Full title (250 characters):

Design, optimization and validation of genes commonly used in expression studies on DMH/AOM rat colon carcinogenesis model

Short title (100 characters):

Validation of genes commonly used in qPCR studies on DMH/AOM rat colon carcinogenesis model

David Bars-Cortina[1,2], Antoni Riera[3], Gemma Gou[4,5], Carme Piñol-Felis[2] *, Maria-José Motilva[1]

[1] Food Technology Department, XaRTA-TPV, Agrotecnio Center, Escola Tècnica Superior d'Enginyeria Agrària, Universitat de Lleida, Avinguda Alcalde Rovira Roure 191, 25198 Lleida, Catalonia, Spain

[2] Department of Medicine, Universitat de Lleida-Institut de Recerca Biomèdica de Lleida (IRBLleida), Avinguda Alcalde Rovira Roure 80, 25198 Lleida, Catalonia, Spain.

[3] Andrology Department, Fundació Puigvert, Universitat Autònoma de Barcelona, Instituto de Investigaciones Biomédicas Sant Pau (IIB-Sant Pau), Barcelona, Catalonia, Spain.

[4] Molecular Physiology of the Synapse Laboratory, Biomedical Research Institute Sant Pau (IIB Sant Pau), Sant Antoni Mª Claret 167, 08025 Barcelona, Spain.

[5] Universitat Autònoma de Barcelona, 08193, Bellaterra, Cerdanyola del Vallès, Spain

*Corresponding author:

Email:pinyol@medicina.udl.cat (CP-F)

Abstract

Colorectal cancer (CRC), also known as colon cancer, is the third most common form of cancer worldwide in men and the second in women and is characterized by several genetic alterations, among them the expression of several genes. 1,2-dimethylhydrazine (DMH) and its metabolite azoxymethane (AOM) are procarcinogens commonly used to induce colon cancer in rats (DMH/AOM rat model). This rat model has been used to study changes in mRNA expression in genes involved in this pathological condition. However, a lack of proper detailed PCR primer design in the literature limits the reproducibility of the published data.

The present study aims to design, optimize and validate the qPCR, in accordance with the MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines, for seventeen genes commonly used in the DMH/AOM rat model of CRC (*Apc, Aurka, Bax, Bcl2, β-catenin, Ccnd1, Cdnk-1A, Cox2, Gsk3beta, IL-33, iNOs, Nrf2, p53, RelA, Smad4, Tnfα* and *Vegfa*) and two reference genes (*Actb* or *β-actin* and *B2m*). The specificity of all primer pairs was empirically validated on agarose gel, and furthermore, the melting curve inspection was checked as was their efficiency (%) ranging from 90 to 110 with a correlation coefficient of $r^2 >$ 0.980. Finally, a pilot study was performed to compare the robustness of two candidate reference genes.

Introduction

Colorectal cancer (CRC) is the third most common form of cancer worldwide in men (surpassed by lung and prostate cancer) and the second in women (overtaken by breast cancer). The incidence of CRC varies significantly between populations, Australia and New Zealand being the countries with the highest rate of new diagnoses, while the countries of western Africa have the lowest incidence[1]. In the United States, CRC represents the fourth most prevalent cancer with 135,430 new cases diagnosed in 2017 and representing 8.0% of all new cancer cases[2]. In Europe, CRC is the second most common cancer in both sexes[3]. In Asia, especially in the industrialized regions, the incidence of CRC has increased over the last decade due to the adoption of the western lifestyle[4]. Interestingly, a similar situation is taking place in Eastern Europe, Latin America and the Caribbean countries[5]. Unfortunately, the global incidence of CRC is expected to increase by about 60% and it is predicted that, in 2030, more than 2.2 million new cases will be diagnosed and 1.1 million people will die from this disease[5].

It is well established that lifestyle and especially eating patterns play an important role in the risk of developing cancer in the digestive tract[6–9]. Hence, several studies have focused on diet as the major strategy to counteract and prevent colon cancer[10]. Different animal models have been used to evaluate the effect of food components on colon cancer prevention. Such models of colon carcinogenesis can be divided into two broad categories: transgenic and chemically-induced [11,12]. A classic example of a transgenic animal CRC model concerns adenomatous polyposis coli (*Apc*) gene mutations. Nevertheless, these animal models have been generated to

78 study familial adenomatous polyposis (FAP) and hereditary nonpolyposis
79 colorectal cancer (HNPCC) syndromes which only account for approximately
80 5% of all cases [11,12].

81 Dimethylhydrazine (DMH) and its metabolite azoxymethane (AOM) are
82 frequently used to generate chemically-induced models for CRC[11,13].
83 These models share many similarities with human sporadic colon cancer
84 since DMH/AOM colon carcinogenesis occurs as a multistep process. The
85 stepwise development of CRC from dysplastic crypts, adenomas to
86 carcinomas provides the opportunity to investigate and identify molecular
87 alterations in each stage of tumour development[11]. Interestingly, genes that
88 have been found mutated in human sporadic colon cancer have also been
89 found to mutate in DMH/AOM-induced colon carcinogenesis[11].
90 Nutrigenomics, the study of the effects of food components on gene
91 expression, is a broad approach used in CRC animal models. In this regard, a
92 wealth of data addresses this issue through different dietary regimes (more
93 than 100 results appear when searching for “rat colon cancer diet gene
94 expression” in Pubmed). However, the data generated is often confusing
95 when it is carefully evaluated. For instance, in some published articles, the
96 accession number of the reference sequence used is not indicated or the
97 qPCR is not well designed or described (i.e. the primer sequence is missing
98 or contains mistakes, primers are not specific or the melting temperature is
99 too low). On the basis of the aforementioned data, in this study we provide a
100 well-designed and described qPCR protocol according to the MIQE (Minimum
101 Information for Publication of Quantitative Real-Time PCR Experiments)
102 guidelines[14] for 17 genes routinely studied in DMH/AOM CRC rat model.

