

## RESEARCH ARTICLE

# *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* Were Identified as Reference Genes in Neuroendocrine Lung Cancer via the nCounter Technology

Robert Fred Henry Walter<sup>1,2\*</sup>, Robert Werner<sup>3</sup>, Claudia Vollbrecht<sup>4,5,6</sup>, Thomas Hager<sup>2</sup>, Elena Flom<sup>1</sup>, Daniel Christian Christoph<sup>7</sup>, Jan Schmeller<sup>2</sup>, Kurt Werner Schmid<sup>2</sup>, Jeremias Wohlschlaeger<sup>2,8</sup>, Fabian Dominik Mairinger<sup>2</sup>

**1** Ruhrlandklinik, West German Lung Center, University Hospital Essen, University of Duisburg-Essen, Essen, Germany, **2** Institute of Pathology, University Hospital Essen, University of Duisburg-Essen, Essen, Germany, **3** Institute of Pathology, Helios Klinikum Emil von Behring, Berlin, Germany, **4** German Cancer Consortium (DKTK), partnersite Berlin, Germany, **5** Charité - Universitaetsmedizin Berlin, Institute of Pathology, Berlin, Germany, **6** German Cancer Research Center (DKFZ), Heidelberg, Germany, **7** Department of Medical Oncology, West German Cancer Center, University Hospital Essen, University of Duisburg-Essen, Essen, Germany, **8** Institute of Pathology, Ev.-Luth. Diakonissenkrankenhaus Flensburg, Flensburg, Germany

\* [robert.walter@ruhrlandklinik.uk-essen.de](mailto:robert.walter@ruhrlandklinik.uk-essen.de)



CrossMark  
click for updates

 OPEN ACCESS

**Citation:** Walter RFH, Werner R, Vollbrecht C, Hager T, Flom E, Christoph DC, et al. (2016) *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* Were Identified as Reference Genes in Neuroendocrine Lung Cancer via the nCounter Technology. PLoS ONE 11(11): e0165181. doi:10.1371/journal.pone.0165181

**Editor:** Giancarlo Troncone, Universita degli Studi di Napoli Federico II, ITALY

**Received:** July 21, 2016

**Accepted:** October 8, 2016

**Published:** November 1, 2016

**Copyright:** © 2016 Walter et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** \*\*\*For PA\*\*\* Please request for contact info for additional data. All relevant data are within the paper and its Supporting Information files. Ethical concerns preclude publication of clinical and pathological data. To prevent conclusion on individuals investigated, only pure expression data and tumor type can be published. We try to avoid releasing information about age of the patient, patient ID, year of treatment etc. as stated in the ethic's approval of the Ethics Committee of the University

## Abstract

### Background

Neuroendocrine lung cancer (NELC) represents 25% of all lung cancer cases and large patient collectives exist as formalin-fixed, paraffin-embedded (FFPE) tissue only. FFPE is controversially discussed as source for molecular biological analyses and reference genes for NELC are poorly established.

### Material and methods

Forty-three representative FFPE-specimens were used for mRNA expression analysis using the digital nCounter technology (NanoString). Based on recent literature, a total of 91 mRNA targets were investigated as potential tumor markers or reference genes. The geNorm, NormFinder algorithms and coefficient of correlation were used to identify the most stable reference genes. Statistical analysis was performed by using the R programming environment (version 3.1.1)

### Results

RNA integrity (RIN) ranged from 1.8 to 2.6 and concentrations from 34 to 2,109 ng/μl. However, the nCounter technology gave evaluable results for all samples tested. *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* were identified as constantly expressed genes with high stability (M-)values according to geNorm, NormFinder and coefficients of correlation.

Hospital Essen, University Duisburg-Essen, Essen, Germany.

**Funding:** The study was financed by the Institute of Pathology and Ruhrlandklinik.

**Competing Interests:** The authors have declared that no competing interests exist.

## Conclusion

FFPE-derived mRNA is suitable for molecular biological investigations via the nCounter technology, although it is highly degraded. *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* are potent reference genes in neuroendocrine tumors of the lung.

## Introduction

Neuroendocrine lung cancer (NELC) represents 25% of all lung cancer cases and is divided into four subgroups (typical (TC) and atypical carcinoids (AC), large- (LCNEC) and small-cell (neuroendocrine) cancer (SCLC)) [1].

