Establishment of suitable reference genes for studying relative gene expression during the transition from trophozoites to cyst-like stages and first evidences of stress-induced expression of meiotic genes in *Trichomonas vaginalis*

Juliana Figueiredo Peixoto, Daniele das Graças dos Santos, Lupis Ribeiro, Vitor Silva Cândido de Oliveira, Rodrigo Nunes da Fonseca, José Luciano Nepomuceno-Silva.

SUPPLEMENTARY MATERIAL

Supplementary table 1: Description of candidate reference genes used in this study

Target gene/ TrichDB accession codes ‡	Cellular function of gene product	References
α-*actin* TVAG_200190 TVAG_249200 TVAG_172680 TVAG_150270 TVAG_160060 TVAG_310030 TVAG_090470 TVAG_149090 TVAG_054030 TVAG_512800 TVAG_485210	Fundamental protein in the eukaryotic cytoskeleton and major component of microfilaments. Actin is involved in cell motility, contraction, architecture maintenance and physical integrity. In *T. vaginalis*, actin microfilaments are related to parasite cytopathogenicity, since morphological changes (in the amoeboid stage) seem to be related to virulence. Widely used as RT-qPCR reference gene in many organisms and experimental conditions.	∗ Bricheux and Brugerolle, 1997†; ∗ Kusdian et al., 2013; • dos Santos et al., 2015†; • Wang et al., 2020.
β-*tubulin* TVAG_456920 TVAG_062880 TVAG_034440 TVAG_008680 TVAG_525430 TVAG_073810	Component of tubulin dimers, that polymerize into microtubules. Microtubules are involved with cellular architecture maintenance, intracellular transport of vesicles and organelles, mitotic chromosome segregation and cilia/flagella assembly. *T. vaginalis* trophozoite morphology is mostly shaped by the microtubule cytoskeleton, which forms the pelta-axostyle complex, the mitotic spindle, basal bodies and flagellar axonemes. The *T. vaginalis* genome revealed several copies for the three genes encoding β-tubulin proteins. Widely used as RT-qPCR reference gene in many organisms and experimental conditions.	∗ Katiyar and Edlind, 1994†; ∗ Lecke et al., 2002; • Yan and Liou, 2006; • Yang et al., 2019.
***Rpb*1** **(RNA pol. II)** TVAG_247430	Main subunit of eukaryotic RNA polymerase II. RNA Pol II catalyzes transcription of all nuclear mRNAs. Sometimes used as RT-qPCR reference gene.	∗ Malik et al., 2011; ∗ Quon et al., 1996; • Liu et al., 2015; • Lv et al., 2020.
GAPDH TVAG_347410 TVAG_476100 TVAG_146910 TVAG_366380	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is the sixth enzyme in the glycolysis pathway. The enzyme catalyzes the conversion of glyceraldehyde 3-phosphate to D-glycerate 1,3-bisphosphate. Furthermore, in *T. vaginalis* GAPDH is a member of the family of surface-associated enzymes that function also as receptors for fibronectin, among other host substrates. Widely used as RT-qPCR reference gene in many organisms and experimental conditions.	∗ Markoš et al., 1993†; ∗ Lama et al., 2009†; ∗ Nicholls et al., 2012; • Valenzuela-Castillo et al., 2017; • Jin et al., 2019; • Zhang et al., 2020.
G6PDH TVAG_414060	Glucose-6-phosphate dehydrogenase (G6PD) is the first enzyme in the oxidative phase of the conserved pentose phosphate pathway. This enzyme catalyzes an irreversible reaction of glucose-6-phosphate (G6P) in 6-phosphoglucone-δ-lactone, with the production of NADPH. In *T. vaginalis*, G6PDH and 6-phosphogluconolactonase (6PGL) genes are fused in a single open reading frame. Commonly used as RT-qPCR reference gene in many organisms and experimental conditions.	∗ Morales-Luna et al., 2020; • Yin et al., 2013; • Zarivi et al., 2015; • Omondi et al., 2015.
ferredoxin-1 TVAG_003900	Central protein in the hydrogenosome metabolism. Ferredoxin is a iron-sulfur electron acceptor protein that couples pyruvate decarboxylation to H^+ reduction into H_2. No available report of its use as an RT-qPCR reference gene.	∗ Quon et al., 1992; ∗ Vidakovic et al., 1996.
***eIF2*α** TVAG_379780	The eukaryotic translation initiation factor 2α (eIF2α) gene encodes the subunit α of eIF2, which escorts initiator Met-$tRNA_i$ to the small ribosomal subunit during formation of the pre-initiation complex. Sometimes used as RT-qPCR reference gene.	∗ Kozak, 1999; • Xia et al., 2014;
***eEF1*α** TVAG_067400 TVAG_013320 TVAG_463940 TVAG_354100	The eukaryotic translation elongation factor (eEF1α) gene encodes the eEF1α protein, which is part of the eEF1 protein and plays a central role in the elongation steps of eukaryotic translation. eEF1α escorts aminoacyl-tRNAs to the 80S ribosome acceptor site. Widely used as RT-qPCR reference gene in many organisms and experimental conditions.	∗ Roger at al., 1999†; ∗ Noble and Song, 2008; • Majerowicz et al., 2011; • Kianianmomeni & Hallmann, 2013; • Zarivi et al., 2015; • Yang et al., 2019.
18S rRNA TVAG_350330 (and 113 additional gene copies)	Sole rRNA component of the small subunit of the eukaryotic 80S ribosome. Widely used as RT-qPCR reference gene in many organisms and experimental conditions.	∗ Mayta et al., 2000; • Majerowicz et al., 2011; • Kianianmomeni & Hallmann, 2013; • Zarivi et al., 2015.