Moreover, we also analyse two reference genes commonly used in this carcinogenesis model.

Materials and Methods

Animals

All animal care and experimental procedures were in accordance with the EU Directive 2010/63/EU guidelines for animal experiments and approved by the Animal Ethics Committee at the University of Lleida (CEEA 02/06-16). Healthy adult male Wistar rats weighing between 200 to 250 g maintained in the animal facilities at the University of Lleida were used. The animals were housed in polyvinyl cages at a controlled temperature (21ºC ± 1ºC) and humidity (55% ± 10% RH), maintained under a constant 12 h light-dark cycle. All the animals were fed with water and a standard diet for rodents (Envigo Teklad Global Diet 2014, batch 3201, Settimo Milanese, Italy) *ad libitum*. Three randomly-selected animals were sacrificed by intracardiac puncture after isoflurane anaesthesia (ISOFlo, Veterinaria Esteve, Bologna, Italy). Distal colon tissue (the most relevant region in CRC studies with DMH/AOM induced models)[13] was extracted and immediately frozen in liquid nitrogen and then stored at -80ºC until it was analysed.

RNA isolation & cDNA synthesis

Tissue Lyser LT (Quiagen, Germany) was used as a tissue homogenizer (4 cycles of 50 Hz for 30 sec. with a 1 min. pause within each cycle). Total RNA was extracted using the Trizol™ Plus PureLink™ Kit RNA Mini Kit (Invitrogen,

USA) following the kit instructions. RNA quantity and purity (260/280 and 260/230 ratios) were assessed with a ND-1000 Nanodrop spectrophotometer (ThermoFisher Scientific, USA). Furthermore, the integrity of the total RNA obtained was evaluated through 1% agarose gel[15]. In all cases, 18S and 28S ribosomal RNA bands were clearly detected and no degraded RNA (illustrated as smear in the gel lane) was identified.

Reverse transcription was performed with the Maxima H Minus First Strand cDNA Synthesis kit with dsDNase (ThermoFisher Scientific, USA) according to the manufacturer's instructions (≤ 5 µg of total RNA as template and using 100 pmol random hexamer primer).

Primer pairs design

Primer pairs for seventeen different CRC related genes (*Apc, Aurka, Bax, Bcl2, β-catenin, Ccnd1, Cdnk1a, Cox2, Gsk3beta, IL-33, iNOs, Nrf2, p53, RelA, Smad4, Tnfα and Vegfa*) and two candidate reference genes (*Actb and B2m*) were designed and evaluated for their suitability through a number of bioinformatics tools summarized in **Fig 1A**. Briefly, we selected the genes to be studied (from a literature search) and obtained their accession number. Then, the nucleotide sequence was retrieved from the PubMed Nucleotide database. Afterwards, if no previously published primers were used, we checked for their splice variants (through Ensembl Release 87)[16]. In the case of spliced genes and also in order to avoid the presence of SNPs (single-nucleotide polymorphisms), multiple sequence alignment was performed to find common regions within the splice variants.[16] The selected sequence was transferred to Primer3Plus software version 2.4.2[17] to pick

153 up primers. The technical parameters used in the design of the primers were
154 based on Thorton et al. 2011[18]. Finally, the primer pairs obtained were
155 submitted to further *in silico* analysis in order to avoid strong primer secondary
156 structures (through OligoAnalyzer 3.1, Integrated DNA Technologies)[19],
157 robust amplicon secondary structures (with the UNAFold tool, Integrated DNA
158 Technologies)[19] and unspecificity (with the In-Silico PCR tool of the UCSC
159 Genome Browser Database[20] and the Nucleotide BLAST tool[21]). The
160 primer pairs selected after these bioinformatics tool tests were acquired from
161 the Sigma-Aldrich custom oligo facilities (Haverhill, UK).
162
163 **Fig 1.** Flowchart indicating the strategy followed to design and validate the
164 candidate primers.
165 (A) In silico validation flowchart. (B) Empirical validation flowchart.
166
167 **PCR**
168 **PCR reaction**
169 PCR reactions were performed in a total reaction volume of 25 µl
170 comprising 2.5 µl of 10X Dream Taq Buffer, 0.5 µl of dNTP mix (R0191
171 ThermoFisher Scientific, USA), 0.5 µl of gene-specific primer pair at 10 µM, 2
172 µl of cDNA template, 0.625 U Dream Taq DNA Polymerase (EP0701
173 ThermoFisher Scientific, USA) and filled up to 25 µl with nuclease free water
174 (BP561-1 Fisher Scientific, USA). The PCR conditions used were 3 min of
175 polymerase activation at 95ºC followed by 35 cycles of denaturation at 95ºC
176 for 30s, an annealing step at 57ºC (or between 51ºC and 61ºC in the case of a

gradient) for 30s and extension at 72ºC for 30s. Final extension (72ºC) was performed for 5 min followed by an infinite 4ºC step.

