



OPEN

## Un-biased housekeeping gene panel selection for high-validity gene expression analysis

Ana I. Casas<sup>1,2,8</sup>✉, Ahmed A. Hassan<sup>2,8</sup>, Quirin Manz<sup>3,8</sup>, Christian Wiwie<sup>4</sup>, Pamela Kleikers<sup>2</sup>, Javier Egea<sup>5,6</sup>, Manuela G. López<sup>6</sup>, Markus List<sup>7,8</sup>, Jan Baumbach<sup>3,8</sup> & Harald H. H. W. Schmidt<sup>2,8</sup>✉

Differential gene expression normalised to a single housekeeping (HK) is used to identify disease mechanisms and therapeutic targets. HK gene selection is often arbitrary, potentially introducing systematic error and discordant results. Here we examine these risks in a disease model of brain hypoxia. We first identified the eight most frequently used HK genes through a systematic review. However, we observe that in both ex-vivo and in vivo, their expression levels varied considerably between conditions. When applying these genes to normalise expression levels of the validated stroke target gene, inducible *Nox4*, we obtained opposing results. As an alternative tool for unbiased HK gene selection, software tools exist but are limited to individual datasets lacking genome-wide search capability and user-friendly interfaces. We, therefore, developed the HouseKeepR algorithm to rapidly analyse multiple gene expression datasets in a disease-specific manner and rank HK gene candidates according to stability in an unbiased manner. Using a panel of de novo top-ranked HK genes for brain hypoxia, but not single genes, *Nox4* induction was consistently reproduced. Thus, differential gene expression analysis is best normalised against a HK gene panel selected in an unbiased manner. HouseKeepR is the first user-friendly, bias-free, and broadly applicable tool to automatically propose suitable HK genes in a tissue- and disease-dependent manner.

Analysing differential gene expression via mRNA levels is a common tool in biomedical sciences to understand cellular regulation<sup>1</sup> or its dysregulation in disease. One application is the identification of target genes for therapeutic intervention and drug discovery<sup>2</sup>. The gold standard and most widely used technology to quantify mRNA levels is real-time quantitative PCR (RT-qPCR) because of its high sensitivity and specificity<sup>3</sup>. However, to be able to draw a reliable conclusion with respect to differential expression, e.g., health vs disease or control vs treatment, normalisation to a stable so-called housekeeping gene (HK) is essential. For this purpose, it is common practice to select genes related to basal cell metabolism as their expression is thought to be stable and thus fulfil the key HK requirement<sup>4</sup>. However, the expression levels of several commonly used HK genes, in fact, strongly vary between different cell types, disease models and drug interventions<sup>4,5</sup>. Since the validity of a differential gene expression result<sup>6</sup> resides entirely on the stability of the chosen HK gene, diametrically opposing results can be obtained if data normalisation is performed using unstable and differentially regulated HK genes<sup>7,8</sup>. Moreover, on top of a disease condition, therapeutic drugs may further modulate HK gene expression introducing further variability. In the worst scenario, such results, if used for cell-based diagnostics, e.g., in liquid biopsies, could mislead subsequent drug intervention.

<sup>1</sup>Department of Neurology and Center for Translational Neuro- and Behavioural Sciences (C-TNBS), University Clinics Essen, Essen, Germany. <sup>2</sup>Department of Pharmacology & Personalised Medicine, MeHNS, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands. <sup>3</sup>Faculty of Mathematics, Informatics and Natural Sciences, University of Hamburg, Hamburg, Germany. <sup>4</sup>Department of Mathematics and Computer Science, University of Southern Denmark, Odense, Denmark. <sup>5</sup>Molecular Neuroinflammation and Neuronal Plasticity Research Laboratory, Hospital Universitario Santa Cristina, Instituto de Investigación Sanitaria-Hospital Universitario de la Princesa, Madrid, Spain. <sup>6</sup>Departamento de Farmacología, Instituto de I+D del Medicamento Teófilo Hernando (ITH), Facultad de Medicina, Universidad Autónoma de Madrid, Madrid, Spain. <sup>7</sup>Chair of Experimental Bioinformatics, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Munich, Germany. <sup>8</sup>These authors contributed equally: Ana I. Casas, Ahmed A. Hassan, Quirin Manz, Markus List, Jan Baumbach and Harald H. H. W. Schmidt. ✉email: anaisabel.casasgujarro@uk-essen.de; hschmidt@ppmlab.net

Hippocampal brain slices—acute model								
Published	Excluded	Included	Genes					
364	353	11	<i>Hprt</i> 1	<i>β-actin</i> 4	<i>18S</i> 1	<i>Gapdh</i> 5		
Organotypic hippocampal culture—chronic model								
Published	Excluded	Included	Genes					
249	222	27	<i>β-actin</i> 9	<i>Hprt</i> 1	<i>Gapdh</i> 12	<i>18S</i> 2	<i>β2-mg</i> 2	<i>Rpl13</i> 1

**Table 1.** Normalization genes used for in vitro gene expression determination in brain ischemia models.

To prevent such analytical errors, several mitigation methods have been suggested. The most basic approach is the comparative delta-Ct method ranking stably expressed genes based on pair-wise comparisons<sup>9</sup>. Other strategies use expression variance of either an individual gene, a subset of genes, or all genes represented as a co-variance matrix<sup>10–12</sup>. However, these options are based on the assumption of a constant composition of all samples, which, however, is not necessarily the case. Similarly, HK genes under different conditions were selected based on microarray data<sup>13–15</sup>, again assuming that the vast number of publicly available data sets resolves pre-selection bias. However, most of these methods were neither automated nor made available for future access<sup>11,16</sup>, with RefGenes<sup>11</sup> and NormFinder<sup>12</sup>, and BestKeeper<sup>10</sup> being three positive examples. However, even these three tools have significant limitations as they do not scale well for use in genome-wide de novo searches for HK gene candidates.

We here address these fundamental gaps in HK gene identification by developing a user-friendly, bias-free, and broadly applicable tool to automatically propose suitable HK genes, e.g., suitable for a given tissue and disease condition. We then validate this approach by comparing classical biased KH gene selection for brain hypoxia to a de novo identified highly stable HK panel and their reliability in reproducing the discovery of *Nox4* as a key inducible gene with a causal role in post-stroke neurodegeneration<sup>17</sup>. We thereby also address whether, for the analysis of differential gene expression, single HK genes suffice or HK gene panels are preferable.

