

1 **How to prevent chick culling in the poultry industry? Discovery of a new**  
2 **biomarker for *in ovo* gender screening**

3  
4 Nicolas DROUIN<sup>a,φ</sup>, Hyung Lim ELFRINK<sup>a,φ</sup>, Wouter BRUINS<sup>b,φ</sup>, Slavik KOVAL<sup>a</sup>, Amy C HARMS<sup>a</sup>, Wil  
5 STUTTERHEIM<sup>b</sup>, Thomas HANKEMEIER<sup>a</sup>

6  
7  
8 a) *Metabolomics and Analytics Centre, Leiden Academic Centre for Drug Research, Leiden*  
9 *University, Einstein weg 55 2333CC, Leiden, the Netherlands*

10  
11 b) *In Ovo, Haagse Schouwweg 12, 2332 KG, Leiden, the Netherlands*

12  
13 *φ: equal contribution*

14  
15 **Corresponding Author:**

16 Prof. Thomas Hankemeier

17 Tel: +31 (0) 71 527 4226

18 E-mail: [hankemeier@lacdr.leidenuniv.nl](mailto:hankemeier@lacdr.leidenuniv.nl)

19

20 **Abstract**

21  
22 Chicken eggs are one of the most consumed foods worldwide. However, the practice of chicken culling in  
23 the poultry industry involves unnecessary animal suffering and finding a way to put an end to this has  
24 become a societal priority. One approach that has been propagated as acceptable is based on the selection  
25 of female eggs early in the incubation process and the devitalization of the male eggs.

26 It is with this objective in mind that we searched for a biomarker for early gender screening in eggs.  
27 Applying an untargeted mass spectrometry approach, we profiled allantoic fluid of different day-old eggs  
28 and identified the feature 3-[(2-aminoethyl)sulfanyl]butanoic acid (ASBA) as a strong biomarker for *in-ovo*  
29 gender prediction for day-9 old embryos. After validation using LC-APCI-MRM with an internal standard,  
30 we found ASBA can predict the female gender with a sensitivity and specificity well above 95% in our  
31 experiments.

32  
33 **Keywords:** Mass spectrometry, metabolomics, biomarker, gender screening, chicken eggs, animal welfare,  
34 poultry industry

35  
36 **Highlights**

- 37
- Discovery of a biomarker of chicken embryo gender in allantoic fluid from eggs
  - Day 9 after laying was determined as optimum for sex prediction and animal welfare
  - 3-[(2-aminoethyl)sulfanyl]butanoic acid was identified using mass spectrometry
  - The biomarker was validated on large cohorts of different chicken species
- 40

41

## 42 1. Introduction

43 Chicken eggs are one of the most popular cooking ingredients in every country over the world. Every year,  
44 more than 1.3 trillion eggs are produced by 4.7-6.5 billion of laying hens, representing a tremendous 80.1  
45 million metric tons of eggs (WATTPoultry, 2019). The productivity of laying hens drops anywhere between  
46 1 to 3 years of age after which they are slaughtered and replaced. This causes a perpetual need for massive  
47 laying hen production. Unlike broiler chicken breeds, males from these egg-laying breeds do not produce  
48 enough meat to have any significant value in the food chain. Hatching males increases the costs in the  
49 laying hen industry. Furthermore, this poses an ethical problem: the need to cull the males. More than 3.2  
50 billion day-old male chickens are culled every year. Currently, due to lack of viable alternatives, chick sexing  
51 is manually performed after hatching and, depending on the breed, is based on different colors of the  
52 down, or observation of the cloaca. While female chicks are kept for egg production, males are killed by  
53 mechanical methods such as maceration, electrocution or asphyxiation (Health, Welfare, Nielsen, Alvarez,  
54 Bicout, Calistri, et al., 2019; UNION, 2009).

55 Interest in animal welfare has been growing over the last decade. Therefore, several countries, such as  
56 France, Germany, Switzerland or USA, are currently designing new and more restricted legislation on this  
57 practice in the poultry industry (Blide & Montel, 2015; Commion de la Science 2019; l'Alimentation, 2020;  
58 Producers, 2020). This adds to the urgency to find economically viable solutions to end male chick culling.  
59 In this context, *in ovo* gender determination to select and sort eggs before hatching is one of the most  
60 promising strategies. In the early stages of the embryo development, the central nervous system is not  
61 fully developed. Consequently, the embryo is not able to feel the sensation of pain (Close, Banister,  
62 Baumans, Bernoth, Bromage, Bunyan, et al., 1997). This would therefore make the early removal of male  
63 eggs a more acceptable practice.

64 As any gendered species, gender can be determined with absolute certainty by genotyping. In that respect,  
65 He *et al.* developed sensitive and robust PCR and qPCR approaches (He, Martins, Huguenin, Van, Manso,  
66 Galindo, et al., 2019). However, PCR based approaches are expensive and too slow to screen each egg in  
67 a hatchery. Some alternative methods have already been developed for this purpose. In 2011, Steiner *et*  
68 *al.* developed a FT-IR spectroscopy method to detect the chromosomic difference between males and  
69 females from the germinal disc (Steiner, Bartels, Stelling, Krautwald-Junghanns, Fuhrmann, Sablinskas, et  
70 al., 2011). However, this method required puncture and collection of cells from the embryo, presenting a  
71 high risk for the fetus. More recently, the same group developed Raman and fluorescence spectroscopy  
72 methods targeting the blood vessels located under the shell (Galli, Preusse, Uckermann, Bartels,  
73 Krautwald-Junghanns, Koch, et al., 2016, 2017). The combination of both spectroscopic measurements

74 attained an overall correct prediction rate of 93% in small-scale experiments (Galli, Preusse, Schnabel,  
75 Bartels, Cramer, Krautwald-Junghanns, et al., 2018). Only a handful of processes have reached a pilot test  
76 scale. Among these figures the method developed by Weissmann *et al.*, which is now allowing for a limited  
77 commercial production by Seleggt GmbH. They identified estrone sulfate, a metabolism product of sex  
78 steroids, as a biomarker for *in ovo* early gender prediction and they developed an ELISA test with a  
79 sensitivity of 86.0% and specificity of 82.9% (Weissmann, Reitemeier, Hahn, Gottschalk, & Einspanier,  
80 2013). However, immunoassays are expensive due to the manufacture of selective antibodies and are  
81 relatively low throughput that makes this method largely unsuitable for a large-scale industrial application.  
82 Also, the method can only be applied at a relatively late stage of the embryonic development due to the  
83 late occurrence of estrogen sulfate in the embryo development.