PCR empirical validation

After the previous *in silico* steps described above, all the primer pairs were submitted to further analysis (**Fig 1B**). Primer specificity was assessed through conventional PCR followed by agarose gel electrophoresis in order to check that unique band with the amplicon size was obtained. The annealing temperature was set at 57ºC by default but, in some cases, an annealing temperature gradient was needed (see above).

qPCR

qPCR reaction

Real-time PCR reactions were performed in a total reaction volume of 20 µl comprising 10 µl of SYBR™ Select Master Mix (2X) (ThermoFisher Scientific, USA), 0.2 to 0.8 µl of each gene-specific primer (depending on the primer pairs concentration required) (**Table 1**), 2 µl of cDNA, and filled up to 20 µl with nuclease free water (BP561-1 Fisher Scientific, USA).

Table 1. qPCR efficiency and correlation coefficient (R^2) obtained for each selected gene.

Gene (Accession No.)	Linearity range in ng cDNA (Ta: annealing temperature)	qPCR efficiency	R^2	[primer], nM
Actb* (NM_031144.3)	2 to 128 (Ta: 57ºC) 22.5 to 114 (Ta: 59.3ºC)	108.9% 90.5%	0.998 0.999	100
Apc (NM_012499.1)	2 to 128	108.5%	0.998	100
Aurka (NM_153296.2)	9 to 243	108.8%	0.998	200
Bax (NM_017059.2)	0.16 to 100	106.3%	0.994	200
Bcl2 (NM_016993.1)	0.5 to 128	102.9%	0.990	200
B2m* (NM_012512.2)	1.6 to 148.8 (Ta:57ºC) 2 to 128 (Ta:59.3ºC)	100.4% 108.7%	0.996 0.997	200
β-catenin (AF_121265)	2 to 128	109.8%	0.980	150
Ccnd1 (NM_1719924)	22.5 to 114	108.5%	0.997	100
Cdnk-1A (NM_080782.3)	0.16 to 100	101.2%	0.994	200
Cox2 (AF233596.1)	0.5 to 100	106.6%	0.998	200
Gsk3beta (NM_032080)	1.56 to 100	106.8%	0.997	200
IL-33 (NM_001014166.1)	2 to 128	110.0%	0.996	100
iNOs (NM_012611.3)	20 to 338.8	97.0%	0.990	200
Nrf2 (NM_031789.2)	0.16 to 100	106.6%	0.996	200
p53 (NM_030989.3)	8 to 128	102.6%	0.997	100
RelA (NM_199267.2)	0.16 to 100	103.0%	0.997	200
Smad4 (AB010954.1)	6.4 to 100	104.5%	0.993	400
Tnfα (NM_012675.3)	10 to 114	90.9%	0.987	100
Vegfa (ENSRNOG00000019598)	0.16 to 100	106.9%	0.992	200

202 * denotes reference gene. nM: nanomolar concentration

The qPCR reactions were carried out on a Bio-Rad CFX96 real-time PCR system (Bio-Rad Laboratories, USA) under the following conditions: 2 min of uracil-DNA glycosylase (UDG) activation at 50ºC, 2 min of polymerase activation at 95ºC, followed by 40 cycles of denaturation at 95ºC for 15 s and annealing/extension at the corresponding annealing temperature (57ºC or 59.3ºC) (**Table 2**) for 1 min. A melting curve analysis was done immediately after the qPCR analysis.

Table 2. List of the primer pairs validated.

Gene (Accession No.)	Primer sequences (5'-3')	Amplicon size	Ta	Reference
Actb * (NM_031144.3)	F: TCTGTGTGGATTGGTGGCT R: TCATCGTACTCCTGCTTGCT	80 bp	57ºC / 59.3ºC	In-house
Apc (NM_012499.1)	F: ACTCCTTACTGCTTCTCACG R: GTCCTTACTTTCTTTGCCCTTT	114 bp	57ºC	In-house
Aurka (NM_153296.2)	F: AGTGCTATCTGTCCATCAACC R: ACCCGCATTTCCAGTCATCT	98 bp	59.3ºC	In-house
Bax (NM_017059.2)	F: AGAGGATGATTGCTGATGTGG R: CCCAGTTGAAGTTGCCGT	93 bp	57ºC	In-house
Bcl2 (NM_016993.1)	F: GATTGTGGCCTTCTTTGAG R: CAGGCTGAGCAGCGTCTTC	232 bp	59.3ºC	Based on [31]
B2m * (NM_012512.2)	F: CCCACCCTCATGGCTACTTC R: GATGAAAACCGCACACAGGC	157 bp	57ºC / 59.3ºC	[32]
β-catenin (AF_121265.1)	F:CAAGTGGGTGGCATAGAGG R: ATGACGAAGAGCACAGATGG	93 bp	57ºC	In-house
Ccnd1 (NM_171992.4)	F: AGTTGCTGCAAATGGAACTG R: TGGAGAGGAAGTGTTCGATG	93 bp	57ºC	Based on [33]
Cdnk-1A (NM_080782.3)	F: ATGTCCGATCCTGGTGATGT R: GCTCAACTGCTCACTGTCCA	90 bp	57ºC	In-house
Cox2 (AF233596.1)	F: TGTATGCTACCATCTGGCTTCGG R: GTTTGGAACAGTCGCTCGTCATC	94 bp	57ºC	[34]
Gsk3beta (NM_032080.1)	F: TGGGTCATTTGGTGTGGT R: GGTTCTTAAATCGCTTGTCCT	95 bp	57ºC	In-house
IL-33 (NM_001014166.1)	F: TTCAGTCCTGCCCTTTCCTT R: TGTGGTGCGTGCTCTTCT	84 bp	57ºC	In-house
iNOs (NM_012611.3)	F: CACCACCCTCCTTGTTCAAC R: CAATCCACAACTCGCTCCAA	132 bp	57ºC	[35]
Nrf2 (NM_031789.2)	F: GTGACTCGGAAATGGAAGAG R: AGAAGAATGTGTTGGCTGTG	83 bp	57ºC	In-house
p53 (NM_030989.3)	F: GCAGAGTTGTTAGAAGGC R: TTGAGAAGGGACGGAAGA	138 bp	57ºC	In-house
RelA (NM_199267.2)	F: TCACCAAAGACCCACCTCA R: GTTCAGCCTCATAGAAGCCA	81 bp	57ºC	In-house
Smad4 (AB010954.1)	F: CCACCAACTTCCCCAACATT R: TGCAGTCCTACTTCCAGTCCAG	191 bp	57ºC	[36]
Tnfα (NM_012675.3)	F: ACCACGCTCTTCTGTCTACTG R: CTTGGTGGTTTGCTACGAC	169 bp	59.3ºC	[37]
Vegfa (ENSRNOG00000019598)	F: GACACACCCACCCACATAC R: TCCAGTGAAGACACCAATAACA	141 bp	57ºC	In-house

