1   **Understanding the Liver Under Heat Stress With Statistical Learning: A**

2   **Multiomics Computational Approach**

3

4   Allen Hubbard[1], Xiaoke Zhang[2], Sara Jastrebski[3], Abhi Singh[4] and Carl Schmidt[3]

5   [1]University of Delaware, Bioinformatics and Systems Biology

6   [2]George Washington University, Statistics

7   [3]University of Delaware, Animal and Food Sciences

8   [4]University of Delaware, Electrical Engineering and Computer Science

9

10   * Allen H. Hubbard
11   E-mail: allenhub@udel.edu
12

13   Xiaoke Zhang
14   E-mail: xkzhang@email.gwu.edu
15
16   Sara Jastrebski
17   E-mail: sjas@udel.edu
18

19   Abhi Singh
20   E-mail: absingh@udel.edu
21

22   Carl J Schmidt
23   E-mail: schmidtc@udel.edu
24

Abbreviation	Term
RNA-seq	Ribonucleic Acid Sequencing
GWA	Genome Wide Association
SNP	Single Nucleotide Polymorphism
PCA	Principal Component Analysis
GTEX	Genotype Tissue Expression
K1	rate constant for forward reaction
K2	rate constant for reverse reaction
F6P	fructose-6-phosphate
G3P	glycerol-3-phosphate
S100Z	S100 Calcium Binding Protein Z
FBP2	Fructose-Bisphosphatase-2
NADKD1	NAD Kinase, mitochondrial
NAD	Nicotinamide Adenine Dinucleotide
NADP	Nicotinamide Adenine Dinucleotide Phoshpate
NADPH	Nicotinamide Adenine Dinucleotide Phosphate, Reduced

31

32

Abstract

34   We present results from a pipeline developed to integrate multi-omics data

35 in order to explore the heat stress response in the liver of the modern broiler

36 chicken. Heat stress is a significant cause of productivity loss in the poultry

37 industry, both in terms of increased livestock morbidity and its negative influence

38 on average feed efficiency. This study focuses on the liver because it is an important

39 regulator of metabolism, controlling many of the physiological processes impacted

40 by prolonged heat stress. Using statistical learning methods, we identify genes and

41 metabolites that may regulate the heat stress response in the liver and adaptations

42 required to acclimate to prolonged heat stress. Our findings provide more detailed

43   context for genomic studies and generates hypotheses about dietary interventions

44   that can mitigate the negative influence of heat stress on the poultry industry.

Background

46   Obtaining biological insight from large-scale transcriptome data is

47   challenging due to biological and technical variance. Careful experimental design

48   can limit unwanted noise. However, when properly harnessed, heterogeneity can be

49   used to detect biological signals that elude traditional enrichment analysis. For

50   example, biological variation relating to a treatment response depends on many

51   variables that are not easily controlled such as allelic or physiological variants. This

52   fact can be informative because many compounds involved in the same process will

53   have similar patterns of heterogeneity. This can be used to identify relationships

54   between elements of the same pathway, even when their scales of expression and

55   variance differ considerably, by relying on statistical learning strategies. This

56   approach allows the combination of transcriptome and metabolome data to gain a

57   more comprehensive biological understanding of a system. This is particularly

58   helpful in identifying significant features from the large, complex datasets now

59   common in multi-omics studies.

60   The modern broiler chicken is a fundamental source of poultry meat. It has

61   been under strong artificial selection during the past several decades for increased

62   breast muscle yield (Tallentire *et al.*, 2016). This is thought to be at the expense of

63   other systems, resulting in decreased heat tolerance and increased mortality during

64   heat stress. The relationship between the altered physiology of the broiler and

65   susceptibility to heat stress is not fully understood, however. It is believed to

66 involve altered appetite and preferential routing of resources to muscles tissue.

67 Such changes are systemic, influenced by both behavior and metabolism.

68   One organ capable of exerting strong influence on both bird growth and

69 thermoregulation is the liver. This organ has recently proved as a subject for

70 studies that leverage multiomics approaches including transcriptomics and

71 metabolomics. Such work has shed light on differentially regulated genes and

72 metabolites. However, a systems level understanding in which fluxes in metabolites

73 are related to gene expression, are lacking. This is partly because computational

74 approaches exploring the totality of a biological response including gene expression

75 and metabolite production is lacking. We combine RNA-seq (Ribonucleic Acid

76 Sequencing) expression and metabolites from the liver to identify genes and

77 compounds that function as biomolecules associated with heat stress. While

78 metabolomics data identifies changes in biologically active compounds, RNA-Seq

79 data identifies genes regulate metabolic changes. We offer a geometric

80 interpretation for our statistical procedures, describing how they recapitulate novel

81 biology.

82   Our analysis applies statistical learning approaches on metabolite and gene

83 expression data restricts transcriptome analysis to a core module of liver enriched

84 genes. These are determined by a definition we propose that proves more stringent

85 than other types of relative expression analysis. Sub-setting in this fashion isolates

86 tissue-enriched genes that reflect unique biology specific to the liver in a tissue

87 diverse dataset. The approach of selecting tissue enriched genes and focusing on

88 patterns of heterogeneity within such a subset provides a framework to integrate

89   metabolite and transcriptome data. This approach of combining data from different

90   high-throughput technologies makes it possible to identify important features of the

91   high dimensional dataset.

92   Finally, extending the work of earlier GWA (genome wide association)

93   studies that sought to model ratios of metabolites as functions of SNP's, (single

94   nucleotide polymorphisms) we model metabolite ratios in terms of other

95   metabolites. The original purpose of these GWA metabolite studies was to detect

96   genetic basis of metabolic changes (Gieger *et al.*, 2008). However, modeling ratios

97   as function of metabolites allows detection of metabolic forks, or small network

98   motifs where precursors are selectively routed to different metabolic fates under

99   heat stress. The compounds used to compose triplets representing possible

100   metabolic forks are selected from hypotheses developed through the combined k-

101   means, random forest and PCA (principal component analysis) pipeline. A triplet is

102   defined as a function of the form cor(A, (B/C)) where A, B and C are any

103   combination of metabolites. Candidates for A, B and C are chosen from amino acids

104   known to be catabolized under heat stress (Jastrebski *et al.*, 2017) and sugar and

105   fat molecules that may incorporate these molecules prioritized by our pipeline.

106   The combination of RNA-Seq with metabolite data identifies novel shifts in

107   gene regulation that reflect pathway changes influencing metabolite levels.

108   Our combined informatics strategy identifies elements under genetic control and

109   which could be targets for selective breeding. Additionally, the identification of heat

110   stress responsive metabolites produces candidates for feed supplementation

111   studies.