∗: Functional description ; **•**: Use/establishment as reference gene in RT-qPCR assays.

†: The *T. vaginalis* paralogous copies described in these studies are indicated together with their respective bibliographic reference in the following pages.

‡: Oligonucleotide pairs were designed based on the 1st accession codes from each list (underlined), but are equally able to amplify the other paralogous gene copies listed above.

References for supplementary table 1:

• **Bricheux G, Brugerolle G** (1997) Molecular cloning of actin genes in *Trichomonas vaginalis* and phylogeny inferred from actin sequences. *FEMS Microbiology Letters* 153(1), 205-213. doi: 10.1111/j.1574-6968.1997.tb10483.x.

†**:** *T. vaginalis* actin paralogues studied in this work: U63122 (TVAG_172680), U63123 (TVAG_249200), U63124 (TVAG_054030), U63125 (TVAG_200190) and U63126 (TVAG_310030).

• **Jin Y, Liu F, Huang W, Sun Q, Huang X** (2019) Identification of reliable reference genes for qRT-PCR in the ephemeral plant *Arabidopsis pumila* based on full-length transcriptome data. *Scientific Reports* 9(1):8408. doi: 10.1038/s41598-019-44849-1.

• **Katiyar SK, Edlind TD** (1994) β-Tubulin genes of *Trichomonas vaginalis*. *Molecular and Biochemical Parasitology* 64(1), 33-42. doi: 10.1016/0166-6851(94)90132-5.

†**:** *T. vaginalis* β-tubulin paralogues studied in this work: L05468 (TVAG_008680), L05469 (TVAG_034440) and L05470 (TVAG_338530).

• **Kianianmomeni A, Hallmann A** (2013) Validation of reference genes for quantitative gene expression studies in *Volvox carteri* using real-time RT-PCR. *Molecular Biology Reports* 40(12), 6691-6699. doi: 10.1007/s11033-013-2784-z.

• **Kozak M** (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene* 234(2), 187-208. doi: 10.1016/s0378-1119(99)00210-3.

• **Kusdian G, Woehle C, Martin WF, Gould SB** (2013) The actin-based machinery of *Trichomonas vaginalis* mediates flagellate-amoeboid transition and migration across host tissue. *Cell Microbiology* 15(10):1707-1721. doi: 10.1111/cmi.12144.

• **Lama A, Kucknoor A, Mundodi V, Alderete JF** (2009) Glyceraldehyde-3-phosphate dehydrogenase is a surface-associated, fibronectin-binding protein of *Trichomonas vaginalis*. *Infection and immunity* 77(7), 2703–2711. doi: 10.1128/IAI.00157-09.

†**:** *T. vaginalis GAPDH* paralogue studied in this work: L11394 (TVAG_366380).

• **Lecke SB, Tasca T, Souto AA, de Carli GA** (2002) *Trichomonas vaginalis*: microtubule cytoskeleton distribution using fluorescent taxoid. *Experimental Parasitology* 102(2), 113-116. doi: 10.1016/s0014-4894(03)00030-4.

• **Liu W, Zhao T, Wang H, Zeng J, Xiang L, Zhu S, Chen M, Lan X, Liao Z** (2015) Reference gene selection in *Artemisia annua* L., a plant species producing anti-malarial artemisinin. *Plant Cell, Tissue and Organ Culture (PCTOC)* 121(1), 141-152. doi: 10.1007/s11240-014-0690-2.

• **Lv Y, Li Y, Liu X, Xu K** (2020) Identification of ginger (*Zingiber officinale* Roscoe) reference genes for gene expression analysis. *Frontiers in Genetics* 11:586098. doi: 10.3389/fgene.2020.586098.

• **Majerowicz D, Alves-Bezerra M, Logullo R, Fonseca-de-Souza AL, Meyer-Fernandes JR, Braz GR, Gondim KC** (2011) Looking for reference genes for real-time quantitative PCR experiments in *Rhodnius prolixus* (Hemiptera: Reduviidae). *Insect Molecular Biology* 20(6), 713-722. doi: 10.1111/j.1365-2583.2011.01101.x.

• **Malik SB, Brochu CD, Bilic I, Yuan J, Hess M, Logsdon JM Jr, Carlton JM** (2011) Phylogeny of parasitic parabasalia and free-living relatives inferred from conventional markers vs. Rpb1, a single-copy gene. *PLoS One* 6(6), e20774. doi: 10.1371/journal.pone.0020774.

• **Mayta H, Gilman RH, Calderon MM, Gottlieb A, Soto G, Tuero I, Sanchez S, Vivar A** (2000) 18S ribosomal DNA-based PCR for diagnosis of *Trichomonas vaginalis*. *Journal of Clinical Microbiology* 38(7), 2683-2687. doi: 10.1128/JCM.38.7.2683-2687.2000.

• **Markoš A, Miretsky A, Müller, M** (1993) A glyceraldehyde-3-phosphate dehydrogenase with eubacterial features in the amitochondriate eukaryote, *Trichomonas vaginalis*. *Journal of Molecular Evolution* 37(6), 631–643. doi: 10.1007/BF00182749.

