



OPEN

## Identification of functional features underlying heat stress response in Sprague–Dawley rats using mixed linear models

Krzysztof Kotlarz<sup>1</sup>, Magda Mielczarek<sup>1,2</sup>, Yachun Wang<sup>3</sup>, Jinhuan Dou<sup>4</sup>, Tomasz Suchocki<sup>1,2</sup> & Joanna Szyda<sup>1,2</sup>✉

Since global temperature is expected to rise by 2 °C in 2050 heat stress may become the most severe environmental factor. In the study, we illustrate the application of mixed linear models for the analysis of whole transcriptome expression in livers and adrenal tissues of Sprague–Dawley rats obtained by a heat stress experiment. By applying those models, we considered four sources of variation in transcript expression, comprising transcripts (1), genes (2), Gene Ontology terms (3), and Reactome pathways (4) and focussed on accounting for the similarity within each source, which was expressed as a covariance matrix. Models based on transcripts or genes levels explained a larger proportion of  $\log_2$  fold change than models fitting the functional components of Gene Ontology terms or Reactome pathways. In the liver, among the most significant genes were *PNKD* and *TRIP12*. In the adrenal tissue, one transcript of the *SUCO* gene was expressed more strongly in the control group than in the heat-stress group. *PLEC* had two transcripts, which were significantly overexpressed in the heat-stress group. *PER3* was significant only on gene level. Moving to the functional scale, five Gene Ontologies and one Reactome pathway were significant in the liver. They can be grouped into ontologies related to DNA repair, histone ubiquitination, the regulation of embryonic development and cytoplasmic translation. Linear mixed models are valuable tools for the analysis of high-throughput biological data. Their main advantages are the possibility to incorporate information on covariance between observations and circumventing the problem of multiple testing.

The most general formulation of heat stress is defined as an increase in ambient temperature above the threshold beyond which body temperature cannot be maintained on the physiologically optimal level<sup>1,2</sup>. Heat stress is among the best characterized environmental stressors and has the most severe detrimental effects. Individuals cope with heat stress by increasing body temperature, reducing feed intake as well as changing of physiological state, which is referred to as physiological adaptive responses<sup>3</sup>. Therefore, heat stress can lead to economic losses due to lower productivity and reproductivity, as well as increased health burden<sup>4</sup>. Unfortunately, global temperature is expected to rise by 2 °C in 2050<sup>5</sup> and furthermore, the average temperature in 2100 may rise by up to 4.8 °C<sup>6</sup>. As a consequence, the maximum daily temperature will be higher, so the intensity and duration of heat stress will be more severe and prolonged, compared to 2021. Therefore, heat stress may become the most severe environmental factor.

Studies have reported that heat stress significantly altered physiological, biochemical, metabolic, and cellular responses in mice and livestock models<sup>3,7,8</sup>. However, animals have developed protective measures against the challenges of heat stress manifested by altered transcript and gene expression levels<sup>9–11</sup>. With the development of next generation sequencing technologies, including RNA-sequencing (RNA-seq), hundreds or thousands of genes involved in heat stress response have been identified<sup>12,13</sup>. Consequently, the main goal of the research was to illustrate how linear mixed models can be used to perform a biologically driven analysis of changes in transcript

<sup>1</sup>Biostatistics Group, Department of Genetics, Wrocław University of Environmental and Life Sciences, Kozuchowska 7, 51-631 Wrocław, Poland. <sup>2</sup>National Research Institute of Animal Production, Krakowska 1, 32-083 Balice, Poland. <sup>3</sup>Key Laboratory of Animal Genetics, Breeding and Reproduction, MARA, National Engineering Laboratory for Animal Breeding, College of Animal Science and Technology, China Agricultural University, Beijing 100193, China. <sup>4</sup>College of Animal Science and Technology, Beijing University of Agriculture, Beijing 102206, China. ✉email: joanna.szyda@upwr.edu.pl

and gene expression levels between the control and the heat-stress groups. In contrast to the vast majority of approaches, when particular transcripts are considered as independent, we incorporated biological information on the correlation between them directly into the statistical model. Furthermore, we also considered higher-order units, expressed by Gene Ontology terms and Reactome pathways, since these are the actual functional components of the physiological response towards heat stress.