84 The present study reports the discovery and characterization of a new biomarker using mass  
85 spectrometry-based analytical tools. The identification of the relevant features was first made using  
86 untargeted RPLC-MS analysis and subsequent structure elucidation of these features was achieved by the  
87 combination of several analytical approaches using MS/MS. The newly discovered biomarker allowed an  
88 early female gender determination at day-9 in the experiments described below. We also demonstrate its  
89 prediction capability in different chicken breeds.

90

## 91 **2. Materials and methods**

### 92 **2.1. Chemicals**

93 S-Propyl-L-Cysteine and 3-[(2-aminoethyl)sulfanyl]butanoic acid (ASBA) and the deuterated version were  
94 synthesized upon request by Enamine (Kiev, Ukraine). S-(2-carboxypropyl)-Cysteamine was synthesized by  
95 the group of Prof. Byeong-Seon Jeong (Gyeongsan, South Korea). MS grade water was produced by a milliQ  
96 water generator. Formic acid, methanol and acetonitrile were purchased from Acros Organics (Geel,  
97 Belgium) and Actu-All Chemicals (Oss, the Netherlands) respectively.

98

### 99 **2.2. Samples**

100 All eggs used in this study came from brown chickens (VB1636 Brown Nick) and white chickens (Hy-Line  
101 CV 24). Depending on the development stage, from 100  $\mu$ L to 1 ml of allantoic fluid has been manually  
102 collected by carefully drilling a hole in the top of the shell. A 0.22 Gauge needle on a syringe was used to  
103 extract the allantoic fluid without inflicting harm to the developing embryo. After collection, the allantoic  
104 fluid samples were snap frozen in liquid nitrogen before transport and were stored at -80 °C prior to  
105 analysis.

106 During this study, 3 different sample cohorts were used: (i) the discovery cohort, (ii) the breed cohort, (iii)  
107 the validation cohort. The first discovery cohort consisted of 50 Brown Nick eggs collected at day-7 and 60  
108 eggs at day-8,9,10 and 11 after the start of the incubation. The breed cohort consisted of a total of 146  
109 Brown Nick and 151 Hy-Line eggs, both collected at day-9, and these were used to assess the prediction  
110 power of the candidate biomarker in different chicken breeds. Finally, the validation cohort included 143  
111 Brown Nick eggs collected at day-9 and these were analyzed with a targeted LC-APCI-MRM method in  
112 order to confirm the identity of the biomarker and its prediction performance.

113

### 114 **2.3. Genetic gender determination by PCR**

115 To determine the gender of the embryo at the early stages of its development, a PCR analysis was made  
116 using the primer 1237L and 1237H (Kahn, St John, & Quinn, 1998). To do so, the embryo toes of each  
117 corresponding allantoic fluid sample have been collected and snap frozen in liquid nitrogen. The samples  
118 were then subjected to PCR analysis.

119

### 120 **2.4. Sample preparation**

121 Allantoic samples from the discovery and breed cohort have been deproteinized with methanol. Briefly,  
122 10 µL of the sample has been mixed with 90 µL of cold methanol and then vortexed. After centrifugation  
123 at 10,000 rpm at 10°C for 10 min, the supernatant was collected. An equal fraction of each supernatant  
124 has been collected for each sample and then mixed together to prepare QC samples which were analyzed  
125 at a regular interval between samples.

126 Prior to targeted analysis, samples from the validation study were spiked with a deuterated internal  
127 standard. Briefly, 100 µL of allantoic fluid has been mixed with 100 µL of a solution made of the deuterated  
128 internal standard at a concentration of 600 ng/mL in water, using a Tecan Freedom Evo liquid handling  
129 platform (Männedorf, Switzerland). QC samples have been prepared by mixing an equal volume of each  
130 raw samples and processed in parallel to individual allantoic fluid samples.

131

### 132 **2.5. RPLC-MS/MSMS**

133 Untargeted metabolite analyses were performed on an ACQUITY UPLC from Waters (Milford, USA) directly  
134 coupled to the ESI jet stream nebulizer of an Agilent 6530 series ESI-Q-TOF high-resolution mass  
135 spectrometer from Agilent Technologies (Waldbronn, Germany). The separation was performed with a  
136 Waters AccQ-Tag Ultra column (1.7 µm, 2.1 mm × 100 mm), thermostated at 60 °C. The LC gradient has  
137 been described previously in the analytical chromatography section of our previous work (Laan, Elfrink,

138 Azadi-Chegeni, Dubbelman, Harms, Jacobs, et al., 2021) but with 2% formic acid added to both mobile  
139 phases. MS acquisitions were done in ESI positive mode using full scan mode with a mass range from 100-  
140 1000 m/z and used a lock mass for improved mass accuracy. MS/MS data were recorded at 5 and 25 eV  
141 using an inclusion list of targeted features.

142  
143 **2.6. HILIC-MS**  
144 HILIC-MS analyses were performed using an Agilent 1260 Infinity LC system directly coupled to a  
145 DuoSpray™ ionization source of a Sciex 6600 TripleTOF (Darmstadt, Germany). The separation was  
146 performed with a Zic-cHILIC column (3 μm, 2.1 mm × 100 mm, 100 Å), purchased from Merck Millipore  
147 and thermostated to 25 °C. The gradient is similar to published HILIC chromatography (Laan, Elfrink, Azadi-  
148 Chegeni, Dubbelman, Harms, Jacobs, et al., 2021) but used 5mM of ammonium acetate instead of  
149 ammonium formate and a flow rate of 0.2 ml/min. MS acquisitions were performed in ESI positive mode  
150 using full-scan acquisition mode with a mass range from 50 to 1500 m/z.