†**:** *T. vaginalis GAPDH* paralogue studied in this work: L11394 (TVAG_366380).

• **Morales-Luna L, Hernández-Ochoa B, Ramírez-Nava EJ, Martínez-Rosas V, Ortiz-Ramírez P, Fernández-Rosario F, González-Valdez A, Cárdenas-Rodríguez N, Serrano-Posada H, Centeno-Leija S, Arreguin-Espinosa R, Cuevas-Cruz M, Ortega-Cuellar D, Pérez de la Cruz V, Rocha-Ramírez LM, Sierra-Palacios E, Castillo-Rodríguez RA, Vega-García V, Rufino-González Y, Marcial-Quino J, Gómez-Manzo S** (2020) Characterizing the fused TvG6PD::6PGL protein from the protozoan *Trichomonas vaginalis*, and effects of the $NADP^+$ molecule on enzyme stability. *International Journal of Molecular Sciences* 21(14), 4831. doi: 10.3390/ijms21144831.

• **Nicholls C, Li H, Liu JP** (2012) GAPDH: a common enzyme with uncommon functions. *Clinical and Experimental Pharmacology and Physiology* 39(8), 674-679. doi: 10.1111/j.1440-1681.2011.05599.x.

• **Noble CG, Song H** (2008) Structural studies of elongation and release factors. *Cellular and Molecular Life Sciences* 65(9), 1335-1346. doi: 10.1007/s00018-008-7495-6.

• **Omondi BA, Latorre-Estivalis JM, Rocha Oliveira IH, Ignell R, Lorenzo MG** (2015) Evaluation of reference genes for insect olfaction studies. *Parasite & Vectors* 8, 243. doi: 10.1186/s13071-015-0862-x.

• **Quon DV, d'Oliveira CE, Johnson PJ** (1992) Reduced transcription of the ferredoxin gene in metronidazole-resistant *Trichomonas vaginalis*. *Proceedings of the National Academy of Sciences of the USA* 89(10), 4402-4406. doi: 10.1073/pnas.89.10.4402.

• **Quon DV, Delgadillo MG, Johnson PJ** (1996) Transcription in the early diverging eukaryote *Trichomonas vaginalis*: an unusual RNA polymerase II and alpha-amanitin-resistant transcription of protein-coding genes. *Journal of Molecular Evolution* 43(3), 253-262. doi: 10.1007/BF02338833.

• **Roger AJ, Sandblom O, Doolittle W F, Philippe H** (1999) An evaluation of elongation factor 1 alpha as a phylogenetic marker for eukaryotes. *Molecular Biology and Evolution* 16(2), 218-233. doi: 10.1093/oxfordjournals.molbev.a026104.

†**:** *T. vaginalis eEF1*α paralogue studied in this work: AF058282 (TVAG_067400).

• **dos Santos O, de Vargas Rigo G, Frasson AP, Macedo AJ, Tasca T** (2015) Optimal reference genes for gene expression normalization in *Trichomonas vaginalis*. *PLoS ONE* 10(9), e0138331. doi: 10.1371/journal.pone.0138331.

†**:** *T. vaginalis actin* paralogues target by the primers used in this work: TVAG_534990, TVAG_249200, TVAG_172680, TVAG_054030 and TVAG_512800.

• **Valenzuela-Castillo A, Mendoza-Cano F, Enríquez-Espinosa T, Grijalva-Chon JM, Sánchez-Paz A** (2017) Selection and validation of candidate reference genes for quantitative real-time PCR studies in the shrimp *Penaeus vannamei* under viral infection. *Molecular and Cellular Probes* 33, 42-50. doi: 10.1016/j.mcp.2017.02.005.

• **Vidakovic MS, Fraczkiewicz G, Germanas JP** (1996) Expression and spectroscopic characterization of the hydrogenosomal [2Fe-2S] ferredoxin from the protozoan *Trichomonas vaginalis*. *Journal of Biological Chemistry* 271(25), 14734-14739. doi: 10.1074/jbc.271.25.14734.

• **Wang Z, Meng Q, Zhu X, Sun S, Liu A, Gao S, Gou Y** (2020) Identification and evaluation of reference genes for normalization of gene expression in developmental stages, sexes, and tissues of *Diaphania caesalis* (Lepidoptera, Pyralidae). *Journal of Insect Science* 20(1), 6. doi: 10.1093/jisesa/iez130.

• **Xia W, Mason AS, Xiao Y, Liu Z, Yang Y, Lei X, Wu X, Ma Z, Peng M** (2014) Analysis of multiple transcriptomes of the African oil palm (*Elaeis guineensis*) to identify reference genes for RT-qPCR. *Journal of Biotechnology* 184, 63-73. doi: 10.1016/j.jbiotec.2014.05.008.

• **Yan HZ, Liou RF** (2006) Selection of internal control genes for real-time quantitative RT-PCR assays in the oomycete plant pathogen *Phytophthora parasitica*. *Fungal Genetics and Biology* 43(6), 430-438. doi: 10.1016/j.fgb.2006.01.010.

• **Yang T, Gu B, Xu G, Shi Y, Shen H, Rao R, Mzuka HL** (2019) Identification of candidate reference genes for qRT-PCR normalization studies of salinity stress and injury in *Onchidium reevesii*. *Peer Journal* 7, e6834. doi: 10.7717/peerj.6834.