The incorporation of functional units into the analysis of RNA-seq data is a known idea. It has been applied either through enrichment analysis as e.g. by Dou et al.<sup>14</sup> to the data set analysed in our study or through the Ingenuity Pathway Analysis as e.g. by Lan et al.<sup>15</sup> to gene expression under heat stress in poultry. Nevertheless, the major difference between our approach and the aforementioned studies is that the former conducts a two-step analysis, by first identifying significantly differentially expressed transcripts and genes by considering them independent (step 1) and then further analysing the selected, significant transcripts and genes in terms of the correlation between them or the enrichment of ontologies or pathways in this significant gene set (step 2). The approach proposed in our study attempted to combine the process within one-step by including the correlation of the functional information directly into the model fitted to the whole (i.e. not only to significant) expression data and is statistically similar to the linear mixed model fitted to gene expression data by Wang et al.<sup>16</sup>, albeit with a different approach to the modelling of the covariance of random effects.

In the current study, we illustrate the application of the mixed models using the whole genome transcript expression data of Sprague–Dawley rats obtained from a heat stress experiment<sup>14</sup>.

## Methods

**Experimental animals.** The data set underlying the analysis is a subset of the material used by Dou et al.<sup>14</sup>. In brief, the analysed individuals comprised eight weeks old female, specific-pathogen-free, Sprague–Dawley rats. Prior to the experiment, all rats were housed in a laboratory at  $22 \pm 1$  °C, 50% relative humidity with 12 h reverse light/dark cycle with feed and water provided ad libitum. After one week, five rats from the heat-stressed group were exposed to 42 °C for 120 min, while five rats from the control group were housed at an initial temperature of 22 °C. After the completion of the experiment, the animals were euthanised and samples of liver and adrenal gland tissues were used as a source of total RNA.

**Bioinformatic analysis.** Illumina HiSeq2000 was used to sequence 150 long reads in the paired-end (PE) data mode. The total number of read pairs per sample was from 51,706,978 to 97,059,004. The detailed description of the RNA isolation and sequencing was also provided by Dou et al.<sup>14</sup>. The bioinformatic pipeline consisted of the following steps: quality control of raw reads, editing of raw reads based on their quality, and quantification of transcripts' expression. In particular, the quality of raw reads was assessed by applying the FastQC software<sup>17</sup>. Then, reads were processed by the Trimmomatic software<sup>18</sup>, which removed adapter sequences, trimmed reads with an average sequencing quality of 4 consecutive reads below 20 (*SLIDINGWINDOW:4:20*), and removed reads shorter than 60 bp (*MINLEN:60*). Next, the Salmon software<sup>19</sup> was used to quantify the abundances of transcripts. This software implements the pseudoalignment process—an approach allowing for rapid identification of the compatibility of reads with transcripts, without the need of a computationally intensive whole genome alignment. In the last step,  $\log_2$  fold changes in transcript expression levels between the control and heat stressed groups were calculated using DESeq2<sup>20</sup>.

**Statistical modelling of expression data.** The  $\log_2$  fold changes ( $\log_2FC$ ) calculated based on the transcript expression levels pooled over the control and heat-stressed animals respectively, were analysed in four mixed linear models.

The *transcript-based model* (M1) is given by:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{Z}_{M1}\mathbf{t} + \mathbf{e}_{M1}, \quad (1)$$

where  $\mathbf{y}$  is the vector of  $\log_2FC$  of transcript expression,  $\boldsymbol{\mu}$  represents the general mean,  $\mathbf{t}$  is the random transcript effect with a predisposed normal distribution defined by  $N(0, \mathbf{V}_{M1}\sigma_t^2)$ ,  $\mathbf{e}_{M1}$  is a vector of residuals distributed as  $N(0, \mathbf{I}\sigma_{e_{M1}}^2)$ ,  $\mathbf{Z}_{M1}$  is an incidence matrix for  $\mathbf{t}$ . In this model, the similarity between transcripts  $i$  and  $j$ , was introduced into the model by incorporating a nondiagonal transcript covariance matrix  $\mathbf{V}_{M1}$ . The covariance between transcripts was expressed by the Jaccard similarity coefficient:  $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$ , which in the case of transcripts was calculated based on the similarity in their exon composition:

$$J(i, j) = \frac{a}{N},$$

where  $a$  represents the number of exons common between transcripts  $i$  and  $j$ , while  $N$  represents the total number of exons of a given gene. Independence was assumed between genes, so the resulting matrix had a block diagonal structure. Transcript information was obtained from the Ensembl database, release100<sup>21</sup> and Jaccard coefficients were calculated using the ADE-4 package<sup>22</sup>.