The majority of these clinical samples are stored as formalin-fixed, paraffin-embedded (FFPE) tissue. FFPE samples can be stored at ambient temperature with minimal logistical effort and costs. The main disadvantage of formalin fixation is the degradation and alteration of nucleic acids by oxidative deamination and cross-linking [2–4]. Besides, temperature shifts, ultraviolet light and reactive oxygen species have a negative impact on nucleic acid stability and integrity [4] leading to highly degraded nucleic acids. Storage of fresh-frozen tissue may help to overcome the mentioned problems, but the costs are high and biobanking is in the early stages of development. An ideal method for the investigation of FFPE is the nCounter technology from NanoString. The technology enables the detection of short target regions (~100 nt) making it an ideal method for analyzing fragmented nucleic acids from FFPE tissue. Fragments of 100 nt were reported to be detectable with nearly 100% efficiency [3, 5–7]. Reis *et al.* compared FFPE versus fresh-frozen samples using the nCounter technique and reported a high comparability of the results [3]. Very short regions of approximately 100 nt are targeted by two specific oligo probes that hybridize to the region of interest [6, 8] without any enzymatic step, hence overcoming limitation of other methods (e.g. PCR), which suffer from enzymatic alterations such as amplification efficiencies [9, 10]. Fluorescent barcodes allow for a simultaneous detection of up to 800 different targets per sample [6, 8]. To compensate for technical or biological variations, normalization of the raw data should be performed. So far, reference gene normalization seemed to be the best choice for quantitative mRNA analysis, although alternative quantification/normalization methods are under discussion: These alternatives include standardization to cell number and RNA mass quantification (most frequently used in Northern Blot analysis), but these methods are hardly applicable for clinical samples with low RNA quality and unknown cell content [9]. Hence, identification of reference genes is need.

But detectable and reliable expression levels as well as robustness of potential reference genes are discussed controversially, because expression is disease- and tissue-type dependent [9, 11]. Using inappropriate internal references can have a drastic impact as it can lead to errors in data acquisition and analysis [10].

Based on recent literature, a custom CodeSet for expression profiling in NELC was designed and tested on the nCounter platform. To identify potent reference genes in neuroendocrine lung cancer, two common algorithms (geNorm and NormFinder) were applied. Here we present reliable reference genes in NELC that enable a proper normalization, which builds the fundament for good quality data and provides the basis for research approaches such as biomarker detection.

## Material and Methods

Representative samples of each tumor subtype (15 TC, 9 AC, 7 LCNEC and 12 SCLC) were used for the mRNA expression analysis using the nCounter platform (NanoString Technologies, Seattle, WA, USA). Specimens from 2005 till 2012 were taken from the tumor bank at the Institute of Pathology, University Hospital Essen, University Duisburg-Essen (Germany). Tumor classification was performed according to the *WHO Classification Of Tumours* guidelines (2004) [12] and TNM-staging was based on the *UICC Classification of Malignant Tumours* [13]. The study was conducted retrospectively for the identification of potential tumor markers and reference genes. The study was approved by the ethical committee of the University Hospital Essen (ID: 13-5382-BO) and conforms to the principles outlined in the declaration of Helsinki.

## RNA Isolation and RNA Integrity Assessment

Three to five paraffin sections with a thickness of 4  $\mu\text{m}$  per sample were deparaffinized with xylene prior to RNA extraction using the RNeasy FFPE kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations with slight adjustments as described earlier [14, 15]. Total RNA concentrations were measured using a Nanodrop 1000 instrument (Thermo Fisher Scientific, Waltham, USA) and an Agilent 2100 Bioanalyzer (eukaryote total RNA Nano, version 2.6, Agilent Technologies, Santa Clara, USA). RNA integrity was assessed using the Agilent 2100 bioanalyzer.

## nCounter CodeSet Design and Expression Analysis

Multiple genes involved in tumor-associated pathways and neuroendocrine differentiation were included in a custom CodeSet using the standard chemistry. The CodeSet was designed to contain a total of 91 genes with different signature genes for each subgroup as described elsewhere [7, 14–16]. The CodeSet was designed and synthesized at NanoString Technologies, Seattle, USA (S1 Table for properties of the CodeSet). Total RNA (100 ng) including miRNA was analyzed at the NanoString nCounter Core Facility at the University of Heidelberg, Germany.

## nCounter Data Processing and Statistical Analysis

For each gene, raw NanoString counts were subjected to a technical background correction. Therefore, the average of the negative controls counts plus two-times the standard deviation was calculated and subtracted from each target count. Afterwards, either the geNorm or NormFinder algorithm was used to calculate the gene expression stability measure (M-value) for all tested genes. In brief, the M-value of a certain gene, calculated by the geNorm algorithm, is the arrhythmic mean of all pairwise variations of potential control genes [9]. The NormFinder algorithm estimates a) the intra- and inter-sample variation leading to a distribution, which is then b) transformed to a one-dimensional value by defining a value of the mean plus one standard deviation [17]. Genes with average and median counts below 500 were not considered as potential reference genes, because mean and median counts  $<500$  indicate that a fraction of the samples present with absent expression for a certain gene. All statistical and graphical analyses were performed using the R statistical programming environment (v 3.1.1). For dichotomous factors such as gender and expression level the Wilcoxon Mann-Whitney rank sum test was applied. The Kruskal-Wallis test was used to correlate tumor type and gene expression. Correlations between gene expression and TNM-stages were analyzed by Spearman's rank correlation test. Scatterplots were generated and coefficients of determination ( $R^2$ ) were calculated

for pairwise comparison of potential reference genes. As control, two previously identified tumor markers (*CDK6* and *TYMS* [14, 15]) were correlated with the reference genes as described above.

The level of statistical significance was defined as  $p \leq 0.05$ .

