Suitable reference genes determination for real-time PCR using induced sputum samples

Catherine Moermans [1,2], Esteban Deliege[3], Dimitri Pirottin[4,5], Christophe Poulet [6], Julien Guiot [1,2], Monique Henket[1,2], Jane da Silva[7] and Renaud Louis[1,2]

Affiliations: [1]Dept of Pneumology-Allergology, CHU of Liege, Liege, Belgium. [2]I3 group, GIGA Research Center, University of Liege, Liege, Belgium. [3]Haute école de la Province de Liège (HEPL), Liege, Belgium. [4]Laboratory of Cellular and Molecular Immunology, GIGA Research, University of Liège, Liège, Belgium. [5]Faculty of Veterinary Medicine, University of Liège, Liège, Belgium. [6]Unit of Human Genetics, GIGA Research Center, University of Liege, Liege, Belgium. [7]Post-graduate Program in Health Sciences, University of Southern Santa Catarina (UNISUL), Tubarão, Brazil.

Correspondence: Catherine Moermans, Dept of Pneumology-Allergology, Avenue de l'hopital, 13, 4000 Liege, Belgium. E-mail: c.moermans@chuliege.be

ABSTRACT Induced sputum is a non-invasive method of collecting cells from airways. Gene expression analysis from sputum cells has been used to understand the underlying mechanisms of airway diseases such as asthma or chronic obstructive pulmonary disease (COPD). Suitable reference genes for normalisation of target mRNA levels between sputum samples have not been defined so far.

The current study assessed the expression stability of nine common reference genes in sputum samples from 14 healthy volunteers, 12 asthmatics and 12 COPD patients.

Using three different algorithms (geNorm, NormFinder and BestKeeper), we identified HPRT1 and GNB2L1 as the most optimal reference genes to use for normalisation of quantitative reverse transcriptase (RT) PCR data from sputum cells. The higher expression stability of HPRT1 and GNB2L1 were confirmed in a validation set of patients including nine healthy controls, five COPD patients and five asthmatic patients. In this group, the RNA extraction and RT-PCR methods differed, which attested that these genes remained the most reliable whatever the method used to extract the RNA, generate complementary DNA or amplify it.

Finally, an example of relative quantification of gene expression linked to eosinophils or neutrophils provided more accurate results after normalisation with the reference genes identified as the most stable compared to the least stable and confirmed our findings.

Introduction

Induced sputum is a non-invasive method used to collect cells from airways, which allows many applications such as measurement of mediators in the supernatant or detailed investigation using the sputum cells [1]. Induced sputum has been used in research to analyse gene expression profile and to better understand the pathophysiology of lung diseases. It has been used specifically to help revealing molecular mechanisms of common lung diseases such as asthma and chronic obstructive pulmonary disease (COPD) [2–6]. The development of reverse transcriptase (RT)-PCR and microarray techniques allowed to detect various RNA-containing infectious agents in induced sputum with high sensitivity [7–9] and to investigate inflammatory mediators [10–12] and microRNA expression [13, 14].

The use of reference genes for normalisation of quantitative RT-PCR (RT-qPCR) data is now the method of choice. In the literature, the most frequently found reference genes when performing RT-qPCR analyses using sputum cells are β-actin [15, 16], glyceraldehyde-3-phosphate dehydrogenase (GAPDH) [17, 18] and ribosomal RNA 18S [19, 20]. However, optimal reference genes for sputum gene expression analysis have not been explored so far. The choice of reference genes is indeed crucial for RT-qPCR data normalisation and should be assessed in each specific experiment or biological sample [21] and it is therefore of prime importance to fill this gap.

For this purpose, we screened nine commonly used reference genes in sputum cells. As their expression levels can vary according to the airway disease or cell type, we assessed their stability in samples obtained from healthy controls, asthmatic patients and patients suffering from COPD. They exhibit different sputum cellular profiles and there is therefore a need for invariant expression of the chosen reference gene(s). Three different algorithms for identifying the best reference genes among a set of candidates were applied (geNorm, NormFinder and BestKeeper). In addition, the experiment was performed in a new set of patients using different RNA extraction and RT-qPCR protocols. Finally, we used an example of relative quantification of target genes (interleukin (IL)-5 and CXCL8) known to be linked to eosinophil [22] and neutrophil recruitment [23] respectively, to attest that the choice of stable endogenous reference genes is crucial to obtain unbiased results from RT-qPCR using sputum cells.

Material and methods

Subjects

The characteristics of the patients are given in table 1. Asthmatic and COPD patients were recruited through the outpatient clinic and pulmonary rehabilitation centre (CHU, Sart-Tilman, Liège, Belgium). Asthma was diagnosed as described in the Global Initiative for Asthma guidelines (http://ginasthma.org/). Mild-to-moderate asthma was defined as patients without maintenance treatment or with a low-to-moderate dose of inhaled corticosteroids (<1000 µg beclomethasone per day) and had forced expiratory volume in 1 s ⩾80% predicted. Severe and refractory asthma were defined according to American Thoracic Society criteria [24]. Diagnosis of COPD was made according to Global Initiative for Chronic Obstructive Lung Disease criteria (http://goldcopd.org/). All asthmatics and COPD patients were recruited during stable state of the disease. Healthy volunteers were enrolled by advertisement among hospital staff. This study was approved by the local ethics committee of CHU Liège and all subjects gave written informed consent for participation.