112

Methods

113

114   The heat stress response is multi-tiered and involves input from multiple

115   tissues. At the cellular level, the heat stress response unfolds across an intricate

116   program of organelle specific changes. Which changes are causal, and which merely

117   correlative with underlying signal or sensing pathways, thus becomes a complex

118   question. However, the variability associated with most basal regulators of the heat

119   stress response should be most closely related to the variation in the heat stress

120   response. By the transitive nature of biological communication, the introduction of

121   noise into the signal diminishes the capacity of downstream molecules, which

122   correlate with, but do not cause the heat stress response, to discriminate between

123   treatment and control samples. From this perspective, the problem of identifying

124   causal molecules from expression profile is well posed as a statistical learning

125   problem that can be addressed through random forests. Random forests can rank

126   candidates on their ability to correctly identify the class of samples as assigned to

127   control or experimental treatment groups. Our approach follows sorting

128   compounds into clusters based off of their mean expression, using k-means

129   clustering, prior to application of the random forest algorithm and finally

130   prioritizing these top biomolecules with PCA. After standardizing, we identify

131   compounds most strongly associated with heat stress among liver enriched genes

132   and metabolites. All reads were mapped to the latest NCBI release of the chicken

133   genome and accompanying annotation, GalGal5. Mapping was done with Tophat2

134   and Cufflinks2, with raw counts quantification by featureCounts and differential

135   expression accomplished with edgeR.

136
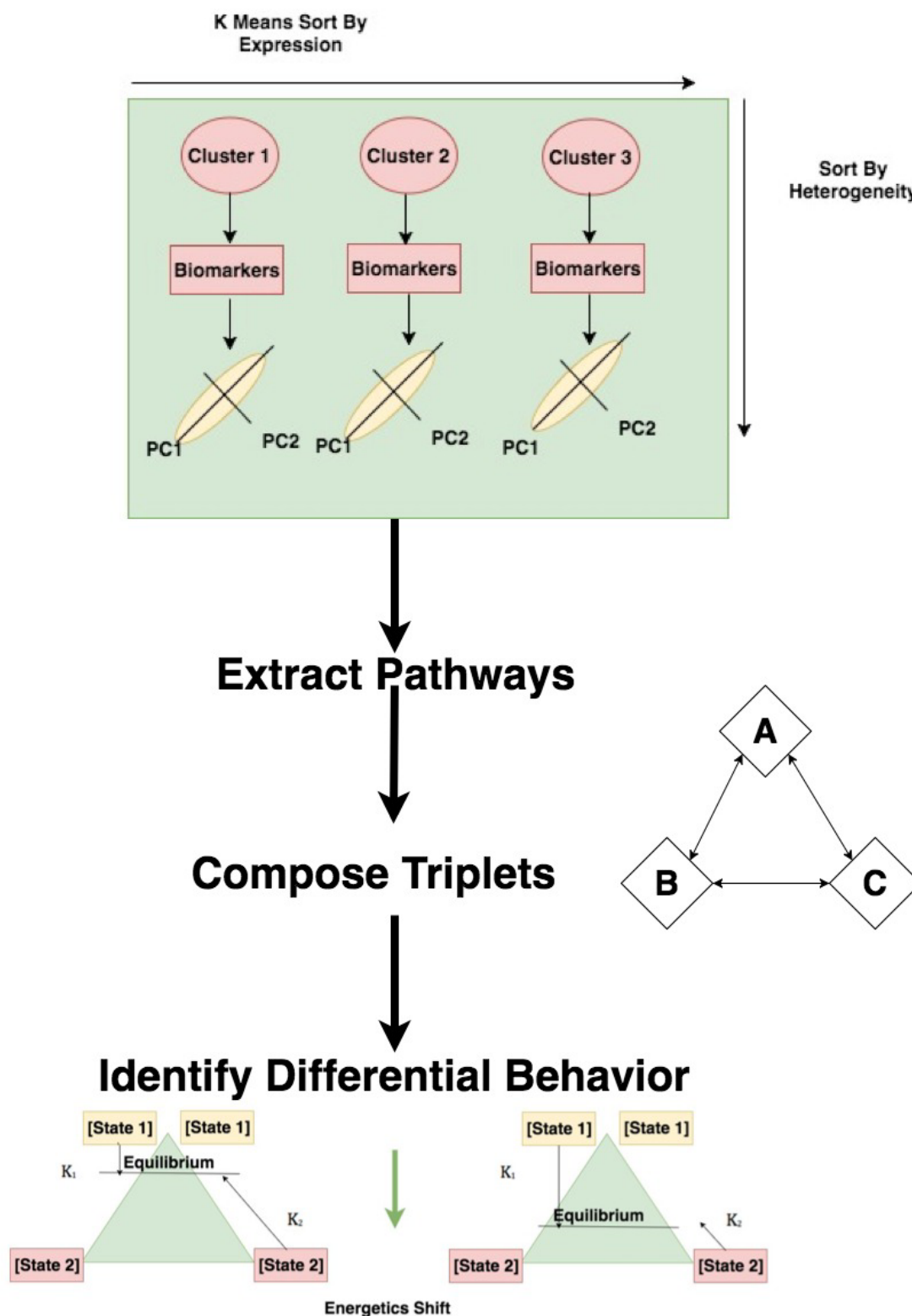
Figure 1: Total pipeline, from data analysis to identifying hypothetical mechanisms.
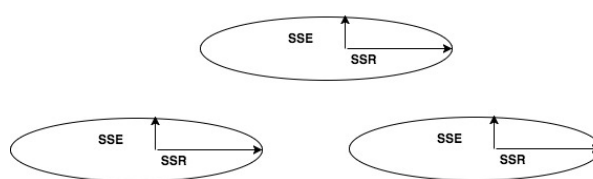
137
138

139   Biomolecules are identified and prioritized to extract pathways from whose

140   elements triplets can be calculated. Triplets showing differential behavior selected,

141   which demonstrate equilibrium shifts at state assumptions and thus indicate

142   behavior of a metabolic fork.

143   **Figure 2A**   **Figure 2B**
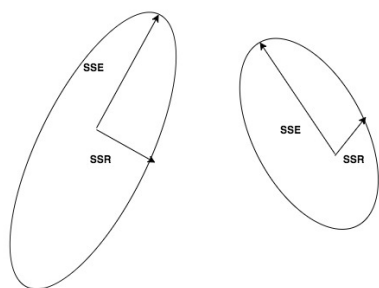

144
145
146   **Figure   2A and 2B:** Example of possible models around specific cluster with

147   different k-means selection, illustrating more uniform clustering results with k = 3

148   (2B) compared to k = 2 (2A).