## Results

Forty-three tumor samples were investigated (16 TC (27%) and 13 AC (22%), 16 LCNEC (27%) and 15 SCLC (25%)). Twenty-five female (58%) and 16 male patients (37%) were analyzed. For two patients the gender remained inconclusive after anonymization.

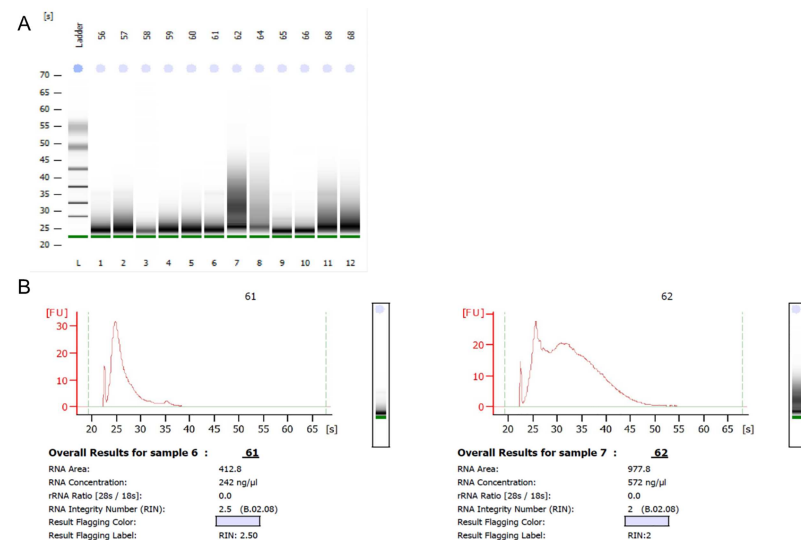
RNA integrity (RIN) ranged from 1.8 to 2.6 and concentrations from 34 to 2,109 ng/ $\mu$ l. Although RIN was considerably low, all tested specimens gave evaluable results with respect to the nCounter expression analysis. Representative results are summarized in [S2 Table](#). Besides, [Fig 1A and 1B](#) presents exemplary data of the smear analysis.

The NormFinder algorithm identified the following genes as most stably expressed: *GRB2* (M-value = 0.422), *CDKN1B* (M-value = 0.4787), *GAPDH* (M-value = 0.5228), *SDCBP* (M-value = 0.6232), *PNN* (M-value = 0.6563) and *ACTB* (M-value = 0.6646).

Whereas, the geNorm algorithm identified these stable reference genes: *CDKN1B* (M-value = 4.45), *GRB2* (M-value = 4.45), *RHOA* (M-value = 4.48), *SDCBP* (M-value = 4.45), *GAPDH* (M-value = 4.58) and *LDHB* (M-value = 4.59).

With respect to the 20 most stable reference genes identified by both algorithms, a consensus for 17 genes (85%) was found. The results for mean and median expression as well as M-values for both algorithms are summarized in [S3 Table](#).

Statistical analysis for TNM-criteria revealed significant correlations of tumor type and gene expression of *ACTB* ( $p = 0.016$ ), *CAT* ( $p = 0.041$ ), *LDHB* ( $p = 0.024$ ) and *SDCBP*



**Fig 1. A and B show gel-smear analysis to assess the RNA quantity and quality (RIN) in total RNA derived from formalin-fixed, paraffin-embedded tissue. Fig 1A depicts a representative smear gel analysis of twelve samples. A ladder was included to allow size calculation. The microfluid analysis shows that RNA from FFPE is highly degraded giving no distinct size patterns. Fig 1B depicts the electropherogram of two representative samples. The rRNA Ratio (28s/18s) is used to calculate the RNA quality according to an algorithm supplied by the manufacturer. Neither 28s nor 18s bands can be found for FFPE-derived RNA leading to considerably low RNA integrity numbers (RIN). RNA concentration is calculated from the area under the curve.**

doi:10.1371/journal.pone.0165181.g001

( $p = 0.044$ ), which would render insignificant after a multivariate analysis correction (e.g. Bonferroni correction or false discovery rate adjustment (FDR)) (S1 Fig). Similarly, *LDHB* ( $p = 0.023$ ) and *PNN* ( $p = 0.044$ ) were significantly associated with grade of the tumor (data not shown). For age of the FFPE sample and gender no significant correlations were noted.

After identification of the six most stable reference genes for each algorithm, the coefficient of determination ( $R^2$ ) was generated for *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *LDHB*, *PNN*, *RHOA* and *SDCBP*. Considering the top-10 of the  $R^2$ -values, *GRB2* (5 times), *RHOA* (4 times), *ACTB* (3 times), *GAPDH* (3 times) and *SDCBP* (3 times) showed the highest correlation frequency. *CDKN1B* and *PNN* showed only one correlation with another gene. Then a correlation matrix was created that confirmed the previous results and also identified *CDKN1B* to have a considerable correlation with *GRB2* and *GAPDH* expression (Fig 2A).

As control, correlation of potential reference gene and suspected tumor markers (*CDK6* and *TYMS*) was performed revealing very low  $R^2$ -values, as expected. The  $R^2$ -values and top-10 ranking are summarized in Table 1. Fig 2A–2C depicts the correlation matrix, a heat map and two exemplary scatterplots of the correlation analysis. The scatterplots for all gene correlations are summarized in S2 Fig.