* denotes reference gene. Ta, annealing temperature

228

229 **qPCR empirical validation & qPCR analysis**

230 Once the unique band had been obtained in the previous PCR step,
231 qPCR efficiency, linearity and specificity (unique and clear melt curve) were
232 assessed taking into account Taylor et al. 2010[22], and therefore the MIQE
233 guidelines[14]. qPCR must be within a range of 90 to 110% and with a
234 standard curve correlation coefficient (R^2) ≥ 0.98[22,23]. Each point on the
235 standard curve was performed in triplicate and covered all potential template
236 concentrations that may be encountered in future studies. Whenever possible,
237 the standard curve comprised three orders of magnitude. C_q values > 38 were
238 not considered for data analysis due their low efficiency[14]. Furthermore, in
239 triplicate, no template control (NTC) was included for each primer pair in every
240 run. The data resulting from the qPCR were analysed using the Bio-Rad CFX
241 Maestro 1.1 software. Baseline correction and threshold setting were
242 performed using the automatic calculation in the CFX Maestro 1.1 software
243 (Bio-Rad).

244

245 **Reference gene selection**

246 In parallel to the primer validation described in this paper, we
247 implemented a related experiment with the aim of addressing the possible
248 effect of dietary supplementation with a particular fruit (white- and red-fleshed
249 apples) and cyanidin galactoside (the main anthocyanin in red-fleshed apple)
250 on these genes in the early phases of rat colon cancer induced by AOM (**S1**

Appendix). For this reason, two reference genes commonly used in DMH/AOM rat model experiments were selected and submitted to check their expression stability in the different experimental groups in the related experiment cited above. In detail, six different experimental groups were compared and two rats per treatment group were analysed with three technical replicates each. The amount of cDNA used in each reaction was 100 ng.

The stability (aptitude) of the candidate reference genes was evaluated with two software tools (web-based RefFinder platform (http://leonxie.esy.es/RefFinder/ [data access: 08/05/2018]) and Bio-Rad CFX Maestro 1.1. software, based on geNorm algorithm).

RESULTS AND DISCUSSION

Genetic material used

As stated in the previous section, three healthy adult male Wistar rats were selected randomly and sacrificed. The distal region of the colon was obtained and immediately frozen. The distal colon samples were pooled prior to total RNA extraction. The quality and quantity of the RNA was good (ratio 260/280 = 1.96, ratio 260/230= 2.09, 423.6 ng/µl).

Primer design and validation through agarose gel

A correct selection of the primer pairs is a critical step for a qPCR experiment in order to obtain a specific amplification of the target gene. In addition, the primer design for SYBR® Green based detection needs to be

more carefully done than for a classic TaqMan® assay since former interacts with double-stranded PCR products and may lead to "false" signal. Hence, the sensitivity of detection with SYBR® Green may be hindered by the lack of specificity of the primers, primer concentration and the formation of secondary structures in the PCR product. The formation of primer-dimers may register false positive fluorescence. However, this can easily be overcome by running a PCR melting curve analysis. As shown in **Fig 1**, all precautions were taken into account in this study.

The primer pairs detailed in **Table 2** passed all the bioinformatic tests (OligoAnalyzer 3.1 and UNAFold from Integrated DNA Technologies; in-silico PCR of the UCSC Genome Browser Database and the Nucleotide Blast tool) and a single band was observed in the agarose gel. These primers were selected for further analysis through qPCR. On the contrary, the primers which do not pass some step in the validation process are shown in **S1 Table.** The majority of the primer pairs (12, if we also consider the reference genes) were in-house designed. The remainder were from published intact sequences or, in some cases, were obtained after some *in silico* mismatch corrections (indicated as "Based on" in **Table 2 and S1 Table**).

Primer validation through qPCR

In order to check whether the primers pairs designed were useful for qPCR analysis, we need to validate primer specificity again through the melting curve and also the qPCR assay, mainly qPCR efficiency, as stated in Derveaux et al. and Taylor et al. [15,22]. One of the most common options to assess the specificity of the primer pairs is the melting curve. This determines

whether the intercalating dye (SYBR green) has produced single and specific products. In this study we checked the melting curve for all the primer pairs and these all demonstrated their specificity as a unique peak was detected among the concentrations used in the standard curve in all cases. Therefore, no interfering and unspecific peaks were detected.