Study design

The objective of the study was to determine the most reliable reference gene(s) to use in RT-qPCR experiments using induced sputum samples. As recommended in the manual of geNorm, NormFinder and BestKeeper, we performed comparisons of more than eight commonly used reference genes in groups of patients >10 subjects including asthmatics, COPD and healthy subjects to obtain confident results. The nine reference genes chosen were β-actin (ACTB), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), β2-microglobulin (B2M), β-glucuronidase (GUSB), hypoxanthine ribosyltransferase 1 (HPRT1), guanine nucleotide-binding protein, b-peptide 2-like 1 (GNB2L1), TATA-box binding protein (TBP), ribosomal protein L13A (RPL13A) and ribosomal RNA 18S (RNA18S), which were selected because they belong to different biological pathways and are then presumably not co-regulated.

Sputum induction and processing

The sputum was induced and processed as described previously [4, 25]. Cell viability was assessed by trypan blue exclusion and the differential leukocyte count was performed on cytospins stained with Rapi Diff II Stain Kit (Atom Scientific, Manchester, UK) on 500 cells. All samples were selected according to following selection criteria: <30% of squamous cells and viability >50%. These criteria have been determined in our lab as the optimal threshold to obtain reliable expression results. The cell pellet (median (interquartile range) 1.8 (1.2–2.2) cells×10^6) was mixed with 5 volumes of RNAprotect cell reagent (Qiagen, Hilden, Germany) and kept at −80°C until RNA extraction.

TABLE 1 Demographic and functional characteristics				
	Control	**COPD**	**Asthma**	**p-value**
Subjects	14	12	12	
Age years	48±15	58±13	55±11	0.15
Male/female	6/8	6/6	4/8	0.71
Tobacco status nonsmoker/ex-smoker/current smoker	9/5/0	0/5/7	5/6/1	<0.001
Smoking pack-years	6±10	61±27	14±11	<0.0001
Treatment				
ICS yes/no	0/14	7/5	11/1	<0.0001
LABA yes/no	0/14	9/3	6/6	<0.001
LAMA yes/no	0/14	6/6	0/12	<0.001
SABA yes/no	0/14	0/12	3/9	<0.05
LTRA yes/no	0/14	1/11	4/8	<0.05
Severity status				
GOLD 1		3		
GOLD 2		8		
GOLD 3		1		
Mild–moderate			7	
Severe–refractory			5	
FEV_1 % pred	108±11	63±14***,#	80±18***	<0.0001
FEV_1 postbronchodilator % pred	111±17	70±15***	82±19**	<0.0001
FVC % pred	114±13	89±13***	92±15***	<0.0001
FEV_1/FVC %	80±6	57±9***,##	71±8*	<0.0001
Sputum				
Squamous cells %	21.0 (4.7–27.5)	2.5 (0.0–10.5)*	4.0 (1.2–13.5)	<0.01
Total non-squamous cells $\times 10^6 \cdot g^{-1}$	1.0 (0.3–1.9)	3.3 (1.2–8.1)	4.6 (2.2–9.8)**	<0.01
Viability %	77.0 (68.5–81.8)	75.0 (64.5–86.5)	75.0 (62.0–81.0)	0.68
Macrophages %	37.6 (23.9–64.7)	18.3 (4.5–27.2)*	20.4 (11.2–25.8)	<0.01
Neutrophils %	53.7 (28.1–68.7)	69.6 (54.6–86.1)	39.0 (24.3–81.0)	<0.05
Lymphocytes %	1.9 (0.9–2.4)	2.4 (0.5–5.0)	0.5 (0.4–3.3)	0.50
Eosinophils %	0.0 (0.0–0.4)	2.3 (0.3–4.4)*	21.4 (1.8–44.2)***	<0.001
Epithelial cells %	2.8 (2.3–6.2)	1.1 (0.4–4.8)	5.8 (3.6–7.7)	0.15

Data are presented as n, mean±SD or median (interquartile range), unless otherwise stated. COPD: chronic obstructive pulmonary disease; ICS: inhaled corticosteroids; LABA: long-acting β-agonist; LAMA: long-acting muscarinic antagonist; SABA: short-acting β-agonist; LTRA: leukotriene receptor antagonist; GOLD: Global Initiative for Chronic Obstructive Lung Disease; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity. *: p<0.05, **: p<0.01, ***: p<0.001 versus healthy subjects; #: p<0.05, ##: p<0.01 versus asthmatic patients.

RNA extraction and RT-qPCR methods

These steps were performed according to the description of DA SILVA et al. [5], except that the TaqMan PCR step was achieved in 96-well plates allowing a sample maximisation approach. All procedure information was given according to the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines for the minimum information required for a qPCR experiment (experimental procedure is included in the supplementary material) [26].