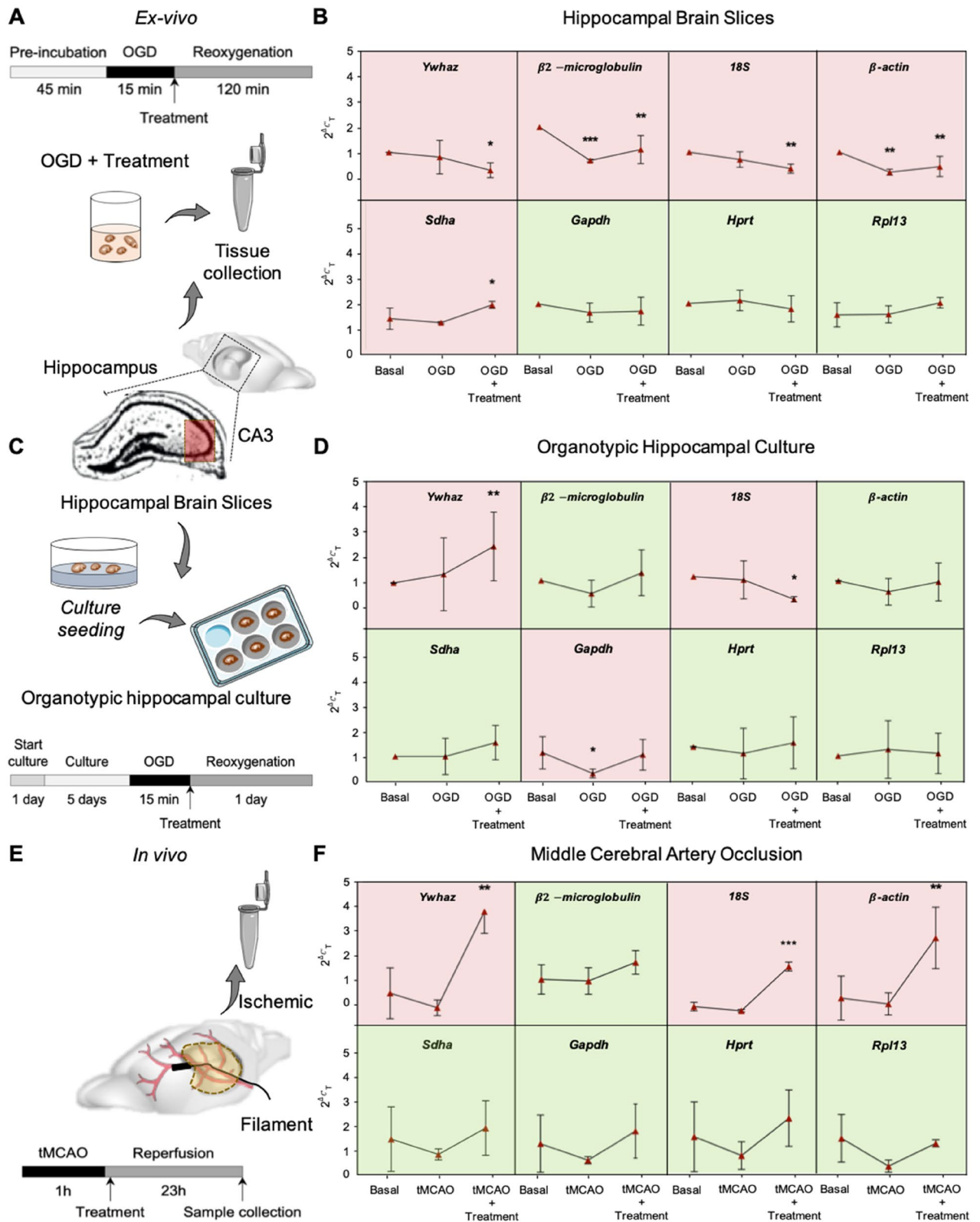
## Results

**Systematic review on the most used HK genes in brain hypoxia.** The selection of HK genes is often perpetuated from one publication to another<sup>18</sup>. We conducted a systematic literature review to identify the most frequently used HK genes in two broadly used ex vivo brain hypoxia models, rat hippocampal brain slices (HBS) and rat organotypic hippocampal culture (OHC). Specifically, 38 and 16 original articles were identified with both ex vivo animal models, respectively (Supplementary Fig. 1), with six potential HK genes of which *β-actin* and *Gapdh* were most frequently used (Table 1). Two additional genes, *Sdha* and *Ywhaz*<sup>19</sup>, were previously suggested as the most suitable for differential gene expression normalisation in vivo stroke models. These eight genes were then validated for stability in three widely used brain hypoxia models in two different species, ex vivo in rat HBS or OHC subjected to oxygen and glucose deprivation (OGD) and in vivo in the mouse middle cerebral artery occlusion (MCAO) model.

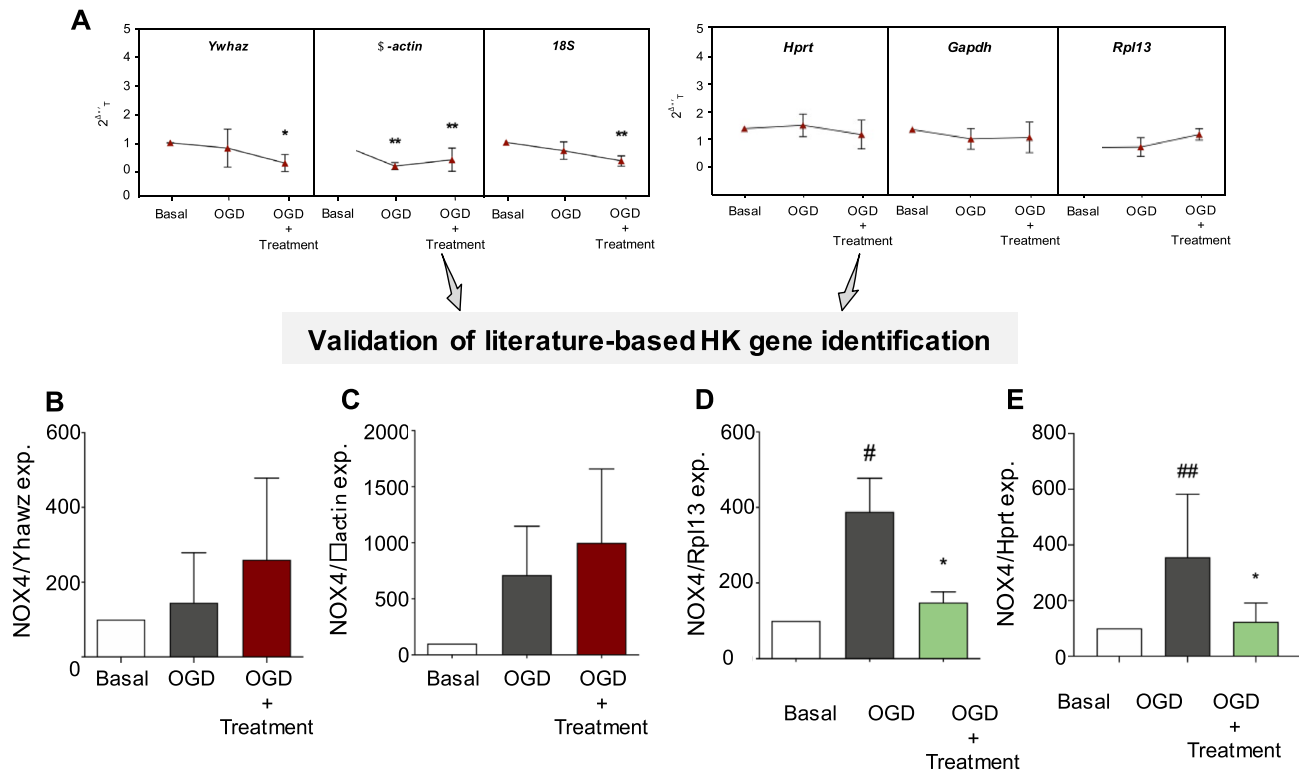
First, rat HBS were subjected to OGD for 15 min, followed by 2 h of re-oxygenation (Re-Ox) with or without different pharmacological interventions (Fig. 1A). Upon RT-qPCR analysis (Supplementary Table 2), five of the eight presumable HK genes significantly varied in expression with or without OGD and were considered unsuitable for this model. Only *Gapdh*, *Hprt* and *Rpl13* remained stable (Fig. 1B). Second, OHCs were subjected to 15 min of OGD followed by 24 h of re-oxygenation with or without different pharmacological treatments (Fig. 1C). Here, three genes were unacceptable, *Ywhaz* and *18S*, but surprisingly, also the most widely used HK gene, *Gapdh* (Fig. 1D). Finally, adult mice were subjected to a 1 h transient MCAO with or without subsequent pharmacological treatments (Fig. 1E). In line with the previous ex vivo findings, *Ywhaz* and *18S* were too variable, but with *β-actin* is also one of the most frequently used HK genes in biomedicine (Fig. 1F). Thus, for brain hypoxia, HK genes are highly variable across different experimental models, species, and pharmacological interventions. Therefore, even for this single disease condition, no standard HK gene recommendation can be inferred from the literature, even for highly frequently used genes.

**Contradicting results on *Nox4* expression when using stable and rejected HK genes.** To test the consequences of such HK gene variability for the analysis of differential gene expression, we used the established<sup>20</sup> and repeatedly confirmed<sup>21</sup> upregulation of NADPH oxidase 4 (*Nox4*) upon hypoxia as a test case. Upon brain hypoxia, *Nox4* is broadly induced in different cell types including blood–brain barrier cells and neurons<sup>22</sup>; genetic deletion of *Nox4* or pharmacological inhibition of the NOX4 enzyme is directly neuroprotective, stabilises the blood–brain barrier, and improves neuro-motor function<sup>22</sup>. Therefore, NOX4 could be considered a prominent therapeutic target in ischemic stroke.