• **Yin Z, Ke X, Huang D, Gao X, Voegele RT, Kang Z, Huang L** (2013) Validation of reference genes for gene expression analysis in *Valsa mali* var. mali using real-time quantitative PCR. *World Journal of Microbiology and Biotechnology*. 29(9), 1563-1571. doi: 10.1007/s11274-013-1320-6.

• **Zarivi O, Cesare P, Ragnelli AM, Aimola P, Leonardi M, Bonfigli A, Colafarina S, Poma AM, Miranda M, Pacioni G** (2015) Validation of reference genes for quantitative real-time PCR in Perigord black truffle (*Tuber melanosporum*) developmental stages. *Phytochemistry* 116, 78-86. doi: 10.1016/j.phytochem.2015.02.024.

• **Zhang E, Wu S, Cai W, Zeng J, Li J, Li G, Liu J (2020)** Validation of superior reference genes for qRT-PCR and Western blot analyses in marine *Emiliania huxleyi*-virus model system. *Journal of Applied Microbiology* (Early view - December). doi: 10.1111/jam.14958.

Supplementary table 2: Cq data for candidate reference genes

	α-*actin*	β-*tubulin*	*rpb1*	*GAPDH*	*G6PDH*	*ferredoxin1*	*eIF2α*	*eEF1α*	18S rRNA
N	32	32	32	32	32	32	32	32	32
Geo. Mean [Cq]	17.78	21.92	28.42	17.04	24.27	20.07	27.23	17.00	8.83
Ar. Mean [Cp]	18.03	22.40	28.90	17.21	24.44	20.23	27.32	17.27	9.31
Min. [Cp]	13.74	17.52	13.11	13.31	17.79	15.69	23.55	13.33	4.92
Max. [Cq]	30.40	36.86	37.84	21.78	33.96	23.98	34.27	29.12	20.93

N: Number of samples; Geo Mean: Geometric mean of Cq values; Ar. Mean: Arithmetic mean of Cq values; Min.: Lower Cq value; Max.: Higher Cq value.
Data were calculated by the BestKeeper program.

Supplementary table 3: Stability values/scores for the algorithms used in this work

	∆*Cq*	*geNorm*	*NormFinder*	*BestKeeper* SD (±Cq)	*BestKeeper* coeff. of corr. (r)	*BestKeeper p value*
α-*actin*	3.608232	1.950537	0.408370	2.32	0.681	< 0.001
β-*tubulin*	5.083149	1.784633	0.376808	3.66	0.567	< 0.001
rpb1	4.840302	2.345323	0.800115	3.98	0.628	< 0.001
GAPDH	3.011740	**0.622130**	0.194674	1.97	0.785	< 0.001
G6PDH	3.126157	1.166116	0.217217	1.96	0.843	< 0.001
ferredoxin 1	3.404484	**0.622130**	**0.176028**	2.27	0.600	< 0.001
eIF2α	**2.984542**	1.119349	0.320583	**1.72**	**0.867**	< 0.001
eEF1α	3.847907	1.477999	0.382490	2.45	0.619	< 0.001
18S rRNA	3.858006	1.383471	0.383195	2.48	0.686	< 0.001

The four more stable transcripts indicated by each algorithm are highlighted in gray and the most stable transcript for each one is in bold.

LEGENDS TO SUPPLEMENTARY FIGURES:

Supplementary figure 1. Calculation of amplification efficiencies for primer pairs by linear regression of Cq values obtained in RT-qPCR reactions using tenfold serial dilutions of *T. vaginalis* trophozoite cDNA.

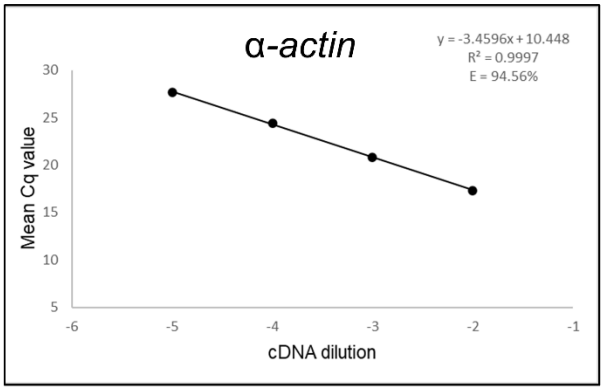
Supplementary figure 2. ΔCq comparison for the complete set of candidate reference genes. Variability in gene expression is represented by differences in Cq values (ΔCq) between each gene pair. Mean ΔCqs are depicted as filled circles and error bars correspond to standard deviations of these means. Shorter standard deviations indicate that compared genes share relatively stable expression or are co-regulated. Dashed vertical lines limit datasets for each compared gene and the solid horizontal line indicates the boundary wherein identical expression levels would converge.

