System: **www.qiagen.com/qiacuity**

Trademarks: QIAGEN®, Sample to Insight®, QIAcuity® (QIAGEN Group).
Registered names, trademarks, etc. used in this document, even when not specifically marked as such, are not to be considered unprotected by law.
© 2026 QIAGEN, all rights reserved. QPRO25395 10/2026