151  
152 **2.7. LC-APCI-MRM**  
153 Targeted analyses were performed on an ACQUITY UPLC from Waters (Milford, USA) directly coupled to  
154 the Multimode source of an Agilent 6460 series QqQ mass spectrometer from Agilent Technologies  
155 (Waldbronn, Germany). The separation was performed with a Waters AccQ-Tag Ultra column (1.7 μm, 2.1  
156 mm × 100 mm), thermostated at 60 °C. The mobile phases were prepared by adding (A) water and (B)  
157 methanol both with 0.1 % formic acid as modifier. The injection volume was 5 μL, and the flow rate of the  
158 mobile phase was set to 0.3 mL/min. Mobile phase B was increased from 0.5 to 30% over 3 min, then was  
159 further increased to 99% in 0.1 min. Finally, mobile phase B was brought back to the initial conditions in  
160 0.9 min over a period of 3 min. MS acquisitions were done in APCI positive mode using the MRM acquisition  
161 mode. The transitions 164.1 → 119.0 and 168.1 → 119.0 with a collision energy of 12 V and a dwell time  
162 of 50 ms were used to monitor 3-[(2-aminoethyl)sulfanyl]butanoic acid and the deuterated version  
163 respectively.

164  
165 **2.8. Data processing**  
166 The raw data from both discovery and breed cohorts were pre-processed using Agilent MassHunter  
167 Workstation Software Profinder (Agilent, Version B.06.00, Build 6.0.606.0. The Batch Recursive Feature  
168 Extraction was performed on the raw files loaded into the software. The following settings were used,  
169 settings not mentioned were as used as default by the software. Molecular feature extraction was

170 performed for peaks between 0 to 8 min, for m/z between 100 and 900 m/z and with a minimum count of  
171 300. Peak binning and RT alignment have been done using a RT window of  $1\% \pm 0.15$  min and a mass  
172 window =  $5.00$  ppm  $\pm 2$  mDa. The maximum number of features was set to 2000. Recursive feature  
173 extraction was made using Find by Ion mode, using matching tolerance of  $\pm 1.5$  min for retention time and  
174 35 ppm of symmetric error on the m/z. The peak area extraction of the features was done using Agile  
175 algorithm. Duplicate features, artefact features from the data preprocessing, and features present in less  
176 than 90% of the samples were manually removed, as well as samples with less than 80% of the total  
177 amount of features present.

178 The raw data from the validation cohort has been processed using Skyline (MacLean, Tomazela, Shulman,  
179 Chambers, Finney, Frewen, et al., 2010). Finally, the statistical treatment of the extracted features has  
180 been done using MetaboAnalyst 5.0 (Xia, Mandal, Sinelnikov, Broadhurst, & Wishart, 2012). and R  
181 packages.

182

### 183 **3. Results**

#### 184 **3.1. Allantoic fluid is an accessible and non-damaging fluid for biomarker discovery and** 185 **routine analysis**

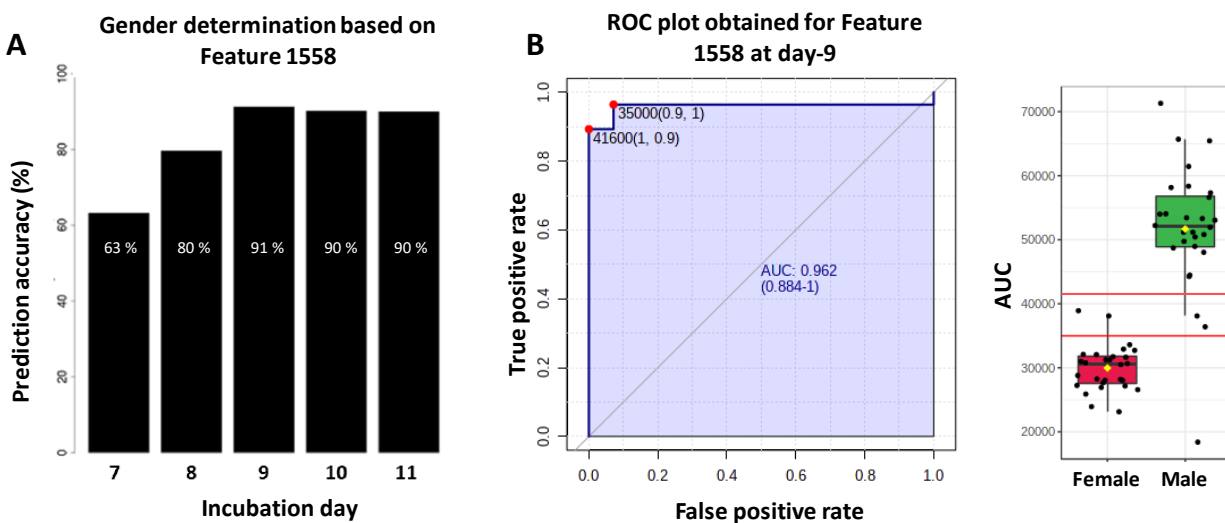
186 The allantois is a hollow-sac structure developed by the embryos and has two main functions. As it is highly  
187 vascularized, the allantois *de facto* acts as the lungs of the embryonic chick supplying oxygen for the red  
188 blood cells that passed through the eggshell. Its second function is to collect the nitrogenous waste  
189 produced by the fetus, thus making the composition of the allantoic fluid close to mammalian urine. It is  
190 also interesting to note that having the lowest fluid density, the allantoic fluid is easily accessible after  
191 puncturing the shell through the air sac located at the apex of the egg. Consequently, the collection of a  
192 reasonable volume, dependent on the development of the fetus, of this biofluid is not detrimental to the  
193 fetus and makes it a good source for potential biomarkers (Weissmann, Reitemeier, Hahn, Gottschalk, &  
194 Einspanier, 2013).