The *gene-based model* (M2) was applied to the same dependent variable ( $\mathbf{y}$ ) as M1 and is given by:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{Z}_{M2}\mathbf{g} + \mathbf{e}_{M2}, \quad (2)$$

where  $\mathbf{g}$  is the random gene effect with a preimposed normal distribution defined by  $N(0, \mathbf{V}_{M2}\sigma_g^2)$ ,  $\mathbf{e}_{M2}$  is a vector of random residuals distributed as  $N(0, \mathbf{I}\sigma_{e_{M2}}^2)$ ,  $\mathbf{Z}_{M2}$  is an incidence matrix for  $\mathbf{g}$ . No covariance between genes was assumed so  $\mathbf{V}_{M2}$  was diagonal.

Furthermore, the *Gene Ontology-based model* (M3):

Tissue	$\frac{\sigma_t^2}{\sigma_y^2}$ (M1) (%)	$\frac{\sigma_g^2}{\sigma_y^2}$ (M2) (%)	$\frac{\sigma_{go}^2}{\sigma_y^2}$ (M3) (%)	$\frac{\sigma_r^2}{\sigma_y^2}$ (M4) (%)
Liver	12.43	10.12	1.48	1.45
Adrenal	18.32	16.45	1.44	1.93

**Table 1.** Variance components estimated by the mixed models (M1–M4) expressed as the percentage of the total variance of  $y$ .  $\sigma_t^2$  the transcript variance,  $\sigma_g^2$  the gene variance,  $\sigma_{go}^2$  the GO term variance,  $\sigma_r^2$  the Reactome pathway variance,  $\sigma_y^2$  the variance of  $\log_2FC$ .

$$y = \mu + Z_{M3}go + e_{M3}, \quad (3)$$

where  $go$  is the random effect of Gene Ontology (GO; geneontology.org) terms assigned to transcripts whose  $\log_2FC$  of transcript expressions are contained in  $y$ . It was assumed that  $go$  follows the normal distribution  $N(0, V_{M3}\sigma_{go}^2)$ ,  $e_{M3}$  is a vector of residuals distributed as  $N(0, I\sigma_{e_{M3}}^2)$ , and  $Z_{M3}$  is an incidence matrix for the  $go$  terms.  $V_{M3}$  describes the covariance between GO terms expressed, as above, by the Jaccard coefficients quantifying genes overlapping between two given GO terms. Each transcript was assigned GO term(s) from the biological process ontology, considering the ontologies from the 2nd hierarchy level.

Finally, a model incorporating the effects of Reactome pathways (M4) was fitted to transcript  $\log_2FC$ :

$$y = \mu + Z_{M4}r + e_{M4}, \quad (4)$$

in this model,  $r$  represents the random effect of Reactome pathways (reactome.org) corresponding to transcripts. The normal distribution pre-imposed on  $r$  is given by  $N(0, V_{M4}\sigma_r^2)$ ,  $e_{M4}$  is a vector of residuals distributed as  $N(0, I\sigma_{e_{M4}}^2)$ , and  $Z_{M4}$  is an incidence matrix assigning Reactome pathways to transcripts. The covariance matrix between Reactome pathways ( $V_{M4}$ ) was also expressed by the Jaccard coefficients quantifying genes overlapping between the two given pathways.