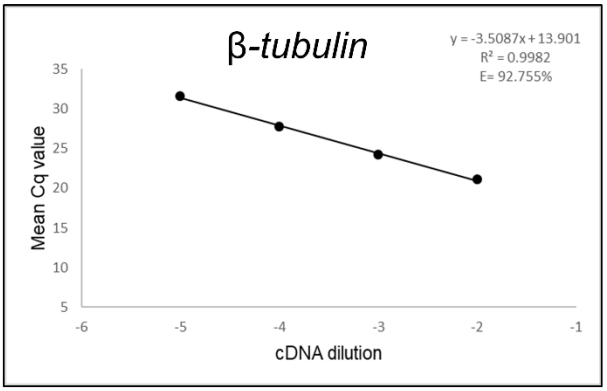
Supplementary figure 3: Expression analysis of the *T. vaginalis* recombinase *Dmc1* and nine other genes from the meiosis detection "toolkit" in published RNA-Seq datasets. **A:** Expression of meiotic mRNAs in RNA-Seq datasets from normal (IR = iron rich) and iron depleted (ID) parasites. Read counts were normalized as reads per kilobase of transcript per million mapped reads (RPKM). Datasets were obtained from Cheng et al. (2015). **B:** Expression of meiotic mRNAs in RNA-Seq datasets from normal trophozoites (1%_12h) and parasites cultured in glucose restricted medium for 12, 24 and 36 hours (GR 12, 24 and 36 h). Read counts were also normalized as RPKM. Datasets were obtained from Huang et al. (2014). TrichDB accession codes as follows: *Dmc1* (TVAG_155030), *Msh4* (TVAG_455180), *Msh5* (TVAG_472000), *Spo11* (TVAG_258950), *Hop1* (TVAG_230730), *Hop2A* (TVAG_058400), *Hop2B* (TVAG_151700), *Mnd1* (TVAG_062830) and *Mer3* (TVAG_292060).

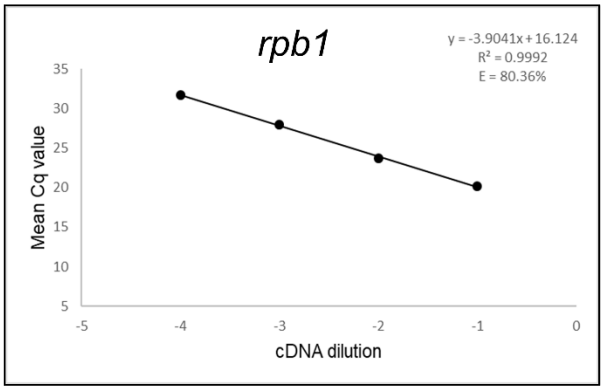
Supplementary figure 4: Expression analysis of the four more stable candidate reference genes in RT-qPCR using previously published EST (expressed sequence tags) and RNA-Seq datasets. **A:** Analysis of *GAPDH* (polled), *G6PDH* and *ferredoxin-1* mRNAs in normal (37 °C; TvE) and cold-stressed (4 °C for 4 hours = TvC) EST libraries of *T. vaginalis*. EST counts were normalized as transcripts per million (tpm). No ESTs were detected for *eIF2α*. Datasets were obtained from Fang et al. (2015). **B:** Analysis of *GAPDH* (polled), *G6PDH, ferredoxin-1* and *eIF2α* mRNAs inRNA-Seq datasets from normal (IR = iron rich) and iron depleted (ID) parasites. Read counts were normalized as reads per kilobase of transcript per million mapped reads (RPKM). Datasets were obtained from Cheng et al. (2015). **C:** Analysis of *GAPDH* (polled), *G6PDH, ferredoxin-1* and *eIF2α* transcripts in RNA-Seq datasets from normal trophozoites (1% 12 h) and parasites cultured in glucose restricted medium for 12, 24 and 36 hours (GR 12, 24 and 36 h). Read counts were normalized as RPKM. Datasets were obtained from Huang et al. (2014).

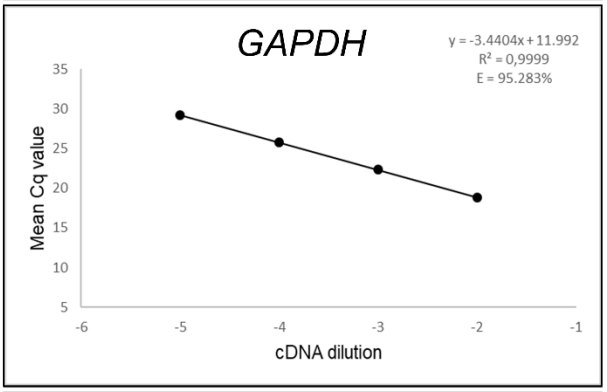
Supplementary figure 5: Comparison of protein products corresponding to the four most stable reference genes observed by RT-qPCR during cold induced cyst-like differentiation of *T. vaginalis*. The comparison was based in quantitative proteomic detected gene products representing *T. vaginalis* trophozoites and induced cyst-like/pseudocyst forms induced by iron depletion. Normalized protein copy number per cell data were obtained from replicate experiments reported by Dias-Lopes *et al.* (2018; supplementary table S1). *GAPDH* is represented as a pool of three different isoforms corresponding to paralogues that can by potentially amplified by our primer pair (TVAG-476100 was not detected in their analysis). Data are presented as mean ± standard deviation of three (cyst-like forms) and four (trophozoites) replicates. Statistically significant differences were not detected for any of the gene products (Mann-Whitney U test, $p > 0.05$).