195

#### 196 **3.2. Untargeted metabolomics approach for the discovery of gender specific biomarkers in** 197 **allantoic fluid**

198 Metabolomics based biomarker discovery methodologies rely on the detection of a very large variety of  
199 compounds in order to maximize the chance of finding interesting hits (Drouin, Pezzatti, Gagnebin,  
200 Gonzalez-Ruiz, Schappler, & Rudaz, 2018; Pezzatti, Gonzalez-Ruiz, Codesido, Gagnebin, Joshi, Guillaume,  
201 et al., 2019). For this reason, an untargeted liquid chromatography with mass spectrometry (LC-MS)

202 approach was chosen using positive electrospray mode on a high-resolution Q-TOF (Noga, Dane, Shi, Attali,  
203 van Aken, Suidgeest, et al., 2012). The discovery cohort comprised 50 eggs for day 7 and 60 at days 8 to  
204 11. In total, 1954 peaks with unique retention time and mass, so called features, were detected in the LC-  
205 MS chromatograms with this method. After manual curation of the data, 1390 feature were considered  
206 and the gender was modeled as a binary outcome with sampling day and measured features as model  
207 predictor. The model was further optimized using backward elimination techniques (Koller & Sahami,  
208 1996)., validated using cross-validation (Steyerberg, Harrell, Borsboom, Eijkemans, Vergouwe, &  
209 Habbema, 2001). and evaluated by the percentage of correctly predicted genders, male and female  
210 combined together. This approach identified a particular feature (F1558) as the most promising biomarker  
211 candidate with a prediction rate up to 91% at day-9 (Figure 1A).

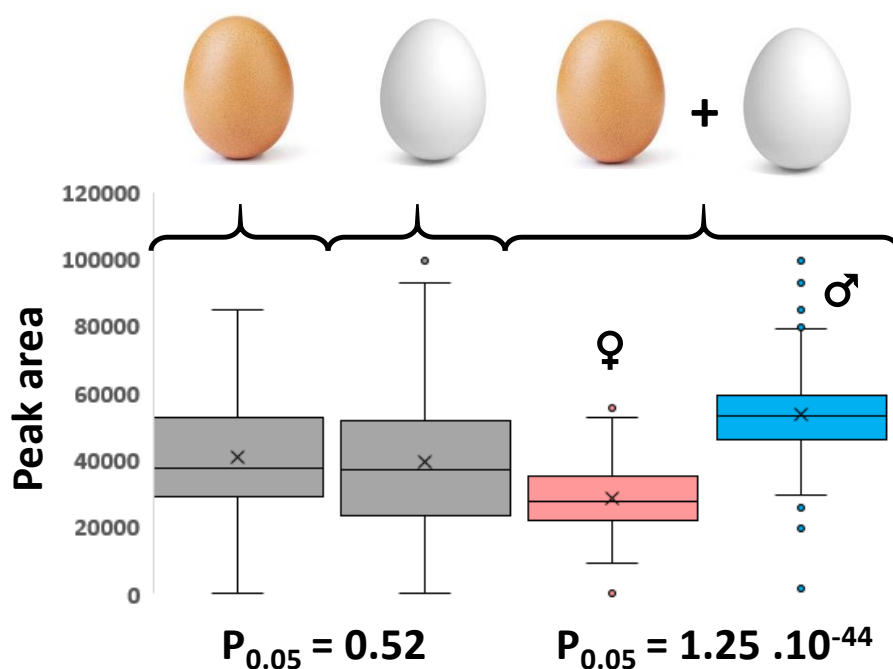


212  
213 *Figure 1: (A) Logistic regression classification model On Single Feature. We evaluated single predictor models for all measured*  
214 *features and observed the best accuracy using Feature 1558 (F1558), which provides >90% accuracy at day-9. (B) ROC plot of*  
215 *F1558 at day-9 and its associated boxplot.*

216  
217 This particular feature presented similar prediction power for days 9 to 11. However, during chicken  
218 embryonic development, the neural tube needs approximately 50% of the 21 days of the total gestation  
219 duration to turn into a functional brain (Bjornstad, Austdal, Roald, Glover, & Paulsen, 2015; Close, et al.,  
220 1997; Mellor & Diesch, 2007). This indicates that chicken embryos can be euthanized up to 11 days after  
221 laying, before they can perceive pain. In addition, we found the collection of allantoic fluid from an  
222 individual egg during day-9 of the development did not lead to a loss of hatchability of the embryos. For  
223 ethical reasons, we choose to focus on day-9 and disregarded later days. A univariate receiver operating



224 characteristic (ROC) analysis for F1558 (shown in Figure 1B) shows a sensitivity of 96.4% ( $IC_{95\%} = 89.3-100\%$ )  
225 and a specificity of 92.9% ( $IC_{95\%} = 85.7-100\%$ ) using a peak area threshold of 35,000 counts.  
226 The first experiment identified a potential biomarker of gender from day-8 to 11 in H&N Brown Nick  
227 chickens. To evaluate the prediction capacity of this feature in different chicken lines, a new experiment  
228 was performed with the analysis of a breed-cohort of allantoic fluid coming from Brown Nick chickens and  
229 white Hy-Line chickens (146 and 151 samples respectively) collected at day-9. As shown in Figure 2, we did  
230 not find any statistical difference of the average biomarker level in the overall population of each chicken  
231 line ( $p_{0.05} = 0.52$ ). When considering brown and white eggs together, we find the endogenous level of the  
232 biomarker was significantly different between male and female eggs ( $p_{0.05} = 1.25 \times 10^{-44}$ ). It has thus been  
233 possible to set a common intensity threshold, independent of the chicken lines, giving a prediction  
234 accuracy above 89.5%.



235  
236  
237

Figure 2: Box-plot comparison of the white and brown eggs populations

238 These results clearly highlight the strength of the discovered feature for early gender determination of  
239 chicken embryos 9 days after the egg-laying regardless of chicken species.