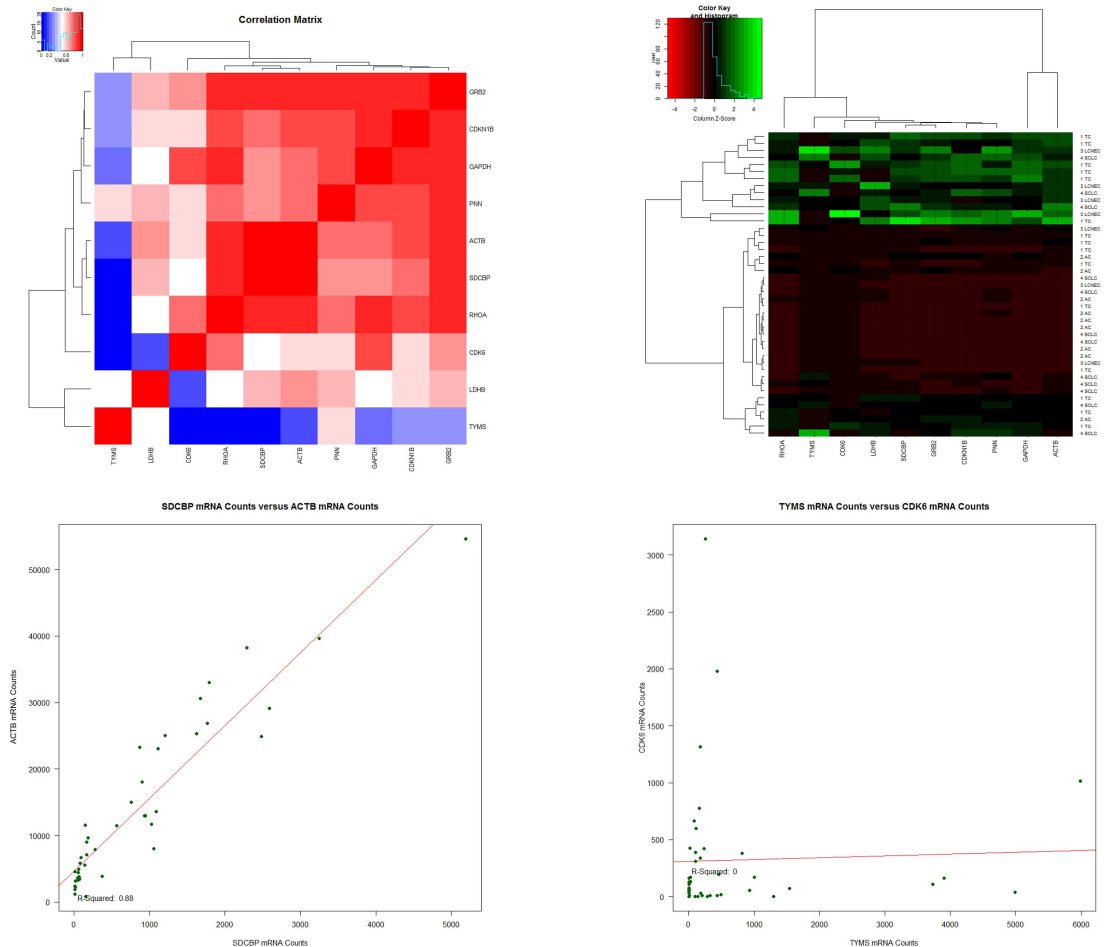
Taking the presented results into consideration, *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* can be considered as reliable and stably expressed reference genes in neuroendocrine lung cancer.

## Discussion

Degradation of nucleic acids in FFPE tissue leads to highly degraded RNA and represents an eminent problem for routine diagnostic and research. Formalin fixation leads to RNA integrity numbers  $<3$  [18]. Nevertheless, a large proportion of RNA fragments of  $\geq 100$  nt remain intact and can be detected reliably [2, 5]. This fragment length is sufficient to analyze such samples with the nCounter technology, which uses probe pairs binding to  $\sim 100$  nt of the target regions [3, 5–7]. This approach makes it an ideal method for the analysis of degraded archival FFPE tissue. In the present study, we were able to specifically and sensitively analyze 91 mRNA-target transcripts in 43 FFPE samples by using 100 ng of total RNA per sample. Tumor specimens from 2005 till 2012 were investigated and age of the FFPE sample did not show a statistically significant influence on the results.

Housekeeping genes are no longer referred to as reliable reference genes for normalization [9–11]. For several reference genes the expression is tissue- and disease-specific, therefore leading to heterogeneous ratings and controversial discussions which genes are suitable as internal references [9–11]. According to our results, NormFinder [17] and/or geNorm [9] algorithms plus calculation of the coefficient of determination and matrix correlation identified *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* to be reliable and stably expressed reference genes in neuroendocrine lung cancer. *ACTB* and *GAPDH* were detectable in 100% of cases and the remaining genes were detected in all but one case. As control, the coefficient of determination was also calculated for the reference genes and tumor markers *CDK6* [15] and *TYMS* [14] and showed considerably low correlations between them.