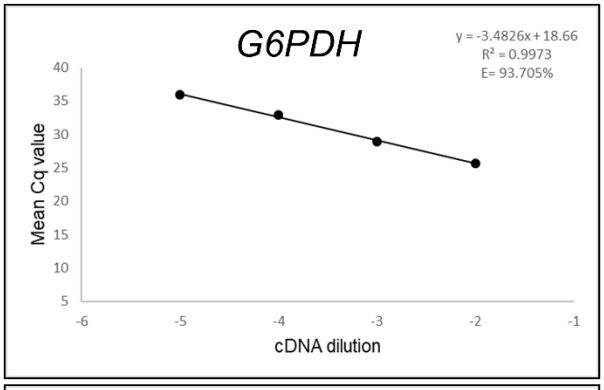
Supplementary figure 1: Amplification efficiencies

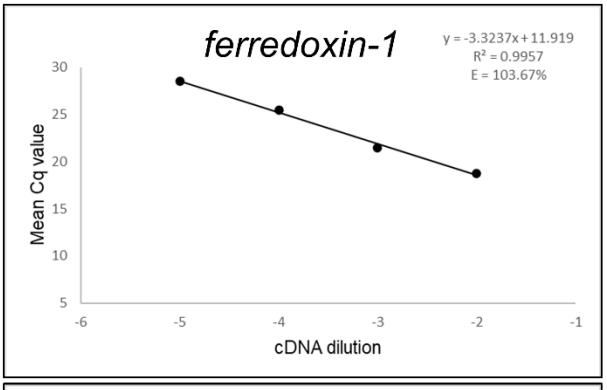


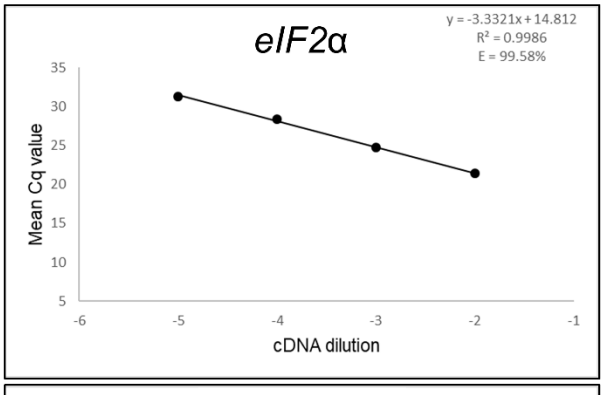