240

### 241 3.3. 3-[(2-aminoethyl)sulfanyl]butanoic acid was identified as a gender-specific biomarker

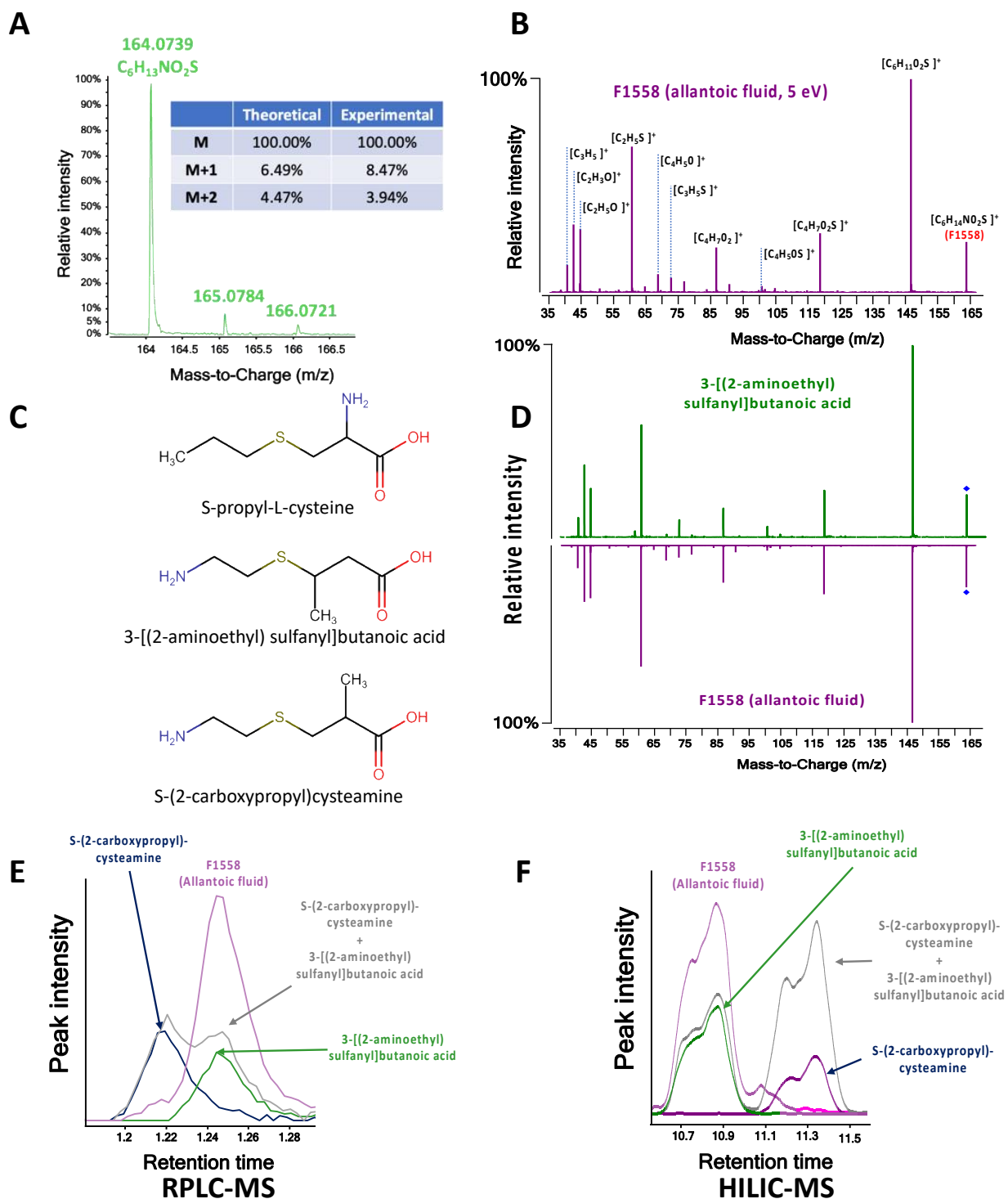
242 Unambiguous identification of the biomarker requires the comparison of the feature of interest to  
243 reference compounds using different characterization approaches to obtain orthogonal information.

244 Therefore, for confident annotation of an unknown feature, at least 2 independent parameters need  
245 to match the reference compound for the feature to be identified, such as liquid chromatography  
246 retention time, accurate mass or mass spectrometry fragmentation pattern (MSMS spectra) (Rochat,  
247 2017; Sumner, Amberg, Barrett, Beale, Beger, Daykin, et al., 2007).

248 The exact mass of the protonated form of F1558 has been accurately measured as m/z 164.0739.  
249 This mass was assigned by Agilent MassHunter as corresponding to the chemical formula C<sub>6</sub>H<sub>13</sub>NO<sub>2</sub>S  
250 which was confirmed by the isotopic distribution showing the specific pattern of sulfur atoms (Figure  
251 3A). In order to identify the biomarker, we undertook the identification of its MSMS fragments  
252 obtained after RPLC-MSMS of a pooled allantoinic sample (Figure 3B). Based on the fragmentation  
253 pattern, we proposed different structural isomers, whereby S-propyl-L-cysteine, S-(2-carboxypropyl)-  
254 cysteamine, and 3-[(2-aminoethyl)sulfanyl] butanoic acid were the most likely to produce the same  
255 fragments (Figure 3C). Since they were not commercially available, these compounds were  
synthesized on our request.

