



HAL
open science

The extent of mRNA editing is limited in chicken liver and adipose, but impacted by tissular context, genotype, age and feeding as exemplified with a conserved edited site in COG3

Pierre-François Roux, Laure Fresard, Morgane Boutin, Sophie Leroux, Christophe C. Klopp, Anis Djari, Diane Esquerre, Pascal G.P. Martin, Tatiana Zerjal, David Gourichon, et al.

► To cite this version:

Pierre-François Roux, Laure Fresard, Morgane Boutin, Sophie Leroux, Christophe C. Klopp, et al.. The extent of mRNA editing is limited in chicken liver and adipose, but impacted by tissular context, genotype, age and feeding as exemplified with a conserved edited site in COG3. *G3*, 2016, 6 (2), pp.321-335. 10.1534/g3.115.022251 . hal-01251260

HAL Id: hal-01251260

<https://hal.science/hal-01251260>

Submitted on 5 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 **The extent of mRNA editing is limited in chicken liver and adipose,**
 2 **but impacted by tissular context, genotype, age and feeding as**
 3 **exemplified with a conserved edited site in COG3**

4 Pierre-François Roux^{*,†,1,3}, Laure Frésard^{‡,§,**,1,4}, Morgane Boutin^{*,†}, Sophie Leroux^{‡,§,**},
 5 Christophe Klopp^{††}, Anis Djari^{‡,§,**,††}, Diane Esquerre^{‡,§,**,‡‡}, Pascal GP Martin^{§§}, Tatiana
 6 Zerjal^{***}, David Gourichon^{†††}, Frédérique Pitel^{‡,§,**,2}, Sandrine Lagarrigue^{*,†,2}

7 * Agrocampus Ouest, UMR1348 Physiologie, Environnement et Génétique pour l'Animal et
 8 les Systèmes d'Élevage (PEGASE), F-35000 Rennes, France

9 † INRA, UMR1348 PEGASE, F-35000 Rennes, France

10 ‡ UMR INRA / § INPT ENSAT / ** INPT ENVT, Génétique, Physiologie et Systèmes d'élevage,
 11 INRA, F-31326 Castanet Tolosan, France

12 †† INRA, Sigenae UR875 Biométrie et Intelligence Artificielle, F-31326 Castanet-Tolosan,
 13 France,

14 ‡‡ INRA, GeT-PlaGe Genotoul, F-31326 Castanet-Tolosan, France,

15 §§ INRA, UMR1331 Research Center in Food Toxicology (ToxAlim), Toulouse, France,

16 *** INRA, UMR1313 Génétique Animale et Biologie Intégrative, F-78350 Jouy-en-Josas,
 17 France,

18 ††† INRA, PEAT Pôle d'Expérimentation Avicole de Tours, F- 37380 Nouzilly, France

19

¹ These authors contributed equally to this work.

² These authors contributed equally to this work.

³ Present address: Laboratory of Nuclear Organization and Oncogenesis, Pasteur Institute, Paris, France

⁴ Present address: Department of Genetics, Stanford University, Stanford, CA, USA

20 **RUNING TITLE**

21 RNA editing in chicken liver and adipose tissue

22 **KEYWORDS**

23 mRNA editing; chicken; RNA-seq; DNA-seq; liver and adipose tissue

24 **CORRESPONDING AUTHORS**

25 Sandrine LAGARRIGUE

26 UMR1349, Laboratoire de Génétique et Génomique, 65 rue de Saint Briec, F-35042

27 Rennes, France.

28 E-mail : sandrine.lagarrigue@agrocampus-ouest.fr

29 Phone : + 33 2 23 48 59 59

30 Frédérique PITEL

31 UMR444, Laboratoire de Génétique Cellulaire, 24 chemin de Borde-Rouge, F-31326

32 Auzeville, France.

33 E-mail : frederique.pitel@toulouse.inra.fr

34 Phone : +33 5 61 28 54 35

35

36 **ABSTRACT**

37 RNA editing is a post-transcriptional process leading to differences between genomic DNA
38 and transcript sequences, potentially enhancing transcriptome diversity. With recent
39 advances in high-throughput sequencing, many efforts have been made to describe mRNA
40 editing at the transcriptome scale, especially in mammals, raising contradictory conclusions
41 regarding the extent of this phenomenon. We show by a detailed description of the 25
42 studies focusing so far on mRNA editing at the whole-transcriptome scale that systematic
43 sequencing artifacts are somehow considered in most studies, whereas biological
44 replication is often neglected and multi-alignment not properly evaluated, which ultimately
45 impairs the legitimacy of results. We recently developed a rigorous strategy to identify
46 mRNA editing using mRNA and genomic DNA sequencings, taking into account sequencing
47 and mapping artifacts, and biological replicates. We applied this method to screen for mRNA
48 editing in liver and white adipose tissue from 8 chickens and confirm the small extent of
49 mRNA recoding in this species. Among the 25 unique edited sites identified, 3 events were
50 previously described in mammals, attesting this phenomenon is conserved through
51 evolution. Deeper investigations on 5 sites revealed the impact of tissular context,
52 genotype, age, feeding condition and sex on mRNA editing levels. More specifically, this
53 analysis highlighted that the editing level at the site located on *COG3* was strongly regulated
54 by 4 out of these factors. By comprehensively characterizing the mRNA editing landscape in
55 chicken, our results highlight how this phenomenon is limited, and suggest a regulation of
56 editing levels by various genetic and environmental factors.

57 INTRODUCTION

58 RNA editing has become a generic term for a wide array of post-transcriptional processes
59 changing the mature RNA sequence relatively to the corresponding encoding genomic DNA
60 matrix. This phenomenon, which is almost limited to eukaryotes with some exceptions, is
61 characterized by nucleotide insertion, deletion or substitution in various types of RNAs,
62 including mRNAs (Knoop 2011), tRNAs (Börner *et al.* 1996, Gott *et al.* 2010, Torres *et al.*
63 2014), miRNAs (Warnefors *et al.* 2014) and rRNAs (Eifler *et al.* 2013, Valach *et al.* 2014) and
64 is likely to contribute to RNA diversity. Until recently this mechanism was considered
65 relatively rare in vertebrates, mainly restricted to brain-specific substrates and repetitive
66 regions of the genome (Bass 2002), and limited to extensively validated ADAR-mediated
67 adenosine to inosine (A-to-I) substitutions and APOBEC-mediated cytosine to uracile (C-to-
68 U) changes (Knoop 2011).

69 Since 2009, the advents of high-throughput sequencing technologies enabled to study this
70 phenomenon at a transcriptome-wide scale and progressively challenged this view with
71 estimates ranging from several hundred (Ju *et al.* 2011, Kleinman *et al.* 2012) to several
72 thousand (Clop *et al.* 2006, Li *et al.* 2011, Bahn *et al.* 2011, Peng *et al.* 2012, Ramaswami *et*
73 *al.* 2012, Park *et al.* 2012, Chen 2013, Kang *et al.* 2015) and even million (Bazak *et al.* 2013)
74 mRNA edited sites throughout mammalian genomes. According to some of these mRNA
75 editing screening studies, mRNA recoding appears as an extremely common process greatly
76 contributing to transcript diversity. Furthermore, most of these studies report mRNA editing
77 events leading to transversions that cannot be explained in the light of our current
78 knowledge regarding the molecular bases of mRNA recoding (Li *et al.* 2011, Ju *et al.* 2011,
79 Bahn *et al.* 2011, Peng *et al.* 2012, Chen 2013, Kang *et al.* 2015), suggesting the existence of
80 so far uncharacterized mRNA editing mechanisms, and novel molecular components implied

81 in gene expression regulation. The conclusions raised by these studies regarding the extent
82 and the nature of mRNA recoding, if further supported, would deeply impact our
83 understandings of gene expression regulation and transcriptional modification.

84 Facing contradictory results regarding the extent of mRNA editing, a large number of studies
85 and comments pointed the requirement of comprehensive and rigorous bioinformatics
86 pipelines to limit technical artifacts in the editome characterizations (Schrider *et al.* 2011,
87 Kleinman and Majewski 2012, Lin *et al.* 2012, Pickrell *et al.* 2012, Kleinman *et al.* 2012, Piskol
88 *et al.* 2013, Lagarrigue *et al.* 2013). Working with short-read sequencing data for the
89 detection of polymorphisms requires to carefully deal with technical artifacts related to
90 mapping on paralogous or repetitive regions (Malhis and Jones 2010, Treangen and Salzberg
91 2012), mapping errors at splice sites (Park *et al.* 2012), or systematic and random
92 sequencing errors (Nakamura *et al.* 2011, Meacham *et al.* 2011). This is especially the case
93 when screening for mRNA editing events, since all of these artifacts are likely to generate
94 artificial discrepancies between genomic DNA and mRNA further interpreted as edited sites.
95 In this context, the huge variation regarding the extent of intra-tissue and intra-species
96 mRNA editing revealed in literature could be in part due to the varying level of stringency of
97 bioinformatics filters used to control these error prone artifacts, and whether biological
98 replication is considered or not.

99 As shown in Table 1, most of the 25 RNA-seq based mRNA editing screening studies
100 performed on vertebrates are not considering matched genomic DNA sequences to detect
101 mRNA recoding, but rather consider either a consensus genomic sequence for the species
102 studied, or EST databases to remove false positives arising from potential genomic
103 polymorphisms, therefore occulting unreferenced individual variations (Bahn *et al.* 2011,

104 Danecek *et al.* 2012, Gu *et al.* 2012, Ramaswami *et al.* 2012, Cattenoz *et al.* 2013, Lagarrigue
105 *et al.* 2013, Chen 2013, Bazak *et al.* 2013, Blanc *et al.* 2014, Toung *et al.* 2014, Sakurai *et al.*
106 2014, Chan *et al.* 2014, Zhang and Xiao 2015). More strikingly, while it is fully admitted that
107 filtering on minor allele frequency is required to select high-quality genomic polymorphisms
108 (Consortium 2010, 1000 Genomes Project Consortium *et al.* 2012), some mRNA editing
109 screening studies still consider that reproducibility across biological replicates is not a
110 mandatory criterion for considering a difference between DNA and RNA as a reliable editing
111 event (Picardi *et al.* 2012, Peng *et al.* 2012, Ramaswami *et al.* 2012, Kleinman *et al.* 2012,
112 Park *et al.* 2012, Cattenoz *et al.* 2013, Chen 2013, Bazak *et al.* 2013, Sakurai *et al.* 2014, Chen
113 *et al.* 2014, Chan *et al.* 2014, Kang *et al.* 2015, Zhang and Xiao 2015). However, as depicted
114 on Figure 1, considering biological replication clearly impacts the total number of editing
115 events detected in high-throughput-based screening studies, since the number of
116 differences between DNA and RNA reported appears to be directly negatively correlated
117 with the number of biological replicates considered. From a methodological point of view,
118 this study proposes a rigorous strategy to identify mRNA editing using both mRNA and
119 genomic DNA high-throughput sequencings, taking into account sequencing and mapping
120 artifacts, as well as biological replicates, to control the false positive rate. The efficiency of
121 this approach has already been validated in our previous study on chicken embryo mRNA
122 editing (Frésard *et al.* 2015). To strictly control multi-mapping, we looked for mRNA
123 sequences spanning edited sites in unmapped genomic DNA sequences, allowing to
124 consider potential errors and gaps in the reference assembly that still represent roughly
125 15% of the chicken genome (Wicker *et al.* 2005, Schmid *et al.* 2015).

126 From a biological perspective, in addition to our recent work screening mRNA editing in
127 chicken whole-embryo (Frésard *et al.* 2015), this study is an answer to the evident lack of
128 transcriptome-wide mRNA editing screening investigations focusing on non-mammalian
129 vertebrates such as birds, that could contribute to our understanding of RNA editing
130 evolutionary bases. Indeed, as depicted in Table 1, except our works, all mRNA editing
131 screening studies in vertebrates are focused on primates or mouse transcriptomes. To date,
132 the origins of RNA editing are still rather obscure, and even if it is currently proposed that it
133 may have arisen several times in different phyla throughout evolution, it remains unclear
134 whether selection was involved or not (Gommans *et al.* 2009, Gray *et al.* 2010, Gray 2012).
135 While chicken is extensively used as a model organism in developmental biology (Davey and
136 Tickle 2007), it also bridges the evolutionary gap between mammals and other vertebrates
137 and therefore stands as an ideal species to explore conservation of mRNA editing events in
138 vertebrates throughout evolution. In addition, our knowledge related to the regulation of
139 mRNA editing level, and to factors enhancing or repressing mRNA recoding is still limited.
140 Hitherto, few studies has been carried out to assess whether the genetic background, the
141 sex, the feeding condition or the age influence mRNA recoding level. Most of these studies
142 targeted the extensively studied APOBEC-mediated C-to-U editing event occurring in
143 mammalian *APOBEC1* mRNA revealing the influence of ethanol intake (Lau *et al.* 1995, Van
144 Mater *et al.* 1998), insulin (Wronski *et al.* 1998), obesity (Phung *et al.* 1996), and diet
145 (Funahashi *et al.* 1995) on *APOBEC1* mRNA editing level. Others focused on previously
146 described ADAR-mediated editing events in primate (Li *et al.* 2013), mouse (Gan *et al.* 2006)
147 or rat (Holmes *et al.* 2013) transcriptomes, highlighting an insulin-dependent activity of
148 ADAR in mouse pancreas (Gan *et al.* 2006), or suggesting influence of aging on ADAR-
149 mediated mRNA editing in human, mouse and pig (Wahlstedt *et al.* 2009, Shtrichman *et al.*

150 2012, Venø *et al.* 2012). A better characterization of environmental and genetic factors
151 influencing the level of mRNA recoding would definitely offer new insights on the role of
152 mRNA editing in vertebrates.