Validation experiment

A new experiment was performed to analyse the stability of seven genes out of the nine previously assessed (ACTB, GAPDH, B2M, HPRT1, GNB2L1, RPL13A and RNA18S) in sputum collected from nine healthy controls, five COPD patients and five asthmatic patients. Ribosomal protein L32 (RPL32) was added to the panel, as it was previously shown to be the most stable reference gene to use for bronchoalveolar lavage (BAL) cells [27]. The characteristics of the patients are given in supplementary table S1.

In contrast to the first experiment where the RNAs were isolated using the TRIzol and phenol–chloroform extraction method followed by washing and elution on an RNA-binding column, the RNAs were directly extracted using RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. (All materials cited hereafter are from ThermoFisher Scientific, Wilmington, MA, USA.) Genomic DNA contamination was eliminated by a treatment with TURBO DNA free™ kit from Ambion. The reverse transcription was performed starting from 500 ng of RNA using a High Capacity cDNA Reverse Transcription kit according to the manufacturer's protocol. The qPCR was achieved with the Taqman Universal master mix II. The

cDNA was loaded onto a custom TaqMan low-density array (384 wells plate) as specified in the manufacturer's instructions. The plates were read using 7900 HT Fast Real-Time PCR System. The primers and probes and the efficiencies and specificities are included in the MIQE checklist file in the supplementary material. The three algorithms were applied as for the primary experiment.

Immune cell correlation

In order to validate the findings in a real application, we proceeded to a relative quantification of genes correlated with eosinophils and neutrophils (IL-5 and CXCL8, respectively). This experiment was performed with samples from six healthy controls, six patients with asthma and six COPD patients (characteristics of the patients are described in supplementary table S2). The COPD patients exhibited an intense neutrophilic inflammation and the asthmatic cohort showed high sputum eosinophil and neutrophil percentages. The RNA extraction and RT-qPCR procedures were performed as described for the first cohort. Relative quantification in gene expression compared to healthy controls was determined using qbase$^+$ qPCR analysis software (Biogazelle, Zwijnaarde, Belgium) in accordance with the target specific amplification efficiency values.

Statistics

Three different algorithms were used to assess the stability of the nine reference genes. GeNorm (implementation in qbase$^+$) uses a normalisation strategy which provides a ranking of the candidate genes according to an average stability value (M) of remaining reference genes calculated during stepwise exclusion of the least stable reference gene [28]. NormFinder, available as an Excel add-in, is based on the analysis of overall gene expression variation and the variation between sample subgroups [29]. BestKeeper software, available as an Excel-based tool determines the optimal reference gene with pairwise correlation analysis of all pairs of candidate genes and calculates the geometric mean of the best ones [30].

The demographic and functional characteristics of the patients were expressed as mean±SD and comparisons between groups were performed by one-way ANOVA followed by Tukey's multiple comparisons test for continuous variables. A Chi-squared test was applied for categorical analyses. Sputum cell counts were expressed as median (interquartile range). Comparisons between groups were performed by Kruskal–Wallis test followed by Dunn's multiple comparisons test. In the application example, the relative expression between groups of subjects was analysed in the same manner and with the Mann–Whitney test when two groups were compared. Correlations were tested with Spearman's rank correlation analysis. Statistical analyses were carried out with GraphPad Prism 7.0 (GraphPad Software, San Diego, CA, USA). Differences were considered statistically significant when a two-sided p-value was <0.05.

Results

Baseline characteristics of the patients

The three groups of patients were well matched according to sex and age, but not for tobacco habits and treatments. Regarding sputum cell counts, the COPD patients exhibited lower percentages of macrophages, but higher proportions of neutrophils than controls. In addition, COPD patients had a higher proportion of eosinophils than controls, but to a lesser extent than asthmatic patients.

Raw quantification cycle distribution of the candidate reference genes

The raw quantification cycle distributions of the candidate reference genes are presented in figure 1. They displayed a wide range of cycles, from 15.7 (15.0–16.6) for RNA18S to 33.0 (32.0–35.2) for HPRT1.

Reference gene expression stability evaluation

The ranking obtained with the three algorithms are combined in table 2. The M value obtained in the geNorm pilot experiment was low for the best reference gene and high for most unstable genes. In addition, the NormFinder algorithm gave a stability value for each candidate gene, the lowest being considered the best. Finally, BestKeeper software combined all the candidate normalisation genes into an index and analysed the correlation of this index with each individual gene. The most appropriate genes had the greatest correlation coefficient values.

The final ranking was computed by the addition of each individual rank obtained with the three algorithms. It appeared that GNB2L1 and HPRT1 were the reference genes identified as the most suitable. These three analyses showed slight differences only and the two genes exhibiting the greater variation were the same for all (namely GAPDH and RNA18S).