256

257



258  
 259  
 260 *Figure 3: Identification process of F1558 (A: determination of the elemental composition, B: MSMS spectra of F1558 and*  
 261 *elemental composition of fragments, C: possible structures, D: comparison of fragmentation pattern of F1558 and 3-[(2-*  
 262 *aminoethyl)sulfanyl] butanoic acid, E-F: retention profile of F1558, S-propyl-L-cysteine, S-(2-carboxypropyl)-cysteamine and 3-[(2-*  
 263 *aminoethyl)sulfanyl] butanoic acid in RPLC and HILIC mode respectively)*  
 264

265 The fragmentation pattern of the biomarker F1558 was then compared to the MSMS spectra of the  
266 three candidate compounds. It quickly appeared that 3-[(2-aminoethyl)sulfanyl] butanoic acid  
267 possessed 11 common fragments with F1558 (Figure 3D) whereas the two other compounds only had 2  
268 (Table S1). Additional confirmation of the identity was obtained by comparing retention time profiles  
269 in two orthogonal chromatography methods. After injection in RPLC mode, S-propyl-cysteine was  
270 quickly discarded from the list since its retention time (RT) differed by more than one minute from  
271 F1558 (data not shown). The RT obtained for 3-[(2-aminoethyl)sulfanyl] butanoic acid and S-(2-  
272 carboxypropyl)-cysteamine were slightly different, and two peaks were clearly observed when co-  
273 injected (Figure 3E). We found the RT of 3-[(2-aminoethyl)sulfanyl] butanoic acid was exactly the same  
274 as F1558. Second, we undertook a HILIC comparison. As shown in Figure 3F, in this mode, 3-[(2-  
275 aminoethyl)sulfanyl] butanoic acid and S-(2-carboxypropyl)-cysteamine gave two baseline separated  
276 peaks. The peak shape and the RT of 3-[(2-aminoethyl)sulfanyl] butanoic acid were strictly identical to  
277 F1558. Therefore, combining all these orthogonal data together, we were able to unequivocally  
278 identify 3-[(2-aminoethyl)sulfanyl]butanoic acid (ASBA) as being the feature of interest F1558.

#### 278 **3.4. ASBA shows high sensitivity and specificity for gender prediction**

279 Based on a small discovery cohort using untargeted metabolomics profiling of allantoic fluid at day-9,  
280 we identified ASBA as a potential biomarker for *in ovo* gender prediction of chicken eggs. In order to  
281 validate our hypothesis and to determine the prediction power of ASBA, we investigated a  
282 validation cohort including 143 allantoic fluid samples collected from Brown Nick eggs at day-9.  
283 Untargeted metabolomics methods are known to have reduced accuracy due to matrix effect and  
284 low peak abundance. For this reason, we decided to validate the candidate biomarker using a targeted  
285 LC-APCI-MRM method and using the deuterated version of the metabolite as internal standard.  
286  
287  
288  
289  
290

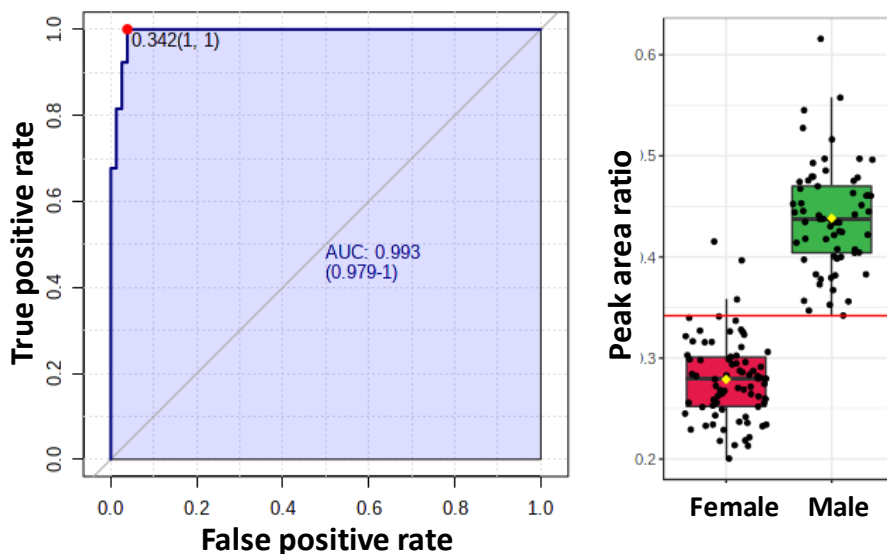


Figure 4: ROC and box plots obtained with targeted LC-APCI-MRM analysis.

291  
292  
293  
294 As highlighted in Figure 4 by the increase of the AUC of the ROC plot, the improvement of the analytical  
295 method led to an increase of the prediction accuracy, leading to a total separation of the male embryos  
296 with a sensitivity of 100% ( $IC_{95\%} = 100-100\%$ ) and a specificity of 96.2% ( $IC_{95\%} = 92.3-100\%$ ) for a peak area  
297 ratio threshold established at 0.342.

298 To study the prediction performance of the model, we so far have used a male/female cut-off determined  
299 *a posteriori* and based on true gender. Obviously, this is not possible in routine analysis and we attempt  
300 to set a fixed cut-off based on QC samples. Indeed, the theoretical distribution of males and females in a  
301 population is a Gaussian curve with an average close to 50%. Therefore, we used the peak area ratio of QC  
302 samples (i.e. 0.345) prepared from equal volumes of every sample as threshold for the gender  
303 determination. Using this approach, the model was able to predict the male gender with a sensitivity of  
304 98.5% ( $IC_{95\%} = 96.1-100\%$ ) and a specificity of 96.2% ( $IC_{95\%} = 92.3-99.4\%$ ).

#### 305 306 **4. Discussion**

307 With the changing mindset and the rapid evolution of legislation, finding alternatives to the culling of day-  
308 old male chicks has become a societal issue. In this context, finding a method to *in ovo* discriminate male  
309 and female chickens, before the full development of the central nervous system appears to be a humane  
310 solution to this problem. Applying untargeted LC-MS approaches on limited sample cohorts, we were able  
311 to identify a candidate biomarker feature, and with it able to accurately predict the gender of chicken  
312 embryos from day-8 to 11 after egg-laying. We have also demonstrated the prediction accuracy of this  
313 candidate biomarker across two different chicken breeds. After its identification, we have been able to

314 confirm that 3-[(2-aminoethyl)sulfanyl]butanoic acid (ASBA) is a powerful biomarker for early gender  
315 screening. Surprisingly, this is the first report of this gender-specific biomarker in chickens or any living  
316 species. Finally, we have validated this biomarker using a targeted LC-APCI-MRM method on more than  
317 140 samples and we showed by only diluting the raw allantoic fluid with the corresponding deuterated  
318 compound, we were able to improve the prediction power of the method to a sensitivity of 100% and a  
319 specificity of 96.2%. Nevertheless, a posteriori threshold is not possible in routine analysis. To circumvent  
320 this issue, we have shown the approach based on a pooled QC sample and the theoretical distribution of  
321 the gender in a population was able to reach a gender prediction accuracy of 97.2 %. Beyond the discovery  
322 of ASBA as a powerful early biomarker of chick gender, we successfully developed a high throughput  
323 gender screening workflow able to gender screen more than 2000 eggs per our with a prediction accuracy  
324 above 95%. This method, using automation and the acoustic droplet ejection hyphenated with mass  
325 spectrometry (ADE-MS) is the object of a separate research article (ref) and is now commercialized by In  
326 Ovo (Leiden, the Netherlands).