153 In this study, we report the results from the first genome-wide characterization of the
154 chicken liver and adipose mRNA editomes, based on both genomic DNA and mRNA high-
155 throughput sequencings. Our results confirm the low extent of mRNA recoding in chicken
156 and the absence of non A-to-I editing event in this species, in agreement with what has
157 already been shown for chicken embryos (Frésard *et al.* 2015). We also highlight that mRNA
158 editing level is impacted by genetic and environmental factors such as tissular context,
159 genotype, age, and, to a minor extent, by feeding condition and sex. As exemplified with the
160 recoding event located on *COG3* and confirmed at other positions, the mRNA editing level is
161 tightly depending on several environmental and genetics factors.

162

163 MATERIALS AND METHODS

164 Ethics Statement

165 Chickens were bred at INRA, UE1295 Pôle d'Expérimentation Avicole de Tours, F-37380
166 Nouzilly in accordance with European Union Guidelines for animal care, following the
167 Council Directives 98/58/EC and 86/609/EEC. Animals were maintained under standard
168 breeding conditions, and subjected to minimal disturbance. The farm is registered with the
169 French Ministry of Agriculture under license number C37-175-1 for animal
170 experimentation. The experiment was performed under authorization 37-002 delivered to
171 D. Gourichon.

172 Tissue collection, and library preparation

173 Two experimental meat-type chicken lines were divergently selected for seven generations
174 using the ratio between abdominal fat weight and whole animal weight at 9 weeks as a
175 fattening index, while maintaining live body weight constant (Leclercq *et al.* 1980). After
176 selection, the two lines were maintained by carefully limiting inbreeding. Four 9 week-old
177 males from the 35th generation in each line were slaughtered by electronarcosis and
178 immediate bleeding. Liver and abdominal adipose tissue were then harvested and stored in
179 nitrogen. Liver genomic DNA and total liver and adipose RNA were concurrently extracted
180 according to the manufacturer's instructions using the AllPrep DNA/RNA Mini Kit (Agilent,
181 Agilent Technologies, Santa Clara, CA). RNA quality was assessed on a BioAnalyzer 1000
182 (Agilent Technologies, Santa Clara, CA) and RIN (RNA Integrity Number) ≥ 9 were required.

183 Sequencing

184 RNA sequencing

185 Libraries with a mean insert size of 200 bp were prepared according to the manufacturer's
186 instructions for RNA-seq library preparation, selecting polyadenylated mRNA using the
187 TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA) from each sample. Samples were
188 tagged using a barcode sequence for subsequent identification, amplified by PCR and
189 quantified by qPCR using the QPCR Library Quantification Kit (Agilent Technologies, Santa
190 Clara, CA). A total of 16 libraries were sequenced in paired-ends 2 x 101 bp in triplicates on 3
191 different lanes on the Illumina HiSeq 2000 sequencer using the TruSeq PE Cluster Kit v3
192 (Illumina, San Diego, CA), the cBot SBS Kit v3 (Illumina, San Diego, CA) and the TruSeq SBS
193 Kit v3 (Illumina, San Diego, CA). After quality checks and adapter trimming using CASAVA
194 1.8, matched libraries for a given sample were merged. Liver and adipose mRNA-seq raw
195 data are available on Sequence Read Archive under accession SRP042257.

196 *DNA sequencing*

197 Liver DNA from the 8 animals was sequenced in paired-ends 2 x 101 bp on 4 lanes on an
198 Illumina HiSeq 2000. Library preparation, DNA quantification and sequencing were
199 performed according to manufacturers' instructions using TruSeq DNA Sample Prep Kit
200 (Illumina, San Diego, CA), Agilent QPCR Library Quantification Kit (Agilent Technologies,
201 Santa Clara, CA), TruSeq PE Cluster Kit v3 (Illumina, San Diego, CA) and cBot TruSeq SBS Kit
202 v3 (Illumina, San Diego, CA). After quality checks, and adapter trimming using CASAVA 1.8,
203 matched libraries for a given sample were merged. Liver genomic DNA-seq raw data are
204 available on Sequence Read Archive under accession SRP042641.

205 **Computational analyses**

206 When not specified, analyses were performed with in-house Perl and R scripts.

207 *Genomic sequences analyses*

208 DNA sequences were aligned to the latest chicken genome assembly (Galgal4) using *BWA*
209 *v0.7.0* (Li and Durbin 2009) (*Command: bwa aln*). Sequences were then filtered based on
210 mapping quality (*Command: samtools view -bS -q 30*). *SAMtools v0.1.19* (Li *et al.* 2009a)
211 *rmdup* (*Command: samtools rmdup*) was used to remove possible PCR and optical
212 duplicates.

213 *mRNA sequences analyses*

214 mRNA sequences were aligned with *Tophat v2.0.5* (Kim *et al.* 2013) on the chicken reference
215 genome Galgal4 as described in (Frésard *et al.* 2014) (*Command: tophat --min-intron-length*
216 *3 --max-intron-length 25000 --max-deletion-length 1 -mate-inner-dist 200 --read-realign-*
217 *edit-dist 0 --microexon-search*). Uniquely mapped unduplicated sequences with a mapping
218 quality greater than 30 were selected, using *SAMtools v.0.1.19* (*Command: samtools view -*
219 *bS -q 30*) and in-house Python script.

220 **Identification of mRNA editing candidates**

221 Sequences were locally realigned and recalibrated before SNP detection, with *GATK v1.6.11*
222 for DNA (Van der Auwera *et al.* 2013) (*Commands: GATK -T RealignerTargetCreator -R; GATK*
223 *-T BaseRecalibrator -R -knownSites; GATK -T PrintReads -R -BQSR*), and *BamUtil* (*Command:*
224 *bam recab*) for RNA.

225 *SAMtools v0.1.19 mpileup* was used to detect SNPs between DNA and RNA samples from
226 each individual (*Command: samtools mpileup -d 10000*). We set a maximum coverage of
227 10,000 reads in pileup for each calling to take into account as many reads as possible in the
228 calling. SNPs were detected independently on each biological replicate. VCF files generated
229 by *SAMtools mpileup* were then used for subsequent analysis. For each biological replicate,
230 only variations where DNA was homozygous either for the reference allele or for the

231 alternative allele (MAF=1), and where RNA was heterozygous or homozygous for the
232 alternative allele, were kept. Finally, we removed positions covered by less than 15 reads in
233 both DNA and RNA alignments as well as tri-allelic sites.

234 **Impact of biases on mRNA editing detection**

235 To explore whether each editing event was likely related to sequencing errors or alignment
236 artifacts, we developed custom R and Perl scripts. We computed information related to: (1)
237 Extremity bias: an editing event was considered as biased if the edited allele was mostly
238 supported by the 10 first or last bases of reads in the RNA-seq read pileup. In accordance
239 with previous studies (Kleinman and Majewski 2012, Frésard *et al.* 2015), we chose to
240 consider only the distribution of the edited nucleotide position, to increase the stringency of
241 the method; (2) Strand bias: an editing event was considered as biased if the proportion of
242 forward and reverse reads supporting it was markedly different ($\Delta > 0.5$); (3) Splice junction
243 bias: an editing event was considered as biased if it was located within the region of a
244 predicted splice site, *i.e.* within 1-3 bases of the exon or 3-8 bases of the intron. To perform
245 this analysis, we determined the annotation and localization of editing events in transcripts
246 using *Ensembl v71 Variant Effect Predictor* (McLaren *et al.* 2010); (4) Homopolymer and low
247 complexity bias: an editing event was considered as biased if the 4 neighboring positions
248 harbored the same nucleotide or if it was falling in a single sequence repeat (SSR). SSR were
249 identified using *SciRoKo* (Kofler *et al.* 2007). The SSR patterns were investigated near
250 candidate edited sites with an offset of ± 3 bases; (5) Multimapping bias: for each editing
251 events, we generated a consensus 40 bp sequence centered on the edited allele, based on
252 the pileup of RNA-seq reads harboring the edited alleles in a given sample. We then used
253 *fuzznuc* (Olson 2002) to search for this sequence throughout the whole genomic DNA-seq
254 reads including unaligned reads from the same sample. An editing event was therefore

255 considered as biased if we found any match between its consensus surrounding sequence
256 and genomic DNA-seq reads.

257 **Impact of biological replication on mRNA editing detection**

258 For each tissue independently, we explored how reproducible each unbiased editing event
259 was across the 8 samples using custom R scripts. At each step, we computed the overall
260 amount of events belonging to each class of substitution from DNA to RNA. Since our
261 sequencing libraries were not strand-aware, the complement substitution of canonical
262 editing events (*i.e.* A-to-G for ADAR-mediated editing and C-to-T for APOBEC-mediated
263 editing) were also considered as canonical (*i.e.* C-to-T and G-to-A respectively).

264 **Validation assays and editing yield quantification**

265 *DNA Sanger sequencing and RNA pyro-sequencing*

266 To assess whether the DNA genotype of candidate editing events was homozygous or not,
267 we performed Sanger sequencing on the liver genomic DNA from the 8 animals. We then
268 assessed the mRNA genotype at these candidate sites on a PyroMark Q24 pyro-sequencer
269 (Qiagen, Valencia, CA). Primers were designed with PyroMark Assay Design software (Table
270 S2). PCR products were prepared using PyroMark PCR Kit (Qiagen, Valencia, CA) following
271 manufacturer's instructions. Data were analyzed with PyroMark Q24 v1.0.10 using default
272 parameters.

273 *Experimental designs used to test the effect of age, sex, genetic background and feeding on* 274 *mRNA editing level*

275 To measure the impact of age, sex, genotype and feeding on mRNA editing level, we used
276 different independent experimental designs with animals contrasted for these factors: (1)

277 Genotype: broilers (N=8) and layers (N=8) in two unrelated experimental designs (*i.e.* N=2 x
278 16 in all); (2) Age: prepuberal (N=8) and postpuberal (N=8) layers; (3) Feeding: broilers
279 slaughtered after a 24h fasting (N=8) or broilers fed *ad libitum* (N=8) in two independent
280 studies (*i.e.* N=2 x 16 in all); (4) Sex: female (N=8) and male (N=8) layers. Liver DNA and liver
281 RNA and/or adipose RNA were extracted and quality checked as described before (see
282 *Tissue collection and library preparation*). We then performed both liver genomic DNA and
283 liver and/or adipose mRNA pyro-sequencing following the aforementioned procedure for 5
284 edited sites (*i.e.* 2 liver specific, 1 adipose specific, 2 common to both tissues) previously
285 validated in our main experimental design. The pyro-sequencing signal at the edited
286 position in mRNA was standardized according to the signal obtained on DNA to avoid
287 amplification and sequencing biases. The editing level was computed as the ratio between
288 signals for the mRNA edited allele and allele on genomic DNA. We finally tested the effect of
289 each factor on RNA editing level at the 5 selected sites using two-sided unpaired
290 homoscedastic Student *t*-test. Statistical analyses were performed on R 3.2.0 using *t.test*
291 (two sided unpaired homoscedastic Student *t*-test) functions from the *stats* package and
292 graphical visualizations plotted using the *ggplot2* package.

293 ***In silico* prediction of RNA editing impact on protein structure and function**

294 To predict the putative effect of RNA editing on protein structure and function, we first
295 identified genomic structures likely to be impacted using Ensembl v71 Variant Effect
296 Predictor (McLaren *et al.* 2010). Focusing on missense coding editing events, we then
297 recovered orthologous protein sequences from *Gallus gallus*, *Bos taurus*, *Rattus norvegicus*,
298 *Mus musculus* and *Homo sapiens* to carry out multi-alignment. We finally used SIFT
299 prediction tool (Kumar *et al.* 2009) that is based on both sequence homology and physical

300 properties of amino acids, to quantify the potential impact of coding editing events on
301 protein structure and function.

302 RESULTS

303 High-throughput sequence analyses

304 Liver DNA and liver and white adipose tissue RNA were extracted from 8 16-week-old male
305 chickens, and paired-end sequenced on an Illumina HiSeq2000. After alignment on the
306 current genome assembly Galgal4, and filtering on mapping quality and multi-mapping, we
307 conserved an average of 157M DNA-seq reads, 30M liver mRNA-seq reads and 38M adipose
308 RNA-seq reads per sample. On average, 93.5% of the genome was covered by at least 15
309 DNA-seq reads, and 14.3% and 18.4% by at least 15 mRNA-seq reads from liver and white
310 adipose tissue, respectively.