**Estimation of variance components and significance testing.** The expectation–maximization algorithm<sup>23</sup> was applied for the estimation of variance components underlying the four above models (i.e.  $\sigma_t^2$ ,  $\sigma_g^2$ ,  $\sigma_{go}^2$ ,  $\sigma_r^2$ ,  $\sigma_{e_{M1}}^2$ ,  $\sigma_{e_{M2}}^2$ ,  $\sigma_{e_{M3}}^2$  and  $\sigma_{e_{M4}}^2$ ) and the mixed model equations<sup>24</sup> were used to obtain solutions of  $\mu$ ,  $t$ ,  $g$ ,  $go$  and  $r$ :

$$\begin{bmatrix} \hat{\mu} \\ \hat{x} \end{bmatrix} = \begin{bmatrix} \mathbf{1}^T R^{-1} \mathbf{1} & \mathbf{1}^T R^{-1} Z_x \\ Z_x^T R^{-1} \mathbf{1} & Z_x^T R^{-1} Z_x + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}^T R^{-1} y \\ Z_x^T R^{-1} y \end{bmatrix}, \text{ where } R = I\hat{\sigma}_{e_{Mi}}^2 \text{ and } G = V_x \hat{\sigma}_x^2,$$

with  $x$  representing  $t$ ,  $g$ ,  $go$  or  $r$  depending on the model, and  $i \in \{M1, M2, M3, M4\}$  being a model indicator. Each element of the solution vectors  $\hat{t}$ ,  $\hat{g}$ ,  $\hat{go}$  and  $\hat{r}$  was transformed to the standard normal distribution and tested for significance by assessing the probability of obtaining a more extreme value based on the standard normal density function.

**Enhancement of the computational efficiency of the estimation.** In order to maximise the computational performance of the estimation of model parameters and its variance components, a custom-written Python program implementing the Numba library<sup>25</sup> was used. Numba compiles a subset of native Python and NumPy code into the machine code. Since all calculations were carried out on a multicore server, the Numba library was also used to parallelize the code, what further improved the computing time compared to a native Python application.

**Ethics approval.** Not applicable, only in silico data was processed in this study.

## Results

Obviously models M1 and M2, fitting thousands of components, i.e. transcripts (M1) or genes (M2) explain a much larger proportion of the observed variability of  $\log_2FC$  than the M3 and M4 fitting the functional components, i.e. GO terms (M3) or Reactome pathways (M4). Still, by taking into account the correlation between transcripts expressed by their exon composition, we explain 12.43% and 18.32% of the total variability of  $\log_2FC$  observed in liver and adrenal tissues, respectively. A similar picture arises when the assumed source of variability of  $\log_2FC$  is considered on a gene level, resulting in 10.12% and 16.45% variance components in liver and adrenal tissues respectively—somewhat lower than for correlated transcripts in M1. By further shrinking the functional units to GO terms and Reactome pathway units we explain “only” 1.48% (GO) and 1.45% (Reactome) of the total variance for liver, as well as 1.44% (GO) and 1.93% (Reactome) for the adrenal tissue (Table 1).

Considering the transcript and gene levels, in the liver, M1 and M2 point at the significant effects of the *PNKD* and *TRIP12* genes. *PNKD* exhibits a significant effect of a transcript ENSRNOT00000046229, which is approximately 2<sup>23</sup> times higher expressed in the control than in the heat-stress group. However, three of the five transcripts of the gene show higher expression in the heat-stress group, reaching 2<sup>5</sup> higher expression of ENSRNOT00000089580. *TRIP12* synthesises one particular transcript (ENSRNOT00000022822) which is 2<sup>22</sup> times higher expressed in the control than in the heat-stress group. Different genes were significantly differentially expressed in the adrenal tissue. One transcript of the *SUCO* gene (ENSRNOG00000026542) is 2<sup>23</sup> times higher