A shortcoming of both algorithms is that they prefer low counts over high counts. Higher counts may lead to a considerably higher deviation. Therefore, very low counts (mean and median  $<500$  mRNA counts) result in  $M$ -values indicating high stability/constant expression. But in fact, most of the genes identified as stably expressed ones showed very low counts in the majority of cases (S3 Table). *ACTB* and *GAPDH* gave high counts in all samples, but showed a higher variability. This leads to slightly weaker stability values than found for low expressed



**Fig 2. A to D show a correlation matrix for gene expression (A), a heatmap for tumor type versus gene expression (B), scatterplots (C and D) for gene versus gene correlations and  $R^2$  calculation. Fig 2A** depicts a correlation matrix of genes that were identified as potential reference genes by geNorm and NormFinder algorithms and previously identified tumor markers (*CDK6* and *TYMS*). High correlations are outlined by red colored squares. Between *CDKN1B*, *GRB2* and *GAPDH* as well as between *ACTB*, *SDCBP* and *RHOA* a high correlation was identified. Low correlations are indicated by blue squares and were found for tumor markers (*CDK6* and *TYMS*) versus reference gene. Fig 2B displays a heatmap. On the x-axis the potential reference genes and tumor markers *CDK6* and *TYMS* are shown. On the y-axis the investigated tumor types are depicted. Differential expression was found between tumor types. Though, the reference genes show a constant expression cluster (either low or high) between the samples investigated. The tumor markers present with differential expression between all samples without showing a specific cluster. Fig 2C and 2D are exemplary scatterplots of gene versus gene correlation, which were created to calculate the coefficient of determination ( $R^2$ ). Fig 2C depicts the highest correlation identified ( $R^2 = 0.88$ ) between two potential reference genes (*ACTB* and *SDCBP*). In D, the weakest correlation is depicted, which was found between the two tumor markers (*CDK6* and *TYMS*).

doi:10.1371/journal.pone.0165181.g002

genes. Still, *ACTB* and *GAPDH* can be considered reliable reference genes due to the ubiquitous and quiet stable expression as well as high-ranked M-values.

To overcome the limitations of both algorithms, calculation of the coefficient of determination was performed and a correlation matrix was created. The additional tests showed that the six identified reference genes have a high conformity with respect to their expression profiles. Usage of multiple robustly expressed internal references is highly encouraged, because a single reference could lead to false-negative/-positive results due to varying expression [10, 11].

**Table 1. Summarizes coefficients of correlation (R-squared) for tested genes and top-10 correlations are highlighted in white.**

Ranking	Gene A	Gene B	R-squared
1	<i>ACTB</i>	<i>SDCBP</i>	0.88
2	<i>GRB2</i>	<i>RHOA</i>	0.84
3	<i>ACTB</i>	<i>GRB2</i>	0.84
4	<i>GRB2</i>	<i>SDCBP</i>	0.82
5	<i>ACTB</i>	<i>RHOA</i>	0.81
6	<i>RHOA</i>	<i>SDCBP</i>	0.8
7	<i>GRB2</i>	<i>PNN</i>	0.8
8	<i>GAPDH</i>	<i>GRB2</i>	0.79
9	<i>GAPDH</i>	<i>RHOA</i>	0.78
10	<i>GAPDH</i>	<i>CDKN1B</i>	0.77
11	<i>CDKN1B</i>	<i>GRB2</i>	0.76
12	<i>CDKN1B</i>	<i>PNN</i>	0.74
13	<i>GAPDH</i>	<i>PNN</i>	0.72
14	<i>CDKN1B</i>	<i>RHOA</i>	0.71
15	<i>ACTB</i>	<i>CDKN1B</i>	0.66
16	<i>GAPDH</i>	<i>CDK6</i>	0.66
17	<i>ACTB</i>	<i>GAPDH</i>	0.63
18	<i>CDKN1B</i>	<i>SDCBP</i>	0.63
19	<i>ACTB</i>	<i>PNN</i>	0.6
20	<i>RHOA</i>	<i>CDK6</i>	0.59
21	<i>PNN</i>	<i>RHOA</i>	0.58
22	<i>ACTB</i>	<i>LDHB</i>	0.53
23	<i>PNN</i>	<i>SDCBP</i>	0.53
24	<i>GAPDH</i>	<i>SDCBP</i>	0.52
25	<i>GRB2</i>	<i>CDK6</i>	0.48
26	<i>GRB2</i>	<i>LDHB</i>	0.44
27	<i>LDHB</i>	<i>SDCBP</i>	0.43
28	<i>LDHB</i>	<i>PNN</i>	0.38
29	<i>ACTB</i>	<i>CDK6</i>	0.36
30	<i>PNN</i>	<i>CDK6</i>	0.36
31	<i>CDKN1B</i>	<i>CDK6</i>	0.34
32	<i>PNN</i>	<i>TYMS</i>	0.34
33	<i>CDKN1B</i>	<i>LDHB</i>	0.29
34	<i>LDHB</i>	<i>RHOA</i>	0.28
35	<i>LDHB</i>	<i>TYMS</i>	0.26
36	<i>GAPDH</i>	<i>LDHB</i>	0.25
37	<i>SDCBP</i>	<i>CDK6</i>	0.25
38	<i>CDKN1B</i>	<i>TYMS</i>	0.11
39	<i>GRB2</i>	<i>TYMS</i>	0.09
40	<i>GAPDH</i>	<i>TYMS</i>	0.07
41	<i>LDHB</i>	<i>CDK6</i>	0.04
42	<i>ACTB</i>	<i>TYMS</i>	0.03
43	<i>SDCBP</i>	<i>TYMS</i>	0.01
44	<i>CDK6</i>	<i>TYMS</i>	0
45	<i>RHOA</i>	<i>TYMS</i>	0