We analysed *Nox4* gene expression in all three above models of brain hypoxia with and without pharmacological treatment. For normalisation, we either used the three commonly used HK genes that we, however, rejected because of instability, i.e., *Ywhaz*, *β-actin*, and *18S*, or three HK genes that we confirmed or found to be sufficiently stable for these models, i.e., *Hprt*, *Gapdh*, and *Rpl13* (Fig. 2A). *Nox4* expressional normalisation using the latter three stable HK gene candidates resulted in the reproduction of a significant *Nox4* upregulation post hypoxia which was prevented by pharmacological intervention. In contrast, normalisation using the former



**Figure 1.** Experimental validation of eight housekeeping genes in three different brain ischemia models. (A) Rat hippocampal brain slices were stabilised for 45 min, followed by 15 min of oxygen and glucose deprivation (OGD) and 2 h of re-oxygenation (Re-Ox). (B) *Gapdh*, *Hprt* and *Rpl13* gene expression remains stable while *Ywhaz*, *18S* and *Sdha* are significantly up- or down-regulated after pharmacological intervention.  $\beta 2$ -microglobulin and  $\beta$ -actin showed instability in all hypoxic conditions ( $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ,  $n = 5$ ). (C) After 5 days of culture stabilisation, organotypic hippocampal cultures (OHCs) were exposed to 15 min of OGD followed by 24 h of Re-Ox. (D) While  $\beta 2$ -microglobulin,  $\beta$ -actin, *Sdha*, *Hprt* and *Rpl13* expression remains stable under all experimental conditions, *Ywhaz*, *Gapdh* and *18S* resulted in unsuitable housekeeping genes for this model ( $*p < 0.05$ ,  $**p < 0.01$ ,  $n = 5$ ). (E) Adult WT mice were subjected to a 1 h-transient occlusion of the middle cerebral artery (tMCAO) followed by 23 h of reperfusion. (F)  $\beta 2$ -microglobulin, *Sdha*, *Gapdh*, *Hprt* and *Rpl13* gene expression remained stable under all interventions while *Ywhaz*, *18S* and  $\beta$ -actin are significantly up-regulated due to treatment ( $*p < 0.05$ ,  $**p < 0.01$ ,  $***p < 0.001$ ,  $n = 5$ ).



**Figure 2.** Oposing gene expression results are based on specific housekeeping gene selection. (A) HK gene validation using acute hippocampal brain slices (HBS) identified *Yhawz*,  $\beta$ -actin, and *18S* as non-stable, while *Hprt*, *Gapdh*, and *Rpl13* resulted in the best candidates for this model. (B) NADPH oxidase 4 (*Nox4*) gene expression was assessed in HBSs subjected to 3 different experimental conditions, (i) basal, (ii) oxygen and glucose deprivation (OGD), and (iii) OGD under treatment conditions. *Nox4* gene expression was normalised against previously detected unstable genes (*Yhawz*,  $\beta$ -actin, and *18S*) and the most stable candidates (*Hprt*, *Gapdh*, and *Rpl13*). Diametrically opposing results were identified depending on the selected HK gene ( $^*p < 0.05$ ,  $^{##}p < 0.01$   $n = 5$ ) later prevented by 2 h of pharmacological treatment ( $^*p < 0.05$ ,  $n = 5$ ).

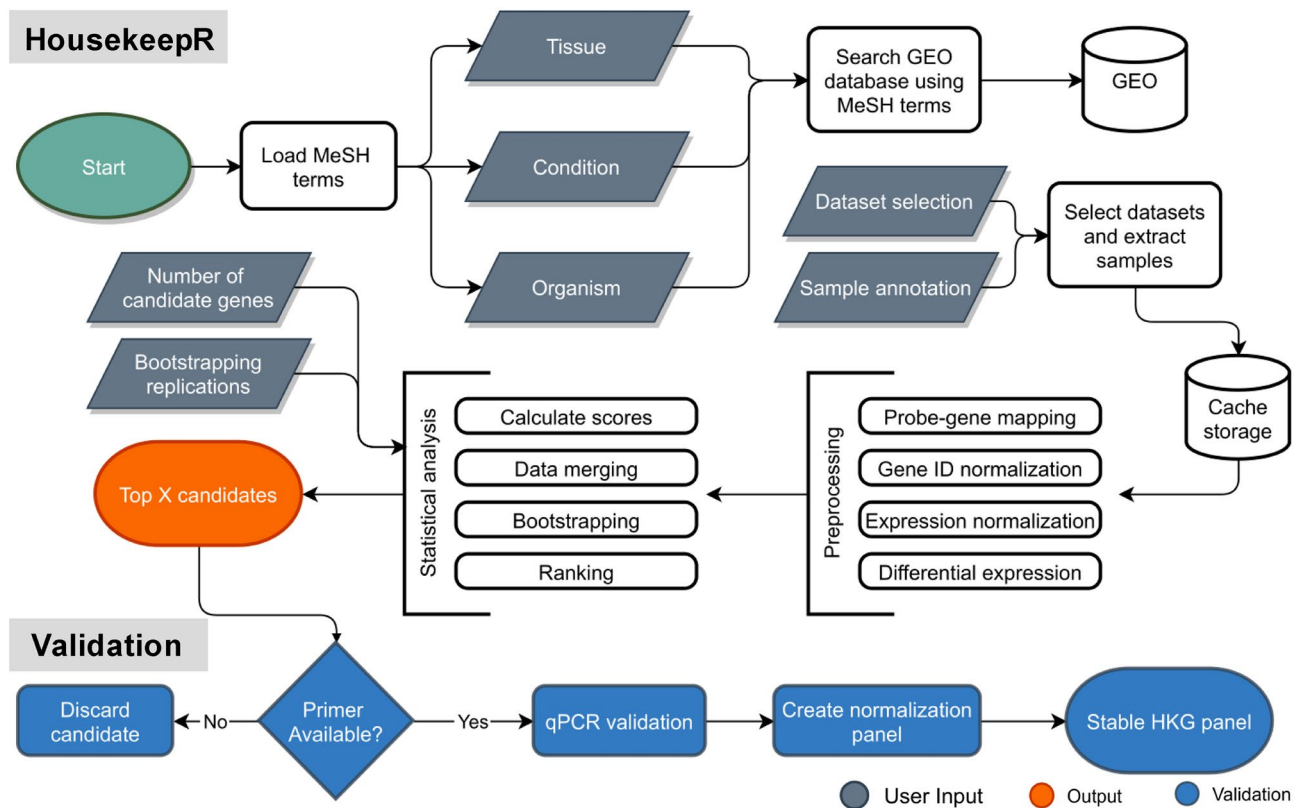
unstable genes resulted in opposing results contradicting previous literature<sup>23</sup> (Fig. 2B–E). Hence, identical therapeutic interventions would have been interpreted as either neuroprotective or ineffective depending on the gene used for expressional normalisation.

**HouseKeepR allows for robust de novo normalisation gene identification.** Our literature-based HK identification approach shows that relying on previous publications, even in the same experimental model, is insufficient for HK gene selection and can lead to grossly misleading results. Thus, experimental validation of gene stability is essential per model, tissue, and condition, leading to additional experiments and, more importantly, would still be biased because the initial gene selection is only based on previous similar and published use. Potentially more stable and useful HK genes may be systematically missed by this approach.

To overcome this knowledge problem and the methodological gap in gene expression normalisation, we developed HouseKeepR, a web tool that robustly and rapidly ranks genes across relevant and publicly available expressional data sets (Fig. 3). Existing approaches for this task rank genes only by variance; HouseKeepR, also for consistently high average expression individually for each data set. Moreover, to ensure that results are robust towards sampling bias, HouseKeepR uses a bootstrapping strategy to obtain a distribution, expected mean rank, and rank variance for each gene.

**HouseKeepR is easily accessible through a user-friendly web interface.** For user-friendliness, HouseKeepR can be applied to specified gene expression data sets, which are automatically retrieved from the public Gene Expression Omnibus (GEO) database<sup>24</sup>. Moreover, to encourage broad applicability also by non-bioinformaticians, we have developed a convenient web interface with integrated search and sample annotation functions (<https://exbio.wzw.tum.de/housekeepR>). HouseKeepR requires only three main input parameters, i.e., tissue, condition, and organism, to automatically identify the most suitable HK gene candidates. Optional parameters include the number of final HK gene candidates to be identified or the number of bootstrap replications (default is 10,000 repetitions).

HouseKeepR then searches the GEO and Ensembl databases and reports back all data sets matching the selected parameters<sup>24</sup>. Next, users must choose at least two of those data sets and annotate condition and control samples to perform the analysis. The server will then download these data sets and run the HouseKeepR



**Figure 3.** HouseKeepR algorithm flow chart. First, MeSH terms are parsed and preloaded to be used as standardised search terms for expressions data sets. User inputs for tissue, condition and organism, encoded in MeSH terms, are used to search and download data sets from Gene Expression Omnibus (GEO). The data sets are then pre-processed and statistically analysed using a scoring formula engineered to find normalisation genes. The top-ranking genes, based on user choice, are selected to be validated. In this scenario, normalisation genes within the first 25 ranked candidates were included in the subsequent validation process. Grey-colored elements show user required input, while the orange elements are the outputs of the user, and the blue elements are external validation steps conducted via wet-lab experiments.

algorithm. Finally, results are displayed as a ranking table (Supplementary Fig. 2) and a rank distribution across data sets (Supplementary Fig. S3). The run time depends on the number of data sets and the response time of the GEO database. Importantly, datasets are subject to updates, possibly affecting reproducibility over time. To address this issue, HouseKeepR allows users to save sessions, which preserve dated results and the parameters used for later retrieval. Furthermore, users can optionally select even older versions from the archive if required for specific questions. With this, HouseKeepR enables every biomedical researcher to quickly and easily identify stable normalisation genes for differential gene expression analysis via a meta-analysis for any tissue, condition, and organism benefiting from a large and continuously growing number of publicly available data sets.