149
150   **Geometric And Biological Consideration Of K-means Step**
151   A goal of first leveraging k-means analysis is to build more biologically

152   interpretable random forests, with compounds initially separated by expression

153   level. This reflects the idea that pathways involving essential biological compounds

154   occur across a spectrum of expression levels. Compounds with different levels of

155   expression levels may have equally important biological roles. Separating out

156   compounds first by this dimension prevents compounds from one expression tier

157   crowding out those from another tier when they have similar capacities for

158   classifying samples as control or heat stress. However, the optimal partitioning

159   should produce clusters that are similar in explanatory power. Selecting k = 3

160   accomplishes this goal by distributing compounds across clusters that are as similar

161   to one another as possible in terms of their explanatory power (Figure 2A and 2B).

162
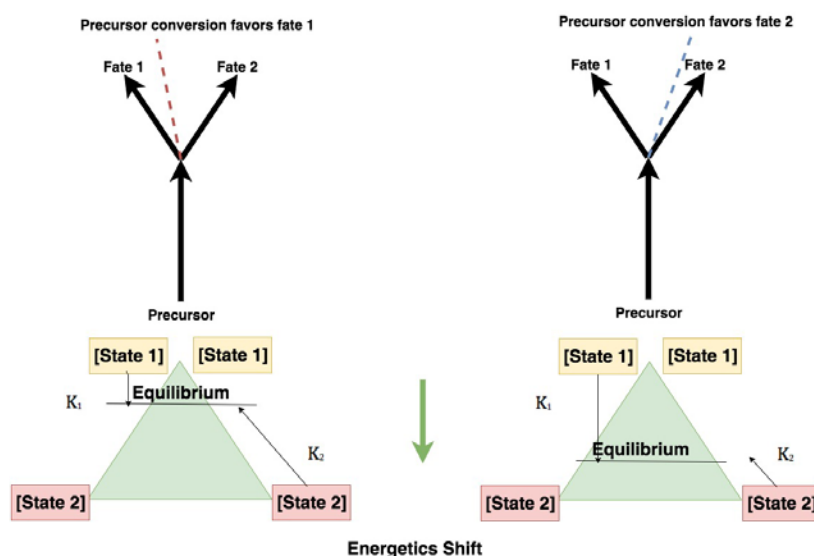
163   **Figure 3:** Under changes in gene expression that alter levels of the regulating

164   enzymes, precursors are preferentially routed to one metabolic fate over another.

165   Shifts in the ratio between metabolites representing fate 1 or fate 2 may represent

166   shifts in biology.

167   **Metabolic Forks**
168   Metabolic forks, in which ratio of metabolites represent activities of

169   competing biological processes are an adaptation of concepts introduced by Giegier

170   et al, in which ratio of metabolites represent biological activity of processes

171   influence by genotype. We refer to these regulatory triplets as such, because they

172   represent divergent fates for metabolites. Candidates for components of metabolic

173   forks are determined via prior knowledge as compounds established in the broiler

174   heat stress response through our previous work (Jastrebski *et al.*, 2017) and which

175   are also biomolecules prioritized by the statistical learning components of the

176   pipeline.

177   Such functions serve as a more realistic description of the biochemistry of

178   pathway steps than simple correlations with raw data. For example, pathway steps

179   where one enzyme regulates the forward reaction and another the reverse, the

180   regulation through gene expression can cause relative increases in one enzyme

181   compared to corresponding enzyme that regulates the opposing reaction. This shifts

182   the favorability of the pathway step towards either the products or reactants,

183   depending on the function of the enzyme. The shift in favorability towards one

184   metabolic fate, at the expense of another, under regulation thus represents a

185   "metabolic fork". Having hypothesized that amino acids from catabolized proteins

186   fuel production of sugar and fats by providing carbon backbones, we calculated

187   "metabolic forks" that included lipids, sugar and amino acids. P-values are

188   determined from the interaction term of the resulting linear model of the metabolic

189   fork, signifying a difference in slope between control and experimental conditions.

190

Results

192   **PCA on Correlation Matrix of Top Biomolecules for Each Cluster Following K-**
193   **means and Random Forest**
194
195   **Cluster 1:**

196
197
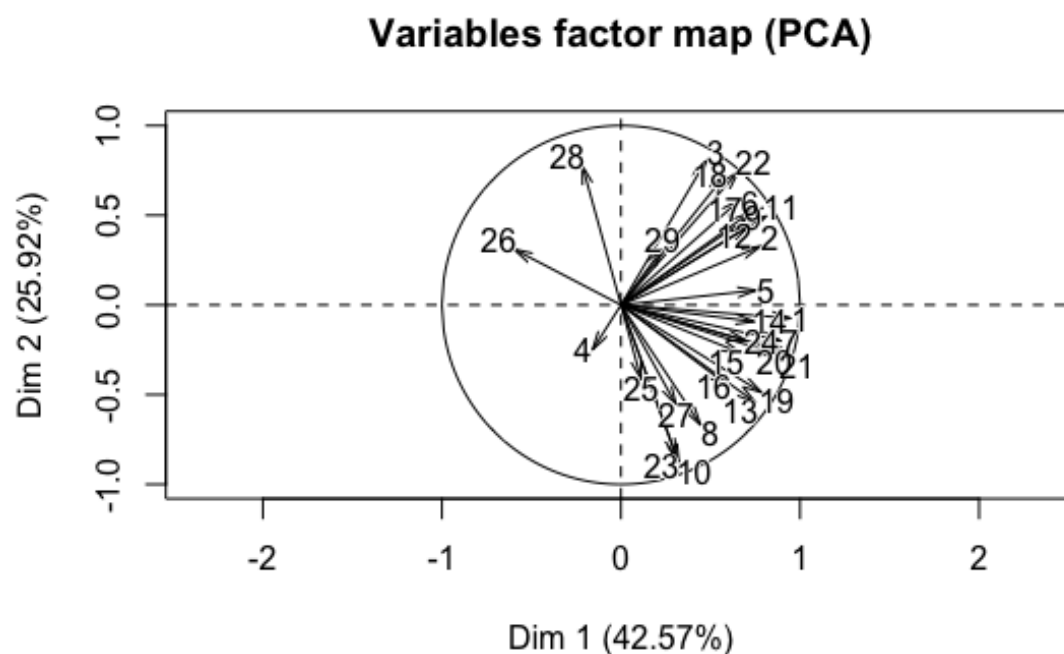
1. 1_2soUnique_1_enyl_palmitoyl_	**16.** picolinate
2 1_stearoyl_GPG__18_0_	**17.** pseudouridine
3. 2__deoxyuridine	**18.** riboflavin_Vitamin_B2
4. "2_hydroxyphenylacetate	**19.** sphingomyelin__d18_1_20_0__d16_1_
5. 2_palmitoyl_GPC__16_0_	**20.** sphingomyelin__d18_1_21_0__d17_1_
6. alanine	**21.** sphingomyelin__d18_1_22_1__d18_2_
7. behenoyl_sphingomyelin__d18_1_22_0_	**22.** thymine
8. dehydroascorbate	**23.** CAPSL
9. dimethylarginine__SDMA___ADMA_	**24.** CPB2
10. glutathione__oxidized__GSSG_	**25.** GFF
11. glycine	**26.** GHR
12. guanosine_2__monophosphate__2__GMP_	**27.** ITIH3
13. hypotaurine	**28.** LOC395933
14. N_formylmethionine	**29.** LOC417848
15. N_methyl_GABA	