Validation experiment

The three algorithms were applied to the data obtained with a new set of patients and the results were combined in supplementary table S3. Even if GAPDH was ranked as the candidate gene with the highest
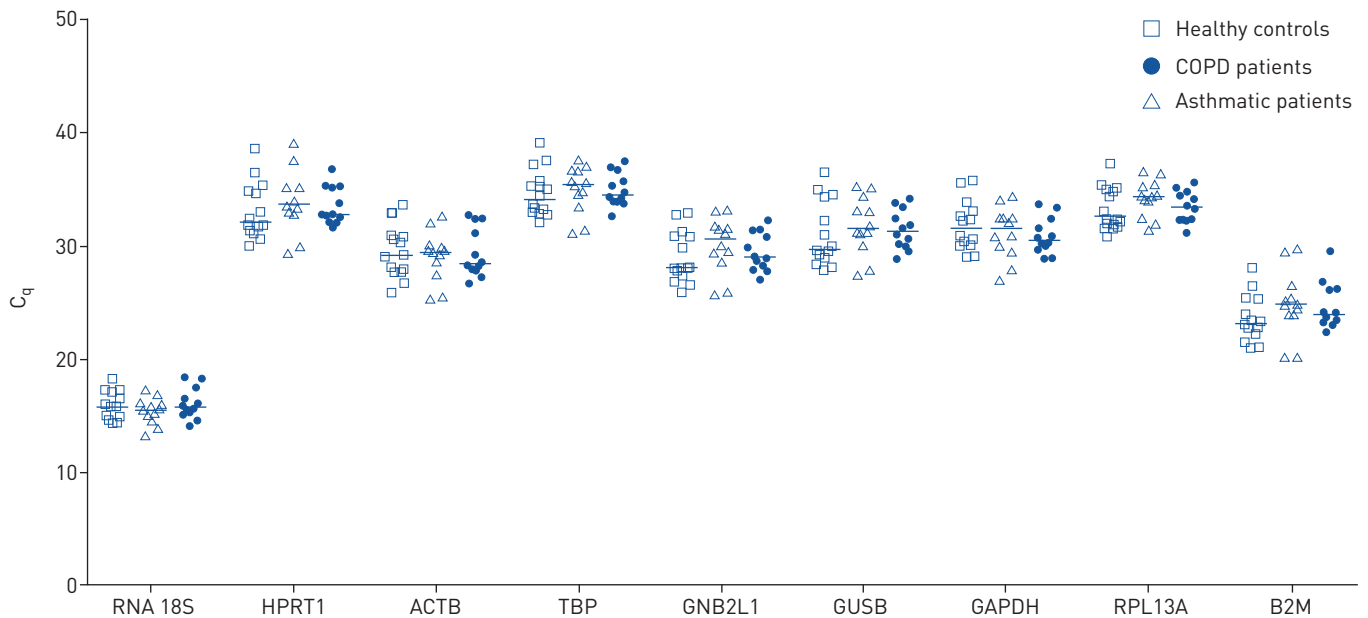
FIGURE 1 Quantification cycle (C_q) values of the nine candidate reference genes in the 38 sputum samples. The median is indicated by a bar. RNA 18S: 18 s ribosomal RNA; HPRT1: hypoxanthine ribosyltransferase 1; ACTB: β-actin; TBP: TATA-box binding protein; GNB2L1: guanine nucleotide-binding protein, b-peptide 2-like 1; GUSB: β-glucuronidase; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; RPL13A: ribosomal protein L13A; B2M: β2 microglobulin.

stability, HPRT1 and GNB2L1 still occupied the next top positions. Ribosomal 18S and RPL32 were classified as the least stable genes.

Immune cell correlation

The relative quantification of IL-5 and CXCL8 gene expression using either HPRT1 (shown as the most suitable) or RNA18S (shown as the most unstable) as reference genes was performed in a new set of patients. As shown in figure 2, the relative expression changed drastically depending on whether the normalisation was made with RNA18S or HPRT1. Indeed, we observed that, even if the quantification using RNA18S did not show any difference between groups (Kruskal–Wallis test p=0.82), the relative quantification based on HPRT1 gave significant differences (Kruskal–Wallis test p<0.05). When the data obtained from controls and asthmatics patients were compared, the results appeared significant (Mann–Whitney test p<0.05). The correlation between IL-5 expression and eosinophil percentage was nonsignificant when RNA18S was used and became significant when HPRT1 was applied (r=0.63, p<0.05).

TABLE 2 Reference gene expression stability obtained with the three algorithms

	geNorm M	Rank	NormFinder	Rank	BestKeeper r+	Rank	Final rank
GNB2L1	0.717	1	0.275	4	0.95	1	1
GUSB	0.725	2	0.260	3	0.91	6	3
HPRT1	0.739	3	0.246	2	0.94	2	2
RPL13A	0.887	4	0.279	5	0.93	4	4
B2M	1.01	5	0.405	7	0.92	5	7
ACTB	1.119	6	0.309	6	0.94	3	5
TBP	1.175	7	0.230	1	0.91	7	6
GAPDH	1.227	8	0.411	8	0.89	8	8
RNA18S	1.336	9	0.423	9	0.73	9	9