Tissue	Model	ID name	Effect	P
Liver	Transcript (M1)	ENSRNOT00000074131 <i>PNKD</i>	2.76	$4.7 \times 10^{-10}$
	Transcript (M1)	ENSRNOT00000093245 <i>TRIP12</i>	2.15	$1.2 \times 10^{-6}$
	Transcript (M1)	ENSRNOT00000093735 <i>TRIP12</i>	1.67	0.00016
	Transcript (M1)	ENSRNOT00000022822 <i>TRIP12</i>	1.65	0.00021
	Transcript (M1)	ENSRNOT00000079452 <i>TRIP12</i>	1.63	0.00023
	Gene (M2)	ENSRNOG00000016963 <i>TRIP12</i>	1.66	$3.5 \times 10^{-5}$
	Gene (M2)	ENSRNOG00000014806 <i>PNKD</i>	1.12	0.00520
	Gene ontology (M3)	GO:1901315 Negative regulation of histone H2A K63-linked ubiquitination	0.52	0.00055
	Gene ontology (M3)	GO:2000780 Negative regulation of double-strand break repair	0.48	0.00130
	Gene ontology (M3)	GO:2000779 Regulation of double-strand break repair	0.37	0.01300
	Gene ontology (M3)	GO:0045995 Regulation of embryonic development	0.30	0.04200
	Gene ontology (M3)	GO:0002181 Cytoplasmic translation	0.29	0.05200
	Reactome pathway (M4)	R-RNO-6791226 Major pathway of rRNA processing in the nucleolus and cytosol	0.29	0.05500
Adrenal	Transcript (M1)	ENSRNOT00000075998 <i>SUCO</i>	3.48	$9.2 \times 10^{-11}$
	Transcript (M1)	ENSRNOT00000084058 <i>SUCO</i>	2.18	$4.9 \times 10^{-5}$
	Transcript (M1)	ENSRNOT00000082271 <i>PLEC</i>	2.10	$8.9 \times 10^{-10}$
	Transcript (M1)	ENSRNOT00000075936 <i>SUCO</i>	1.95	0.00029
	Transcript (M1)	ENSRNOT00000088945 <i>PLEC</i>	1.70	0.00160
	Gene (M2)	ENSRNOG00000026542 <i>SUCO</i>	2.26	$8.4 \times 10^{-6}$
	Gene (M2)	ENSRNOG00000018413 <i>PER3</i>	1.63	0.00130
	Reactome pathway (M4)	R-RNO-212436 Generic Transcription Pathway	0.49	0.00500

**Table 2.** Top 5 significant transcripts from M1, top 2 significant genes from M2, significant Gene Ontology terms from M3, and significant Reactome pathways from M4. *ID* the Ensembl transcript ID, *Ensembl gene ID* Gene Ontology database ID, or Reactome ID, depending on the effect considered, *Effect* represents the estimate from the corresponding mixed model M1–M4, *P* P value corresponding to the normal probability density function with mean zero and the variance estimated by the corresponding model.

expressed in the control than in the heat-stress group, while the three other transcripts synthesised from this gene have an opposite effect, i.e. show higher expression in heat-stressed individuals. *PLEC* has two transcripts which are significantly overexpressed in the heat-stress group. *PER3* was significant only on the gene level (M2) with all of its transcripts being under expressed during heat stress (Table 2).

While moving to a functional level, we did not identify any significant GO terms differentially expressed in the adrenal tissue and a single Generic Transcription Reactome Pathway (R-RNO-212436). However, in the liver, five ontologies and one Reactome pathway were significant. Functionally, they can be grouped into ontologies related to DNA repair through the regulation of double-strand break repair (GO:2000780 and GO:2000779) and histone ubiquitination (GO:1901315), to the regulation of embryonic development (GO:0045995) and to cytoplasmic translation (GO:0002181, R-RNO-6791226).

## Discussion

Various scopes of information considered in our study explain different amounts of the observed variation of transcript differential expressions across the genome (Table 1). The difference in the estimated variance components related to transcripts, genes, GO terms and Reactome pathways associates with the number of levels of those independent variables. Since we are not aware of the corresponding estimates reported by other studies, a discussion with the literature is not possible. Therefore, at this stage, we can only hypothesise that on the one

hand, the differences in variance components can be due to the technical nature of the model, since many more transcript effects were estimated as Reactome pathway effects. On the other hand, those effects were modelled as random and thus their variation was constrained by the predefined co-variance structure so that the differences in variance components express the biological phenomenon—Reactome pathways aggregate effects of many transcript/genes.

By considering the significant transcripts, genes and pathways identified in our study, a consistent picture emerges that DNA stability and its repair mechanisms are affected by heat stress. This phenomenon has already been reported by other authors, e.g.<sup>26</sup> and was reviewed by Kantidze et al.<sup>27</sup>. On a gene level, *TRIP12* was reported as encoding a protein which is associated with chromatin and plays a role in the maintenance of genome integrity<sup>28</sup> and *PER3* influences DNA damage repair by correlating with checkpoint kinase 2 gene<sup>29</sup>. Mutations in the *SUCO* gene induce hypoglycaemia [[www.ensembl.org](http://www.ensembl.org)] in humans, which is the cause of hypothermia in diabetic patients<sup>30</sup>. *PER3* was differentially expressed in broiler chicken maintained under control and heat stress conditions<sup>31</sup>. In general, it has been proven that *PER3* is associated with behavioural differences towards stress and was influenced by stress and ethanol treatment in BXD strain mice<sup>32</sup>.