**HK genes candidates are robust across non-overlapping data sets.** As described, users have the option to select at least two or more databases for HK gene identification. Since different users may select different databases, we next examined whether the selection of non-overlapping databases may affect the robustness of HouseKeepR. Through bootstrapping, HouseKeepR allows studying the stability of the results under different random samplings from the included data sets. As additional consistency validation, we executed HouseKeepR 10 times with the same data sets and bootstrap parameters but using different random seeds. In this context, we computed the reproducibility measure (R) as the average of all pairwise overlap coefficients calculated between each pair of results lists (with the top 50 HK genes) produced by each run. The ten runs achieved an R-value > 0.99, which demonstrates the robustness of our algorithm. Another important measure for HouseKeepR is the reproducibility of HK gene candidates when comparing different non-overlapping data sets for the same tissue, organism and condition. To double-check this functionality, we selected two non-overlapping sets of 12 and two data sets, respectively. The top 20 HK gene candidates of both groups were compared, yielding an overlap coefficient of 0.5 (p-value =  $9.75 \times 10^{-29}$ ) (Supplementary Fig. S4)<sup>25</sup>. The significant overlap indicates that HouseKeepR successfully predicts stable HK genes across different expression platforms and experimental setups. Contrary, as a negative control, we ran HouseKeepR on a group of two data sets of different organisms and conditions. Importantly, the overlap coefficient of the top 20 HK genes to the previously selected ones was 0 (Supplementary Fig. S4), confirming that predictions are condition- and organism-dependent, which underlines the necessity to use HouseKeepR for HK gene selection.

**NormFinder confirms the stability of HouseKeepR candidate genes.** As mentioned, HouseKeepR is not the first software of its kind, although it has major advantages concerning usability and breadth of application. Nevertheless, we next cross-evaluated the results of HouseKeepR with a similar computational approach, NormFinder. The latter employs a more elaborate statistical model to test gene stability which is computationally expensive and thus not suited for systematic screening across several data sets (Supplementary Fig. 5). Stable genes were defined using NormFinder's stability score with a cut-off value of 0.15). 80% of the top 10 HK genes (Supplementary Fig. 6) generated by HouseKeepR were confirmed to be stable by NormFinder, highlighting that HouseKeepR reported genes are also considered robust by independent statistical evaluation.

**HouseKeepR suggested HK genes are consistent in *Nox4* target validation.** Once we demonstrated the reliability and reproducibility of the HouseKeepR tool, we performed an ultimate web lab validation step to confirm literature-based or find de novo and unbiased more suitable HK genes for this condition. We selected 12 data sets (Supplementary Table S3) associated with brain hypoxia/ischemia experiments in rats and mice (see Methods for details). HouseKeepR then calculated an overall ranking of 19,878 genes over 10,000 bootstrap samples. From the top 100 most stable genes, ten were randomly selected (#11-#96 in Fig. 4A,B). Since normalisation with multiple genes has been reported to outperform single-gene normalization<sup>9</sup>, we then assessed the reproducibility of *Nox4* induction under hypoxic conditions normalised against each of the ten genes individually (Fig. 4C) in a panel of the top two (#11 and #12, *Rps23* and *Cst3*), four, six, eight and ten (Fig. 4D). Importantly, only a panel of four HK genes ensured reliable normalisation and a statistically significant upregulation of *Nox4*. Reproducing differential *Nox4* expression in brain hypoxia based on single genes or a panel of two resulted in increased values that did not, however, reach statistical significance.

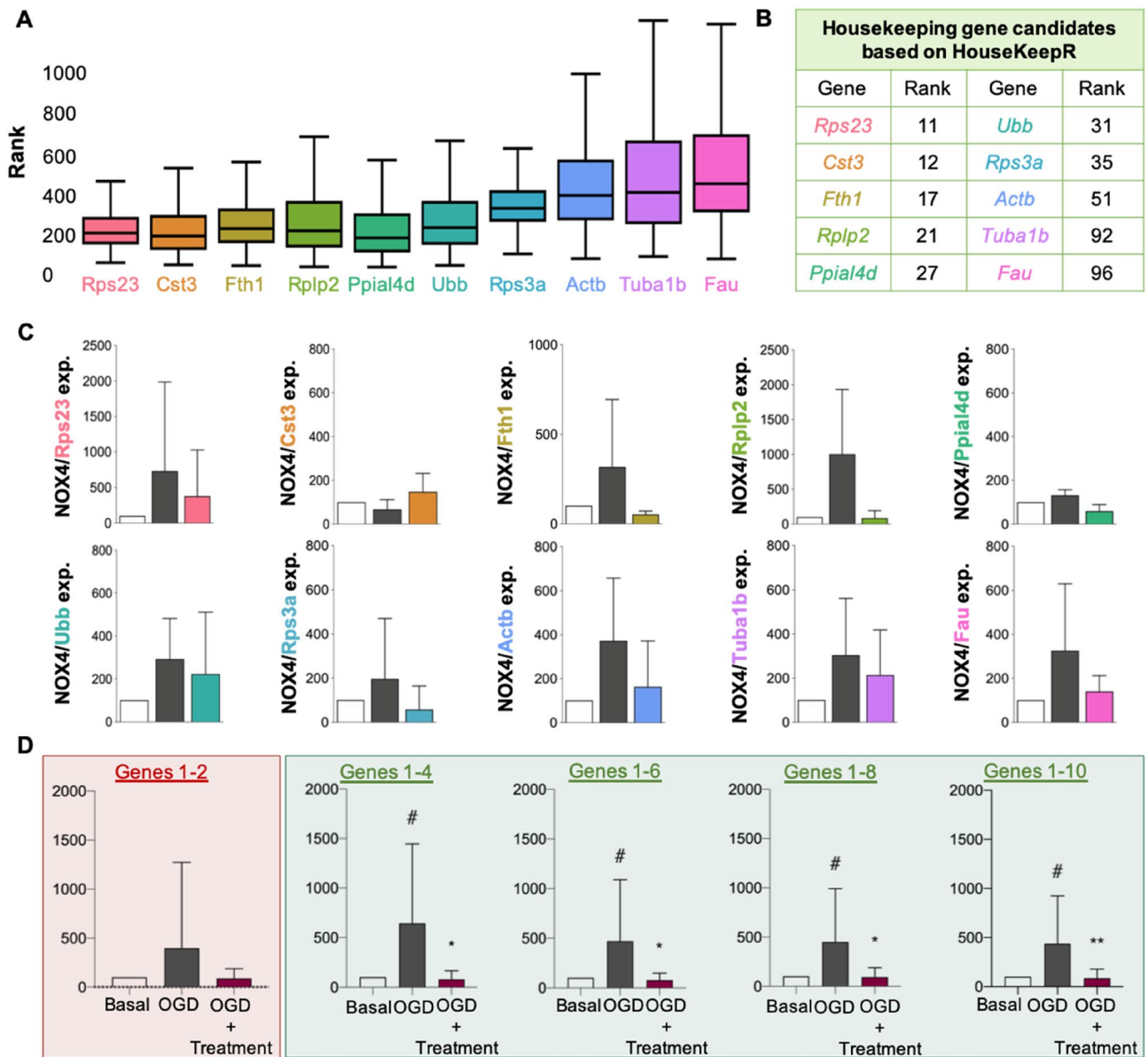
## Discussion

Choosing a reliably stable gene as HK gene is obviously critical for any differential gene expression study. Lack of stability can lead to false conclusions<sup>26,27</sup>. Here we demonstrate for both ex-vivo and in vivo models of brain hypoxia that literature-derived, repeatedly used HK genes display surprisingly considerable variance that can result in opposing conclusions on whether to consider a gene a differentially expressed or not. We demonstrate this for inducible *Nox4*, a promising therapeutic target in brain ischemia, currently under clinical testing (REPO-STROKE I and II)<sup>17,22</sup>.