198

199 **Figure 4:** Biplot of PCA analysis on correlation matrix of the top biomolecules in

200 cluster 1. See keys to match numbers to compounds.

201
202
203

204   **Cluster 2:**

Variables factor map (PCA)
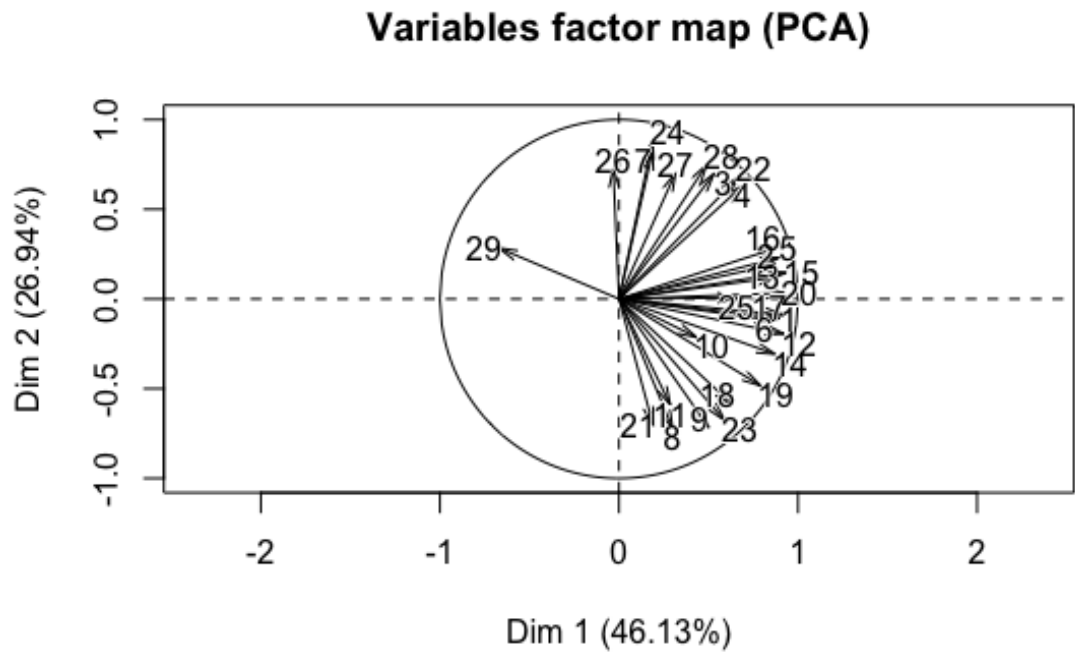

205
206

1. 1_palmitoleoyl_3_oleoyl_glycerol__16_1_18_1_	16. glucose_6_phosphate
2. 1_palmitoyl_2_linoleoyl_GPE__16_0_18_2_	17. glutathione__reduced__GSH_
3. 1_palmitoyl_2_linoleoyl_GPS__16_0_18_2_	18. glycerophosphoethanolamine
4. 1_palmitoyl_2_oleoyl_GPI__16_0_18_1_	19. N_acetylglucosaminylasparagine
5. 1_palmitoyl_2_palmitoleoyl_GPC__16_0_16_1_	20. N6_succinyladenosine
6. 1_stearoyl_2_linoleoyl_GPE__18_0_18_2_	21. phosphopantetheine
7. 1_stearoyl_2_linoleoyl_GPI__18_0_18_2_	22. pterin
8. 3__dephosphocoenzyme_A	23. UDP_glucuronate
9. adenosine	24. C5
10. adenosine_5__monophosphate__AMP_	25. LOC421560
11. coenzyme_A	26. LYG2
12. cysteinylglycine	27. METTL7A
13. fructose_6_phosphate	28. NTSR1
14. gamma_glutamylcysteine	29. TF
15. glucosamine_6_phosphate	

207
208

209   **Figure 5:** Biplot of PCA analysis on correlation matrix of the top biomolecules in

210   cluster 2. See keys to match numbers to compounds.