311 mRNA editing detection and initial filtering

312 For each biological replicate, a base modification A (DNA base) → B (RNA base) was
313 considered as a candidate mRNA editing event if: (1) the genotype inferred for the genomic
314 DNA was homozygous AA with a minor allele frequency equal to 1 (*i.e.* all the reads are
315 supporting a unique allele on the genomic DNA sequence); (2) the genotype inferred for the
316 mRNA sequence was bi-allelic heterozygous AB or homozygous BB; (3) the position was
317 covered by at least 15 reads of both genomic DNA and mRNA sequences; and (4) the mRNA
318 editing event did not imply an insertion or deletion event. A total of 3,229 and 2,305
319 positions met these criteria in white adipose tissue (WAT) and liver, respectively (Figure 2).

320 Impact of sequencing and mapping artifacts on false editing discovery rate

321 In order to increase accuracy in the detection of editing events and to reduce the amount of
322 false positives, a standard procedure consists to apply different *ad hoc* filters to remove
323 suspicious candidates presenting error-prone splice junction bias, strand bias, extremity
324 bias, splice bias, or repetitive region bias. We conducted a first analysis aimed at assessing

325 the amount of candidate mRNA editing events that were spurious with respect to each of
326 these biases. This revealed that splice junction bias concerned about 15% of candidates in
327 both WAT and liver (Figure 2). While the amount of repetitive region biased mRNA editing
328 sites exceeded 30% of the candidate positions, this value increased to more than 45%
329 considering strand or extremity biased positions (Figure 2) in both tissues. Considering
330 altogether these sequencing and mapping artifacts, we showed that more than 85%
331 candidates were subject to at least one source of bias. Thus, using these classical filters (*i.e.*
332 usually applied in editing screening studies), the amount of candidate mRNA editing events
333 dropped from 3,229 and 2,305 to 448 and 342 in WAT and liver respectively

334 **Correction for multi-mapping based on DNA-seq raw sequences**

335 Even if somehow taken into account during or after mRNA-seq read mapping procedure
336 with splice-aware aligners such as TopHat or STAR, the reads multi-mapping may still have a
337 great impact on mRNA editing false discovery rate because of gaps and miss-assemblies in
338 the reference genome. To carefully control this artifact, we used the approach we first
339 introduced in our previous study (Frésard *et al.* 2015), consisting in aligning back mRNA-seq
340 reads harboring a candidate mRNA editing sites on corresponding individual DNA-seq reads,
341 independently if they were aligned or not onto the genome. We revealed that among the
342 448 and 342 remaining candidates, 4.2% and 7.8% were multimapping-related false
343 positives in WAT and liver respectively (Figure 2). Finally, considering both previously
344 described filters and this last filter dealing with multi-mapping, we ended up with 429 and
345 315 unbiased mRNA editing sites in WAT and liver, respectively.

346 **Impact of biological replication on canonical mRNA editing event identification**

347 As previously mentioned in the introduction, the number of biological replicates (N) taken
348 into account in mRNA editing screening studies based on RNA-seq data is highly variable,
349 with N ranging from 1 to 6 for the study conducted by Hu and collaborators (Hu *et al.* 2015).
350 As depicted on Figure 1, when the number of biological replicates is lower than 2, the total
351 number of editing events is extremely variable, ranging from 15 (Picardi *et al.* 2012) to
352 1,586,270 (Bazak *et al.* 2013). This number decreases drastically between 40 (Frésard *et al.*
353 2015) and 253 (Gu *et al.* 2012) when $N \geq 3$, suggesting that considering biological replicates
354 partly counteracts the lack of filters dealing with sequencing and mapping artifacts.
355 Nevertheless, a substantial variability remains between studies, likely related to the
356 differences in artifacts considered and stringency of bioinformatics filters and to the
357 biological context (*e.g.* tissue, species) of each study.

358 With this observation in mind, we characterized the 429 and 315 mRNA editing events
359 previously detected according to the number of biological replicates they were detected in,
360 and the class of base substitutions they belong to. We differentiated canonical mRNA
361 editing events (A-to-I and C-to-U interpreted as A-to-G and C-to-T by genome analyzers, and
362 corresponding to editing events catalyzed by ADARs and APOBECs) from non-canonical
363 events that are not explained by any of the two known editing mechanisms. As the RNA-seq
364 libraries used in this study were not strand-specific, we considered complement bases of
365 canonical changes as canonical editing events too (*i.e.* T-to-C and G-to-A). As can be seen in
366 Figure 3, in the sets of 429 and 315 unbiased mRNA editing events detected in at least one
367 individual described above, the amount of transversions (*i.e.* pyrimidine-to-purine, and
368 purine-to-pyrimidine, 54.1% in WAT and 58.0% in liver) was greater than the amount of
369 transitions (*i.e.* pyrimidine-to-pyrimidine, and purine-to-purine). Adding restriction based on

370 the number of biological replicates these events must be detected in, the amount of
371 transversions progressively decreased from more than 50% considering no replication (N=1),
372 to 25% (N=2) and finally to 10% (N=3) requiring editing events to be detected in at least 3
373 replicates. Considering that mRNA editing events were reproducible across at least 3
374 biological replicates, we finally conserved 19 and 11 positions in WAT and liver respectively,
375 comprising 1 non-canonical transversion event in each tissue (Table 2), distributed across 13
376 chromosomes (Figure 4). Among these 27 unique events, 3 were common to both tissues.
377 While most of these events were spatially isolated from each other, some of them were
378 clustered in short genomic regions spanning a few bp (Figure 4), especially on chromosome
379 1 (3 mRNA editing events in a window of 1.391 bp downstream *NOX4* in WAT, and 2 in a
380 window of 26 bp downstream *MPZL1* and *BRP44* in liver) and chromosome 12 (2 mRNA
381 editing events in a window of 951 bp in *FLNB*).

382 **Validations of candidate mRNA editing events**

383 To assess the validity of some of the edited positions detected using the high-throughput
384 screening approach, we performed Sanger sequencing on DNA to confirm their homozygous
385 genotype and pyrosequencing on RNA to validate the mRNA base recoding at the
386 corresponding position on the transcriptome. The validation revealed that non-canonical
387 mRNA editing events were false positives either related to genomic SNPs undetected using
388 genomic DNA-seq data for the one in WAT, or to unbalanced allelic expression not detected
389 through mRNA-seq data for the one in liver (Table 2). Following the same approach, we
390 selected 5 canonical events for validation: 2 out of the 3 candidates detected in both
391 tissues, 2 specific to WAT, and 1 liver-specific. We first confirmed their homozygous status
392 on genomic DNA and the mRNA recoding at these positions was then confirmed by mRNA
393 pyrosequencing (Table 2) using samples from the tissue they were detected in.

394 **Functional characterization of liver and white adipose editomes**

395 Functional annotation of the 25 unique canonical mRNA editing events using Variant Effect
396 Predictor revealed that most of them were located in non-coding regions, since 81% and
397 70% were situated in either 5kb gene flanking regions, intronic regions or intergenic regions
398 in WAT and liver, respectively (Figure 5). Only 5 unique mRNA editing events were
399 annotated as coding. Among these, 4 were likely to impact the mature protein: 1 common
400 to both tissues on *COG3*, 2 WAT-specific on *CES1* and *FLNB*, and 1 liver-specific on *KCMA1*
401 (Table 2). Focusing on these 4 missense mRNA editing events, we conducted a fine
402 functional annotation analysis using the SIFT software to assess the impact of the amino
403 acid substitution on these proteins. This revealed that none of these mRNA editing events
404 was likely to be deleterious. Nevertheless, after carrying out multiple species protein
405 alignments considering *Gallus gallus*, *Bos taurus*, *Rattus norvegicus*, *Mus musculus* and
406 *Homo sapiens*, we showed that except for *CES1*, these missense mRNA editing events were
407 impacting highly conserved amino acid residues (Figure 6).

408 **Impact of genetic background, age, sex, feeding and tissular context on editing level**

409 To test the impact of genetic background, age, feeding condition and sex on the mRNA
410 editing level, we considered the aforementioned subset of 5 validated canonical mRNA
411 editing events. We tested the effect of (1) the genotype, comparing mRNA editing level in
412 liver and WAT between broilers and layers in two independent experimental designs (Figure
413 7A); (2) the age, comparing mRNA editing level in liver between prepuberal and postpuberal
414 chickens (Figure 7B); (3) the feeding condition comparing mRNA editing level in both tissues
415 between chickens slaughtered after a 24h fasting and feeding *ad libitum*, in two
416 independent experimental designs (Figure 7C); (4) the sex, comparing mRNA editing level in
417 liver between roosters and hens (Figure 7D). To test each effect, we performed both

418 genomic DNA Sanger sequencing and mRNA-derived cDNA pyro-sequencing on 8
419 independent biological replicates in each group. Our analysis on liver revealed a significant
420 effect of genotype on mRNA editing level in the first design for 3 different positions among
421 the 4 tested (p -values: 1.26×10^{-5} , 1.52×10^{-3} and 1.33×10^{-6} , Figure 7A and Table S1). In
422 WAT, we highlighted 1 out of the 3 tested positions for which the editing level was
423 significantly different between broilers and layers (p -value: 9.44×10^{-3} , Figure 7A and Table
424 S1). Interestingly, while the general tendency is a greater editing level in broilers in
425 comparison to layers in the liver (*COG3*, *PLA1A* and *MYO1B*), this trend is reversed for the
426 highlighted site in WAT (*NDUFS6*). Regarding the age, we observed a significant effect on
427 editing levels for two sites in liver (Figure 7B and Table S1). Concerning the effect of sex, one
428 mRNA editing event showed a significant increase of mRNA editing level between males and
429 females in liver (p -value: 1.90×10^{-3} , Figure 7D and Table S1). Finally, by analyzing the effect
430 of the feeding condition, we highlighted a significant increase of the mRNA editing level
431 after a 24h fasting for one site in liver in two independent experimental designs (Figure 7C
432 and Table S1). Noticeably, for the edited position located on *COG3*, the mRNA recoding level
433 was significantly impacted by genetic background, age and feeding.

434 **DISCUSSION**

435 To achieve whole-transcriptome screening for mRNA editing events in chicken liver and
436 adipose tissue, we detected discrepancies between genomic DNA and mRNA sequences
437 using matched genomic DNA-seq and mRNA-seq data in several biological replicates. Since
438 2009 (Li *et al.* 2009b), similar approaches based on mRNA-seq have been extensively used to
439 characterize mouse, human and chimpanzee editomes in different tissues. According to the
440 literature, the extent of mRNA editing is highly variable with estimates ranging from dozens
441 to millions, even when comparing studies focusing on the same tissue in the same species.
442 After an in-depth reading of mRNA editing screening studies, we highlighted that, despite
443 recommendations for the use of rigorous bioinformatics pipelines to characterize editomes
444 (Kleinman and Majewski 2012, Lin *et al.* 2012, Pickrell *et al.* 2012), many recent studies
445 neglected the most reviewed sequencing and mapping artifacts related to mRNA-seq, such
446 as strand bias, read extremity bias, splice junction bias and low complexity region bias
447 (Table 1). In our study, these random or systematic biases were each impacting between
448 13% and 50% of the initial set of differences between DNA and RNA we detected. Overall,
449 we showed that almost 90% of the candidate mRNA editing events initially detected were
450 likely to be false positives arising from one of these artifacts, revealing how huge their
451 impact is on false discovery rate as previously reported (Pickrell *et al.* 2012, Lagarrigue *et al.*
452 2013). Interestingly, the false positive rate we report is in agreement with the observations
453 of Pickrell and collaborators (Pickrell *et al.* 2012), suggesting that among the 28,766 editing
454 events detected in the study of Li and collaborators (Li *et al.* 2011), roughly 90% were likely
455 false positives emerging from sequencing errors and mapping artifacts.

456 Even if taken into account during mRNA-seq read mapping procedures with splice-aware
457 aligners, multi-mapping artifact could still be of a great impact on mRNA editing false

458 discovery rate. Indeed, current genome assemblies used to map short reads from high-
459 throughput sequencing experiments, even of high quality for human, mouse or chicken, are
460 still presenting missing sequences as well as misassembled regions (Groenen *et al.* 2011,
461 Genovese *et al.* 2013a, 2013b). Since these regions might harbor sequences that are
462 paralogous to properly assembled parts of genomes, this ultimately leads to shallow
463 identification of multimapped short reads when only considering the reference sequence.
464 To by-pass this error-prone issue for the identification of mRNA edited sites, we used in this
465 study the approach first introduced in our previous study (Frésard *et al.* 2015) consisting in
466 looking for mRNA sequences spanning edited sites in raw genomic DNA sequences, and
467 confirmed its efficiency. In our study, up to 8% of the initial candidate differences between
468 RNA and DNA were false positives related to errors and assemblies issues in the chicken
469 genome. This result suggests that thousands of edited sites reported in primates and mice
470 mRNA editing screening studies could be partly attributed to false positives resulting from a
471 spurious handling of multimapped mRNA-seq short reads. More strikingly, a huge
472 proportion of mRNA editing screening studies are based on mRNA-seq data solely, and do
473 not consider individual matched genomic DNA-seq data for the samples analyzed. In these
474 studies, candidate edited sites are filtered using the positions of known SNPs referenced in
475 databases such as dbSNP, rather than considering individual polymorphisms, thereby fully
476 occulting inevitable individual specific genomic variations. Even in the case of clonal mice
477 strains, considering a consensus strain-specific genomic sequence as exposed by Danecek
478 and collaborators (Danecek *et al.* 2012) indubitably leads to edited events false calls that
479 arise from somatic mutations, as we previously shown in our study on mice, invalidating
480 25% of mRNA editing candidates arising because of unreferenced genomic SNP (Lagarrigue
481 *et al.* 2013).