327

### 328 **Acknowledgment**

329 In Ovo Holding B.V. acknowledges support from the Dutch Ministry of Economic Affairs. Alida Kindt-  
330 Dunjko is gratefully acknowledged for her support with the data processing.

331

### 332 **CrediT authorship contribution statement**

333 Nicolas DROUIN: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Data  
334 curation, Writing – original draft, Writing – Review & Editing, Visualization

335 Hyung Lim ELFRINK: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Data  
336 curation, Writing – original draft, Writing – Review & Editing, Visualization

337 Slavik KOVAL: Data curation

338 Amy HARMS: Conceptualization, Methodology, Resources, Data curation, Supervision, Writing – original  
339 draft, Writing – Review & Editing, Visualization, Project administration, Funding acquisition

340 Will STUTTERHEIM: Resources, Project administration, Funding acquisition

341 Wouter BRUINS: Resources, Project administration, Funding acquisition

342 Thomas HANKEMEIER: Conceptualization, Methodology, Resources, Data curation, Supervision, Writing –  
343 original draft, Writing – Review & Editing, Visualization, Project administration, Funding acquisition

344

345

## 346 References

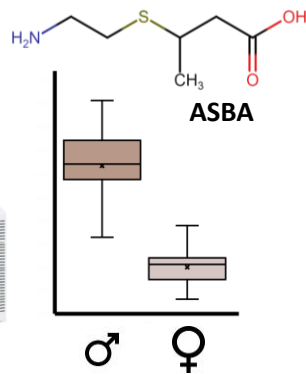
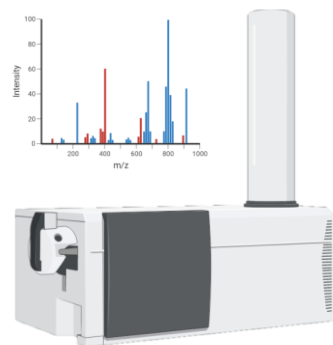
- 347 Bjornstad, S., Austdal, L. P., Roald, B., Glover, J. C., & Paulsen, R. E. (2015). Cracking the Egg: Potential of  
348 the Developing Chicken as a Model System for Nonclinical Safety Studies of Pharmaceuticals. *J*  
349 *Pharmacol Exp Ther*, 355(3), 386-396. <https://doi.org/10.1124/jpet.115.227025>
- 350 Blide, D., & Montel, S., (2015), MOTION FOR A RESOLUTION, European Parliament, B8-0625/2015
- 351 Close, B., Banister, K., Baumans, V., Bernoth, E. M., Bromage, N., Bunyan, J., Erhardt, W., Flecknell, P.,  
352 Gregory, N., Hackbarth, H., Morton, D., & Warwick, C. (1997). Recommendations for euthanasia  
353 of experimental animals: Part 2. DGXT of the European Commission. *Laboratory animals*, 31(1), 1-  
354 32. <https://doi.org/10.1258/002367797780600297>
- 355 Commion de la Science , d. l. E. e. d. l. C. C., (2019), Arrêtons le broyage des poussins vivants, Montion  
356 19.3003
- 357 Drouin, N., Pezzatti, J., Gagnebin, Y., Gonzalez-Ruiz, V., Schappler, J., & Rudaz, S. (2018). Effective mobility  
358 as a robust criterion for compound annotation and identification in metabolomics: Toward a  
359 mobility-based library. *Anal Chim Acta*, 1032, 178-187. <https://doi.org/10.1016/j.aca.2018.05.063>
- 360 Galli, R., Preusse, G., Schnabel, C., Bartels, T., Cramer, K., Krautwald-Junghanns, M. E., Koch, E., & Steiner,  
361 G. (2018). Sexing of chicken eggs by fluorescence and Raman spectroscopy through the shell  
362 membrane. *PLoS One*, 13(2), e0192554. <https://doi.org/10.1371/journal.pone.0192554>
- 363 Galli, R., Preusse, G., Uckermann, O., Bartels, T., Krautwald-Junghanns, M. E., Koch, E., & Steiner, G. (2016).  
364 In Ovo Sexing of Domestic Chicken Eggs by Raman Spectroscopy. *Anal Chem*, 88(17), 8657-8663.  
365 <https://doi.org/10.1021/acs.analchem.6b01868>
- 366 Galli, R., Preusse, G., Uckermann, O., Bartels, T., Krautwald-Junghanns, M. E., Koch, E., & Steiner, G. (2017).  
367 In ovo sexing of chicken eggs by fluorescence spectroscopy. *Anal Bioanal Chem*, 409(5), 1185-  
368 1194. <https://doi.org/10.1007/s00216-016-0116-6>
- 369 He, L., Martins, P., Huguenin, J., Van, T. N., Manso, T., Galindo, T., Gregoire, F., Catherinot, L., Molina, F.,  
370 & Espeut, J. (2019). Simple, sensitive and robust chicken specific sexing assays, compliant with  
371 large scale analysis. *PLoS One*, 14(3), e0213033. <https://doi.org/10.1371/journal.pone.0213033>
- 372 Health, E. P. o. A., Welfare, Nielsen, S. S., Alvarez, J., Bicout, D. J., Calistri, P., Depner, K., Drewe, J. A., Garin-  
373 Bastuji, B., Gonzales Rojas, J. L., Gortazar Schmidt, C., Miranda Chueca, M. A., Roberts, H. C.,  
374 Sihvonen, L. H., Spoolder, H., Stahl, K., Velarde Calvo, A., Viltrop, A., Winckler, C., Candiani, D.,  
375 Fabris, C., Van der Stede, Y., & Michel, V. (2019). Killing for purposes other than slaughter: poultry.  
376 *EFSA J*, 17(11), e05850. <https://doi.org/10.2903/j.efsa.2019.5850>
- 377 Kahn, N. W., St John, J., & Quinn, T. W. (1998). Chromosome-specific intron size differences in the avian  
378 CHD gene provide an efficient method for sex identification in birds. *Auk*, 115(4), 1074-1078.  
379 <https://doi.org/10.2307/4089527>
- 380 Koller, D., & Sahami, M. (1996). Toward Optimal Feature Selection. In S. InfoLab (Ed.).
- 381 l'Alimentation, M. d. l. A. e. d., (2020), Annonces de Didier Guillaume 15 mesures pour renforcer la lutte  
382 contre la maltraitance animale et améliorer le bien-être des animaux domestiques, Press  
383 Communicate
- 384 Laan, T. van der, Elfrink, H., Azadi-Chegeni, F., Dubbelman, A-C., Harms, A. C., Jacobs, D. M., Braumann, U.,  
385 Velders, A. H., Duynhoven, J. van, Hankemeier, T., (2021). Fractionation platform for target  
386 identification using off-line directed two-dimensional chromatography, mass spectrometry and  
387 nuclear magnetic resonance. *Analytica Chimica Acta*, 1142, 28-37.  
388 <https://doi.org/10.1016/j.aca.2020.10.054>.
- 389 MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L.,  
390 Liebler, D. C., & MacCoss, M. J. (2010). Skyline: an open source document editor for creating and  
391 analyzing targeted proteomics experiments. *Bioinformatics*, 26(7), 966-968.  
392 <https://doi.org/10.1093/bioinformatics/btq054>