GNB2L1: guanine nucleotide-binding protein, b-peptide 2-like 1; GUSB: β-glucuronidase; HPRT1: hypoxanthine ribosyltransferase 1; RPL13A: ribosomal protein L13A; B2M: β2 microglobulin; ACTB: β-actin; TBP: TATA-box binding protein; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; RNA 18S: 18 s ribosomal RNA.
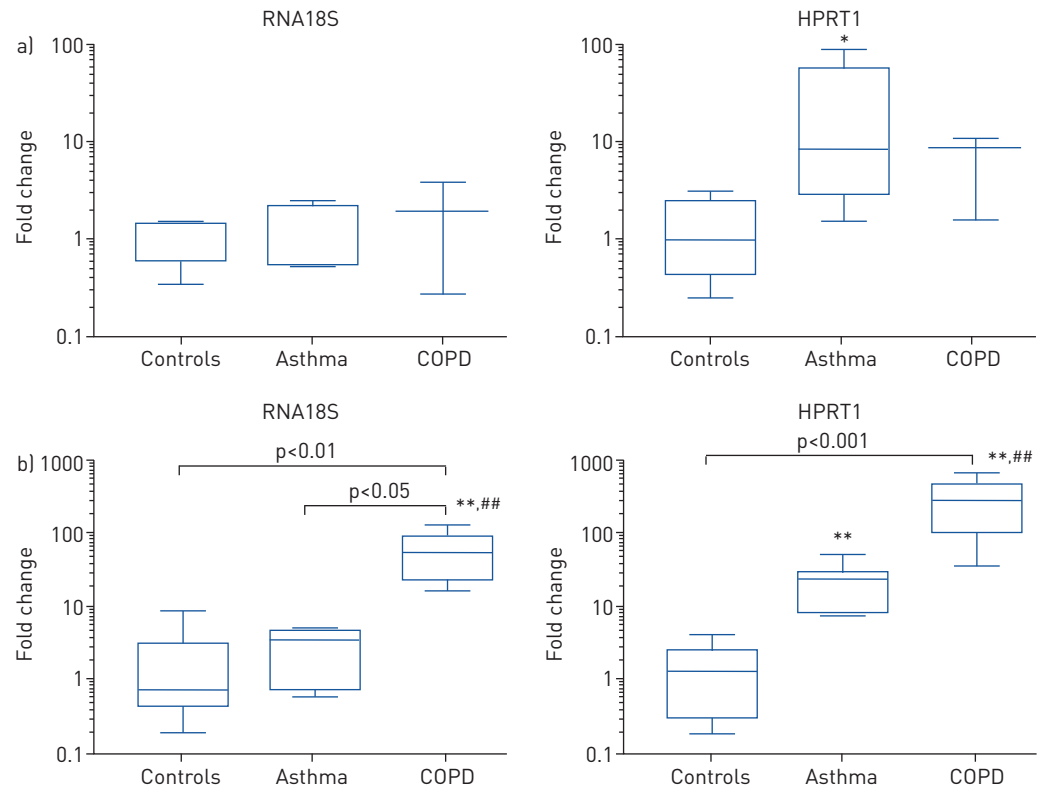
FIGURE 2 Comparison of hypoxanthine ribosyltransferase 1 (HPRT1) and 18 s ribosomal RNA (RNA18S) for normalising a) interleukin-5 or b) CXCL8 gene expression. Control patients n=4, asthmatic patients n=5, chronic obstructive pulmonary disease (COPD) patients n=3. The comparisons were made with Kruskal–Wallis test followed by Dunn's multiple comparisons test (bars) and with Mann–Whitney test when two groups were compared. *: p<0.05, **: p<0.01 *versus* healthy subjects; #: p<0.05, ##: p<0.01 *versus* asthmatic patients.

As for CXCL8 gene expression normalised with RNA18S, the Kruskal–Wallis test gave a p<0.001 and the Dunn's multiple comparison tests were significant for controls *versus* COPD (p<0.01) and asthmatics *versus* COPD (p<0.05). In contrast, when the normalisation was done with HPRT1, the Kruskal–Wallis test gave a p<0.0001 and the Dunn's multiple comparison tests were significant for controls *versus* COPD only (p<0.001). When the data obtained from controls and asthmatics patients were compared by Mann–Whitney test, the p-value was significant (p<0.01), as well as when the asthmatics and COPD patients were compared (p<0.01). The correlation between CXCL8 expression and the neutrophil percentage was more pronounced once the normalisation was made with HPRT1 (r=0.9, p<0.0001) instead of RNA18S (r=0.77, p<0.001). We observed a trend for a positive correlation between CXCL8 expression and the eosinophil percentage only when the data were quantified with HPRT1 (r=0.37, p=0.13).

When the normalisation was done with both HPRT1 and GNB2L1 compared to RNA18S and GAPDH (figure 3), the Kruskal–Wallis test was significant for IL-5 and Dunn's test gave p<0.05 for the comparison between controls and asthmatic patients.

Regarding CXCL8, the results were similar when normalised with HPRT1 and GNB2L1 compared to RNA18S and GAPDH, and did not differ from those obtained with HPRT1 alone.