482 To further limit the amount of false positives among mRNA edited sites and to focus on
483 biologically meaningful mRNA editing events, an obvious approach consists in considering
484 biological replication. Surprisingly, a lot of mRNA editing screening studies report events
485 without considering reproducibility across samples (Table 1, Figure 1). Although it is likely
486 that mRNA editing is a partly individual-specific phenomenon (Gommans *et al.* 2009), short
487 read sequencing technologies are error-prone when it comes to focus on slight variations
488 and are not mature enough to allow investigation of private editing events. Therefore,
489 biological reproducibility is uncontestedly required in this scope. In our study, even after
490 filtering to properly account for systematic and random sequencing artifacts as well as
491 multimapping, we were still detecting more than 50% of non canonical mRNA editing
492 events. Hitherto, the attempts by other group to validate such type of mRNA recoding using
493 targeted Sanger sequencing were unsuccessful (Piskol *et al.* 2013), clearly ascertaining that
494 they arise from unconsidered artifacts. When focusing only on editing events detected in at
495 least 3 biological replicates, the proportion of canonical events raised 90% in the present
496 work. If non-canonical recoding events were not related to artifacts, we would have not
497 expected such an enrichment, which further ascertains they are false positives. With respect
498 to this hypothesis, we invalidated the 2 non-canonical events we were still detecting after
499 considering biological replication. Overall, our results suggest that filtering to consider false
500 positives arising from mapping artifacts and sequencing errors, even if mandatory, is not
501 sufficient to remove all spurious editing events. While allowing focusing on the most
502 biologically meaningful recoding events that are shared between individuals, considering
503 biological replications is decisive to contain the amount of false positives in RNA editing
504 screening studies based on current high-throughput sequencing technologies.

505 Altogether, considering filters dealing with systematic and random sequencing errors,
506 multimapping, mapping artifacts and biological replication, the number of edited sites
507 dramatically felt from 3.229 and 2.305 candidates to 19 and 11 robust events in WAT and
508 liver respectively. And even using highly drastic filters, two non-canonical false positives
509 were still detected, once again suggesting that hard filtering and biological replication are
510 still mandatory when working with current short-sequencing technologies. Even if a greater
511 sequencing depth would have allowed to detect a slightly higher number of edited sites, our
512 work reveals that the extent of mRNA editing is, at least in chicken, far below what has been
513 previously shown in most screening studies on human, mouse and chimpanzee.
514 Interestingly, most of the studies reporting an amount of mRNA editing events close to what
515 we record have been conducted on healthy tissues rather than immortalized cell lines or
516 tumors, and considered only sites edited in at least 2 biological replicates (Gu *et al.* 2012,
517 Lagarrigue *et al.* 2013, Frésard *et al.* 2015). The huge variation in mRNA editing extent
518 between our study and other screening studies in literature could be explained in different
519 ways. First of all, it is likely related to the differences in the stringency of filters applied and
520 in the false positive rate. Second, most of the mRNA editing screening studies were carried
521 out using transformed cell lines or cancer tissues (Table 1) and the extensive mRNA editing
522 reported may reflect real biological changes. Indeed, ADARs and APOBECs may become
523 more active during tumorigenesis, and may consequently increase mRNA editing, as it has
524 been highlighted in some cancer cell lines (Galeano *et al.* 2012). Third, it could also be
525 explained by the huge structural differences between mammalian and sauropsidian
526 genomes. Actually, ADAR-mediated A-to-I mRNA editing occurs in regions of double
527 stranded RNA (dsRNA). Yet, approximately half of typical mammalian genome contains
528 highly repetitive sequences (de Koning *et al.* 2011) such as retrotransposons, short

529 interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs).
530 While these sequences are often repeated in reverse tandems, they may generate dsRNA
531 structures that could be subsequently edited by ADARs (Nishikura 2010). In chicken, since
532 the amount of repetitive sequences across the genome falls below 15% (Wicker *et al.* 2005,
533 Schmid *et al.* 2015), it is expected that less A-to-I editing events occur.

534 At the end, among the 25 unique canonical mRNA editing events we report, only 3 (*i.e.* 13%)
535 located downstream *MPZL1* and *BRP44*, in *COG3* and upstream *NDUSF6*, are common to
536 both tissues. Comparisons with our previous study highlighted that only 4 mRNA edited sites
537 detected in chicken whole-embryo were also found in mature WAT or liver. Surprisingly,
538 only the edited sites located in *COG3* and upstream *NDUSF6* were common to WAT, liver
539 and whole embryo. These results are comparable to those reported by Danecek *et al.* and
540 Lagarrigue *et al.*, revealing a significant amount of tissue-specific edited sites (Danecek *et al.*
541 2012, Lagarrigue *et al.* 2013). Interestingly, while no homologue of *APOBEC1* has been
542 characterized in the chicken genome (Conticello *et al.* 2005), all the APOBEC-mediated C-to-
543 U mRNA editing candidate sites we initially detected were discarded along the filtering
544 pipeline, confirming that this specific mRNA editing mechanism is missing in chicken. We
545 also found that some of the mRNA editing events we detected were localized on mRNA
546 editing clusters, spanning regions of a few kilobases. This observation is supported by our
547 current knowledge regarding the mechanistic basis of ADAR-mediated A-to-I mRNA editing,
548 which occurs unspecifically in dsRNA, and doesn't involve specific mooring sequence as it is
549 the case for APOBEC-mediated C-to-U mRNA editing (Nishikura 2010).

550 Overall, most of the mRNA editing events we detected are falling in non-coding regions (*i.e.*
551 10 kb upstream or downstream genes, in introns, or in intergenic regions). Since these

552 regions are expressed, they are either corresponding to poorly annotated genomic regions,
553 non-mature mRNAs, or unannotated non-coding RNAs in which RNA editing is known to
554 occur (Picardi *et al.* 2014). Out of these 25 unique mRNA editing events, 5 are located in
555 coding sequences and only 4 are non-synonymous, impacting the sequence of *COG3*, *CES1*,
556 *FLNB* and *KCNMA1*. Interestingly, the edited sites falling in *COG3* and *FLNB* were already
557 described in mammalian species (Levanon *et al.* 2005, Shah *et al.* 2009, Danecek *et al.* 2012,
558 Holmes *et al.* 2013, Stulic and Jantsch 2013), revealing that some edited positions are
559 conserved throughout evolution between birds and mammals. Except for *CES1*, our analyses
560 show that each of these coding mRNA editing events are impacting highly conserved regions
561 in the protein sequence, as well as highly conserved amino-acid residues.

562 Our analysis finally shows that mRNA editing level is impacted by various genetic and
563 environmental factors such as the genetic background, the age, the feeding and the sex.
564 While the genetic background and the age influence the editing level at almost all the mRNA
565 editing sites we tested, the feeding and the gender tend to have effect on less positions. The
566 impact of aging on mRNA editing has already been reported in a few studies on mammals
567 (Wahlstedt *et al.* 2009, Shtrichman *et al.* 2012, Venø *et al.* 2012), as well as in our previous
568 study on chicken embryo (Frésard *et al.* 2015). In agreement with most of these studies, we
569 confirm that the level of edited transcripts increases with age, whatever the recoded site
570 considered. Nevertheless, an in-depth unbiased whole-transcriptome exploration of spatio-
571 temporal regulation bases of mRNA editing in vertebrates is still needed. We also observed
572 a significant effect of genotype on the liver mRNA editing level indicating another level of
573 regulation. Indeed, in one of the designs we used to assess the effect of genetic background,
574 3 out of 4 edited isoforms (in *COG3*, *PLA1A* and *MYO1B*) are 1.5- up to 2-fold more frequent

575 in broilers' liver compared to layers' liver. *PLA1A* encodes for the phosphatidylserine-
576 specific phospholipase A1, mostly synthesized in liver, and implied in the release of free
577 fatty acids and lysophosphatidic acid that acts as a lipid mediator in cell signaling. While it is
578 established that this lipase does not catabolize triglycerides, its implication in global cellular
579 process is still poorly understood (Aoki *et al.* 2002). *MYO1B* encodes for the widely
580 expressed myosin 1B motors that functions in endocytosis, membrane trafficking,
581 membrane retraction, and mechano-signal transduction. Even if the physiological landscape
582 of myosin 1B is not fully depicted yet, some authors hypothesized its potential role on
583 myogenesis (Wells *et al.* 1997, Redowicz 2007). While muscle mass stands as one of the
584 most divergent phenotypic traits between layers and broilers, the mRNA editing in *MYO1B*
585 could be part of the transcriptional bases leading to differences in muscle development
586 between these strains, but it remains to test if this site is also differentially edited in
587 muscular tissues between broilers and layers. Finally, *COG3* has a general cellular function
588 related to the structure and function of the Golgi as further described below. Since the
589 editing level at these sites located in *COG3*, *PLA1A* and *MYO1B* are differential between
590 broilers' and layers' liver and because liver is a multi-function organ involved in many
591 physiological processes, we hypothesized they might be implied in cellular and
592 developmental processes leading to physiological differences between these two chicken
593 strains. The genetic regulation of mRNA editing level at these sites could be linked to
594 mechanisms acting in *trans*, involving ADARs, and further analyses comparing ADARs
595 expression and activity between these two chicken genetic backgrounds are mandatory to
596 investigate this hypothesis. But it could also be regulated by *cis*-acting mutations impacting
597 surrounding mRNA sequence and secondary structure, as it has been recently suggested
598 regarding mRNA recoding in drosophila (Sapiro *et al.* 2015).

599 Whereas the extent of mRNA editing appears limited at the transcriptome scale in chicken
600 liver and white adipose, our results suggest that this phenomenon could be tightly
601 regulated. Indeed, the I/V non-synonymous recoding event impacting *COG3* is not only
602 conserved in mammals but is also under the influence of the genetic background, the age
603 and the feeding condition. For this last factor, this edited site was the only one impacted, in
604 two independent designs, which suggests that this observation is highly reliable. It is also
605 noticeable that *COG3* mRNA is almost exclusively edited in white adipose tissue with in
606 average 95% (in the “genotype” design) or 100% (in the “feeding” design) of the edited
607 isoform, in contrast with what we observed in the liver transcriptome. This further suggests
608 that these different isoforms are likely harboring different physiological functions and that
609 ADAR-mediated mRNA editing could act in a highly tissue-specific manner as previously
610 shown (Song *et al.* 2004), in a way that is similar to APOBEC1-mediated *APOB* mRNA editing,
611 which ultimately leads to the synthesis of two *APOB* isoforms – *APOB100* in the liver, and
612 *APOB48* in the small intestine – with distinct physiological functions. Since the edited site in
613 *COG3* is conserved through the evolution of vertebrates, and since it is tightly regulated by
614 multiple genetic and environmental factors, it is likely to have a functional role on the
615 encoded protein. *COG3* is one of the eight proteins of the oligomeric Golgi (*COG*) complex.
616 The *COG* complex is involved in intra-Golgi retrograde trafficking and in membrane
617 trafficking in eukaryotic cells (Loh and Hong 2004, Zolov and Lupashin 2005). Mutations
618 affecting *COG* subunits disturb both structure and function of the Golgi (Ungar *et al.* 2002)
619 and have been reported in congenital disorders of glycosylation (Kodera *et al.* 2015). These
620 different studies show an important role of the *COG* complex in eukaryotic cells. It is known
621 to be an evolutionarily conserved multi-subunit protein complex, but its exact cellular
622 function remains elusive. While this edited site in *COG3* is conserved across stages (embryos

623 and adult stages in chicken) and species (human, mouse, rat and chicken), it would require
624 additional work to decipher its potential role on *COG3* functions and potentially on
625 membrane trafficking pathways.

626 This study, which is complementary to our previous study conducted on chicken embryos, is
627 the first describing the mRNA editing landscape in adult chicken. From a methodological
628 point of view, we show how huge can be the impact of sequencing biases and mapping
629 artifacts on the discovery of mRNA editing events if not properly considered. Moreover, we
630 show the importance to consider biological replication with high-throughput sequencing
631 data to filter spurious candidates, allowing focusing on the most biologically meaningful
632 mRNA editing events. From a biological point of view, even if we cannot claim we are
633 exhaustive, our results support the evidence that the extent of mRNA editing is limited in
634 chicken, restricted to ADAR-mediated events. We also ascertain that some of editing sites
635 are conserved throughout the evolution of vertebrates. Our study finally shows that mRNA
636 editing level is strongly affected by the genetic background and age and - to a minor extent -
637 by the feeding condition and the sex, which provides new insights into the comprehension
638 of mRNA editing functions in vertebrates in relation to genetics and environmental
639 components.