211
212
213   **Cluster 3:**

214
215
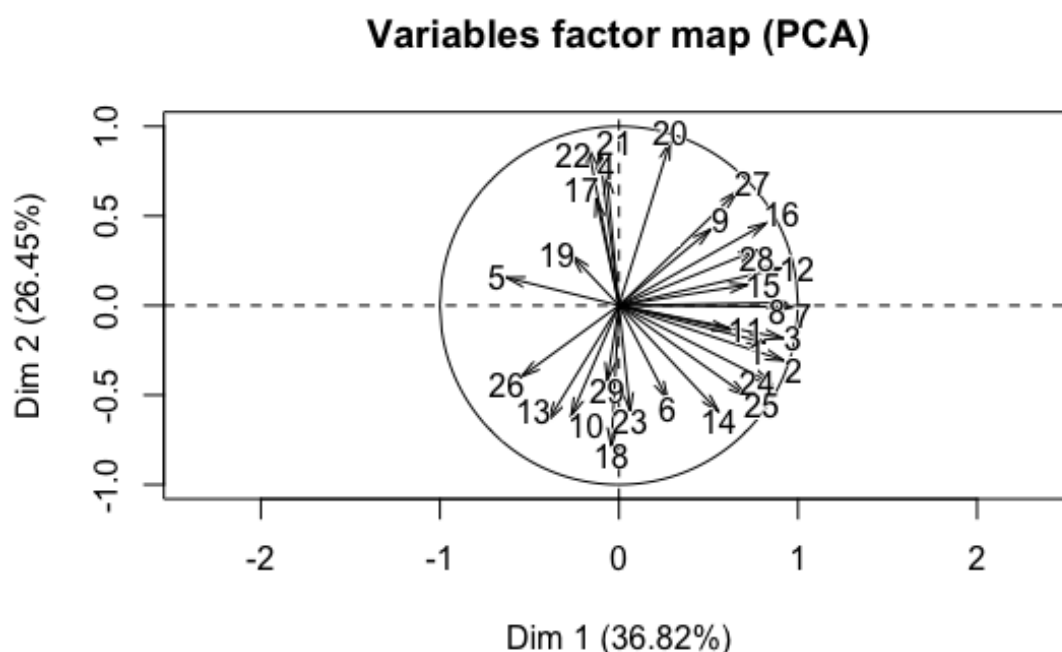
1. 1__6soUnique_enyl_stearoyl_	**16.** docosahexaenoate__DHA;_22_6n3
2. 1__6soUnique_enyl_stearoyl_	**17.** linoleate__18_2n6_
3. 1_arachidonoyl_GPC__20_4n6_	**18.** linoleoylcarnitine
4. 1_arachidonoyl_GPE__20_4n6_	**19.** margarate_17_0
5. 1_palmitoyl_2_linoleoyl_glycerol__16_0_18_2_	**20.** N_acetyltaurine
6. 1_palmitoyl_2_stearoyl_GPC__16_0_18_0_	**21.** N_palmitoyltaurine
7. 1_stearoyl_2_arachidonoyl_GPE__18_0_20_4_	**22.** N_stearoyltaurine
8. 1_stearoyl_2_arachidonoyl_GPI__18_0_20_4_	**23.** oleoylcarnitine
10. adipoylcarnitine	**24.** sphingomyelin__d18_1_24_1__d18_2_24_0_
11. arachidate__20_0_	**25.** sphingomyelin__d18_2_24_1__d18_1_24_2_
12. arachidonate__20_4n6_	**26.** stearoyl_ethanolamide
13. beta_guanidinopropanoate	**27.** tartronate__hydroxymalonate_
14. betaine_aldehyde	**28.** taurine

15. bilirubin_Z_Z	**29.** thiamin_diphosphate

216
217
218 **Figure 6:** Biplot of PCA analysis on correlation matrix of the top biomolecules in
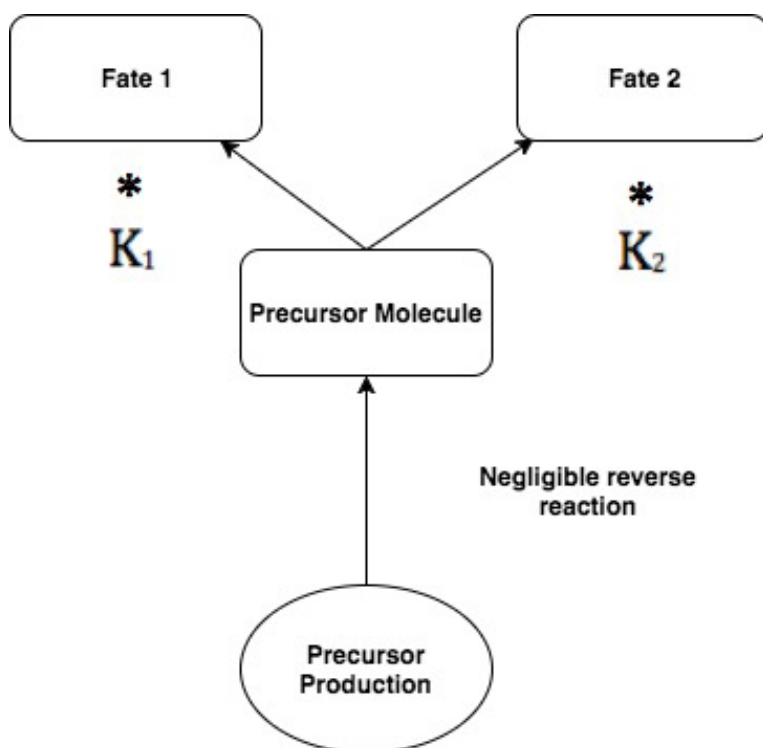
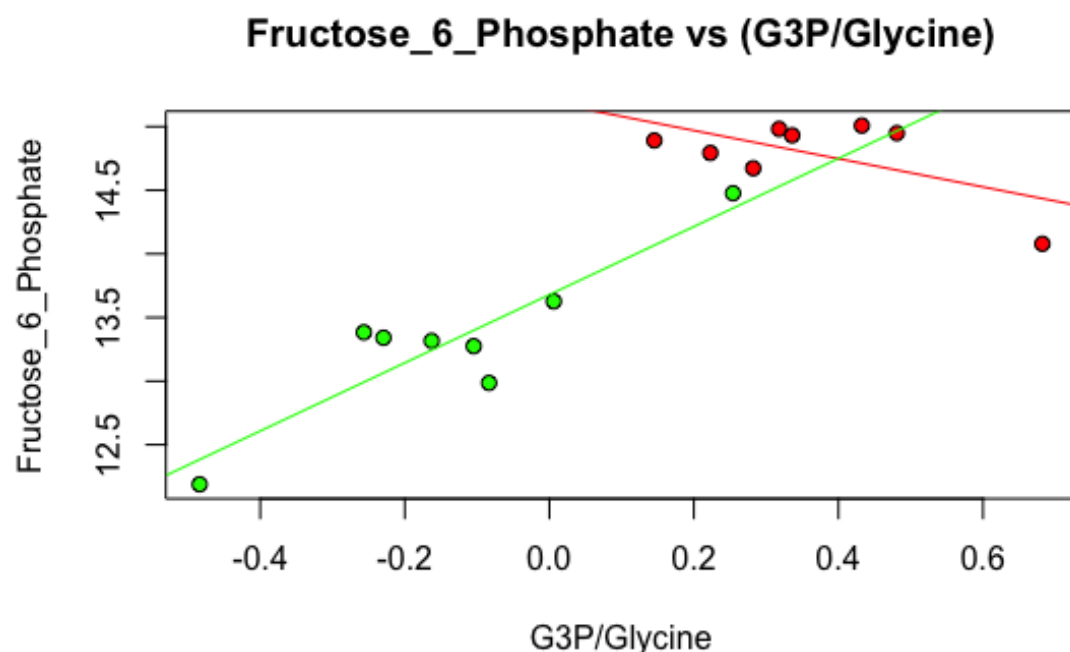219 cluster 3. See keys to match numbers to compounds.