Discussion

Comparisons of gene expressions from sputum samples of controls and asthmatic and COPD patients are frequent. However, until now, information about the most suitable reference genes to normalise these kinds of data are missing. To the best of our knowledge, this study is the first to investigate the most appropriate reference genes to use when performing RT-qPCR analysis using sputum cells. For this purpose, nine common reference genes known to be involved in distinct functions were chosen. Using the three algorithms, we found that GNB2L1 and HPRT1 were the most suitable reference genes to use in this context. Both were validated in another independent group of patients where the RNA extraction and RT-qPCR methods differed. They were already shown to be the most stably expressed reference genes in alveolar macrophages of COPD patients, whatever the disease severity [31], and in isolated human
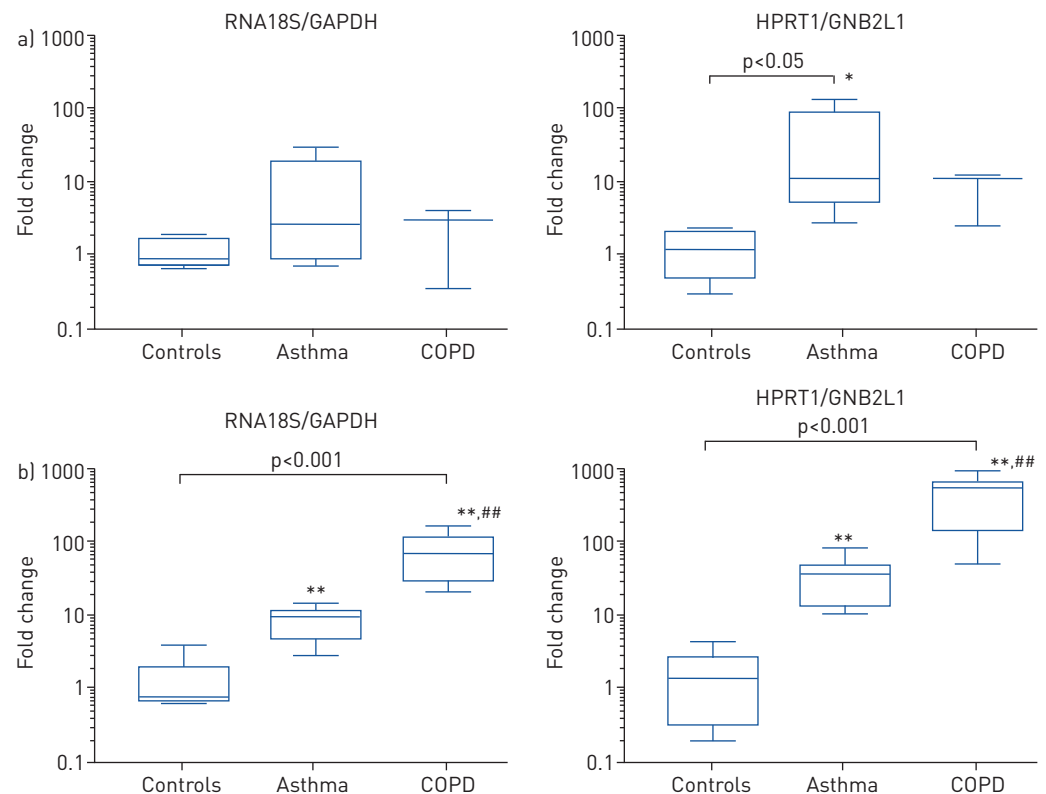
FIGURE 3 Comparison of hypoxanthine ribosyltransferase 1 (HPRT1)/guanine nucleotide-binding protein, b-peptide 2-like 1 (GNB2L1) and 18 s ribosomal RNA (RNA18S)/glyceraldehyde-3-phosphate dehydrogenase (GAPDH) for normalising a) interleukin-5 or b) CXCL8 gene expression. Control patients n=4, asthmatic patients n=5, chronic obstructive pulmonary disease (COPD) patients n=3. The comparisons were made with Kruskal–Wallis test followed by Dunn's multiple comparisons test (bars) and with Mann–Whitney test *: p<0.05, **: p<0.01 versus healthy subjects; #: p<0.05, ##: p<0.01 versus asthmatic patients when two groups were compared.

neutrophils [32]. HPRT1 was already shown to be the most stably expressed reference gene in other systems, but data regarding GNB2L1 appeared limited (supplementary table S4). Based on the expression level, HPRT1 would be more suitable for low-abundance transcripts in induced sputum, GNB2L1 being more appropriate in case of higher abundance transcripts.

Even if commonly used in the context of sputum cells, GAPDH, β-actin and RNA18S did not appear as good candidate reference genes. In a previous study, GAPDH and β-actin were shown as unstable in BAL and bronchial biopsies from asthmatic patients due to different cellular profiles and activation status [33]. It is interesting to note that GAPDH is classified as one of the most variable reference gene in the first cohort and as the most stable in the validation cohort. This discrepancy may be explained by a number of reasons. Indeed, GAPDH is implicated in many cellular process and has many functions in addition to its glycolytic activity. Furthermore, the use of inhaled corticosteroids appeared to influence the expression level of GAPDH [33] and the proportions of patients treated with inhaled corticosteroids is different in each of our two cohorts. Finally, the fact that the primer sequences are different between the two experiments could also explain the variability of our results.