doi:10.1371/journal.pone.0165181.t001

Alternative quantification/normalization methods are under discussion including standardization to cell number and RNA mass quantification (most frequently used in Northern Blot analysis), but these methods are hardly applicable for clinical samples with low RNA quality and unknown cell content [9]. Therefore, normalization to reference genes seems to be the best choice for quantitative mRNA analysis until alternative approaches are found. In the present study, *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* were identified as suitable reference genes in neuroendocrine tumors of the lung.

Based on the presented choice of reference genes, we have identified diagnostic, prognostic and therapy-relevant biomarkers in NELC. With respect to apoptosis and cell cycle, NELC show a distinct expression and mutation pattern. Carcinoids use *CDK4/6* and *CCND1* to drive cell cycle progression by controlling RB1 phosphorylation. On the other hand, carcinomas control this signaling cascade by using the opposite way (*CDK2* and *CCNE1*) [15]. Of note, *TP53* mutations are exclusively found in carcinomas, but not in carcinoids [19] and on the miRNA level NELC subtypes showed significantly different expression of miRNAs regulating proliferation and apoptosis [20].

Besides, our results revealed why the large randomized phase III clinical trial GALES (global analysis of pemetrexed in SCLC extensive stage) failed. Low expression of *FOLR1* and *FPGS*, but high expression of *TYMS* is the predominant phenotype found in SCLC rendering pemetrexed therapy useless in this entity [14]. Similar results were found in squamous cell carcinoma of the lung where high expression of *TYMS* precludes response to pemetrexed [21].

Furthermore, poor outcome correlated with elevated expression of the transcription factors *SOX11* and *PAX6* [16]. Mediators of angiogenesis (*HIF1A*, *CRHR2*, *KDR* and *FIGF*) correlated with decreased progression-free survival, poor outcome and more aggressive clinical behavior of NELC revealing potential therapy targets for tyrosine-kinase inhibitors and antibodies [7]. Another potent therapy target could be the 26S proteasome, because it showed significantly higher expression in NELC than in benign controls [22]. Several potent proteasome inhibitors are currently tested in clinical phase trials or have received clinical clearance [23].

## Conclusion

FFPE-derived mRNA can be reliably used for expression analysis using the nCounter technology (NanoString). The method uses two specific oligo-probes, which hybridize to ~100 nt in the region of interest, enabling the analysis of the highly degraded nucleic acids. *ACTB*, *CDKN1B*, *GAPDH*, *GRB2*, *RHOA* and *SDCBP* were identified as constantly expressed genes with high stability (*M*-)values according to geNorm and NormFinder and coefficients of correlation making them suitable reference genes for expression analysis/biological normalization in neuroendocrine tumors of the lung.

## Supporting Information

**S1 Fig. Boxplots for gene expression versus tumor subtype.** S1 Fig shows boxplots for gene expression of potential reference genes, which were identified by either geNorm or NormFinder algorithm, in correlation with the four investigated tumor subtypes. On the x-axis the four investigated tumor subtypes are shown. The y-axis depicts the mRNA counts for each gene. A Kruskal-Wallis test was performed and significantly differential expression is outlined by p-values. In most cases, atypical carcinoids showed much lower expression than the other three entities. Also SCLC showed relatively low counts.

(PDF)



**S2 Fig. Scatterplots and coefficients of correlation for gene versus gene correlations.** S2 Fig depicts scatterplots and the coefficient of correlation for gene versus gene correlation. Potential reference genes were identified by applying the geNorm and NormFinder algorithm. Potential reference genes were correlated and the coefficient of determination ( $R^2$ ) was calculated to identify the conformity between them. As control, tumor markers (*CDK6* and *TYMS*) were included in the analysis and showed very low correlation with the potential reference genes. (PDF)

**S1 Table. The gene name, gene ID and properties of the custom CodeSet that was used for the nCounter analysis.** (XLS)

**S2 Table. Patient data, RNA concentration and quality (RIN) and exemplarily gene mRNA counts.** (XLSX)

**S3 Table. Summarizes gene names, calculated stability values by NormFinder or geNorm algorithm, mean and median mRNA counts per gene.** Grey stained rows indicate genes that show mean and/or median expression of >500 mRNA counts. Median for all genes is 99 counts and mean for all genes is 1,132 counts. (XLSX)

## Acknowledgments

We thank the nCounter Core Facility Heidelberg for providing the nCounter system and related services.