640 **ACKNOWLEDGMENTS**

641 We thank the members of the INRA experimental poultry unit (UE1295 PEAT, Nouzilly,
642 France) for producing and rearing the chickens. This work was funded by INRA (programs
643 SOSrnaSeq – 2010, F1rnaSEQ – 2012, ELASETIC – 2013) and the French National Research
644 Agency (ANR, program Fatinteger 2012 – 2015). Pierre-François Roux’s graduate work was
645 co-funded by INRA (www.inra.fr/en) and Région Bretagne (www.bretagne.fr). Laure
646 Frésard’s graduate work was co-funded by Animal Genetics Division at INRA
647 (www.inra.fr/en) and Région Midi-Pyrénées (www.midipyrenees.fr).

648

650 **LITERATURE CITED**

- 651 1000 Genomes Project Consortium, G. R. Abecasis, A. Auton, L. D. Brooks, M. A. DePristo, *et al.* 2012 An integrated map of genetic variation from 1,092 human genomes. *Nature*
652 491:56–65.
653
- 654 Aoki, J., Y. Nagai, H. Hosono, K. Inoue, and H. Arai. 2002 Structure and function of
655 phosphatidylserine-specific phospholipase A1. *Biochim. Biophys. Acta* 1582:26–32.
- 656 Van der Auwera, G. A., M. O. Carneiro, C. Hartl, R. Poplin, G. del Angel, *et al.* 2013 From
657 FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices
658 Pipeline. *Current Protocols in Bioinformatics*. John Wiley & Sons, Inc.
- 659 Bahn, J. H., J. H. Lee, G. Li, C. Greer, G. Peng, *et al.* 2011 Accurate Identification of A-to-I RNA
660 editing in human by transcriptome sequencing. *Genome Res.:*gr.124107.111.
- 661 Bass, B. L. 2002 RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.*
662 71:817–846.
- 663 Bazak, L., A. Haviv, M. Barak, J. Jacob-Hirsch, P. Deng, *et al.* 2013 A-to-I RNA editing occurs at
664 over a hundred million genomic sites, located in a majority of human genes. *Genome*
665 *Res.:*gr.164749.113.
- 666 Blanc, V., E. Park, S. Schaefer, M. Miller, Y. Lin, *et al.* 2014 Genome-wide identification and
667 functional analysis of Apobec-1-mediated C-to-U RNA editing in mouse small intestine and
668 liver. *Genome Biol.* 15:R79.
- 669 Börner, G. V., M. Mörl, A. Janke, and S. Pääbo. 1996 RNA editing changes the identity of a
670 mitochondrial tRNA in marsupials. *EMBO J.* 15:5949–5957.
- 671 Cattenoz, P. B., R. J. Taft, E. Westhof, and J. S. Mattick. 2013 Transcriptome-wide
672 identification of A > I RNA editing sites by inosine specific cleavage. *RNA N. Y. N* 19:257–270.
- 673 Chan, T. H. M., C. H. Lin, L. Qi, J. Fei, Y. Li, *et al.* 2014 A disrupted RNA editing balance
674 mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular
675 carcinoma. *Gut* 63:832–843.
- 676 Chen, J.-Y., Z. Peng, R. Zhang, X.-Z. Yang, B. C.-M. Tan, *et al.* 2014 RNA editome in rhesus
677 macaque shaped by purifying selection. *PLoS Genet.* 10:e1004274.
- 678 Chen, L. 2013 Characterization and comparison of human nuclear and cytosolic editomes.
679 *Proc. Natl. Acad. Sci. U. S. A.* 110:E2741–2747.
- 680 Clop, A., F. Marcq, H. Takeda, D. Pirottin, X. Tordoir, *et al.* 2006 A mutation creating a
681 potential illegitimate microRNA target site in the myostatin gene affects muscularity in
682 sheep. *Nat. Genet.* 38:813–818.
- 683 Consortium, T. 1000 G. P. 2010 A map of human genome variation from population-scale
684 sequencing. *Nature* 467:1061–1073.
- 685 Conticello, S. G., C. J. F. Thomas, S. K. Petersen-Mahrt, and M. S. Neuberger. 2005 Evolution
686 of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases. *Mol. Biol. Evol.*
687 22:367–377.

- 688 Danecek, P., C. Nellåker, R. E. McIntyre, J. E. Buendia-Buendia, S. Bumpstead, *et al.* 2012
689 High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. *Genome*
690 *Biol.* 13:26.
- 691 Davey, M. G., and C. Tickle. 2007 The chicken as a model for embryonic development.
692 *Cytogenet. Genome Res.* 117:231–239.
- 693 Eifler, T., S. Pokharel, and P. A. Beal. 2013 RNA-Seq analysis identifies a novel set of editing
694 substrates for human ADAR2 present in *Saccharomyces cerevisiae*. *Biochemistry (Mosc.)*
695 52:7857–7869.
- 696 Frésard, L., S. Leroux, P.-F. Roux, C. Klopp, S. Fabre, *et al.* 2015 Genome-Wide
697 Characterization of RNA Editing in Chicken Embryos Reveals Common Features among
698 Vertebrates. *PLOS ONE* 10:e0126776.
- 699 Frésard, L., S. Leroux, B. Servin, D. Gourichon, P. Dehais, *et al.* 2014 Transcriptome-wide
700 investigation of genomic imprinting in chicken. *Nucleic Acids Res.:*gkt1390.
- 701 Funahashi, T., F. Giannoni, A. M. DePaoli, S. F. Skarosi, and N. O. Davidson. 1995 Tissue-
702 specific, developmental and nutritional regulation of the gene encoding the catalytic subunit
703 of the rat apolipoprotein B mRNA editing enzyme: functional role in the modulation of apoB
704 mRNA editing. *J. Lipid Res.* 36:414–428.
- 705 Galeano, F., S. Tomaselli, F. Locatelli, and A. Gallo. 2012 A-to-I RNA editing: the “ADAR” side
706 of human cancer. *Semin. Cell Dev. Biol.* 23:244–250.
- 707 Gan, Z., L. Zhao, L. Yang, P. Huang, F. Zhao, *et al.* 2006 RNA editing by ADAR2 is metabolically
708 regulated in pancreatic islets and beta-cells. *J. Biol. Chem.* 281:33386–33394.
- 709 Genovese, G., R. E. Handsaker, H. Li, N. Altemose, A. M. Lindgren, *et al.* 2013a Using
710 population admixture to help complete maps of the human genome. *Nat. Genet.* 45:406–
711 414, 414e1–2.
- 712 Genovese, G., R. E. Handsaker, H. Li, E. E. Kenny, and S. A. McCarroll. 2013b Mapping the
713 human reference genome’s missing sequence by three-way admixture in Latino genomes.
714 *Am. J. Hum. Genet.* 93:411–421.
- 715 Gommans, W. M., S. P. Mullen, and S. Maas. 2009 RNA editing: a driving force for adaptive
716 evolution? *BioEssays News Rev. Mol. Cell. Dev. Biol.* 31:1137–1145.
- 717 Gott, J. M., B. H. Somerlot, and M. W. Gray. 2010 Two forms of RNA editing are required for
718 tRNA maturation in *Physarum* mitochondria. *RNA N. Y. N* 16:482–488.
- 719 Gray, M. W. 2012 Evolutionary origin of RNA editing. *Biochemistry (Mosc.)* 51:5235–5242.
- 720 Gray, M. W., J. Lukes, J. M. Archibald, P. J. Keeling, and W. F. Doolittle. 2010 Cell biology.
721 Irremediable complexity? *Science* 330:920–921.
- 722 Groenen, M. A., H.-J. Megens, Y. Zare, W. C. Warren, L. W. Hillier, *et al.* 2011 The
723 development and characterization of a 60K SNP chip for chicken. *BMC Genomics* 12:274.
- 724 Gu, T., F. W. Buaas, A. K. Simons, C. L. Ackert-Bicknell, R. E. Braun, *et al.* 2012 Canonical A-to-
725 I and C-to-U RNA Editing Is Enriched at 3’UTRs and microRNA Target Sites in Multiple Mouse
726 Tissues. *PLoS ONE* 7:e33720.
- 727 Holmes, A. P., S. H. Wood, B. J. Merry, and J. P. de Magalhães. 2013 A-to-I RNA editing does
728 not change with age in the healthy male rat brain. *Biogerontology* 14:395–400.

- 729 Hu, X., S. Wan, Y. Ou, B. Zhou, J. Zhu, *et al.* 2015 RNA over-editing of BLCAP contributes to
 730 hepatocarcinogenesis identified by whole-genome and transcriptome sequencing. *Cancer*
 731 *Lett.* 357:510–519.
- 732 Ju, Y. S., J.-I. Kim, S. Kim, D. Hong, H. Park, *et al.* 2011 Extensive genomic and transcriptional
 733 diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean
 734 individuals. *Nat. Genet.* 43:745–752.
- 735 Kang, L., X. Liu, Z. Gong, H. Zheng, J. Wang, *et al.* 2015 Genome-wide identification of RNA
 736 editing in hepatocellular carcinoma. *Genomics* 105:76–82.
- 737 Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, *et al.* 2013 TopHat2: accurate
 738 alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
 739 *Genome Biol.* 14:R36.
- 740 Kleinman, C. L., V. Adoue, and J. Majewski. 2012 RNA editing of protein sequences: a rare
 741 event in human transcriptomes. *RNA N. Y. N* 18:1586–1596.
- 742 Kleinman, C. L., and J. Majewski. 2012 Comment on “Widespread RNA and DNA Sequence
 743 Differences in the Human Transcriptome.” *Science* 335:1302–1302.
- 744 Knoop, V. 2011 When you can’t trust the DNA: RNA editing changes transcript sequences.
 745 *Cell. Mol. Life Sci. CMLS* 68:567–586.
- 746 Kodera, H., N. Ando, I. Yuasa, Y. Wada, Y. Tsurusaki, *et al.* 2015 Mutations in COG2 encoding
 747 a subunit of the conserved oligomeric golgi complex cause a congenital disorder of
 748 glycosylation. *Clin. Genet.* 87:455–460.
- 749 Kofler, R., C. Schlötterer, and T. Lelley. 2007 SciRoKo: a new tool for whole genome
 750 microsatellite search and investigation. *Bioinforma. Oxf. Engl.* 23:1683–1685.
- 751 de Koning, A. P. J., W. Gu, T. A. Castoe, M. A. Batzer, and D. D. Pollock. 2011 Repetitive
 752 elements may comprise over two-thirds of the human genome. *PLoS Genet.* 7:e1002384.
- 753 Kumar, P., S. Henikoff, and P. C. Ng. 2009 Predicting the effects of coding non-synonymous
 754 variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4:1073–1081.
- 755 Lagarrigue, S., F. Hormozdiari, L. J. Martin, F. Lecerf, Y. Hasin, *et al.* 2013 Limited RNA editing
 756 in exons of mouse liver and adipose. *Genetics* 193:1107–1115.
- 757 Lau, P. P., D. J. Cahill, H. J. Zhu, and L. Chan. 1995 Ethanol modulates apolipoprotein B mRNA
 758 editing in the rat. *J. Lipid Res.* 36:2069–2078.
- 759 Leclercq, B., J. C. Blum, and J. P. Boyer. 1980 Selecting broilers for low or high abdominal fat:
 760 initial observations. *Br. Poult. Sci.* 21:107–113.
- 761 Levanon, E. Y., M. Hallegger, Y. Kinar, R. Shemesh, K. Djinovic-Carugo, *et al.* 2005
 762 Evolutionarily conserved human targets of adenosine to inosine RNA editing. *Nucleic Acids*
 763 *Res.* 33:1162–1168.
- 764 Li, H., and R. Durbin. 2009 Fast and accurate short read alignment with Burrows-Wheeler
 765 transform. *Bioinforma. Oxf. Engl.* 25:1754–1760.
- 766 Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, *et al.* 2009a The Sequence
 767 Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.
- 768 Li, J. B., E. Y. Levanon, J.-K. Yoon, J. Aach, B. Xie, *et al.* 2009b Genome-Wide Identification of