The determination of the efficiency of a qPCR should be among the first things to do when setting up a qPCR assay. The efficiency of a qPCR reaction is defined as the ability of the polymerase reaction to convert reagents (dNTPs, oligos and template cDNA) into amplicon. Ideally, an efficient qPCR reaction achieves a twofold increase in amplicon per cycle[22]. In detail, PCR amplification efficiency must be established by means of standard curves and is determined from the slope of the log-linear portion of the calibration curve[14]. qPCR efficiency values must be within a range from 90 to 110% and with a standard curve correlation coefficient (R^2) ≥0.98[22,23]. As can be seen in **Table 1,** the efficiency of all the primer pairs designed ranged from 90.5% to 109.8% with an R^2 ranging from 0.980 to 0.999, which fulfils the requirements previously defined. Furthermore, repeatability and reproducibility has been assessed (**Fig 2**).

Fig 2. Intra-assay (repeatability) and inter-assay (reproducibility) precision of the validated genes.

In addition, the two reference genes demonstrated their ability to work properly at the two annealing temperatures (Ta) used. This feature is

desirable in order to normalize the results in qPCR studies because the gene of interest and reference gene should share a common Ta.

Although in the vast majority of the literature focused mainly on PCR efficiency[15,22,24], the establishment of the limit of detection (LOD) is also recommended by the MIQE guidelines[14]. Although we did not address this issue specifically, we indirectly came up against it. In some cases (e.g. *Aurka* and *p53* gene), apart from no signal detected at concentrations lower than 5 ng cDNA, the absence of a fluorescence signal and/or anomalous C_q variation was detected within technical replicates in these low concentrations. Accordingly, the lowest standard curve concentration was increased. These arrangements improved the qPCR efficiency without the need for an undesirable increase in the minimum cDNA concentration (note that rat distal colon is a rich source of total RNA and abundant tissue is available).

Pilot reference gene validation

Normalizing the data by choosing the appropriate reference genes is fundamental for obtaining reliable results in reverse transcription-qPCR (RT-qPCR. This process enables different mRNA concentrations across different samples to be compared [14]. Normalization involves the use of stably expressed endogenous reference genes in relation to the expression levels of the gene(s) of interest. However, the expression levels of the reference genes may change between tissues and species and might be influenced by pathological conditions and therapies[25–27]. Hence, an inappropriate choice of reference genes could lead to erroneous interpretations of results [28]. Therefore, the selection and validation of the reference genes is a crucial step

before planning any expression analysis. In this study, we selected two reference genes commonly used in DMH/AOM rat model experiments (*β-actin or Actb* and *B2m*). To our knowledge, this is the first study to address the exploration of valid reference genes in rat colon tissue after dietary interventions (**S1 Appendix**).

As mentioned above, the two reference genes were submitted to the same tests summarized in **Fig 1** as the other seventeen genes of interest (*in silico* design, specificity through PCR, qPCR assay validation and specificity through melting curve analysis). Nevertheless, for the reference genes, one further step was done: to check their stability. As detailed previously in the Material and Methods section, rat colon tissue samples from a parallel dietary intervention in the AOM rat model experiment performed at the same time as this validation were used (**S1 Appendix**).

In order to calculate the stability of the reference gene between the different experimental treatments, we used two software tools: Bio-Rad CFX Maestro 1.1. software (based on geNorm algorithm[29]) and the output of the three software packages using the web-based RefFinder platform[30]. In general, the two reference genes studied presented an analogous pattern with good expression stability (M values<1.0) according to the geNorm algorithm, indicating excellent stability for both genes. Nevertheless, taking into account the RefFinder output (which considers the NormFinder and BestKeeper algorithms apart from the geNorm strategy) promotes the use of *B2m* towards to *Actb* gene. The gene with the lowest geomean value is viewed as the most stable reference gene. In detail, the software gave values of 1.19 and 1.41 for

374 *B2m* and *Actb* genes, respectively. Accordingly, although the differences were
375 minimal, *B2m* was established as the more appropriate reference gene.
376
377 **Conclusions**
378 qPCR is one of the methods of choice for gene expression analysis
379 given its high sensitivity and because it works with very low nucleic acid
380 concentrations. Nonetheless, there is a lack of qPCR validation information in
381 the literature consulted. A lack of validation of the gene expression from the
382 DMH/AOM rat model by qPCR is in line with the literature reviewed by Jacob
383 et al. [27], who concluded that compliance with the MIQE guidelines continues
384 to be an ongoing issue in the scientific community. Specifically, such essential
385 information as the RNA integrity, the amount of cDNA, the linearity range and
386 the efficiency of the qPCR is frequently missed.
387 In this study, with the aim of overcoming the lack of qPCR validation in
388 the rodent CRC model, 17 rat genes related to human/rodent CRC were
389 designed and validated following the MIQE guidelines[14,22]. Furthermore,
390 two reference genes commonly used in colon cancer studies were tested for
391 their stability. Overall, this study provides a detailed list of 17 primer pairs for
392 rat-related human/rodent CRC genes and demonstrates the proper stability in
393 dietary approaches in the rat CRC model of two reference genes frequently
394 used in such studies.
395
396 **Acknowledgements**
397 We sincerely thank Isabel Sánchez and M. Alba Sorolla (UdL SCT-
398 Proteomics & Genomics Unit) for their advice on the empirical qPCR

validation performed. The lead author would like to express gratitude to its family and Dra. Cristina Falip Centellas, Dr. Francesc Pujol Aymerich and Dr. Josep M. Pericay Hosta for helping me to continue my PhD journey positively.