In the same manner, even if the ribosomal RNA level variation is supposed to be low compared to mRNA, it is also regulated according to the cell type, the functional state and it varies between different individuals. In addition, its high abundance and nature limit its use for mRNA normalisation. Moreover, ribosomal RNA is thought to be less affected by RNA degradation compared to other genes and may not serve as a good endogenous control in this regard [34].

Finally, a practical example of relative quantification using RNA18S (classified as the most unstable gene by the three algorithms) compared with HPRT1 (identified as the most suitable) showed contrasting results and highlighted the importance of the reference gene choice. Indeed, some differences can be hidden and the interpretation of the results may be mistaken when the normalisation is done using genes with variable expression. The expression of IL-5 is known to be linked with eosinophil recruitment and

highly expressed in sputum of asthmatic patients [12, 35] compared to healthy volunteers. CXCL8, for its part, was found to be increased in COPD [36, 37], but also in patients with asthma where it participates in neutrophil [23, 38] and eosinophil chemotaxis [39] as shown previously and attested by our positive correlations. Although normalisation against a single reference gene is acceptable when there is a validation of its stability under the experimental conditions [26], the use of two reference genes is recommended to limit errors and increase the accuracy of results [40]. In our experiment, even if the use of two reference genes gave the same results for CXCL8 when the normalisation was performed with the two most stable genes compared to the two most unsuitable, the results for IL-5 were still remarkably different.

These results could be confirmed using other technologies such as Droplet Digital PCR assay, NanoString, microarray or massive parallel RNA sequencing. Indeed, the use of reference genes is not mandatory, as they provide additional normalisation strategies. However, the applicability of these emerging technologies is restrained, as not all research centres are equipped. The other issue is the extensive bioinformatic analyses linked to these techniques. Nevertheless, it is important to note that qPCR remains the gold standard for expression analysis and is used to confirm results from high-throughput analyses.

The limitation of our study is the low number of patients, which did not allow subgroup analyses. For this reason, the authors would recommend the validation of the reference genes prior to their use in an experimental protocol comparing patients with different treatments, tobacco habits or disease severity.

In conclusion, GNB2L1 and HPRT1 are the most ideal reference genes to use for RT-qPCR data normalisation when working with induced sputum and are not affected by airway diseases, sputum cellular composition, RNA extraction and RT-qPCR methods.

References

1   Guiot J, Demarche S, Henket M, et al. Methodology for sputum induction and laboratory processing. *J Vis Exp* 2017; 130: 10.3791/56612.
2   Louis R, Lau LCK, Bron AO, et al. The relationship between airways inflammation and asthma severity. *Am J Respir Crit Care Med* 2000; 161: 9–16.
3   Bakakos P, Schleich F, Alchanatis M, et al. Induced sputum in asthma: from bench to bedside. *Curr Med Chem* 2011; 18: 1415–1422.
4   Moermans C, Heinen V, Nguyen M, et al. Local and systemic cellular inflammation and cytokine release in chronic obstructive pulmonary disease. *Cytokine* 2011; 56: 298–304.
5   da Silva J, Hilzendeger C, Moermans C, et al. Raised interferon-β, type 3 interferon and interferon-stimulated genes – evidence of innate immune activation in neutrophilic asthma. *Clin Exp Allergy* 2017; 47: 313–323.
6   Ronchi MC, Piragino C, Rosi E, et al. Role of sputum differential cell count in detecting airway inflammation in patients with chronic bronchial asthma or COPD. *Thorax* 1996; 51: 1000–1004.
7   Fujisawa T, Suda T, Matsuda H, et al. Real-time PCR is more specific than conventional PCR for induced sputum diagnosis of *Pneumocystis* pneumonia in immunocompromised patients without HIV infection. *Respirology* 2009; 14: 203–209.
8   Simpson JL, Moric I, Wark PAB, et al. Use of induced sputum for the diagnosis of influenza and infections in asthma: a comparison of diagnostic techniques. *J Clin Virol* 2003; 26: 339–346.
9   Yamazaki T, Narita M, Sasaki N, et al. Comparison of PCR for sputum samples obtained by induced cough and serological tests for diagnosis of *Mycoplasma pneumoniae* infection in children. *Clin Vaccine Immunol* 2006; 13: 708–710.
10 Guiot J, Henket M, Corhay JL, et al. Sputum biomarkers in IPF: evidence for raised gene expression and protein level of IGFBP-2, IL-8 and MMP-7. *PLoS One* 2017; 12: e0171344.
11 Gelder CM, Thomas PS, Yates DH, et al. Cytokine expression in normal, atopic, and asthmatic subjects using the combination of sputum induction and the polymerase chain reaction. *Thorax* 1995; 50: 1033–1037.
12 Truyen E, Coteur L, Dilissen E, et al. Evaluation of airway inflammation by quantitative Th1/Th2 cytokine mRNA measurement in sputum of asthma patients. *Thorax* 2006; 61: 202–208.
13 Van Pottelberge GR, Mestdagh P, Bracke KR, et al. MicroRNA expression in induced sputum of smokers and patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011; 183: 898–906.
14 Maes T, Cobos FA, Schleich F, et al. Asthma inflammatory phenotypes show differential microRNA expression in sputum. *J Allergy Clin Immunol* 2016; 137: 1433–1446.
15 Baines KJ, Fu JJ, McDonald VM, et al. Airway gene expression of IL-1 pathway mediators predicts exacerbation risk in obstructive airway disease. *Int J Chron Obstruct Pulmon Dis* 2017; 12: 541–550.