220

221 **Figure 7:** Illustration of the components of a metabolic fork.

Fructose_6_Phosphate vs (G3P/Glycine)

222

223   **Figure 8:**Fructose-6-phosphate (F6P), a sugar produced by gluconeogenesis, as a

224   function of the ratio of glycerol-3-phosphate (G3P) and glycine. All values are log

225   transformed. P-value of interaction term = 0.0005.

226

Discussion

228   Our complete pipeline, which combines statistical learning techniques with

229   hypothesis-free modeling of metabolite ratios, is able to propose novel hypotheses

230   while recapitulating significant known biology from the liver metabolome and

231   transcriptome (Figure 1). Importantly, this perspective identifies changes in

232   compounds with roles across organelles that are increasingly thought to have

233   important functions in the heat stress response.

234   Much interesting biology, for example, relates to changes in the cell

235   membrane. There are widespread shifts in levels of constituent lipids, for example.

236   The exact mechanisms by which these shifts occur remain unclear, but accumulating

237   evidence suggests these changes in the cell membrane exert important downstream

238   effects on heat stress responsive genes and metabolites. At least some of these may

239   be driven by dietary changes. One such example is the essential fat linoleic acid,

240   which is a precursor to arachidonic acid and emerges as a strong heat stress

241   associated biomolecule and whose detected levels are lower under heat stress. The

242   compound also correlates with two principal components among the heat stress

243   associated biomolecules among its cluster (Additional File 1, Tables 2 and 3).

244   Downstream arachidonic acid derivatives are similarly decreased, many of which

245   have roles in inflammatory response.

246   Other biomolecules prioritized through correlation with the same principal

247   component include other lipids, related to signaling and fatty acid oxidation – such

248   as adipoylcarnitine and the taurine related endocannabinoids N-oleoyl taurine and

249   N-Stearoyl taurine (Additional File 1, Tables 2 and 3). These compounds represent

250   a possible intersection between signaling lipids and sulfur metabolism via coupling

251   with taurine. All of these compounds occur at lower concentrations under heat

252   stress. While the mechanisms of such regulation remain unclear, there is much

253   evidence that suggests lipid changes influence cell state and, potentially, bird

254   metabolism.   Lipid changes, in fact, are increasingly recognized as potential

255   regulators of heat stress at a fundamental level (Balogh *et al.,* 2013).

256   Recent studies have focused on nuances of the heat stress response by

257   revising the model that it is primarily triggered by the presence of unfolded proteins

258   (Hoffman, 2007). For example, lipids in the cell membrane may detect membrane

259 disorder and other physical consequences of heat stress and trigger signal cascades

260 (Balogh *et al.*, 2013). The evolutionary value of using a thermo-sensitive organelle

261 such as the cell membrane to refine the heat stress response is the advantage of

262 being able to regulate homeostasis through sensitive adjustments that have

263 meaningful influences on cell fate (Balogh *et al.,* 2013). The inflammatory response

264 may be a significant component in the transition from heat stress to heat stroke.

265 **Heat Stress, Membranes and Lipids**

266   The sophisticated signaling environment created by the cell membrane is

267 comprised of a diverse set of lipids and proteins. Among these is an abundance of

268 sphingolipids that form rafts in the membrane and possess important signaling

269 roles (Simons and Ikonen, 1997). The organization of the cell membrane is intricate

270 and becomes dynamic under stress response. Important structural changes occur

271 through interactions with membrane proteins, the gating of which possess thermal

272 sensitivity (Torok *et al.*, 2014). Additionally, heat causes changes in physical

273 attributes such as diffusion and dimerization rates. Measurements suggest these

274 characteristics change in a predictable fashion during even mild heat stress events

275 (Torok *et al.,* 2014). Thus, the cell membrane is well equipped to sense relative

276 temperature changes.

277   Not surprisingly, among the compounds prioritized by our pipeline are many

278 lipids. These shifts suggest mixture of changes in compounds with signaling and

279 structural roles. Alterations in lipid content are important in thermal shifts

280 associated with both heat stress and extreme cold. For example, a key adaptation to

281 cold is the increase in membrane fluidity mediated by elevating the fraction of cis-

282   unsaturated fatty-acyl groups in membrane lipids (Vigh *et al.,* 1998). Alternatively,

283   during episodes of heat stress mechanisms to endure temperature shifts focus

284   generally on maintaining the integrity of the cellular processes and such pathways

285   can be causally regulated by changes in cell membrane disorder (Vigh *et al.,* 1998).

286   Regulation of heat shock factors can be influenced by addition of saturated and

287   unsaturated fatty acids, with the former inducing expression and the latter

288   suppressing it (Carratu *et al.,* 1996).

289   The possibility that the qualities of the cellular membrane make it an ideal

290   substrate in which to store 'memory' or serve as a 'control center' for a physiological

291   response in terms of the composition of density and sensors could be extremely

292   interesting biologically. This could prove extremely important in terms of

293   identifying mechanistic regulators of the general response. Indeed, changes in

294   membrane fluidity induced via alcohols triggers systemic responses paralleling

295   those caused by heat stress, albeit in the absence of any thermal activation. Such

296   changes include hyperpolarization of the mitochondrial membrane (Balogh *et al.,*

297   2005). Such experimental work confirms the role of lipids from a regulatory

298   perspective and the influence of the heat stress response across organelles.

299   Among the cell membrane lipids influenced by heat stress and which are

300   prioritized among their respective clusters is a number of sphingomyelin species.

301   These are substantially down regulated under heat stress, and emerge as strong

302   classifiers in clusters one and two. This is a potentially significant observation in the

303   context that sphingolipids are up-regulated in the early phases of acute heat stress

304   in studies of yeast (Jenkins *et al.*, 1997). Many of these sphingomyelin species

305   correlate with principal components among their clusters that include the

306   downregulated inflammatory arachidonic acid derivatives (Additional File 1, Tables

307   7 and 8). Their general attenuation may be an important aspect of physiological

308   adaptation to the long term heat stress experienced by the birds, with the pattern of

309   hetereogeneity in their levels indicative of bird acclimatization.

310   **Anti-Oxidants and Energy Burden**
311
312   Heat stress entails a number of challenges that endanger cell function and

313   which must be addressed in order to preserve homeostasis. The management and

314   deployment of protective systems can be quite independent from the initial sensory

315   capacity of the cell membrane. These, for example, can respond to states of cellular

316   stress that could be ongoing in a state of heat stress. Such pathways are essential to

317   the heat stress response, as they relate to management of general consequences of

318   oxidative damage. Several precursors of anti-oxidants, as well as such compounds

319   themselves, are identified as strong classifiers of treatment assignment within each

320   cluster. These compounds manage the effects of toxic intermediates resulting from

321   increased energy production, mitigating their ability to damage DNA or organelles.