The most widely used method for HK gene selection, i.e., literature-based and hence biased, may lead to reproducibility issues contributing to the quality crisis in biomedical research<sup>28</sup>. Thus, a systematic, unbiased method for HK gene selection is evidently needed. Moreover, for practical reasons, this method should be tissue- and condition-specific and user-friendly so biomedical scientists can widely adopt it without affording programming and coding skills. We have examined methods that select HK based on gene expression data and found limitations that hinder their adoption. RefGenes, for instance, which is part of the free version of the Genevestigator (<https://genevestigator.com/>) platform, relies on the standard deviation of HK genes to define housekeeping gene candidates. However, by focusing on a single gene expression data set, it neglects variations between conditions, which can be misleading in cases of low signal-to-noise ratio. While RefGenes offers a graphical user interface, it does not support exporting results for downstream analysis. NormFinder, which is available as Excel and R script, is another widely adopted method for validating HK genes in RT-qPCR data. In contrast to RefGenes, it considers not only the variance within a group of samples but also between groups, allowing for analysis across data sets, organisms, or conditions. While NormFinder is in principle applicable to microarray expression data, it is typically used to validate a set of candidate genes and is not suited for a genome-wide search for de novo HK gene candidates, where the runtime of the NormFinder algorithm becomes the limiting factor. Similarly, BestKeeper is available as an Excel spreadsheet that calculates the geometric mean of a set of input genes as an index which can be used to select the best combination of reference genes. The tool is limited, however, to 10 input genes and 100 samples while requiring expression values to be presented as Crossing Points<sup>29</sup>. Since existing tools do not satisfy one or more of the aforementioned criteria, we developed HouseKeepR, a web application for de novo HK gene discovery which fulfils, to our knowledge for the first time, all these criteria.

HouseKeepR suggests HK gene candidates that are highly reproducible both in silico and in vitro. Using our model system and target gene, *Nox4*, expression changes could be detected with single HK genes; the more subtle response of *Nox4* expression to pharmacological treatment was only detectable when using a panel of four HK genes for normalisation. Thus multiple HK gene-based normalisation results in more solid and uniform fold-changes correlations compared to using single gene expressions, providing consistency across data sets<sup>5</sup>. Thus, unbiased panel-based normalisation with condition-specific and cross-data set validated HK genes should become the future gold standard for normalising RT-PCR data. The number of HK genes for normalisation should be at least two (according to the MIQE guidelines)<sup>30</sup>, while the optimal number can be as high as 17, as shown for studying cellular senescence in human Endothelial Colony Forming Cells<sup>31</sup>. In this study, two reference genes did not achieve stable results, but a panel of at least four genes was necessary. Therefore, for a common experimental setup used in the biomedical field, we suggested that a panel of at least 2 to 4 would be required.

As a possibly perceived limitation, HouseKeepR is currently limited to microarray data in the GEO database. These data are not uniformly processed, and, hence, HousekeepR cannot expect data sets to be normalised for technical biases. Neglecting technical biases can lead to erroneous results when looking for a biological variation in, e.g., differential gene expression analysis. For HouseKeepR's objective of finding suitable reference genes, however, it may even be advantageous to identify candidates that are robust to technical bias. Moreover, as RNA-seq data are becoming more prevalent, they can be incorporated as additional data sources, e.g., ARCHS4<sup>32</sup> or GEMMA<sup>33</sup>. For large scale analysis, RNA-seq will eventually supersede qRT-PCR based gene expression analysis



**Figure 4.** Un-biased de novo housekeeping gene panel generation and validation. **(A)** Ranking distribution of 10 randomly selected housekeeping gene candidates within the 100 top ranking generated over 10,000 bootstrap runs over all the samples in the 12 data sets selected for analysis. The order of the plot follows the final ranking in **(B)**. **(B)** Final ranking of the selected normalisation gene candidates: *Ubb*, *Fth1*, *Fau*, *Ppia4d*, *Cst3*, *Rps23*, *Actb*, *Tuba1*, *Rps3*, *Rplp2*. This ranking is calculated by taking the mean bootstrap rank from **A** and penalising it for the absence of genes from each of the 12 data sets selected for analysis. **(C)** *Nox4* gene expression was normalised against the selected 10 genes of the ranking under 3 different experimental conditions, i.e. Basal, oxygen and glucose deprivation (OGD) and pharmacologically treated hippocampal tissue post-OGD. **(D)** 5 housekeeping genes panels were generated and validated: (i) Panel 1; 2 genes (*Ubb* and *Fth1*); (ii) Panel 2, 4 genes (*Ubb*, *Fth1*, *Fau* and *Ppia4d*); (iii) Panel 3, 6 genes (*Ubb*, *Fth1*, *Fau*, *Ppia4d*, *Cst3* and *Rps23*); (iv) Panel 4, 8 genes (*Ubb*, *Fth1*, *Fau*, *Ppia4d*, *Cst3*, *Rps23*, *Actb* and *Tuba1*); (v) Panel 5, 10 genes (*Ubb*, *Fth1*, *Fau*, *Ppia4d*, *Cst3*, *Rps23*, *Actb*, *Tuba1*, *Rps3* and *Rplp2*). Panel 2, 3, 4 and 5 showed stable *Nox4* gene expression results, while Panel 1 remained statistically not significant ( $^{\#}p < 0.05$ ,  $n = 6$ ;  $n = 6$ ;  $^*p < 0.05$ ,  $n = 6$ ;  $^{**}p < 0.01$ ,  $n = 6$ ).

as the gold standard. However, HK gene-based normalisation will also need to be applied in this context<sup>34,35</sup>, and the need for unbiased and robust HK gene panels remains just as relevant.

## Methods

**Systematic review.** A literature review focused on two in vitro models of brain ischemia: (i) hippocampal brain slices (HBS) and (ii) organotypic hippocampal culture (OHC) was performed. PubMed was searched for original papers and conference abstracts where these ex-vivo models appeared. No terms for RT-qPCR were

included since PubMed only screens abstracts, titles and keywords, and RT-qPCR details are frequently mentioned only in the methods section. No language restriction was used. Our search strategy for the HBS model identified 364 records. First, these hits were screened based on title, abstract and results, excluding other, non-related ex-vivo models. Publications without RT-qPCR experiments or inaccessible full-text, i.e., only title/abstract published, were excluded. Finally, 16 articles were included for full-text screening. In 1 article, no normalisation gene was used, and in 4, RT-qPCR was conducted using other species. After a full-text assessment, 11 articles were considered. Similarly, our search strategy for model 2 (OHCs) identified 249 records in PubMed. Following the first screening, 38 articles were included for full-text screening. 2 of these did not use normalisation genes and 9 conducted RT-qPCR experiments using tissue from other species. Therefore, in total, 27 articles were considered for the OHC model. Studies were included if (i) the specific ex-vivo ischemia model was used; (ii) specific conditions were considered within the experiment; (iii) RT-qPCR experiments were conducted.

**Animals.** Rats used for ex-vivo experiments were handled following the Guide for the Care and Use of Laboratory Animals and were previously approved by the Institutional Ethics Committee of Universidad Autónoma de Madrid, Spain, according to the European guidelines for the use and care of research animals by the European Union Directive of 22 September 2010 (2010/63/UE) and the Spanish Royal Decree of 1 February 2013 (53/2013). Similarly, in vivo experiments in mice strictly followed the Dutch law on animal experiments and were approved by the local animal experimental committee (Maastricht University, DEC2011-106). Both mice and rats were housed under controlled conditions of temperature (22 °C), humidity (55–65%), light (12 h light–dark cycles) and free access to water and standard laboratory chow. Male and female C57/Bl6 mice (8–12 weeks old), Sprague–Dawley adult rats (8–12 weeks old) and Sprague–Dawley pups (7–10 days old) were used.

**Ex-vivo acute model: preparation of hippocampal brain slices and induction of oxygen and glucose deprivation.** Experiments were performed using hippocampal brain slices from adult male Sprague–Dawley rats (8–12 weeks old) as previously described in<sup>36,37</sup>. Briefly, rats were quickly decapitated, and forebrains were rapidly removed from the skull and placed into ice-cold Krebs bicarbonate dissection buffer (pH 7.4), containing (in mM): NaCl 120, KCl 2, CaCl<sub>2</sub> 0.5, NaHCO<sub>3</sub> 26, MgSO<sub>4</sub> 10, KH<sub>2</sub>PO<sub>4</sub> 1.18, glucose 11 and sucrose 200. At least 20 min before starting the experiment, chamber solutions were bubbled with either 95% O<sub>2</sub>/5% CO<sub>2</sub> or 95% N<sub>2</sub>/5% CO<sub>2</sub> gas mixtures to ensure O<sub>2</sub> and N<sub>2</sub> saturation, respectively. The hippocampus was quickly dissected and subsequently cut into transverse slices 300 µm thick using a Tissue Chopper Mcllwain. To recover from slicing trauma, slices were incubated in Krebs buffer for 45 min at 34 °C (stabilisation period). Then, control slices were incubated for 15 min in a Krebs-bicarbonate solution without sucrose (control solution). Oxygen and glucose deprivation was induced by incubating the slices for 15 min in a glucose-free Krebs-bicarbonate solution in which glucose was replaced by 2-deoxyglucose (OGD solution). Both solutions were pre-bubbled for 30 min with a 95% O<sub>2</sub>/5% CO<sub>2</sub> or 95% N<sub>2</sub>/5% CO<sub>2</sub>, respectively. All experiments were performed at 37 °C. Following the OGD period, slices were returned to an oxygenated Krebs-bicarbonate solution containing glucose for 120 min (Re-Ox period). During the re-oxygenation period, either a pharmacological treatment (OGD + Treatment), or no extra measures (OGD) were taken. Specifically, treated hippocampal brain slices were exposed to 0.1 µM GKT136901 (NOX4 inhibitor). After the Re-Ox period, slices were collected and quickly shock-frozen.