On the functional level, we observed a significantly under expression of genes with ontologies related to double-strand break repair (GO:2000780 and GO:2000779) and histone ubiquitination. Furthermore, outside of the nucleus, a significant effect of the cytoplasmic translation ontology (GO:0002181) and rRNA processing in the nucleolus and cytosol pathway (R-RNO-6791226) can be linked to the phenomenon of the aggregation of proteins in the cytoplasm of yeast cells subjected to heat stress, which consequence was impaired cytoplasmic translation<sup>33</sup>.

Noteworthy, our list of significant transcripts, genes, GO terms, and Reactome pathways did not overlap with the results reported for the same material by Dou et al.<sup>12</sup>. Only the *PER3* gene was significantly associated with differential expression in adrenal tissue in both studies. The most emerging difference concerned the number of significant effects reported. While Dou et al.<sup>12</sup> estimated 3909 and 4953 significantly differentially expressed genes for liver and adrenal tissues respectively, our study pointed at only two significant genes for each tissue. Similarly, the number of significant GO terms was 193 for liver and 79 for adrenal tissue in Dou et al.<sup>12</sup>, while only five and zero in our results. The observed differences are caused by the following factors: different raw reads editing criteria, different approaches to estimate transcript expression, and different statistical modelling of expression data. Dou et al.<sup>12</sup> used the Cufflinks software<sup>34</sup> for bioinformatic processing of the expression data, which implements Tophat2<sup>35</sup> for the alignment of reads to the reference genome and a single gene hypothesis test with multiple testing correction of P values via the FDR for the assessment of differential expression. In our analysis, an alignment-free approach and a random effect model incorporating all transcripts/genes simultaneously were applied. Already<sup>36</sup> pointed at differences in the statistical inference based on fixed and random effect models, indicating that the former tend to underestimate, and the latter—overestimate a residual variance. As a consequence, a type I-error is often elevated in fixed-effect models, which was analytically demonstrated by<sup>37</sup> in the context of meta-analysis. Moreover, significant differences in the quantification of expression between alignment-based and alignment-free approaches were recently demonstrated in our unpublished analysis of *Sus scrofa* RNA-seq data by Hoffman et al.

## Conclusions

Mixed models, i.e. statistical models fitting random effects, are a valuable tool for the analysis of high-throughput biological data. Their major advantages comprise: (1) the possibility to incorporate information on covariance between observations, which is often neglected while applying simple, fixed effect models, and (2) circumventing the problem of multiple testing, by simultaneous fitting all effects. We see the major limitation of the proposed approach in the varying quality of functional annotation of genomes available for different species. While genomes of humans and experimental species, such as rats, are very well functionally annotated with the most of transcripts/genes assigned to GO terms and metabolic pathways, less well-studied species have less complete annotation, which would enforce the incorporation of phantom GO and phantom pathway effects (similarly to phantom parent groups in livestock genetic evaluation).

From the biological perspective, *PER3* and *SUCO* genes as well as DNA repair and translation were indicated as factors playing a significant role in heat stress response.

## Data availability

The datasets generated and/or analysed during the current study are not publicly available due to institutional constraints, but are available from the Yachun Wang on reasonable request.

## Code availability

Python code is available in principle upon request from the corresponding author.

Received: 8 September 2021; Accepted: 25 April 2022

Published online: 10 May 2022

## References

1. Berman, A. et al. Upper critical temperatures and forced ventilation effects for high-yielding dairy cows in a subtropical climate. *J. Dairy Sci.* **68**(6), 1488–1495. [https://doi.org/10.3168/jds.S0022-0302\(85\)80987-5](https://doi.org/10.3168/jds.S0022-0302(85)80987-5) (1985).
2. Kadzere, C. T., Murphy, M. R., Silanikove, N. & Maltz, E. Heat stress in lactating dairy cows: A review. *Livest Prod. Sci.* **77**(1), 59–91. [https://doi.org/10.1016/S0301-6226\(01\)00330-X](https://doi.org/10.1016/S0301-6226(01)00330-X) (2002).
3. Gonzalez-Rivas, P. A. et al. Effects of heat stress on animal physiology, metabolism, and meat quality: A review. *Meat Sci.* <https://doi.org/10.1016/j.meatsci.2019.108025> (2020).