- 769 Human RNA Editing Sites by Parallel DNA Capturing and Sequencing. *Science* 324:1210–
770 1213.
- 771 Li, M., I. X. Wang, Y. Li, A. Bruzel, A. L. Richards, *et al.* 2011 Widespread RNA and DNA
772 Sequence Differences in the Human Transcriptome. *Science* 333:53–58.
- 773 Lin, W., R. Piskol, M. H. Tan, and J. B. Li. 2012 Comment on “Widespread RNA and DNA
774 Sequence Differences in the Human Transcriptome.” *Science* 335:1302–1302.
- 775 Li, Z., H. Bammann, M. Li, H. Liang, Z. Yan, *et al.* 2013 Evolutionary and ontogenetic changes
776 in RNA editing in human, chimpanzee, and macaque brains. *RNA N. Y. N* 19:1693–1702.
- 777 Loh, E., and W. Hong. 2004 The binary interacting network of the conserved oligomeric Golgi
778 tethering complex. *J. Biol. Chem.* 279:24640–24648.
- 779 Malhis, N., and S. J. M. Jones. 2010 High quality SNP calling using Illumina data at shallow
780 coverage. *Bioinformatics* 26:1029–1035.
- 781 Van Mater, D., M. P. Sowden, J. Cianci, J. D. Sparks, C. E. Sparks, *et al.* 1998 Ethanol
782 increases apolipoprotein B mRNA editing in rat primary hepatocytes and McArdle cells.
783 *Biochem. Biophys. Res. Commun.* 252:334–339.
- 784 McLaren, W., B. Pritchard, D. Rios, Y. Chen, P. Flicek, *et al.* 2010 Deriving the consequences
785 of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* 26:2069–
786 2070.
- 787 Meacham, F., D. Boffelli, J. Dhahbi, D. I. Martin, M. Singer, *et al.* 2011 Identification and
788 correction of systematic error in high-throughput sequence data. *BMC Bioinformatics*
789 12:451.
- 790 Mo, F., A. W. Wyatt, Y. Sun, S. Brahmabhatt, B. J. McConeghy, *et al.* 2014 Systematic
791 Identification and Characterization of RNA Editing in Prostate Tumors. *PLoS ONE* 9:e101431.
- 792 Nakamura, K., T. Oshima, T. Morimoto, S. Ikeda, H. Yoshikawa, *et al.* 2011 Sequence-specific
793 error profile of Illumina sequencers. *Nucleic Acids Res.:*gkr344.
- 794 Nishikura, K. 2010 Functions and Regulation of RNA Editing by ADAR Deaminases. *Annu.*
795 *Rev. Biochem.* 79:321–349.
- 796 Olson, S. A. 2002 EMBOSS opens up sequence analysis. *European Molecular Biology Open*
797 *Software Suite. Brief. Bioinform.* 3:87–91.
- 798 Park, E., B. Williams, B. J. Wold, and A. Mortazavi. 2012 RNA editing in the human ENCODE
799 RNA-seq data. *Genome Res.* 22:1626–1633.
- 800 Peng, Z., Y. Cheng, B. C.-M. Tan, L. Kang, Z. Tian, *et al.* 2012 Comprehensive analysis of RNA-
801 Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* 30:253–
802 260.
- 803 Phung, T. L., M. P. Sowden, J. D. Sparks, C. E. Sparks, and H. C. Smith. 1996 Regulation of
804 hepatic apolipoprotein B RNA editing in the genetically obese Zucker rat. *Metabolism.*
805 45:1056–1058.
- 806 Picardi, E., A. M. D’Erchia, A. Gallo, A. Montalvo, and G. Pesole. 2014 Uncovering RNA
807 Editing Sites in Long Non-Coding RNAs. *Front. Bioeng. Biotechnol.* 2.
- 808 Picardi, E., A. Gallo, F. Galeano, S. Tomaselli, and G. Pesole. 2012 A novel computational

- 809 strategy to identify A-to-I RNA editing sites by RNA-Seq data: de novo detection in human
810 spinal cord tissue. *PLoS One* 7:e44184.
- 811 Pickrell, J. K., Y. Gilad, and J. K. Pritchard. 2012 Comment on “Widespread RNA and DNA
812 sequence differences in the human transcriptome.” *Science* 335:1302; author reply 1302.
- 813 Piskol, R., Z. Peng, J. Wang, and J. B. Li. 2013 Lack of evidence for existence of noncanonical
814 RNA editing. *Nat. Biotechnol.* 31:19–20.
- 815 Ramaswami, G., W. Lin, R. Piskol, M. H. Tan, C. Davis, *et al.* 2012 Accurate identification of
816 human Alu and non-Alu RNA editing sites. *Nat. Methods* 9:579–581.
- 817 Redowicz, M. J. 2007 Unconventional myosins in muscle. *Eur. J. Cell Biol.* 86:549–558.
- 818 Sakurai, M., H. Ueda, T. Yano, S. Okada, H. Terajima, *et al.* 2014 A biochemical landscape of
819 A-to-I RNA editing in the human brain transcriptome. *Genome Res.* 24:522–534.
- 820 Sapiro, A. L., P. Deng, R. Zhang, and J. B. Li. 2015 Cis Regulatory Effects on A-to-I RNA Editing
821 in Related *Drosophila* Species. *Cell Rep.* 11:697–703.
- 822 Schmid, M., J. Smith, D. W. Burt, B. L. Aken, P. B. Antin, *et al.* 2015 Third Report on Chicken
823 Genes and Chromosomes 2015. *Cytogenet. Genome Res.* 145:78–179.
- 824 Schrider, D. R., J.-F. Gout, and M. W. Hahn. 2011 Very Few RNA and DNA Sequence
825 Differences in the Human Transcriptome. *PLoS ONE* 6:e25842.
- 826 Shah, S. P., R. D. Morin, J. Khattri, L. Prentice, T. Pugh, *et al.* 2009 Mutational evolution in a
827 lobular breast tumour profiled at single nucleotide resolution. *Nature* 461:809–813.
- 828 Shtrichman, R., I. Germanguz, R. Mandel, A. Ziskind, I. Nahor, *et al.* 2012 Altered A-to-I RNA
829 Editing in Human Embryogenesis. *PLoS ONE* 7:e41576.
- 830 Song, W., Z. Liu, J. Tan, Y. Nomura, and K. Dong. 2004 RNA editing generates tissue-specific
831 sodium channels with distinct gating properties. *J. Biol. Chem.* 279:32554–32561.
- 832 Stulic, M., and M. F. Jantsch. 2013 Spatio-temporal profiling of Filamin A RNA-editing reveals
833 ADAR preferences and high editing levels outside neuronal tissues. *RNA Biol.* 10:1611–1617.
- 834 Torres, A. G., D. Piñeyro, L. Filonava, T. H. Stracker, E. Batlle, *et al.* 2014 A-to-I editing on
835 tRNAs: biochemical, biological and evolutionary implications. *FEBS Lett.* 588:4279–4286.
- 836 Toung, J. M., N. Lahens, J. B. Hogenesch, and G. Grant. 2014 Detection theory in
837 identification of RNA-DNA sequence differences using RNA-sequencing. *PLoS One*
838 9:e112040.
- 839 Treangen, T. J., and S. L. Salzberg. 2012 Repetitive DNA and next-generation sequencing:
840 computational challenges and solutions. *Nat. Rev. Genet.* 13:36–46.
- 841 Ungar, D., T. Oka, E. E. Brittle, E. Vasile, V. V. Lupashin, *et al.* 2002 Characterization of a
842 mammalian Golgi-localized protein complex, COG, that is required for normal Golgi
843 morphology and function. *J. Cell Biol.* 157:405–415.
- 844 Valach, M., S. Moreira, G. N. Kiethega, and G. Burger. 2014 Trans-splicing and RNA editing of
845 LSU rRNA in *Diplonema* mitochondria. *Nucleic Acids Res.* 42:2660–2672.
- 846 Venø, M. T., J. B. Bramsen, C. Bendixen, F. Panitz, I. E. Holm, *et al.* 2012 Spatio-temporal
847 regulation of ADAR editing during development in porcine neural tissues. *RNA Biol.* 9:1054–
848 1065.

- 849 Wahlstedt, H., C. Daniel, M. Ensterö, and M. Ohman. 2009 Large-scale mRNA sequencing
850 determines global regulation of RNA editing during brain development. *Genome Res.*
851 19:978–986.
- 852 Warnefors, M., A. Liechti, J. Halbert, D. Valloton, and H. Kaessmann. 2014 Conserved
853 microRNA editing in mammalian evolution, development and disease. *Genome Biol.* 15:R83.
- 854 Wells, C., D. Coles, A. Entwistle, and M. Peckham. 1997 Myogenic cells express multiple
855 myosin isoforms. *J. Muscle Res. Cell Motil.* 18:501–515.
- 856 Wicker, T., J. S. Robertson, S. R. Schulze, F. A. Feltus, V. Magrini, *et al.* 2005 The repetitive
857 landscape of the chicken genome. *Genome Res.* 15:126–136.
- 858 Wronski, M. A. von, K. I. Hirano, L. M. Cagen, H. G. Wilcox, R. Raghov, *et al.* 1998 Insulin
859 increases expression of apobec-1, the catalytic subunit of the apolipoprotein B mRNA
860 editing complex in rat hepatocytes. *Metabolism.* 47:869–873.
- 861 Zhang, Q., and X. Xiao. 2015 Genome sequence-independent identification of RNA editing
862 sites. *Nat. Methods* 12:347–350.
- 863 Zolov, S. N., and V. V. Lupashin. 2005 Cog3p depletion blocks vesicle-mediated Golgi
864 retrograde trafficking in HeLa cells. *J. Cell Biol.* 168:747–759.
- 865

7 Table 1. Whole transcriptome mRNA editing screening studies in vertebrates.

Study	Species	Cells	Edited sites (N)	Matched DNA ¹	Replicates ²	Potential biases
(Ju <i>et al.</i> 2011)	<i>H. sapiens</i>	Immortalized B cells	1,809	Yes	2/17	Splice, homopolymer, strand, extremity
(Li <i>et al.</i> 2011)	<i>H. sapiens</i>	Immortalized B cells	28,766	Yes	2/27	Splice, homopolymer, multimapping, strand, extremity
(Bahn <i>et al.</i> 2011)	<i>H. sapiens</i>	Glioblastoma cells	10,000	No	2/2	Splice, homopolymer, multimapping, strand, extremity
(Peng <i>et al.</i> 2012)	<i>H. sapiens</i>	Immortalized B cells	22,688	Yes	1/1	Splice, homopolymer, extremity
(Park <i>et al.</i> 2012)	<i>H. sapiens</i>	14 ENCODE cell lines	5,695	No	1/1	Homopolymer, extremity, multimapping
(Kleinman <i>et al.</i> 2012)	<i>H. sapiens</i>	Immortalized B cells	1,503	Yes	1/2	Homopolymer
(Ramaswami <i>et al.</i> 2012)	<i>H. sapiens</i>	ENCODE cell lines	150,865	No	1/2	Strand, multimapping
(Picardi <i>et al.</i> 2012)	<i>H. sapiens</i>	Spinal cord cells	15	Yes (Exome)	1/1	Homopolymer, multimapping, strand
(Bazak <i>et al.</i> 2013)	<i>H. sapiens</i>	16 tissues	1,586,270	No	1/1	Homopolymer, multimapping, strand
(Chen 2013)	<i>H. sapiens</i>	7 ENCODE cell lines	259,385	No	1/2	Multimapping
(Chan <i>et al.</i> 2014)	<i>H. sapiens</i>	Liver cells	20 007	No	1/3	Splice, homopolymer, multimapping, strand, extremity
(Mo <i>et al.</i> 2014)	<i>H. sapiens</i>	Prostate cancer cells	16,194	Yes	2/10	Homopolymer, multimapping, strand, extremity
(Sakurai <i>et al.</i> 2014)	<i>H. sapiens</i>	Brain cells	19,791	No	1/1	Splice, homopolymer, multimapping, strand, extremity
(Toung <i>et al.</i> 2014)	<i>H. sapiens</i>	Immortalized B cells	5,997	No	2/2	Multimapping, strand, extremity
(Zhang and Xiao 2015)	<i>H. sapiens</i>	Immortalized B cells	22,715	No	1/1	Multimapping, extremity
(Hu <i>et al.</i> 2015)	<i>H. sapiens</i>	Hepatocellular carcinoma cells	900	Yes	6/6	Splice, multimapping, strand
(Kang <i>et al.</i> 2015)	<i>H. sapiens</i>	Liver cells	485,684	Yes	1/9	Splice, multimapping, strand
(Danecek <i>et al.</i> 2012)	<i>M. musculus</i>	Brain cells	7,389	No	2/2	Homopolymer, multimapping
(Gu <i>et al.</i> 2012)	<i>M. musculus</i>	Liver, adipose and bone cells	253	No	3/3	Homopolymer, multimapping, extremity
(Lagarrigue <i>et al.</i> 2013)	<i>M. musculus</i>	Liver and adipose cells	63 and 188	No	4/6	Multimapping
(Cattenoz <i>et al.</i> 2013)	<i>M. musculus</i>	Brain cells	665	No	1/1	Splice, homopolymer, multimapping, strand, extremity
(Blanc <i>et al.</i> 2014)	<i>M. musculus</i>	Intestine and liver cells	500	No	1/1	Homopolymer, multimapping, extremity
(Chen <i>et al.</i> 2014)	<i>R. macaque</i>	Prefrontal cortex, cerebellum, muscle, kidney, heart, testis and lung cells	31,250	Yes	1/1	Homopolymer, multimapping, extremity
Frésard <i>et al.</i> 2015	<i>G. gallus</i>	Whole embryo	40	Yes	2/8	-
Roux <i>et al.</i> (The present study)	<i>G. gallus</i>	Liver and adipose cells	11 and 17	Yes	3/8	-

¹ If “yes”: individual genomic DNA information is used to account for potential private individual genomic polymorphisms. If “no”: potential private genomic polymorphisms are defined considering either genomic

Comment citer ce document :

Roux *et al.*

38

Roux, P.-F., Frésard, L., Boutin, M., Leroux, S., Klopp, C., Djari, A., Esquerre, D., Martin, P., Zerjal, T., Gourichon, D., Pitel, F., Lagarrigue, S. (2016). The extent of mRNA editing is limited in chicken liver and adipose, but impacted by tissular context, genotype, age and feeding as exemplified with a conserved edited site in COG3. G3 - Genes Genomes Genetics. sous

0 variants databases such as dbSNPs, or strain-specific consensus genomic sequence in the case of studies
1 based on clonal mouse strains.