**Ex-vivo chronic model: preparation of organotypic hippocampal slices and induction of oxygen and glucose deprivation.** Hippocampal brain slices for cultures were obtained from brains of 7- to 10-days-old Sprague–Dawley rats. Organotypic cultures were prepared based on the methods previously described in<sup>26</sup>. Briefly, pups were quickly decapitated and brains removed from the skull and dissected. The hippocampus was cut into 300 µm-thick slices using a Tissue Chopper Mcllwain. Then, they were separated in sterile ice-cold Hank's balanced salt solution (HBSS, Biowest, Madrid, Spain) containing (in mM): glucose 15, CaCl<sub>2</sub> 1.3, KCl 5.36, NaCl 137.93, KH<sub>2</sub>PO<sub>4</sub> 0.44, Na<sub>2</sub>HPO<sub>4</sub> 0.34, MgCl<sub>2</sub> 0.49, MgSO<sub>4</sub> 0.44, NaHCO<sub>3</sub> 4.1, HEPES 25, 100 U/ml penicillin, and 0.100 mg/ml gentamicin. Six slices were placed on each Millicell-0.4 µm culture insert (Millipore, Madrid, Spain) within each well of a six-well culture plate. Specific neurobasal medium (Invitrogen, Madrid, Spain) enriched with 10% of fetal bovine serum (Sigma-Aldrich, Madrid, Spain) was used for the next 24 h (1 ml/well). 24 h later, B27 supplement and antioxidants were added to the culture medium. Slices were in culture for 4d before inducing the OGD period. On day 6, inserts were placed in 1 ml of OGD solution composed of (in mM): NaCl 137.93, KCl 5.36, CaCl<sub>2</sub> 2, MgSO<sub>4</sub> 1.19, NaHCO<sub>3</sub> 26, KH<sub>2</sub>PO<sub>4</sub> 1.18, and 2-deoxyglucose 11 (Sigma-Aldrich, Madrid, Spain). The cultures were then placed in an airtight chamber (Billups-Rothenberg Inc., USA) and exposed for 5 min to 95% N<sub>2</sub>/5% CO<sub>2</sub> gas flow to ensure oxygen removal. Then, the chamber was sealed for 15 min and placed at 37 °C (OGD period). At the same time, control cultures were maintained under a normoxic atmosphere in a solution with the same composition as previously described but containing glucose (15 mM) instead of 2-deoxyglucose. Pharmacological treatment was added to the cultures before returning them to normal oxygen and glucose concentrations for 24 h (re-oxygenation period). Specifically, treated organotypic brain slices were exposed to either a single treatment or a combination of 0.1 µM GKT136901 (NOX4 inhibitor) after 15 min of OGD. After the Re-Ox period, slices were collected and quickly shock-frozen.

**In vivo model: transient occlusion of the middle cerebral artery (tMCAO) in mice.** Stroke surgery was conducted as previously described in<sup>38</sup>. After administering a painkiller, animals were anaesthetised with isoflurane (induction 4–5% in air, maintenance 2–2.5% in air) and placed on a heating pad that maintained the rectal temperature at 37.0 °C using a feedback-controlled infrared lamp. Using a surgical microscope (Wild M5A, Wild Heerbrugg, Gais, CH), a midline neck incision was made, and both the right common and external



carotid arteries were isolated and permanently ligated while a microvascular clip was temporarily placed on the internal carotid artery. A small incision into the common carotid artery was performed where the silicon rubber-coated 6.0 nylon monofilament (602312PK10, Doccol Corporation, Sharon, MA, USA) was inserted until resistance was felt. The monofilament tip should be specifically located intracranially at the origin of the right middle cerebral artery, thereby interrupting blood flow. The filament was fixed with a tourniquet suture to prevent dislocation. 1 h after occlusion of the middle cerebral artery, reperfusion was initiated by monofilament removal. Wounds were carefully sutured, and animals could recover in a temperature-controlled cupboard. Pharmacological treatment (GKT136901, 10 mg/kg) was given via i.p. injections, 2 h and 12 h after the start of ischemia. Animals were sacrificed 24 h after induction of ischemia by cervical dislocation. Brains were quickly removed, and shock froze.

**RNA extraction, quantification and reverse transcription.** Hippocampal brain slices and brain tissue from ex-vivo and in vivo models were crushed and homogenised using TRI Reagent<sup>®</sup> (Sigma-Aldrich, The Netherlands). 100  $\mu$ l of chloroform was added to the samples, followed by a 15 min centrifugation at 11,000 rpm and 4 °C. After centrifugation, 250  $\mu$ l of isopropanol was added to the upper phase (mRNA) and then kept for 1 h at – 20 °C. After incubation, samples were centrifuged during 10 min at 13,000 rpm, and 4 °C. 200  $\mu$ l ethanol 80% was added to the supernatant followed by 10 min centrifugation at 13,000 rpm. After ethanol removal, the mRNA was dissolved in RNase free water. mRNA was quantified spectrophotometrically using the Nanodrop 2000 device. 0.08  $\mu$ g of total mRNA was reverse transcribed to cDNA with the High-Capacity Reverse Transcription Kit (Applied Biosystems, The Netherlands) according to the manufacturer's protocol.

**Real-time PCR.** mRNA levels of studied genes were quantified using the fluorescent Taqman<sup>®</sup> technology. We used TaqMan<sup>®</sup> gene expression arrays (TaqMan<sup>®</sup> Universal PCR Master Mix, ThermoFisher Scientific, The Netherlands) for all species: (i) For rat:  $\beta$ 2-microglobulin (Rn00560865\_1, ThermoFisher Scientific, The Netherlands),  $\beta$ -actin (Rn00667869\_m1, ThermoFisher Scientific, The Netherlands), *Rpl13* (Rn00821946\_m1, ThermoFisher Scientific, The Netherlands), *18S* (Hs99999901\_s1, ThermoFisher Scientific, The Netherlands), *Hprt* (Rn01527840\_m1, ThermoFisher Scientific, The Netherlands), *Sdha* (Rn00590475\_m1, ThermoFisher Scientific, The Netherlands), *Ywhaz* (Rn00755072\_m1, ThermoFisher Scientific, The Netherlands), *Gapdh* (Rn01775763\_g1, ThermoFisher Scientific, The Netherlands), *Nox4* (Rn01506793\_m1, ThermoFisher Scientific, The Netherlands). (ii) For mice:  $\beta$ 2-microglobulin (Mm00437762\_m1, ThermoFisher Scientific, The Netherlands),  $\beta$ -actin (Mm02619580\_g1, ThermoFisher Scientific, The Netherlands), *Rpl13* (Mm02526700\_g1, ThermoFisher Scientific, The Netherlands), *18S* (Hs99999901\_s1, ThermoFisher Scientific, The Netherlands), *Hprt* (Mm03024075\_m1, ThermoFisher Scientific, The Netherlands), *Sdha* (Mm01352366\_m1, ThermoFisher Scientific, The Netherlands), *Ywhaz* (Mm03950126\_s1, ThermoFisher Scientific, The Netherlands), *Gapdh* (Mm99999915\_g1, ThermoFisher Scientific, The Netherlands), *Ppia4d* (Mm01191872\_g1, ThermoFisher Scientific, The Netherlands), *Fth1* (Mm00850707\_g1, ThermoFisher Scientific, The Netherlands), *Cst3* (Mm00438347\_m1, ThermoFisher Scientific, The Netherlands), *Rps23* (Mm03019701\_g1, ThermoFisher Scientific, The Netherlands), *Rplp2* (Mm00782638\_s1, ThermoFisher Scientific, The Netherlands), *Ubb* (Mm01622233\_g3, ThermoFisher Scientific, The Netherlands), *Fau* (Mm02601595\_u1, ThermoFisher Scientific, The Netherlands), *Tuba1* (Mm00846967\_g1, ThermoFisher Scientific, The Netherlands), *Rps3* (Mm00656272\_m1, ThermoFisher Scientific, The Netherlands) (Table S2). Water controls were included to ensure specificity, and the comparative  $2^{-\Delta\Delta C_t}$  method was used for relative quantification of gene expression. When pre-designed primers are not commercially available, manual design is strongly recommended based on the required experimental scenario. In case of multiple comparisons, experimental biological samples were normalised to either 1 specific gene or different ones of the proposed housekeeping panel with the same final number of normalisation rounds.