- 393 Mellor, D. J., & Diesch, T. J. (2007). Birth and hatching: key events in the onset of awareness in the lamb  
394 and chick. *N Z Vet J*, 55(2), 51-60. <https://doi.org/10.1080/00480169.2007.36742>
- 395 Noga, M. J., Dane, A., Shi, S., Attali, A., van Aken, H., Suidgeest, E., Tuinstra, T., Muilwijk, B., Coulier, L.,  
396 Luider, T., Reijmers, T. H., Vreeken, R. J., & Hankemeier, T. (2012). Metabolomics of cerebrospinal  
397 fluid reveals changes in the central nervous system metabolism in a rat model of multiple sclerosis.  
398 *Metabolomics*, 8(2), 253-263. <https://doi.org/10.1007/s11306-011-0306-3>
- 399 Pezzatti, J., Gonzalez-Ruiz, V., Codesido, S., Gagnebin, Y., Joshi, A., Guillarme, D., Schappler, J., Picard, D.,  
400 Bocard, J., & Rudaz, S. (2019). A scoring approach for multi-platform acquisition in metabolomics.  
401 *J Chromatogr A*, 1592, 47-54. <https://doi.org/10.1016/j.chroma.2019.01.023>
- 402 Producers, U. E., (2020), United Egg Producers updated statement on male chick culling, Retrived from  
403 <https://unitedegg.com/united-egg-producers-updated-statement-on-male-chick-culling/>.  
404 Accessed February, 14th 2020
- 405 Rochat, B. (2017). Proposed Confidence Scale and ID Score in the Identification of Known-Unknown  
406 Compounds Using High Resolution MS Data. *J Am Soc Mass Spectrom*, 28(4), 709-723.  
407 <https://doi.org/10.1007/s13361-016-1556-0>
- 408 Steiner, G., Bartels, T., Stelling, A., Krautwald-Junghanns, M. E., Fuhrmann, H., Sablinskas, V., & Koch, E.  
409 (2011). Gender determination of fertilized unincubated chicken eggs by infrared spectroscopic  
410 imaging. *Anal Bioanal Chem*, 400(9), 2775-2782. <https://doi.org/10.1007/s00216-011-4941-3>
- 411 Steyerberg, E. W., Harrell, F. E., Jr., Borsboom, G. J., Eijkemans, M. J., Vergouwe, Y., & Habbema, J. D.  
412 (2001). Internal validation of predictive models: efficiency of some procedures for logistic  
413 regression analysis. *Journal of clinical epidemiology*, 54(8), 774-781.  
414 [https://doi.org/10.1016/s0895-4356\(01\)00341-9](https://doi.org/10.1016/s0895-4356(01)00341-9)
- 415 Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W., Fiehn, O., Goodacre,  
416 R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J.  
417 C., Marriott, P., Nicholls, A. W., Reily, M. D., Thaden, J. J., & Viant, M. R. (2007). Proposed minimum  
418 reporting standards for chemical analysis Chemical Analysis Working Group (CAWG)  
419 Metabolomics Standards Initiative (MSI). *Metabolomics*, 3(3), 211-221.  
420 <https://doi.org/10.1007/s11306-007-0082-2>
- 421 UNION, T. C. O. T. E., (2009), Council Regulation (EC) No 1099/2009 of 24 September 2009 on the  
422 protection of animals at the time of killing 32009R1099
- 423 WATTPoultry. (2019). *Poultry trends 2019: The Statistical Reference for Poultry Executives*: Watt Executive  
424 Guide to World Poultry Trends.
- 425 Weissmann, A., Reitemeier, S., Hahn, A., Gottschalk, J., & Einspanier, A. (2013). Sexing domestic chicken  
426 before hatch: a new method for in ovo gender identification. *Theriogenology*, 80(3), 199-205.  
427 <https://doi.org/10.1016/j.theriogenology.2013.04.014>
- 428 Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D., & Wishart, D. S. (2012). MetaboAnalyst 2.0--a  
429 comprehensive server for metabolomic data analysis. *Nucleic Acids Research*, 40(Web Server  
430 issue), W127-W133. <https://doi.org/10.1093/nar/gks374>

431

432 **Graphical abstract**





433