Literature cited

1. International Agency for Research on Cancer. Fact Sheets by Cancer [Internet]. [cited 21 Mar 2018]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.
2. National Cancer Institute. Colorectal Cancer - Cancer Stat Facts [Internet]. [cited 21 Mar 2018]. Available from: https://seer.cancer.gov/statfacts/html/colorect.html.
3. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. Eur J Cancer. 2013; 49: 1374–1403. doi:10.1016/j.ejca.2012.12.027
4. Koo JH, Leong RWL, Ching J, Yeoh K-G, Wu D-C, Murdani A, et al. Knowledge of, attitudes toward, and barriers to participation of colorectal cancer screening tests in the Asia-Pacific region: a multicenter study. Gastrointest Endosc. 2012; 76: 126–135. doi:10.1016/j.gie.2012.03.168
5. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017; 66: 683–691. doi:10.1136/gutjnl-2015-310912
6. Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. Am J Epidemiol. 1998; 148: 4–16. Available:

http://www.ncbi.nlm.nih.gov/pubmed/9663397

7. Pan P, Yu J, Wang L-S. Colon Cancer. Surg Oncol Clin N Am. 2018; 27: 243–267. doi:10.1016/j.soc.2017.11.002

8. Kurotani K, Budhathoki S, Man Joshi A, Yin G, Toyomura K, Kono S, et al. Dietary patterns and colorectal cancer in a Japanese population: The Fukuoka Colorectal Cancer Study. Br J Nutr. 2010; 104: 1703–1711. doi:10.1017/S0007114510002606

9. Willett WC. Diet and health: what should we eat? Science 1994; 264: 532–7.

10. Grosso G, Bella F, Godos J, Sciacca S, Del Rio D, Ray S, et al. Possible role of diet in cancer: systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. Nutr Rev. 2017; 75: 405–419. doi:10.1093/nutrit/nux012

11. Perše M, Cerar A. Morphological and molecular alterations in 1,2 dimethylhydrazine and azoxymethane induced colon carcinogenesis in rats. J Biomed Biotechnol. 2011; Vol. 2011: 473964. doi:10.1155/2011/473964

12. Femia A, Caderni G. Rodent Models of Colon Carcinogenesis for the Study of Chemopreventive Activity of Natural Products. Planta Med. 2008; 74: 1602–1607. doi:10.1055/s-2008-1074577

13. Megaraj V, Ding X, Fang C, Kovalchuk N, Zhu Y, Zhang Q-Y. Role of hepatic and intestinal p450 enzymes in the metabolic activation of the colon carcinogen azoxymethane in mice. Chem Res Toxicol. 2014; 27: 656–62. doi:10.1021/tx4004769

14. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et

449 al. The MIQE Guidelines: Minimum Information for Publication of
450 Quantitative Real-Time PCR Experiments. Clin Chem. 2009; 55: 611–
451 622. doi:10.1373/clinchem.2008.112797
452 15. Derveaux S, Vandesompele J, Hellemans J. How to do successful gene
453 expression analysis using real-time PCR. Methods 2010; 50: 227–230.
454 doi:10.1016/j.ymeth.2009.11.001
455 16. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al.
456 Ensembl 2018. Nucleic Acids Res. 2018; 46: D754–D761.
457 doi:10.1093/nar/gkx1098
458 17. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm
459 M, et al. Primer3--new capabilities and interfaces. Nucleic Acids Res.
460 2012; 40: e115. doi:10.1093/nar/gks596
461 18. Thornton B, Basu C. Real-time PCR (qPCR) primer design using free
462 online software. Biochem Mol Biol Educ. 2011; 39: 145–54.
463 doi:10.1002/bmb.20461
464 19. Owczarzy R, Tataurov A V, Wu Y, Manthey JA, McQuisten KA,
465 Almabrazi HG, et al. IDT SciTools: a suite for analysis and design of
466 nucleic acid oligomers. Nucleic Acids Res. 2008; 36: W163-W169.
467 doi:10.1093/nar/gkn198
468 20. Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, et
469 al. The UCSC Genome Browser database: update 2010. Nucleic Acids
470 Res. 2010; 38: D613–D619. doi:10.1093/nar/gkp939
471 21. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S,
472 Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Res.
473 2008; 36: W5–W9. doi:10.1093/nar/gkn201

474 22. Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M. A practical
475 approach to RT-qPCR—Publishing data that conform to the MIQE
476 guidelines. Methods 2010; 50: S1–S5. doi:10.1016/j.ymeth.2010.01.005
477 23. Kennedy S, Oswald N. PCR troubleshooting and optimization : the
478 essential guide. 1st ed. Caister Academic Press; 2011.
479 24. Sun M, Lu M-X, Tang X-T, Du Y-Z. Exploring Valid Reference Genes for
480 Quantitative Real-Time PCR Analysis in Sesamia inferens (Lepidoptera:
481 Noctuidae). Zhang Y, editor. PLoS One 2015; 10: e0115979.
482 doi:10.1371/journal.pone.0115979
483 25. van Rijn SJ, Riemers FM, van den Heuvel D, Wolfswinkel J, Hofland L,
484 Meij BP, et al. Expression Stability of Reference Genes for Quantitative
485 RT-PCR of Healthy and Diseased Pituitary Tissue Samples Varies
486 Between Humans, Mice, and Dogs. Mol Neurobiol. 2014; 49: 893–899.
487 doi:10.1007/s12035-013-8567-7
488 26. Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A.
489 Validation of housekeeping genes for normalizing RNA expression in
490 real-time PCR. Biotechniques 2004; 37: 112–119.
491 27. Jacob F, Guertler R, Naim S, Nixdorf S, Fedier A, Hacker NF, et al.
492 Careful Selection of Reference Genes Is Required for Reliable
493 Performance of RT-qPCR in human normal and cancer cell lines. PLoS
494 One 2013; 8: e59180. doi:10.1371/journal.pone.0059180
495 28. Dheda K, Huggett JF, Chang JS, Kim LU, Bustin SA, Johnson MA, et al.
496 The implications of using an inappropriate reference gene for real-time
497 reverse transcription PCR data normalization. Anal Biochem. 2005; 344:
498 141–143. doi:10.1016/J.AB.2005.05.022