**NormFinder stability score.** The NormFinder algorithm<sup>12</sup> can be used to assess the stability of normalisation genes across groups based on RT-qPCR measurements. Briefly, NormFinder models the variation within and between sample groups to estimate a HK gene candidate's expression stability in the form of a distribution

$$f_{ig} = z_{ig} - \theta_g - \alpha_i$$

where  $z_{ig}$  is the mean, log-transformed expression level of the gene  $i$  for all samples in group  $g$ ,  $\theta_g$  is the average amount of mRNA for group  $g$ , and  $\alpha_i$  is the mean expression level for gene  $i$  over all groups. Hence, the resulting distribution is an additive measure of variance within and between groups. NormFinder transforms the resulting stability distribution to an easier-to-interpret stability value  $\rho_{ig}$  by taking the absolute value of the mean + 1 standard deviation:

$$\rho_{ig} = |\text{mean}(f_{ig}) + \text{sd}(f_{ig})|$$

**HouseKeepR method.** Normalisation genes should show stable expression levels across tissues and conditions of interest, i.e., a low variance and log fold change are desirable. This first characteristic is captured by the coefficient of variation (CV):

$$CV = \frac{\text{Standard Deviation}}{\text{Mean Expression}}$$

The CV captures the relation of standard deviation and mean expression and is thus a unitless measure of noise and overdispersion commonly used for describing the robustness of a measurement<sup>39</sup>. As a second characteristic, we consider the fold change between conditions to avoid selecting normalisation genes affected by the condition of interest. As was previously shown, the CV tends to be higher for genes with low expression and plateaus for higher expression levels<sup>40</sup>. In addition, fold changes tend to be inflated for lowly expressed genes<sup>41</sup>. Thus, ranking normalisation genes by both CV and fold change will help select genes with comparably high expression that are not affected by measurement bias and robustly expressed between conditions. We compute both the CV and the fold change based on raw expression values for each gene across the samples of each data set. Then the fold change is inverted if it is < 1 to get the absolute value of change. Since most GEO data sets are provided as log-transformed expressions, exponentiation is applied to retrieve raw values. To combine CV and fold change into a single quality score, we first rank fold changes and CVs independently before computing a joint rank product. An ideal normalisation gene candidate consequently achieves a low rank across various data sets. To avoid that the selected HK genes exhibit high-rank stability by chance, we use bootstrapping, i.e. sampling with replacement. As the bootstrap samples will always have a different composition, we can judge the effect of sampling bias on rank stability. For each bootstrap repetition, the genes are ranked by the mean rank across the sampled data sets. Genes with stable ranks over a large number of bootstrap repetitions, e.g., 10,000, are selected as the top-performing normalisation gene candidates. The combined use of ranks and bootstrapping enables the processing of different array data sets without the need for batch effects removal or cross-platform normalisation, enabling HouseKeepR to perform largely automated meta-analyses for normalisation gene identification.

**HouseKeepR R shiny web interface.** Existing tools such as NormFinder and BestKeeper do not offer a user-friendly interface and do not afford genome-wide coverage. RefGenes has a user interface but is not capable of integrative analysis across publicly available gene expression data, leading to unstable results when using different expression platforms (Supplementary Fig. S7). This motivated us to make the HouseKeepR method easily accessible to the community via an R shiny web interface which leads users through the process of de-novo detection of suitable normalization genes for user-selected tissues and conditions. The basis for this analysis is the Gene Expression Omnibus (GEO)<sup>42</sup>, a widely used repository currently offering access to more than a hundred thousand gene expression profiles of more than three million samples. HouseKeepR allows users to directly query GEO for data sets related to specific organisms, conditions and tissues. HouseKeepR shows the query's result and provides additional information about available studies. Next, users can select studies relevant to de-novo HK gene discovery and assign conditions and control labels for each sample. Once the user starts an analysis, HouseKeepR will (i) download expression data using the 'GEOquery' R package<sup>43</sup>, (ii) map microarray probe identifiers to gene identifiers such as Entrez or Ensembl, (iii) identify and map gene homologs across organisms, and (iv) compute the differential expression between condition and control samples using the 'limma' Bioconductor package<sup>44</sup> primer. Once the analysis is complete, HouseKeepR reports the top-performing HK genes together with informative visualisations of their ranks across data sets. Figure 3 provides an overview of these steps. HouseKeepR is released under an open-source license (<https://github.com/biomedbigdata/housekeepr>) and available for local use as a docker container or an online web application at <https://exbio.wzw.tum.de/housekeepr>.

**HouseKeepR application to mouse and rat models of ischemic stroke.** We demonstrate the practical application of HouseKeepR by validating de-novo identified HK genes with RT-qPCR. To this end, we selected rats and mice as organisms, ischemia, ischemic or stroke as conditions and brain as tissue. These parameters returned 163 data sets, from which we excluded those that did not satisfy our model criteria or that showed special model attributes that could affect predictions of normalisation genes, such as genetically modified organisms. Finally, 12 high-quality data sets (Supplementary Table 1) were selected for further analysis. An AIME report<sup>45</sup> for reproducible machine learning was created and deposited at <https://aime.report/F9IKDV>.

**Statistical analysis.** Experimental results were presented as means ± SEM. Differences between groups were determined by applying a one-way ANOVA followed by two-way ANOVA followed by Dunnett's Multiple Comparison test or Student's two-tailed t-test and Mann-Whitney test experiments when appropriate. For repeated measurements, a two-way ANOVA was used. Statistical analysis was conducted using GraphPad Prism version 5.00. The level of statistical significance was set at  $p < 0.05$ .

### Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 9 December 2021; Accepted: 4 July 2022

Published online: 19 July 2022

### References

- Atkinson, T. J. & Halfon, M. S. Regulation of gene expression in the genomic context. *Comput. Struct. Biotechnol. J.* **9**, e201401001. <https://doi.org/10.5936/csbj.201401001> (2014).
- Fannon, M. R. Gene expression in normal and disease states—identification of therapeutic targets. *Trends Biotechnol.* **14**, 294–298. [https://doi.org/10.1016/0167-7799\(96\)10041-X](https://doi.org/10.1016/0167-7799(96)10041-X) (1996).
- Kubista, M. *et al.* The real-time polymerase chain reaction. *Mol. Aspects Med.* **27**, 95–125. <https://doi.org/10.1016/j.mam.2005.12.007> (2006).