16 Berthon BS, Gibson PG, Wood LG, *et al.* A sputum gene expression signature predicts oral corticosteroid response in asthma. *Eur Respir J* 2017; 49: 1700180.

17 Bjerregaard A, Laing IA, Poulsen N, *et al.* Characteristics associated with clinical severity and inflammatory phenotype of naturally occurring virus-induced exacerbations of asthma in adults. *Respir Med* 2017; 123: 34–41.

18 Maneechotesuwan K, Wongkajornsilp A, Adcock IM, *et al.* Simvastatin suppresses airway IL-17 and upregulates IL-10 in patients with stable COPD. *Chest* 2015; 148: 1164–1176.

19 Fu JJ, Baines KJ, Wood LG, *et al.* Systemic inflammation is associated with differential gene expression and airway neutrophilia in asthma. *OMICS* 2013; 17: 187–199.

20 Simpson JL, Phipps S, Baines KJ, *et al.* Elevated expression of the NLRP3 inflammasome in neutrophilic asthma. *Eur Respir J* 2014; 43: 1067–1076.

21 De Spiegelaere W, Dern-Wieloch J, Weigel R, *et al.* Reference gene validation for RT-qPCR, a note on different available software packages. *PLoS One* 2015; 10: e0122515.

22 Sanderson CJ. The biological role of interleukin 5. *Int J Cell Cloning* 1990; 8: 147–154.

23 Paplińska-Goryca M, Nejman-Gryz P, Górska K, *et al.* Expression of inflammatory mediators in induced sputum: comparative study in asthma and COPD. *Adv Exp Med Biol* 2018; 1040: 101–112.

24 Wenzel SE, Fahy JV, Irvin C, *et al.* Proceedings of the ATS workshop on refractory asthma: current understanding, recommendations, and unanswered questions. *Am J Respir Crit Care Med* 2000; 162: 2341–2351.

25 Delvaux M, Henket M, Lau L, *et al.* Nebulised salbutamol administered during sputum induction improves bronchoprotection in patients with asthma. *Thorax* 2004; 59: 111–115.

26 Bustin SA, Benes V, Garson JA, *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 2009; 55: 611–622.

27 Kriegova E, Arakelyan A, Fillerova R, *et al.* PSMB2 and RPL32 are suitable denominators to normalize gene expression profiles in bronchoalveolar cells. *BMC Mol Biol* 2008; 9: 69.

28 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: research0034.1.

29 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: 5245–5250.

30 Pfaffl MW, Tichopad A, Prgomet C, *et al.* Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnol Lett* 2004; 26: 509–515.

31 Ishii T, Wallace AM, Zhang X, *et al.* Stability of housekeeping genes in alveolar macrophages from COPD patients. *Eur Respir J* 2006; 27: 300–306.

32 Zhang X, Ding L, Sandford AJ. Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Mol Biol* 2005; 6: 4.

33 Glare EM, Divjak M, Bailey MJ, *et al.* β-actin and GAPDH housekeeping gene expression in asthmatic airways is variable and not suitable for normalising mRNA levels. *Thorax* 2002; 57: 765–770.

34 Lossos IS, Czerwinski DK, Wechser MA, *et al.* Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. *Leukemia* 2003; 17: 789–795.

35 Marra SMG, Borges RO, Alves R, *et al.* Spirometric parameters and levels of interferon gamma and IL-5 in induced sputum from patients with allergic rhinitis or asthma. *Am J Rhinol Allergy* 2011; 25: e196–e199.

36 Maneechotesuwan K, Kasetsinsombat K, Wongkajornsilp A, *et al.* Decreased indoleamine 2,3-dioxygenase activity and IL-10/IL-17A ratio in patients with COPD. *Thorax* 2013; 68: 330–337.

37 Moermans C, Bonnet C, Willems E, *et al.* Sputum cytokine levels in patients undergoing hematopoietic SCT and comparison with healthy subjects and COPD: a pilot study. *Bone Marrow Transplant* 2014; 49: 1382–1388.

38 Bullens DMA, Truyen E, Coteur L, *et al.* IL-17 mRNA in sputum of asthmatic patients: linking T cell driven inflammation and granulocytic influx? *Respir Res* 2006; 7: 135.

39 Gordon JR, Swystun VA, Li F, *et al.* Regular salbutamol use increases CXCL8 responses in asthma: relationship to the eosinophil response. *Eur Respir J* 2003; 22: 118–126.

40 Thellin O, Zorzi W, Lakaye B, *et al.* Housekeeping genes as internal standards: use and limits. *J Biotechnol* 1999; 75: 291–295.