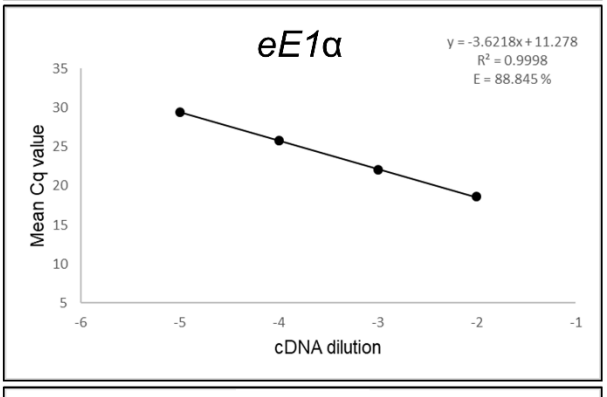


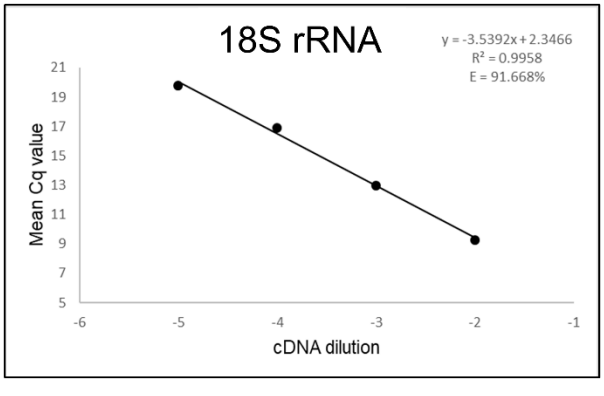


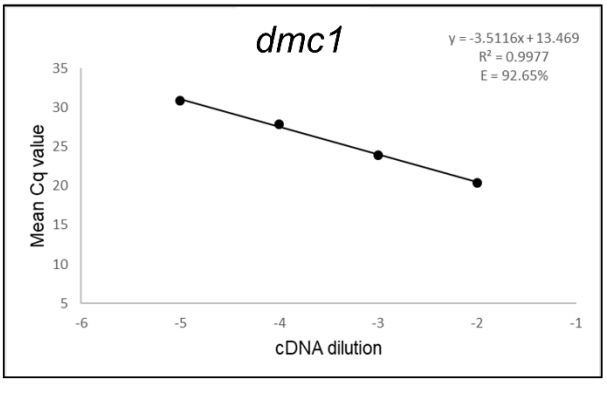




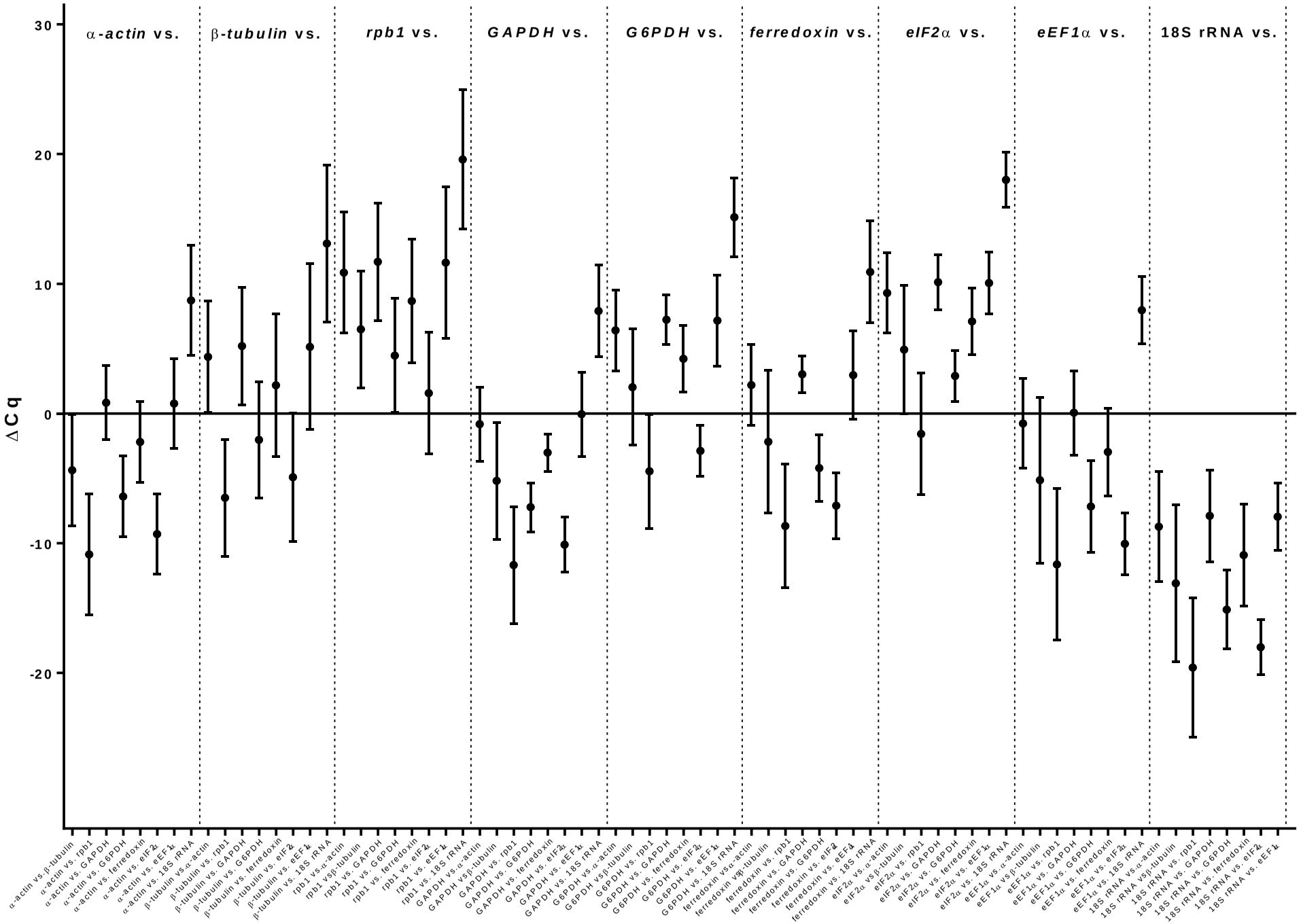