4. Thellin, O. *et al.* Housekeeping genes as internal standards: Use and limits. *J. Biotechnol.* **75**, 291–295. [https://doi.org/10.1016/S0168-1656\(99\)00163-7](https://doi.org/10.1016/S0168-1656(99)00163-7) (1999).
5. Huggett, J., Dheda, K., Bustin, S. & Zumla, A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* **6**, 279–284. <https://doi.org/10.1038/sj.gene.6364190> (2005).
6. Jain, N., Vergish, S. & Khurana, J. P. Validation of house-keeping genes for normalization of gene expression data during diurnal/circadian studies in rice by RT-qPCR. *Sci. Rep.* **8**, 3203. <https://doi.org/10.1038/s41598-018-21374-1> (2018).
7. Tricarico, C. *et al.* Quantitative real-time reverse transcription polymerase chain reaction: Normalization to rRNA or single house-keeping genes is inappropriate for human tissue biopsies. *Anal. Biochem.* **309**, 293–300. [https://doi.org/10.1016/S0003-2697\(02\)00311-1](https://doi.org/10.1016/S0003-2697(02)00311-1) (2002).
8. Dheda, K. *et al.* The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Anal. Biochem.* **344**, 141–143. <https://doi.org/10.1016/j.ab.2005.05.022> (2005).
9. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* <https://doi.org/10.1186/gb-2002-3-7-research0034> (2002).
10. Pfaffl, M. W., Tichopad, A., Prgommet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–515. <https://doi.org/10.1023/b:bile.0000019559.84305.47> (2004).
11. Hruz, T. *et al.* RefGenes: Identification of reliable and condition specific reference genes for RT-qPCR data normalization. *BMC Genomics* **12**, 156. <https://doi.org/10.1186/1471-2164-12-156> (2011).
12. Andersen, C. L., Jensen, J. L. & Orntoft, T. F. 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.* **64**, 5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496> (2004).
13. Chia, C. Y., Lim, C. W., Leong, W. T. & Ling, M. H. High expression stability of microtubule affinity regulating kinase 3 (MARK3) makes it a reliable reference gene. *IUBMB Life* **62**, 200–203. <https://doi.org/10.1002/iub.295> (2010).
14. Keng, B. M., Chan, O. Y., Heng, S. S. & Ling, M. H. Transcriptome analysis of *Spermophilus lateralis* and *Spermophilus tridecemlineatus* liver does not suggest the presence of spermophilus-liver-specific reference genes. *ISRN Bioinform.* **2013**, 361321. <https://doi.org/10.1155/2013/361321> (2013).
15. Too, I. H. & Ling, M. H. Signal peptidase complex subunit 1 and hydroxyacyl-CoA dehydrogenase beta subunit are suitable reference genes in human lungs. *ISRN Bioinform.* **2012**, 790452. <https://doi.org/10.5402/2012/790452> (2012).
16. Xie, F., Xiao, P., Chen, D., Xu, L. & Zhang, B. miRDeepFinder: A miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Mol. Biol.* <https://doi.org/10.1007/s11103-012-9885-2> (2012).
17. Kleinschnitz, C. *et al.* Post-stroke inhibition of induced NADPH oxidase type 4 prevents oxidative stress and neurodegeneration. *PLoS Biol.* <https://doi.org/10.1371/journal.pbio.1000479> (2010).
18. Mane, V. P., Heuer, M. A., Hillyer, P., Navarro, M. B. & Rabin, R. L. Systematic method for determining an ideal housekeeping gene for real-time PCR analysis. *J. Biomol. Tech.* **19**, 342–347 (2008).
19. Gubern, C. *et al.* Validation of housekeeping genes for quantitative real-time PCR in in-vivo and in-vitro models of cerebral ischaemia. *BMC Mol. Biol.* **10**, 57. <https://doi.org/10.1186/1471-2199-10-57> (2009).
20. Vallet, P. *et al.* Neuronal expression of the NADPH oxidase NOX4, and its regulation in mouse experimental brain ischemia. *Neuroscience* **132**, 233–238. <https://doi.org/10.1016/j.neuroscience.2004.12.038> (2005).
21. Mittal, M. *et al.* Hypoxia-dependent regulation of nonphagocytic NADPH oxidase subunit NOX4 in the pulmonary vasculature. *Circ. Res.* **101**, 258–267. <https://doi.org/10.1161/CIRCRESAHA.107.148015> (2007).
22. Casas, A. I. *et al.* NOX4-dependent neuronal autotoxicity and BBB breakdown explain the superior sensitivity of the brain to ischemic damage. *Proc. Natl. Acad. Sci. USA* **114**, 12315–12320. <https://doi.org/10.1073/pnas.1705034114> (2017).
23. Casas, A. I. *et al.* From single drug targets to synergistic network pharmacology in ischemic stroke. *Proc. Natl. Acad. Sci. USA* **116**, 7129–7136. <https://doi.org/10.1073/pnas.1820799116> (2019).
24. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210. <https://doi.org/10.1093/nar/30.1.207> (2002).
25. Rivals, I., Personnaz, L., Taing, L. & Potier, M. C. Enrichment or depletion of a GO category within a class of genes: which test?. *Bioinformatics* **23**, 401–407. <https://doi.org/10.1093/bioinformatics/btl633> (2007).
26. Silver, N., Best, S., Jiang, J. & Thein, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* **7**, 33. <https://doi.org/10.1186/1471-2199-7-33> (2006).
27. Waxman, S. & Wurmbach, E. De-regulation of common housekeeping genes in hepatocellular carcinoma. *BMC Genomics* **8**, 243. <https://doi.org/10.1186/1471-2164-8-243> (2007).
28. Iqbal, S. A., Wallach, J. D., Khoury, M. J., Schully, S. D. & Ioannidis, J. P. Reproducible research practices and transparency across the biomedical literature. *PLoS Biol.* **14**, e1002333. <https://doi.org/10.1371/journal.pbio.1002333> (2016).
29. Rasmussen, R. Quantification on the lightcycler. *Rapid Cycle Real-Time PCR* [https://doi.org/10.1007/978-3-642-59524-0\\_3](https://doi.org/10.1007/978-3-642-59524-0_3) (2001).
30. Bustin, S. A. *et al.* The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* **55**, 611–622. <https://doi.org/10.1373/clinchem.2008.112797> (2009).
31. McLoughlin, K. J., Pedrini, E., MacMahon, M., Guduric-Fuchs, J. & Medina, R. J. Selection of a real-time PCR housekeeping gene panel in human endothelial colony forming cells for cellular senescence studies. *Front. Med.* **6**, 33. <https://doi.org/10.3389/fmed.2019.00033> (2019).
32. Lachmann, A. *et al.* Massive mining of publicly available RNA-seq data from human and mouse. *Nat. Commun.* **9**, 1366. <https://doi.org/10.1038/s41467-018-03751-6> (2018).
33. Zoubarov, A. *et al.* Gemma: A resource for the reuse, sharing and meta-analysis of expression profiling data. *Bioinformatics* **28**, 2272–2273. <https://doi.org/10.1093/bioinformatics/bts430> (2012).
34. Evans, C., Hardin, J. & Stoebel, D. M. Selecting between-sample RNA-Seq normalization methods from the perspective of their assumptions. *Brief. Bioinform.* **19**, 776–792. <https://doi.org/10.1093/bib/bbx008> (2018).
35. Antoniadou, C. *et al.* Association of plasma asymmetrical dimethylarginine (ADMA) with elevated vascular superoxide production and endothelial nitric oxide synthase uncoupling: Implications for endothelial function in human atherosclerosis. *Eur. Heart J.* **30**, 1142–1150. <https://doi.org/10.1093/eurheartj/ehp061> (2009).
36. Egea, J. *et al.* Neuroprotection afforded by nicotine against oxygen and glucose deprivation in hippocampal slices is lost in alpha7 nicotinic receptor knockout mice. *Neuroscience* **145**, 866–872. <https://doi.org/10.1016/j.neuroscience.2006.12.036> (2007).
37. Buendia, I. *et al.* Neuroprotective mechanism of the novel melatonin derivative Neu-P11 in brain ischemia related models. *Neuropharmacology* **99**, 187–195. <https://doi.org/10.1016/j.neuropharm.2015.07.014> (2015).
38. Gob, E. *et al.* Blocking of plasma kallikrein ameliorates stroke by reducing thromboinflammation. *Ann. Neurol.* **77**, 784–803. <https://doi.org/10.1002/ana.24380> (2015).
39. Alemu, E. Y., Carl, J. W. Jr., Corrada Bravo, H. & Hannenhalli, S. Determinants of expression variability. *Nucleic Acids Res.* **42**, 3503–3514. <https://doi.org/10.1093/nar/gkt1364> (2014).
40. Silander, O. K. *et al.* A genome-wide analysis of promoter-mediated phenotypic noise in *Escherichia coli*. *PLoS Genet.* **8**, e1002443. <https://doi.org/10.1371/journal.pgen.1002443> (2012).
41. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).

42. Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–995. <https://doi.org/10.1093/nar/gks1193> (2013).
43. Sean, D. & Meltzer, P. S. GEOquery: A bridge between the gene expression omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846–1847. <https://doi.org/10.1093/bioinformatics/btm254> (2007).
44. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47. <https://doi.org/10.1093/nar/gkv007> (2015).
45. Matschinske, J. *et al.* The AIMe registry for artificial intelligence in biomedical research. *Nat. Methods.* <https://doi.org/10.1038/s41592-021-01241-0> (2021).

## Acknowledgements

J.B. and H.H.H.W.S. also received support from the H2020 project 777111-REPO-TRIAL. The REPO-TRIAL project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777111. This reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains. The Spanish Ministry of Economy and Competence Ref. SAF2015-63935R (to M.G.L.); the Fondo de Investigaciones Sanitarias (Instituto de Salud Carlos III/ Fondo Europeo de Desarrollo Regional) (Programa Miguel Servet: CPII19/00005, PI16/00735, PI19/00082), Fundación Mutua Madrileña (to J.E.); DFG Walter Benjamin Program (CA 2642/1-1) and the Förderprogramm der Corona-Stiftung im Stifterverband (to A.C.); JB's work was also financially supported by VILLUM Young Investigator grant nr. 13154.

## Author contributions

M.L., J.B., and H.H.H.W.S. performed the conceptualisation and supervision of the study. J.E. and M.G.L. helped with sample collection. A.I.C. performed all in vivo and in vitro studies. A.A.H., Q.M., and C.W. conducted the computational analysis. A.I.C., A.A.H., Q.M., M.L., H.H.H.W.S. wrote, reviewed, and edited the final manuscript. The results and their interpretation were discussed by all the authors. All authors read and approved the final manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-15989-8>.

**Correspondence** and requests for materials should be addressed to A.I.C. or H.H.H.W.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022