499 29. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J.
500 qBase relative quantification framework and software for management
501 and automated analysis of real-time quantitative PCR data. Genome
502 Biol. BioMed Central 2007; 8: R19. doi:10.1186/gb-2007-8-2-r19
503 30. Xie F, Xiao P, Chen D, Xu L, Zhang B. miRDeepFinder: a miRNA
504 analysis tool for deep sequencing of plant small RNAs. Plant Mol Biol.
505 2012; 80: 75–84. doi:10.1007/s11103-012-9885-2
506 31. Zucchini N, de Sousa G, Bailly-Maitre B, Gugenheim J, Bars R, Lemaire
507 G, et al. Regulation of Bcl-2 and Bcl-xL anti-apoptotic protein
508 expression by nuclear receptor PXR in primary cultures of human and
509 rat hepatocytes. Biochim Biophys Acta. 2005; 1745: 48–58.
510 doi:10.1016/j.bbamcr.2005.02.005
511 32. Tan BL, Norhaizan ME, Huynh K, Yeap SK, Hazilawati H, Roselina K.
512 Brewers' rice modulates oxidative stress in azoxymethane-mediated
513 colon carcinogenesis in rats. World J Gastroenterol. 2015; 21: 8826–35.
514 doi:10.3748/wjg.v21.i29.8826
515 33. Wu K, Li S, Bodhinathan K, Meyers C, Chen W, Campbell-Thompson
516 M, et al. Enhanced expression of Pctk1, Tcf12 and Ccnd1 in
517 hippocampus of rats: Impact on cognitive function, synaptic plasticity
518 and pathology. Neurobiol Learn Mem. 2012; 97: 69–80.
519 doi:10.1016/j.nlm.2011.09.006
520 34. Peinnequin A, Mouret C, Birot O, Alonso A, Mathieu J, Clarençon D, et
521 al. Rat pro-inflammatory cytokine and cytokine related mRNA
522 quantification by real-time polymerase chain reaction using SYBR
523 green. BMC Immunol. 2004; 5: 3. doi:10.1186/1471-2172-5-3

35. Nergiz I, Başeskioğlu B, Yenilmez A, Erkasap N, Can C, Tosun M. Effects of rotenone on inducible nitric oxide synthase and cyclooxygenase-2 mRNA levels detected by real-time PCR in a rat bladder ischemia/reperfusion model. Exp Ther Med. 2012; 4: 344–348. doi:10.3892/etm.2012.596

36. Kensara OA, El-Shemi AG, Mohamed AM, Refaat B, Idris S, Ahmad J. Thymoquinone subdues tumor growth and potentiates the chemopreventive effect of 5-fluorouracil on the early stages of colorectal carcinogenesis in rats. Drug Des Devel Ther. 2016; 10: 2239–53. doi:10.2147/DDDT.S109721

37. Li H-B, Qin D-N, Cheng K, Su Q, Miao Y-W, Guo J, et al. Central blockade of salusin β attenuates hypertension and hypothalamic inflammation in spontaneously hypertensive rats. Sci. Rep. 2015; 5: 11162. doi:10.1038/srep11162

Supporting information captions

S1 Appendix. LLARS study flowchart.

S1 Table. List of the primer pairs not validated successfully.

A

PubMed

Select genes described in DMH/AOM rodent CRC models

NCBI Resources How To

PubMed.gov Nucleotide

US National Library of Medicine National Institutes of Health

Obtain Fasta sequences for each accession number selected

Crosstalk

Ensembl

Take into account all the transcripts described for each gene.

Clustal Omega

Multiple Sequence Alignment

If necessary, search for the homologue regions among all possible transcripts

Primer3Plus

pick primers from a DNA sequence

Pick primers from the homologous region selected in the previous step (if necessary)

Checking their aptitude

Secondary primer structures (hairpin, self-dimer and cross-dimer)

OligoAnalyzer 3.1

IDT INTEGRATED DNA TECHNOLOGIES

Secondary amplicon structures

UNAFold

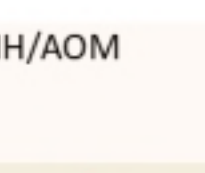
In silico specificity

UNIVERSITY OF CALIFORNIA SANTA CRUZ UCSC

Nucleotide BLAST

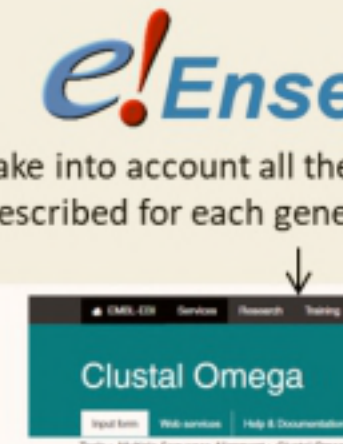
B

Primer pair selected



Wistar rat colon cDNA obtained

Check its specificity through conventional PCR + agarose gel. Optimization annealing temperature (if necessary)