2 ² Ratio between the number of biological replicates considered for reporting a candidate difference
3 between DNA and mRNA as a true mRNA editing event and the total number of biological replicates
4 available in the study for a given cell type.

5

Table 2. mRNA editing screening in adult chicken liver and adipose tissue.

	Chromosome	Position	DNA allele	RNA allele	Canonical	Validation in Sanger	Validation in PyroMark	Replicates ¹	Gene names ²	Localisation ³
WAT	1	79605543	T	C	yes			3	PLA1A, POPDC2	Downstream gene
	1	<i>90873162</i>	T	C	yes			3	MPZL1, BRP44	Downstream gene
	1	103511385	T	C	yes			6	GRIK1	Intron
	1	<u>167109833</u>	A	G	yes	yes	yes	3	COG3	Exon (missense)
	1	169769193	A	G	yes			3	THSD1	Downstream gene
	1	187056174	A	G	yes			3	NOX4	Downstream gene
	1	187056183	A	G	yes			3	NOX4	Downstream gene
	1	187057565	A	G	yes			3	-	Intergenic
	2	<u>86000926</u>	T	C	yes	yes	yes	5	NDUFS6	Upstream
	4	17996546	T	C	yes			4	-	Intergenic
	5	22958596	T	C	yes			3	DGKZ	Intron
	11	<u>10634278</u>	A	G	yes			5	CES1	Exon (missense)
	11	19169664	A	C	no	no	no	3	DHODH, IST1	Upstream, Intron
	12	<u>8909946</u>	A	G	yes	yes	yes	6	FLNB	Exon (missense)
	12	8910897	A	G	yes			5	FLNB	Intron
	17	4705	T	C	yes			3	-	Intron
	17	94999	C	T	yes			6	-	Intergenic
	18	28378	T	C	yes			6	ZNF302	Upstream
	19	4812610	C	T	yes			4	CCL18	Intron
Liver	1	<u>79608260</u>	T	C	yes	yes	yes	7	POPDC2, PLA1A	Downstream gene
	1	<i>90873162</i>	T	C	yes			3	MPZL1, BRP44	Downstream gene
	1	90873188	T	C	yes	yes	yes	4	MPZL1, BRP44	Downstream gene
	1	<u>167109833</u>	A	G	yes	yes	yes	5	COG3	Exon (missense)
	1	193343226	A	G	yes			6	MADPRT, ART7B	Upstream
	2	<u>86000926</u>	T	C	yes	yes	yes	5	NDUFS6	Upstream
	6	<u>13236424</u>	A	G	yes			3	KCNMA1	Exon (missense)
	7	<u>7564791</u>	A	G	yes	yes	yes	5	MYO1B	Intron
	8	25588113	T	C	yes			3	-	Exon (synonymous)
	28	518606	A	C	no	yes	no	3	HNRNPM	Missense
LGE64	615592	T	C	yes			4	-	Intron	

¹ Number of samples in which the mRNA editing event is detected.

² Name of the gene impacted by the mRNA editing event or name of the closest genes (< 10kb) if the mRNA editing event is falling in an intergenic region.

³ Localization of the mRNA editing event inside genomic features, as predicted by Variant Effect Predictor (McLaren *et al.* 2010). If the mRNA editing event is falling inside a coding region, its impact on gene product is given in brackets.

In bold: mRNA editing candidate events subjected to Sanger genomic DNA sequencing and cDNA pyrosequencing; Underlined: mRNA editing events annotated as "coding - missense" and located in 4 different genes. In italics: mRNA editing events common to both tissues.

Comment citer ce document :

Roux *et al.*

Roux, P.-F., Frésard, L., Boutin, M., Leroux, S., Klopp, C., Djari, A., Esquerre, D., Martini, P., Zerjal, T., Gourichon, D., Pitel, F., Lagarrigue, S. (2016). The extent of mRNA editing is limited in chicken liver and adipose, but impacted by tissular context, genotype, age and feeding as exemplified with a conserved edited site in COG3. G3 - Genes Genomes Genetics. sous

884 **FIGURE LEGENDS**

885 **Figure 1. mRNA editing screening studies based on high throughput sequencing in the**
886 **literature.**

887 This graph describes the numbers of mRNA editing events detected (\log_{10}) across RNA-seq-
888 based screening studies as a function of the numbers of biological replicates considered to
889 declare an event as a true positive.

890 **Figure 2. Impact of sequencing and mapping biases on mRNA editing discovery.**

891 Contribution of random or systematic sequencing biases and mapping artifacts to the false
892 discovery of mRNA editing events using combined mRNA and DNA sequencings.

893 **Figure 3. Impact of biological replication on mRNA editing discovery.**

894 Distribution (in %) of unbiased mRNA editing events across the 12 classes of substitution
895 according to the number of replicates they are detected in, ranging from N=1 to N=3, in
896 white adipose tissue (WAT) and liver. The first two classes (AtoG and TtoC) are associated to
897 ADAR-mediated RNA editing, and the next two ones (CtoT and GtoA) to APOBEC-mediated
898 RNA editing. At the top-right of each graph, the total number of RNA editing events
899 detected for a given number of replicates is shown.

900 **Figure 4. Position of mRNA editing events across the chicken genome.**

901 **Figure 5. Distribution of mRNA editing events across genomic features.**

902 Annotations were assessed using Ensembl v71 Variant Effect Predictor (McLaren *et al.*
903 2010).

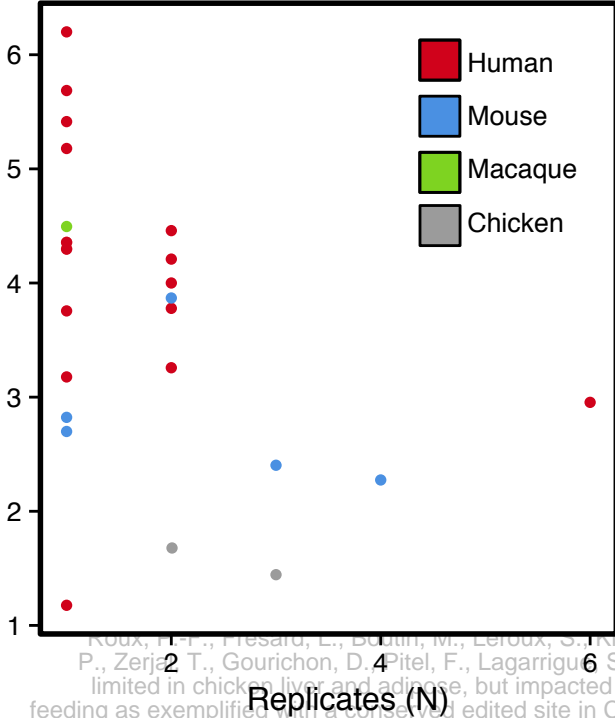
904 **Figure 6. Multispecies protein sequence alignments for coding mRNA editing events.**

905 The red stars indicate the position of the amino acid impacted by coding mRNA events. The
906 overall conservation across sequences is depicted below each alignment. The mRNA editing
907 event impacting *COG3* was detected in both white adipose tissue and liver, while the ones
908 impacting *CES1* and *FLNB* were WAT-specific, and the one impacting *KCMA1* was specific to
909 liver

910 **Figure 7. Impact of genetic background, age, feeding condition and sex on mRNA editing**
911 **level.**

912 Editing level (in %) at 5 genomic positions, in white adipose tissue (WAT) and liver according
913 to (A) genetic background, (B) age, (C) feeding condition and (D) sex. Each boxplot shows
914 the distribution of editing levels across N=8 biological replicates. * $p < 0.05$, ** $p < 0.01$, ***
915 $p < 0.001$, unpaired two-tailed Student *t*-test. Pre-pub.: prepuberal animals; Post-pub.:
916 postpuberal animals; M: Males; F: Females.

Number of editing events detected (\log_{10})



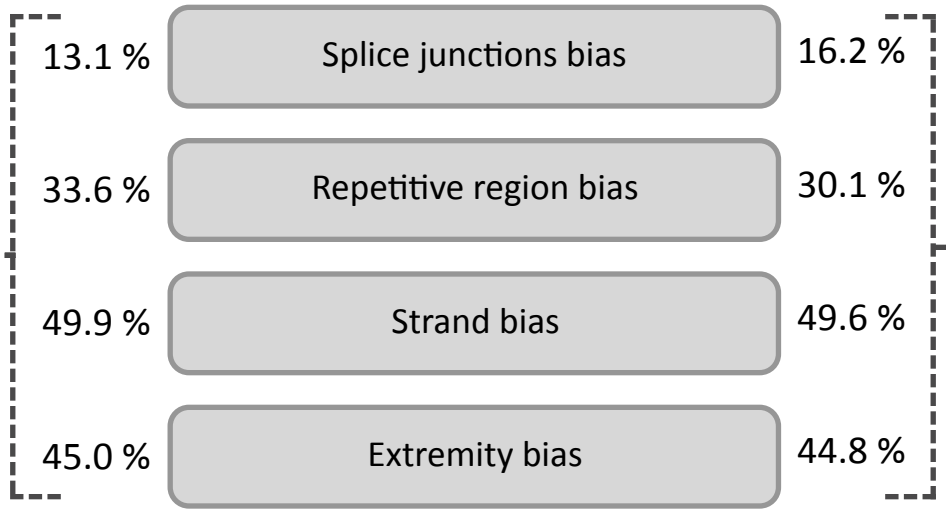
WAT

3229

Liver

2305

Candidate RDDs



Unbiased RDDs
classical filters

Mapping bias

Unbiased RDDs
classical filters + mapping filter

448

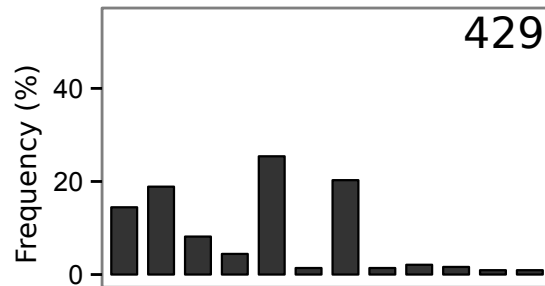
342

429

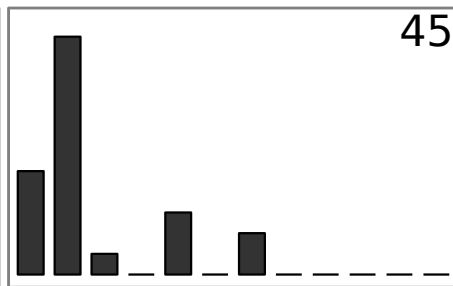
315

Roux, P.-F., Frésard, D., Pilon, P., Lagarde, S. (2016). The extent of miRNA editing is limited in chicken liver and adipose, but impacted by tissular context, genotype, age and feeding as exemplified with a conserved edited site in COG3. G3 - Genes Genomes Genetics. sous

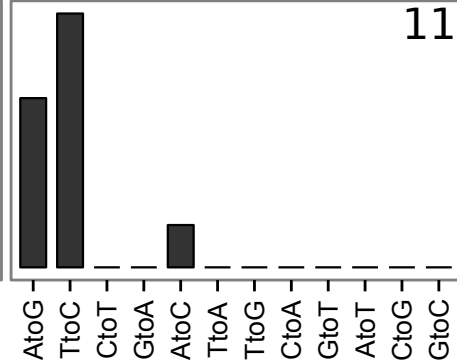
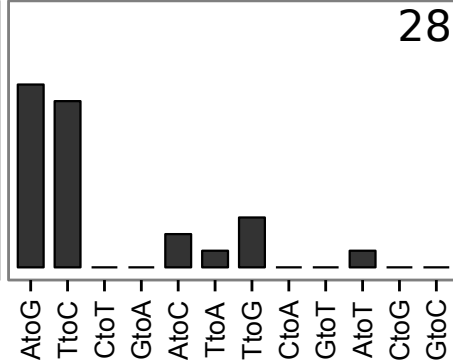
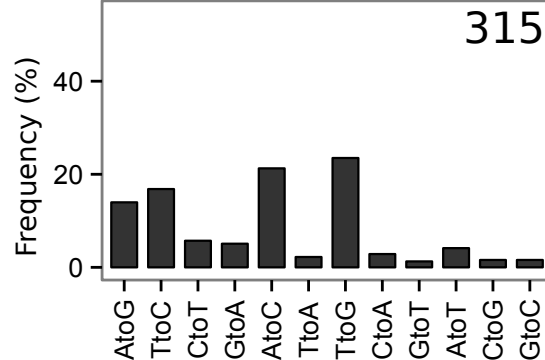
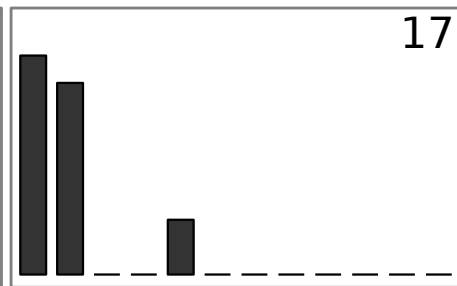
N=1