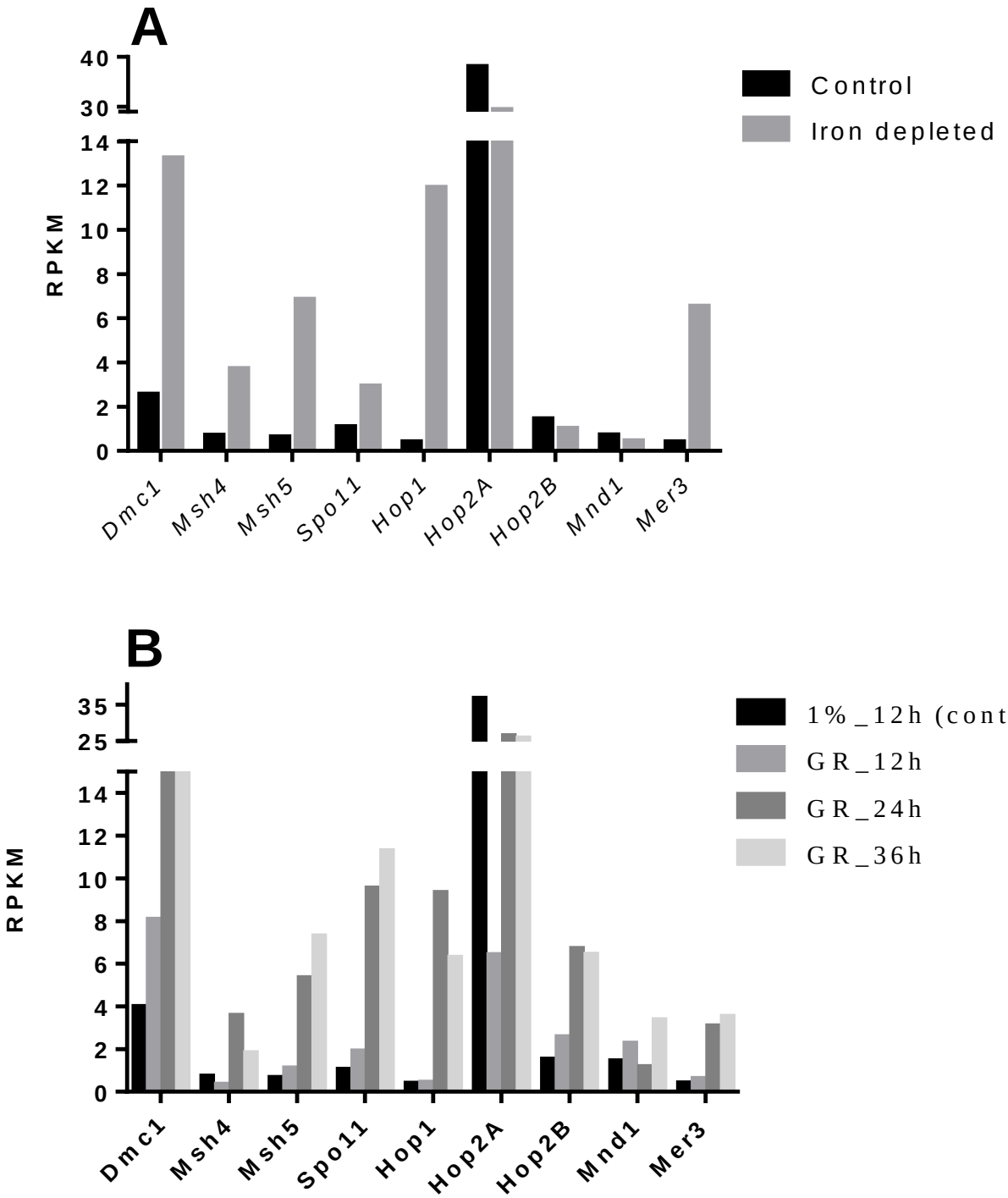




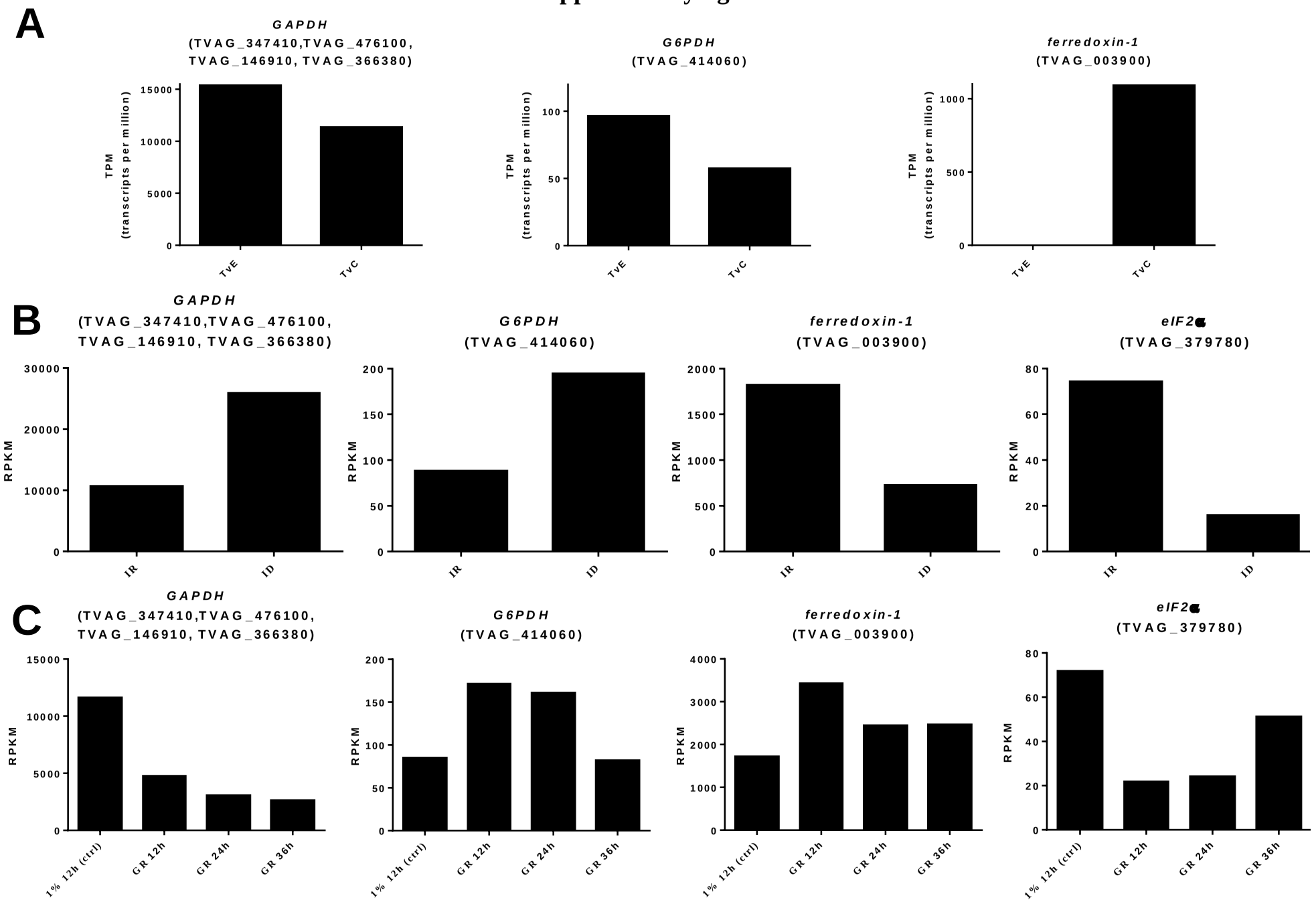
Supplementary figure 2: ΔCq comparison



Supplementary figure 3



Supplementary figure 4



Supplementary figure 5