322   Their production may exploit the carbon backbones of amino acids released by

323   catabolized protein.

324   Not surprisingly, given the relationship between oxidation and energy

325   production, some of the classifiers suggest changes in mitochondrial activity. Even

326   slight changes in cell resting state can have dramatic changes on the production of

327   reactive oxygen species and the behavior of the mitochondria (Akbarian, 2016).

328   Molecules associated with mitochondrial performance are computationally

329   recognized as potential biomarkers of the heat stress response. This suggests that

330   mitochondrial conditions are closely related to general heat stress, and that the cell

331   adjusts antioxidant levels accordingly.

332   At the same time that sugars and other energy-related metabolites show

333   upregulation, an important class of lipids involved in the carnitine shuttle system

334   that transports fatty acids to the mitochondria shows consistent downregulation.

335   These carnitine species (linoleoylcarnitine, stearoylcarnitine, adipoylcarnitine) are

336   identified as strong heat stress associated biomolecules among their clusters and

337   correlate strongly with resulting principal components (Additional File 1, Tables 1,

338   2 and 3). Such patterns suggest sweeping downregulation of fatty acid oxidation

339   pathways, as metabolism is increasingly driven by gluconeogenesis. Transcriptome

340   changes support a coordinated shift in lipid and sugar management (Jastrebski *et al.*,

341   2017).

342   Genes that correlate most highly with the principal components that emerge

343   from the k-means cluster containing gluconeogenesis biomolecules include

344   Neurotensin Receptor 1 (NTSR1) (Additional File 1, Tables 4 and 5). However, the

345   correlation of this transcript with the first principal component, at .48, is relatively

346   weak compared to the main metabolites associated with gluconeogenesis, i.e. .48.

347   Glucosamine-6-phosphate and glucose-6-phosphate have correlations of .89 and .91,

348   for comparison. However, NTSR1 correlates much more strongly with the second

349   principal component, at .75.   NTSR1 is a neurotensin receptor, and is associated

350   with appetite regulation and blood sugar levels. Another strong classifying gene

351   associated with this cluster is the methyltransferase METTL7A, though it only

352   correlates significantly with the second principal component (.703) (Additional File

353   1, Table 5). Upregulated NTSR1 may be important to maintaining blood sugar

354   homeostasis in the presence of enhanced gluconeogenesis, whereas METTL7A may

355   be manage sulfur metabolism associated with enhanced antioxidant production. In

356   addition to

357

358   This shift towards gluconeogenesis is supported strongly from a mechanistic

359   standpoint by the metabolic fork (Figure 10). The metabolic fork provides evidence

360   of large-scale redirection of carbon resources released from the catabolized glycine.

361   This complements the statistical learning pipeline, which prioritizes biomolecules

362   without determining whether they are causal or merely collinear to biological

363   changes.   The detection of a potential metabolic fork also provides kinetic

364   information about biochemical pathways that may be consequences of gene

365   expression changes. Directionality can be inferred through prior knowledge. For

366   example, the metabolic fork suggests that carbon released from catabolized glycine

367   is preferentially routed to F6P through gluconeogenesis, as opposed to G3P, under

368   heat stress. This is supported by increased transcription of the rate-limiting gene

369   controlling this step, FBP2, under heat stress. The ability to identify concrete

370   mechanisms which exhibit differential behavior under heat stress, and which are

371   consistent with transcriptome changes, makes it possible to complement purely

372   correlation-based strategies with mechanistic hypotheses.

373   **Metabolic   Forks   Resulting   From   Gene   Regulation**

374   One of the top differentially regulated triplets contains two compounds

375   prioritized through PCA on top biomolecules on a k-means cluster.   This is

376   consistent with gene important expression changes, such as those involving FBP2.

377   The three members of the triplet span gluconeogenesis (fructose-6-phosphate),

378   glyceroneogenesis (glycerol-3-phosphate) and amino acid catabolism (glycine).

379   Pairwise correlations between each node are provided on the edge corresponding

380   edge. A proposed mechanism for the observed pattern is that catabolized glycine is

381   preferentially   shunted   towards   gluconeogenesis   under   heat   stress,   thus

382   contributing to F6P production. Increasingly fueled by carbon backbones provided

383   by   amino   acids   from   catabolized   proteins,   gluconeogenesis   decouples   from

384   glyceroneogenesis under heat stress.

385   The ratio of G3P to glycine represents the tendency of catabolized amino

386   acids to become backbones for fats, as opposed to sugars. This changes as a function

387   of increased demand for sugar under heat stress and is corroborated by increase in

388   the gene Fructose-Bisphosphatase-2 (FBP2) encoding the rate-limiting gene for

389   gluconeogenesis.

390   **Summary and Future Work**

391   Interest in the heat stress response is broad, stretching from plant

392   physiology to human clinical research, with insights potentially applicable across

393   taxa due to the deep conservation of cell signaling pathways. Next generation

394   sequencing technologies provide new experimental perspectives from which to

395   explore such systems. During the past several years, the advent of next generation

396   sequencing tools has produced a deluge of data. However, methods to process that

397   data have been lacking. Combining the information from multiomics and multi

398   organ datasets compounds this challenge. The capacity to link patterns of

399   heterogeneity to pathway importance is an approach that can ease the burden of

400   prioritizing compounds in such a setting. Here, we do so and leverage a

401   combination of relative tissue enrichment and statistical learning approaches to

402   prioritize compounds based on their ability to identify samples as belonging to heat

403   stress or control conditions. We demonstrate signatures of the heat stress response

404   across several important systems. Importantly, this is a very general strategy that

405   works with any type of continuous data, rendering it applicable to both metabolome

406   and transcriptome data and flexible enough to accommodate future "-omics" data.

407   From both types of "-omics" data, we identify a diverse range of important

408   mechanisms that may be influencing shifting biology.

409   We have leveraged a flexible method of analysis to process a complex dataset

410   and identify important components of the heat stress response. While

411   recapitulating known biology, our analysis also proposes new hypotheses about

412   heat stress regulation that relates to systems controlled by a diverse range of

413   organelles. These can be explored through future experimentation. Additionally,

414   the metabolic fingerprint of heat stress provides candidates for feed

415   supplementation studies. Thus, this study proposes a general workflow to integrate

416   high dimensional, complex datasets in order to yield testable hypotheses about

417   biology.