## Author Contributions

**Conceptualization:** RFHW FDM DCC KWS JW.

**Formal analysis:** CV RW RFHW FDM JS TH EF DCC.

**Investigation:** RFHW RW FDM.

**Resources:** KWS JW TH DCC.

**Supervision:** JW KWS.

**Writing – original draft:** RFHW FDM EF JW CV JS.

**Writing – review & editing:** JW KWS DCC.

## References

1. Rekhtman N. Neuroendocrine tumors of the lung: an update. *Archives of pathology & laboratory medicine*. 2010; 134(11):1628–38. doi: [10.1043/2009-0583-RAR.1](https://doi.org/10.1043/2009-0583-RAR.1) PMID: [21043816](https://pubmed.ncbi.nlm.nih.gov/21043816/).
2. Lewis F, Maughan NJ, Smith V, Hillan K, Quirke P. Unlocking the archive—gene expression in paraffin-embedded tissue. *The Journal of pathology*. 2001; 195(1):66–71. doi: [10.1002/1096-9896\(200109\)195:1<66::AID-PATH921>3.0.CO;2-F](https://doi.org/10.1002/1096-9896(200109)195:1<66::AID-PATH921>3.0.CO;2-F) PMID: [11568892](https://pubmed.ncbi.nlm.nih.gov/11568892/).
3. Reis PP, Waldron L, Goswami RS, Xu W, Xuan Y, Perez-Ordóñez B, et al. mRNA transcript quantification in archival samples using multiplexed, color-coded probes. *BMC biotechnology*. 2011; 11:46. doi: [10.1186/1472-6750-11-46](https://doi.org/10.1186/1472-6750-11-46) PMID: [21549012](https://pubmed.ncbi.nlm.nih.gov/21549012/); PubMed Central PMCID: [PMC3103428](https://pubmed.ncbi.nlm.nih.gov/PMC3103428/).
4. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA quality from FFPE samples. *PloS one*. 2007; 2(12):e1261. doi: [10.1371/journal.pone.0001261](https://doi.org/10.1371/journal.pone.0001261) PMID: [18060057](https://pubmed.ncbi.nlm.nih.gov/18060057/); PubMed Central PMCID: [PMC2092395](https://pubmed.ncbi.nlm.nih.gov/PMC2092395/).