4. St-Pierre, N. R., Cobanov, B. & Schnitkey, G. Economic losses from heat stress by US livestock industries1. *J. Dairy Sci.* [https://doi.org/10.3168/jds.S0022-0302\(03\)74040-5](https://doi.org/10.3168/jds.S0022-0302(03)74040-5) (2003).
5. Trnka, M., Olesen, J. E., Kersebaum, K. C., et al. *Agroclimatic Conditions in Europe under Climate Change*. Vol 17; 2011. <https://doi.org/10.1111/j.1365-2486.2011.02396.x>.
6. Barros, V. R., Field, C. B., Dokken, D. J., et al. *Climate Change 2014 Impacts, Adaptation, and Vulnerability Part B: Regional Aspects: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (2014).
7. Hu, Y. et al. Effects of chronic heat stress on immune responses of the foot-and-mouth disease DNA vaccination. *DNA Cell Biol.* **26**(8), 619–626. <https://doi.org/10.1089/dna.2007.0581> (2007).
8. Lu, Q., Wen, J. & Zhang, H. Effect of chronic heat exposure on fat deposition and meat quality in two genetic types of chicken. *Poult. Sci.* **86**(6), 1059–1064. <https://doi.org/10.1093/ps/86.6.1059> (2007).
9. Sonna, L. A., Fujita, J., Gaffin, S. L. & Lilly, C. M. Invited review: Effects of heat and cold stress on mammalian gene expression. *J. Appl. Physiol.* **92**(4), 1725–1742. <https://doi.org/10.1152/jappphysiol.01143.2001> (2002).
10. He, Y., Maltecca, C., Tiezzi, F., Soto, E. L. & Flowers, W. L. Transcriptome analysis identifies genes and co-expression networks underlying heat tolerance in pigs. *BMC Genet.* **21**, 1. <https://doi.org/10.1186/s12863-020-00852-4> (2020).
11. Stallings, J. D. et al. Patterns of gene expression associated with recovery and injury in heat-stressed rats. *BMC Genom.* **15**(1), 1058. <https://doi.org/10.1186/1471-2164-15-1058> (2014).
12. Dou, J. et al. Comprehensive RNA-Seq profiling reveals temporal and tissue-specific changes in gene expression in Sprague-Dawley rats as response to heat stress challenges. *Front Genet.* <https://doi.org/10.3389/fgene.2021.651979> (2021).
13. Wang, L. C. et al. Transcriptome profiling of the fifth-stage larvae of *Angiostrongylus cantonensis* by next-generation sequencing. *Parasitol. Res.* **112**(9), 3193–3202. <https://doi.org/10.1007/s00436-013-3495-z> (2013).
14. Dou, J. et al. Heat stress impairs the physiological responses and regulates genes coding for extracellular exosomal proteins in rat. *Genes (Basel)* **11**(3), 306. <https://doi.org/10.3390/genes11030306> (2020).
15. Lan, X., Hsieh, J. C. F., Schmidt, C. J., Zhu, Q. & Lamont, S. J. Liver transcriptome response to hyperthermic stress in three distinct chicken lines. *BMC Genom.* **17**, 1. <https://doi.org/10.1186/s12864-016-3291-0> (2016).
16. Wang, T. & Zeng, Z. Contribution of genetic effects to genetic variance components with epistasis and linkage disequilibrium. *BMC Genet.* <https://doi.org/10.1186/1471-2156-10-52> (2009).
17. Andrews S, others. FastQC: A quality control tool for high throughput sequence data. 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
18. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> (2014).
19. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 4. <https://doi.org/10.1038/nmeth.4197> (2017).
20. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 12. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
21. Yates, A. D. et al. Ensembl 2020. *Nucleic Acids Res.* **48**(D1), D682–D688. <https://doi.org/10.1093/nar/gkz966> (2020).
22. Thioulouse, J., Chessel, D., Dolédec, S. & Olivier, J. M. ADE-4: A multivariate analysis and graphical display software. *Stat. Comput.* **7**(1), 75–83. <https://doi.org/10.1023/A:1018513530268> (1997).
23. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **39**(1), 1–22. <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x> (1977).
24. Henderson, C. *Applications of Linear Models in Animal Breeding* (CABI, 1984).
25. Lam, S. K., Pitrou, A., & Seibert, S. Numba. In: *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC-LLVM '15*. ACM Press; 2015:1–6. <https://doi.org/10.1145/2833157.2833162>.
26. Takahashi, A. et al. Evidence for the involvement of double-strand breaks in heat-induced cell killing. *Cancer Res.* **64**(24), 8839–8845. <https://doi.org/10.1158/0008-5472.CAN-04-1876> (2004).
27. Kantidze, O. L., Velichko, A. K., Luzhin, A. V. & Razin, S. V. Heat stress-induced DNA damage. *Acta Nat.* **8**(2), 75–78. <https://doi.org/10.32607/20758251-2016-8-2-75-78> (2016).
28. Larriue, D. et al. The E3 ubiquitin ligase TRIP12 participates in cell cycle progression and chromosome stability. *Sci. Rep.* **10**, 1. <https://doi.org/10.1038/s41598-020-57762-9> (2020).
29. Im, J. S., Jung, B. H., Kim, S. E., Lee, K. H. & Lee, J. K. Per3, a circadian gene, is required for Chk2 activation in human cells. *FEBS Lett.* **584**(23), 4731–4734. <https://doi.org/10.1016/j.febslet.2010.11.003> (2010).
30. Tran, C. et al. Hypothermia is a frequent sign of severe hypoglycaemia in patients with diabetes. *Diabetes Metab.* **38**(4), 370–372. <https://doi.org/10.1016/j.diabet.2012.03.005> (2012).
31. Jastrebski, S. F., Lamont, S. J. & Schmidt, C. J. Chicken hepatic response to chronic heat stress using integrated transcriptome and metabolome analysis. *PLoS One* **12**(7), e0181900. <https://doi.org/10.1371/journal.pone.0181900> (2017).
32. Wang, X. et al. A promoter polymorphism in the Per3 gene is associated with alcohol and stress response. *Transl. Psychiatry* **2**, 1. <https://doi.org/10.1038/tp.2011.71> (2012).
33. Gallardo, P., Real-Calderón, P., Flor-Parra, I., Salas-Pino, S. & Daga, R. R. Acute heat stress leads to reversible aggregation of nuclear proteins into nucleolar rings in fission yeast. *Cell Rep.* **33**(6), 108377. <https://doi.org/10.1016/j.celrep.2020.108377> (2020).
34. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**(5), 511–515. <https://doi.org/10.1038/nbt.1621> (2010).
35. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**(9), 1105–1111. <https://doi.org/10.1093/bioinformatics/btp120> (2009).
36. Overton, R. C. A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychol. Methods* **3**(3), 354–379. <https://doi.org/10.1037/1082-989X.3.3.354> (1998).
37. Hunter, J. E. & Schmidt, F. L. Fixed effects vs random effects meta-analysis models: Implications for cumulative research knowledge. *Int. J. Sel. Assess.* **8**(4), 275–292. <https://doi.org/10.1111/1468-2389.00156> (2000).

## Acknowledgements

Computations were carried at the Wrocław Centre for Networking and Supercomputing (mixed models) and the Poznań Supercomputing and Networking Center (bioinformatic pipelines).

## Author contributions

Y.W., and J.S., designed the research and interpreted the data. K.K., M.M., J.D., and T.S., and performed data analysis. J.S. drafted the manuscript. All authors contributed to the writing of the manuscript, reviewed and approved its final version.

## Funding

The research was funded by the Polish National Science Foundation (NCN) under grant number 2019/35/O/NZ9/00237 and supported by the tutoring programme of the Wroclaw University of Environmental and Life Sciences. The APC is financed by the Wroclaw University of Environmental and Life Sciences.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to J.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