418

419
420
421
422

423

424

425

426

427

428

429

430

431

432

433

434

435

References

436

437   1) Akbarian, A., Michiels, J., Degroote, J., Majdeddin, M., Golian, A., De Smet, S.
438   Association between heat stress and oxidative stress in poultry; mitochondrial
439   dysfunction and dietary interventions with phytochemicals. Journal of Animal Science
440   and Biotechnology. 2016. 7: 37 DOI: 10.1186/s40104-016-0097-5.
441   PMCID: PMC4924307.

442

443   2) Balogh, G., Peter, M., Glatz, A., Gombos, I, Torok, Z., Horvath, I., Harwood, J.L.,
444   Vigh, L. Key Role of Lipids in Heat Stress Management. 2013. 587(13):1970-80
445   DOI: https://doi.org/10.1016/j.febslet.2013.05.016. PubMed PMID: 23684645.

446

447   3) Balogh, G., Horvath, I., Nagy, E., Hoyk, Z., Benko, S., Bensaude, O., Vigh, L. "The
448   hyperfluidization of mammalian cell membranes acts as a signal to initiate the heat shock
449   protein response". 2005. FEBS J. 272(23):6077-86. DOI:
450   https://doi.org/10.1016/j.febslet.2013.05.016. PubMed PMID: 16302971.

451

452   4) Carratu, L., Franceschelli, S., Pardini, C., Kobayashis, G., Horvath, I., Vigh, L.,
453   Maresca.   Biochemistry Membrane Lipid Pertrubation Modifies the Set Point of the
454   Temperature of Heat Shock Response in Yeast (A9-desaturase/membrane physical state).
455   1996. Proc Natl Acad Sci U S A. 93(9): 3870–3875. PubMed PMID: 39451.

456

457   5) Ermak, G., Davies, K.J. Calcium and Oxidative Stress: From Cell Signaling to Cell
458   Death. 2002. Mol Immunol. 38(10):713-21. PubMed PMID: 11841831.

459

460   6) Gieger, C., Geistlinger, L., Altmaier, E., Hrabe de Angelis., Kronenberg, F.,
461   Meitinger, T., Mewes, H.W., Wichmann, H.E., Weinberger, K.M., Adamski, J., Illig, T.,
462   Suhre, K. Genetics Meets Metabolomics: A Genome-Wide Association Study of

463   Metabolite Profiles in Human Serum. 2008. PLoS Genet. 4(11):e1000282.
464   DOI:10.1371/journal.pgen.1000282. PubMed PMID: 19043545.
465
466   7) Hoffman, N.R. The Plasma Membrane as First Responder to Heat Stress. 2009. Plant
467   Cell. 21(9):2544. DOI: 10.1105/tpc.109.210912. PubMed PMID: PMC2768915.
468
469   8) Ishikura, K., Miyazaki, T., Song-Gyu, R., Endo, S., Nakamura, Y., Matsuzaka, T.,
470   Miyakawa, S., Ohmori, H. Effect of Taurine Supplementation of the Alterations in
471   Amino Aid Content in Skeletal Muscle with Exercise in Rat. 2011. J Sports Sci Med.
472   10(2):306-314. PMCID: PMC3761861.
473
474
475   9) Jastrebski, S.F., Lamont, S.J., Schmidt, C.J. Chicken hepatic response to chronic heat
476   stress using integrated transcriptome and metabolome analysis. 2017. PloS One.
477   12(7):e0181900. DOI: 10.1371/journal.pone.0181900. PubMed PMID: 28759571.
478
479   10) Jenkins, G.M., Richards, A., Wahl, T., Mao, C., Obeid, L., Hannun, Y. Involvement
480   of yeast sphingolipids in the heat stress resonse of Saccharomyces cerevisae. 1997.
481   272(51):32566-72. PubMed PMID:9405471.
482
483   11) Mazella, J., Beraud-Dufour, S., Devader, C., Massa, F., Coppola, T. Neurotensin and
484   its receptors in the control of glucose homeostasis. 2012. Front Endocrinol (Lausanne).
485   3(143). doi: 10.3389/fendo.2012.00143. PMCID: PMC3515879.
486
487   12) Pollak, N., Dolle, C., Ziegler, M. The Power to Reduce: Pyridine Nucleotides –
488   small Molecules with a Multitude of Functions. 2007. Biochem J. 402(2):205-18.
489   DOI:10.1042/BJ20061638. PubMed PMID: 17295611.
490
491   13) Tallentire, C.W., Leinonen., I., Kyriazakis, I. Breeding for Efficieincy in the Broiler
492   Chicken: A Review. 2016. 36:66. Agronomy for Sustainable Development.
493
494   14) Torok, Z., Crul, T., Maresca, B., Schutz, G.J., Viana, F., Dindia, L., Piotto, S.,
495   Brameshuber, M., Balogh, G., Peter, M., Porta, A., Trapani,A., Gombos, I., Glatz, A.,
496   Gungor, B., Peksel, B., Vigh, L., Csoboz, Z., Horvath, I., Vijahan, M., Hooper, P.,
497   Harwood, J.L., Vigh, L. Plasma membranes as heat stress sensors: From Lipid controll
498   molecular switches to Therapeutic Applications. 2014.
499   Biochim Biophys Acta. 1838(6):1594-618. DOI: 10.1016/j.bbamem.2013.12.015.
500   PubMed PMID:24374314.
501
502   15) Vigh, L., Horvath, I., Maresca, B., Harwood, J.L. Can the stress protein response be
503   controlled by 'membrane-lipid' Therapy. 2007. Trends Biochem Sci. 2007 32(8):357-63.
504   DOI: 10.1016/j.tibs.2007.06.009. PubMed PMID:17629486.
505
506   16) Saghatelian, A.,McKinney, M.K.,Bandell, M. A FAAH-regulated class of N-acyl
507   taurines that activates TRP ion channels. 2006. Biochemistry. 45(30):9007-15.
508   DOI:10.1021/bi0608008. PMID:16866345.
509

510   17) Simons, K., Ikonen, E. Functional Rafts in Cell Membranes. 1997. Nature.
511   387(6633):569-72. DOI: 10.1038/42408. PubMed PMID:9177342.
512
513   18) Vigh, L., Maresca, B., Harwood, J.L. Does the Membrane's Physical State Control
514   the Expression of Heat Shock and other Genes? 1998. Trends Biochem Sci. 23(10):369-
515   74. PubMed PMID:9810221.
516
517