5. Walter RF, Mairinger FD, Wohlschlaeger J, Worm K, Ting S, Vollbrecht C, et al. FFPE tissue as a feasible source for gene expression analysis—a comparison of three reference genes and one tumor marker. *Pathology, research and practice*. 2013; 209(12):784–9. doi: [10.1016/j.prp.2013.09.007](https://doi.org/10.1016/j.prp.2013.09.007) PMID: [24156825](https://pubmed.ncbi.nlm.nih.gov/24156825/).
6. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature biotechnology*. 2008; 26(3):317–25. doi: [10.1038/nbt1385](https://doi.org/10.1038/nbt1385) PMID: [18278033](https://pubmed.ncbi.nlm.nih.gov/18278033/).
7. Mairinger FD, Walter RF, Werner R, Christoph DC, Ting S, Vollbrecht C, et al. Activation of angiogenesis differs strongly between pulmonary carcinoids and neuroendocrine carcinomas and is crucial for carcinoid tumourgenesis. *Journal of Cancer*. 2014; 5(6):465–71. doi: [10.7150/jca.9235](https://doi.org/10.7150/jca.9235) PMID: [24959299](https://pubmed.ncbi.nlm.nih.gov/24959299/); PubMed Central PMCID: [PMC4066358](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC4066358/).
8. Beaume M, Hernandez D, Docquier M, Delucinge-Vivier C, Descombes P, Francois P. Orientation and expression of methicillin-resistant *Staphylococcus aureus* small RNAs by direct multiplexed measurements using the nCounter of NanoString technology. *Journal of microbiological methods*. 2011; 84(2):327–34. doi: [10.1016/j.mimet.2010.12.025](https://doi.org/10.1016/j.mimet.2010.12.025) PMID: [21195730](https://pubmed.ncbi.nlm.nih.gov/21195730/).
9. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology*. 2002; 3(7):RESEARCH0034. PMID: [12184808](https://pubmed.ncbi.nlm.nih.gov/12184808/); PubMed Central PMCID: [PMC126239](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC126239/).
10. Bustin SA, Benes V, Garson JA, Hellems J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry*. 2009; 55(4):611–22. doi: [10.1373/clinchem.2008.112797](https://doi.org/10.1373/clinchem.2008.112797) PMID: [19246619](https://pubmed.ncbi.nlm.nih.gov/19246619/).
11. Li R, Shen Y. An old method facing a new challenge: re-visiting housekeeping proteins as internal reference control for neuroscience research. *Life sciences*. 2013; 92(13):747–51. doi: [10.1016/j.lfs.2013.02.014](https://doi.org/10.1016/j.lfs.2013.02.014) PMID: [23454168](https://pubmed.ncbi.nlm.nih.gov/23454168/); PubMed Central PMCID: [PMC3614345](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC3614345/).
12. Travis WD, World Health Organization., International Agency for Research on Cancer., International Association for the Study of Lung Cancer., International Academy of Pathology. *Pathology and genetics of tumours of the lung, pleura, thymus and heart*. Lyon Oxford: IARC PressOxford University Press; 2004. 344 p. p.
13. Sobin LH, Gospodarowicz MK, Wittekind C, International Union against Cancer. *TNM classification of malignant tumours*. 7th ed. Chichester, West Sussex, UK; Hoboken, NJ: Wiley-Blackwell; 2010. xx, 309 p. p.
14. Walter RF, Mairinger FD, Werner R, Vollbrecht C, Hager T, Schmid KW, et al. Folic-acid metabolism and DNA-repair phenotypes differ between neuroendocrine lung tumors and associate with aggressive subtypes, therapy resistance and outcome. *Oncotarget*. 2016. doi: [10.18632/oncotarget.7737](https://doi.org/10.18632/oncotarget.7737) PMID: [27064343](https://pubmed.ncbi.nlm.nih.gov/27064343/).
15. Walter RF, Werner R, Ting S, Vollbrecht C, Theegarten D, Christoph DC, et al. Identification of deregulation of apoptosis and cell cycle in neuroendocrine tumors of the lung via NanoString nCounter expression analysis. *Oncotarget*. 2015; 6(28):24690–8. doi: [10.18632/oncotarget.3992](https://doi.org/10.18632/oncotarget.3992) PMID: [26008974](https://pubmed.ncbi.nlm.nih.gov/26008974/); PubMed Central PMCID: [PMC4694788](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC4694788/).
16. Walter RF, Mairinger FD, Werner R, Ting S, Vollbrecht C, Theegarten D, et al. SOX4, SOX11 and PAX6 mRNA expression was identified as a (prognostic) marker for the aggressiveness of neuroendocrine tumors of the lung by using next-generation expression analysis (NanoString). *Future oncology*. 2015; 11(7):1027–36. doi: [10.2217/fon.15.18](https://doi.org/10.2217/fon.15.18) PMID: [25804118](https://pubmed.ncbi.nlm.nih.gov/25804118/).
17. Andersen CL, Jensen JL, Orntoft 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 research*. 2004; 64(15):5245–50. doi: [10.1158/0008-5472.CAN-04-0496](https://doi.org/10.1158/0008-5472.CAN-04-0496) PMID: [15289330](https://pubmed.ncbi.nlm.nih.gov/15289330/).
18. Kalmar A, Wichmann B, Galamb O, Spisak S, Toth K, Leiszter K, et al. Gene expression analysis of normal and colorectal cancer tissue samples from fresh frozen and matched formalin-fixed, paraffin-embedded (FFPE) specimens after manual and automated RNA isolation. *Methods*. 2013; 59(1):S16–9. doi: [10.1016/j.ymeth.2012.09.011](https://doi.org/10.1016/j.ymeth.2012.09.011) PMID: [23036325](https://pubmed.ncbi.nlm.nih.gov/23036325/).
19. Vollbrecht C, Werner R, Walter RF, Christoph DC, Heukamp LC, Peifer M, et al. Mutational analysis of pulmonary tumours with neuroendocrine features using targeted massive parallel sequencing: a comparison of a neglected tumour group. *British journal of cancer*. 2015; 113(12):1704–11. doi: [10.1038/bjc.2015.397](https://doi.org/10.1038/bjc.2015.397) PMID: [26645239](https://pubmed.ncbi.nlm.nih.gov/26645239/); PubMed Central PMCID: [PMC4701994](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC4701994/).
20. Mairinger FD, Ting S, Werner R, Walter RF, Hager T, Vollbrecht C, et al. Different micro-RNA expression profiles distinguish subtypes of neuroendocrine tumors of the lung: results of a profiling study. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc*. 2014; 27(12):1632–40. doi: [10.1038/modpathol.2014.74](https://doi.org/10.1038/modpathol.2014.74) PMID: [24875640](https://pubmed.ncbi.nlm.nih.gov/24875640/).

21. Joerger M, Omlin A, Cerny T, Fruh M. The role of pemetrexed in advanced non small-cell lung cancer: special focus on pharmacology and mechanism of action. *Curr Drug Targets*. 2010; 11(1):37–47. PMID: [19839929](#).
22. Mairinger FD, Walter RF, Theegarten D, Hager T, Vollbrecht C, Christoph DC, et al. Gene Expression Analysis of the 26S Proteasome Subunit PSMB4 Reveals Significant Upregulation, Different Expression and Association with Proliferation in Human Pulmonary Neuroendocrine Tumours. *Journal of Cancer*. 2014; 5(8):646–54. doi: [10.7150/jca.9955](#) PMID: [25157275](#); PubMed Central PMCID: PMCPMC4142326.
23. Dhakal B, Girnius S, Hari P. Recent advances in understanding multiple myeloma. *F1000Res*. 2016; 5. doi: [10.12688/f1000research.8777.1](#) PMID: [27610224](#); PubMed Central PMCID: PMCPMC4995679.