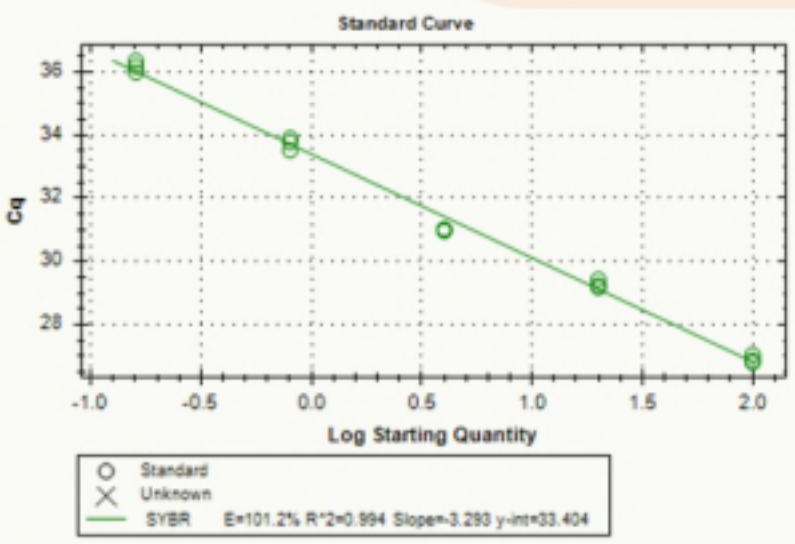


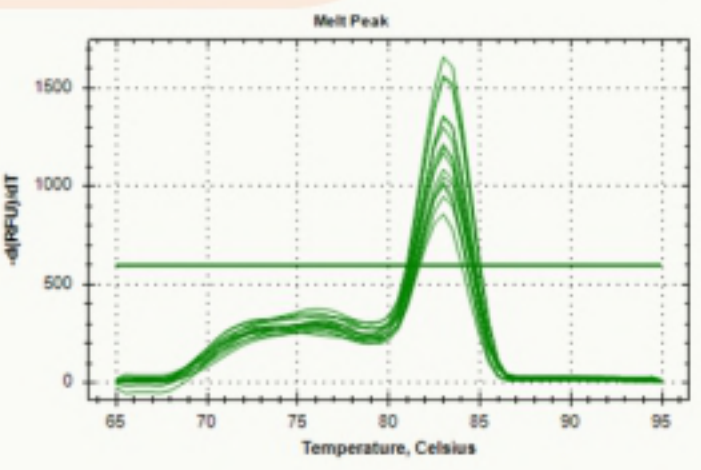
No or unspecific bands

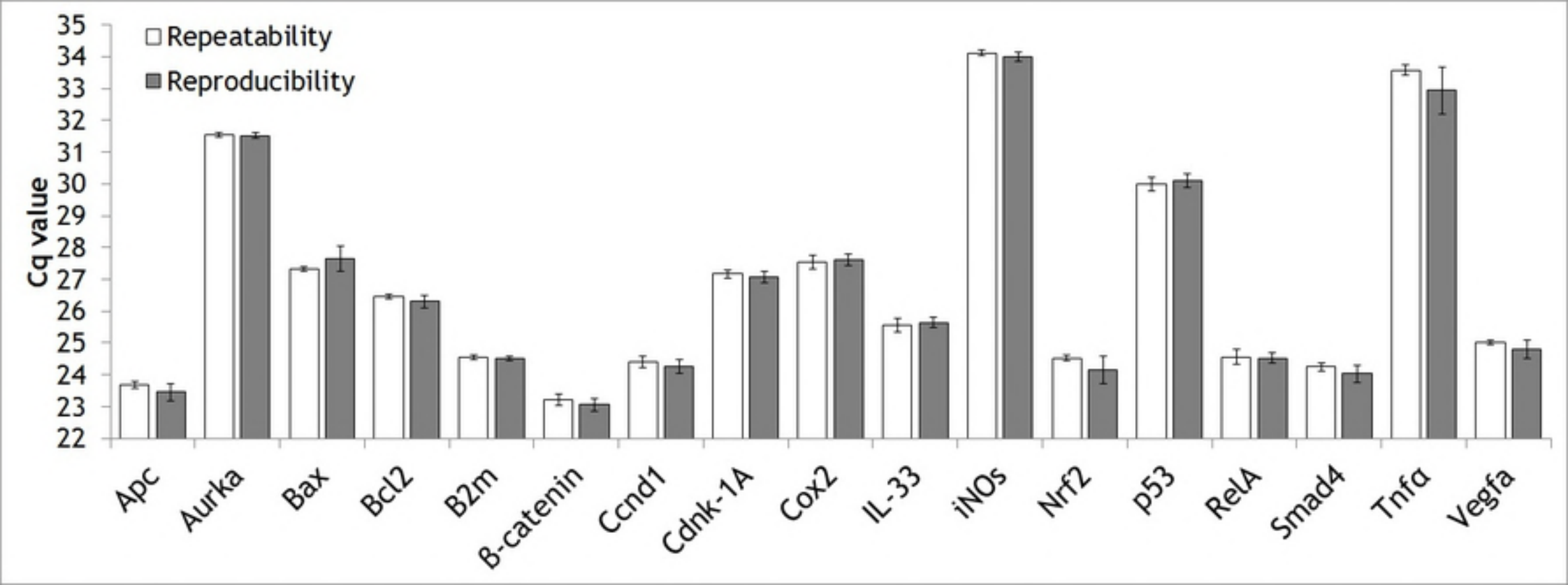
Go to Part 1 and select another primers pair

Unique band

qPCR assays assessed for MIQE guideline premises: efficiency, linearity, specificity (unique and clear melt curve)






Repeatability
Reproducibility
Cq value
35
34
33
32
31
30
29
28
27
26
25
24
23
22
Apc
Aurka
Bax
Bcl2
B2m
B-catenin
Ccnd1
Cdnk-1A
Cox2
IL-33
iNOs
Nrf2
p53
RelA
Smad4
Tnfa
Vegfa