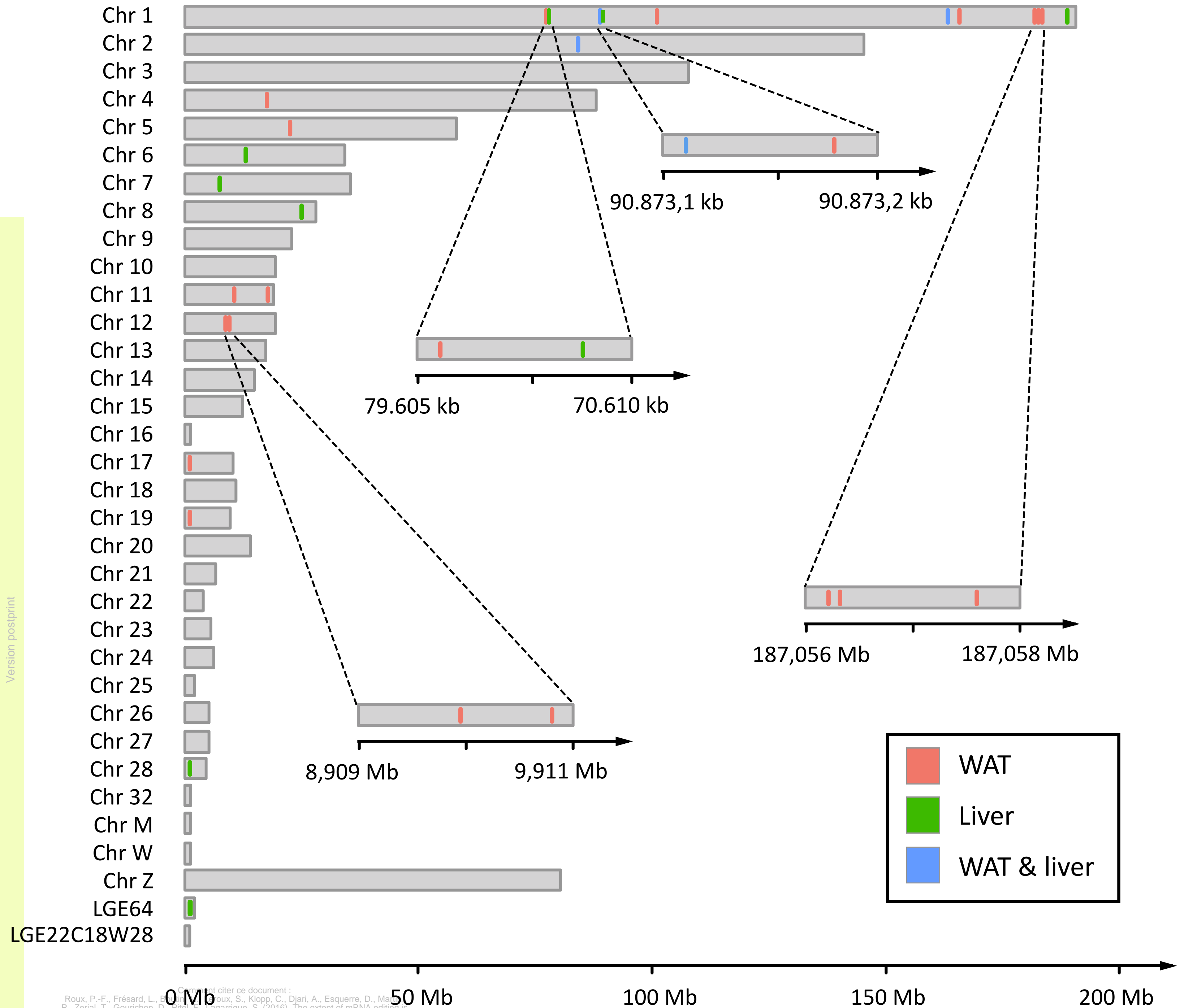
N=2

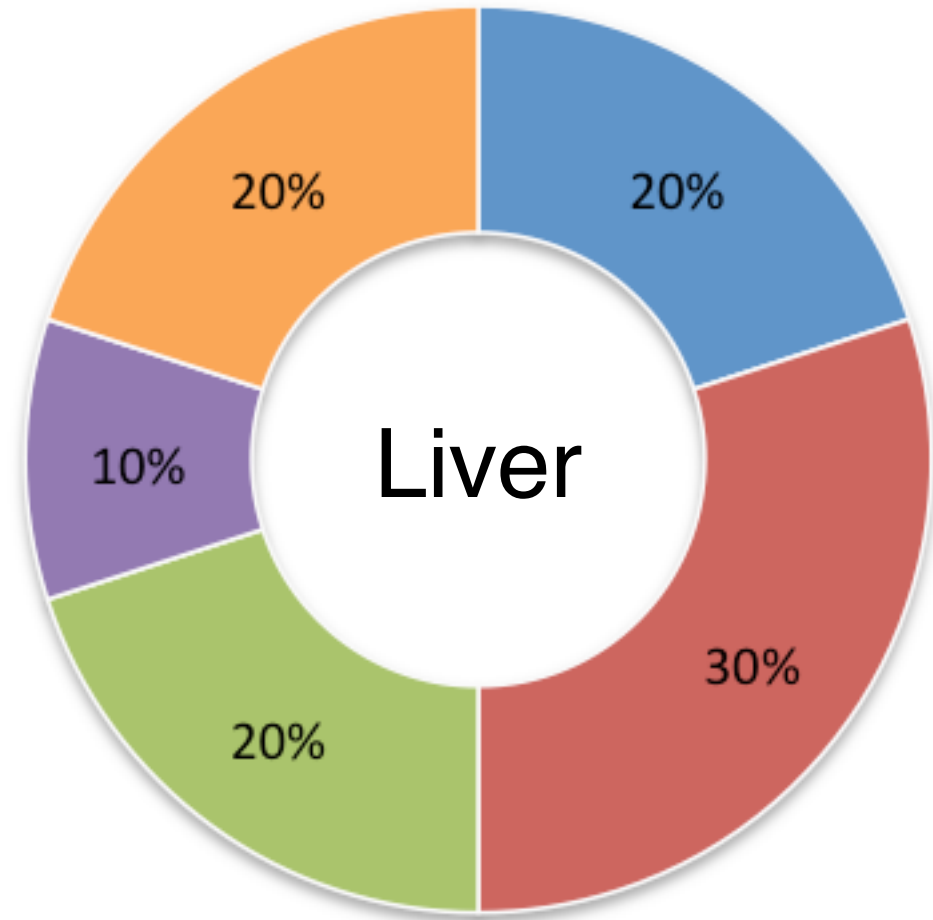
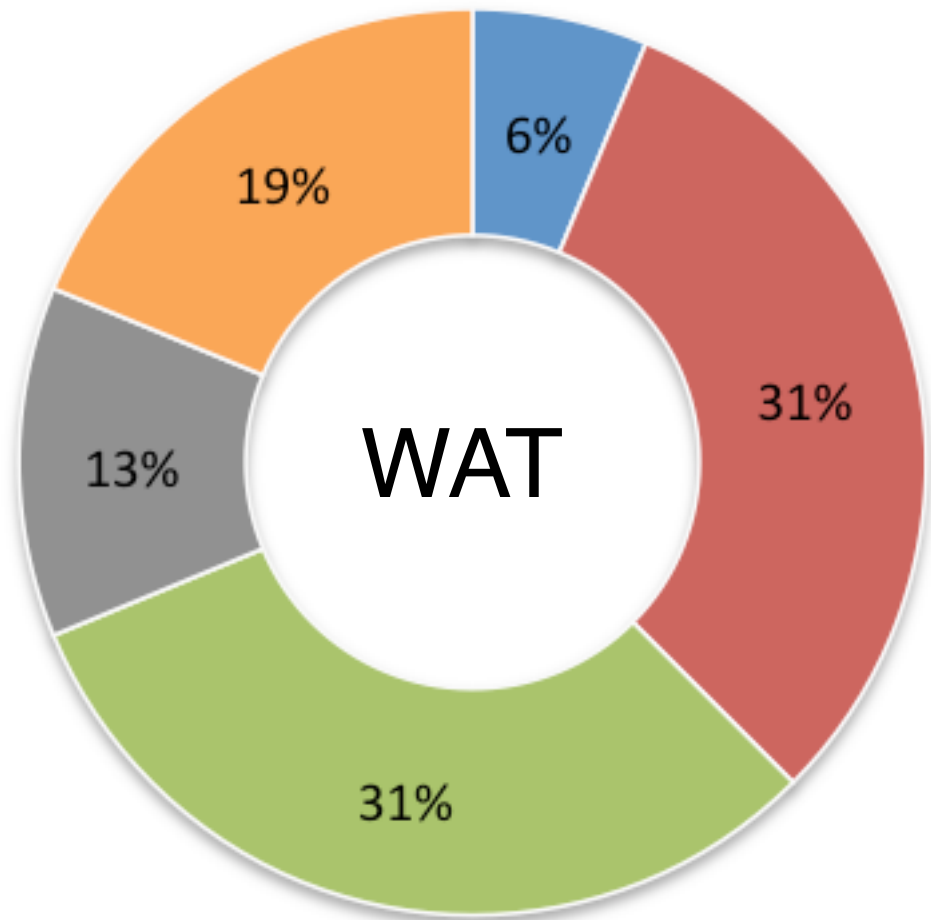


N=3



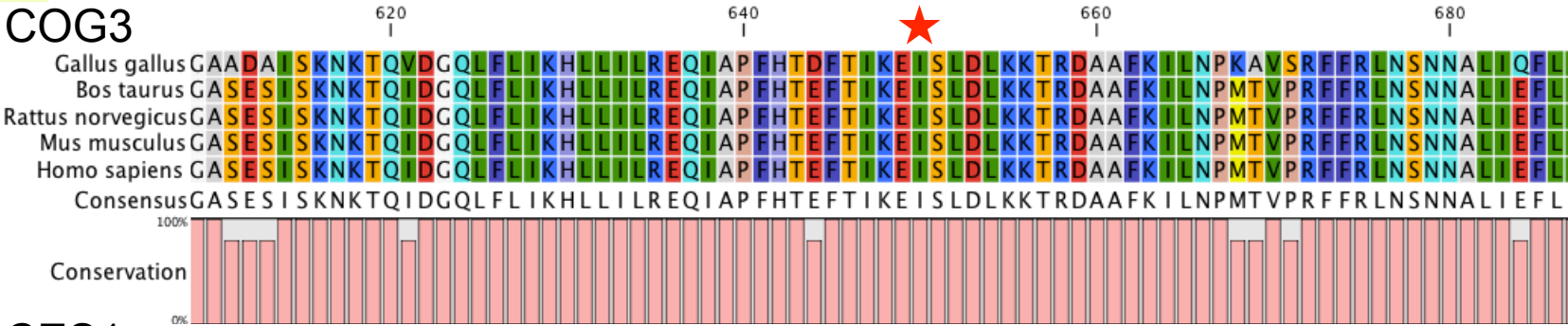
RDD type



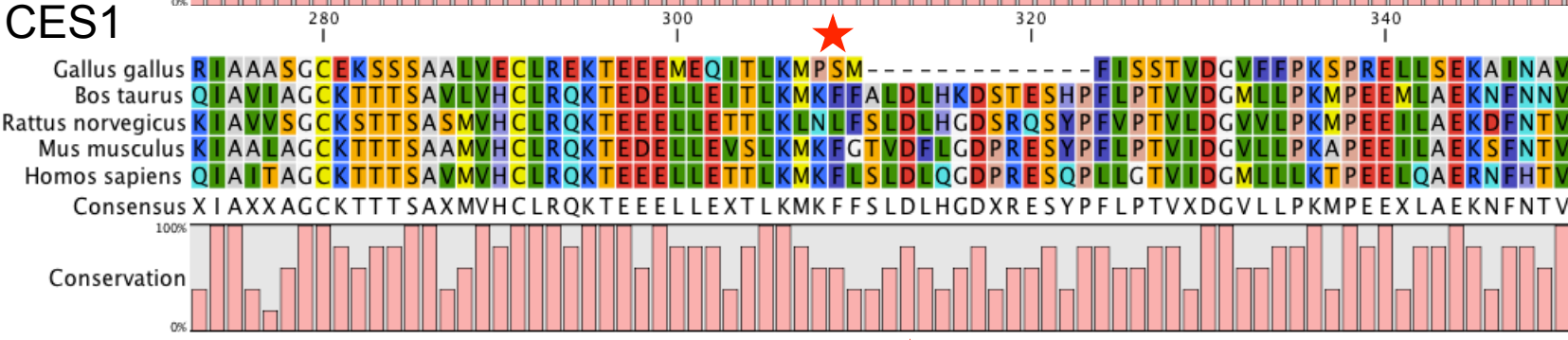


- Upstream gene (- 5 kb)
- Downstream gene (+ 5 kb)
- Intronic region
- Intergenic region
- Coding region - Synonymous
- Coding region - Missense

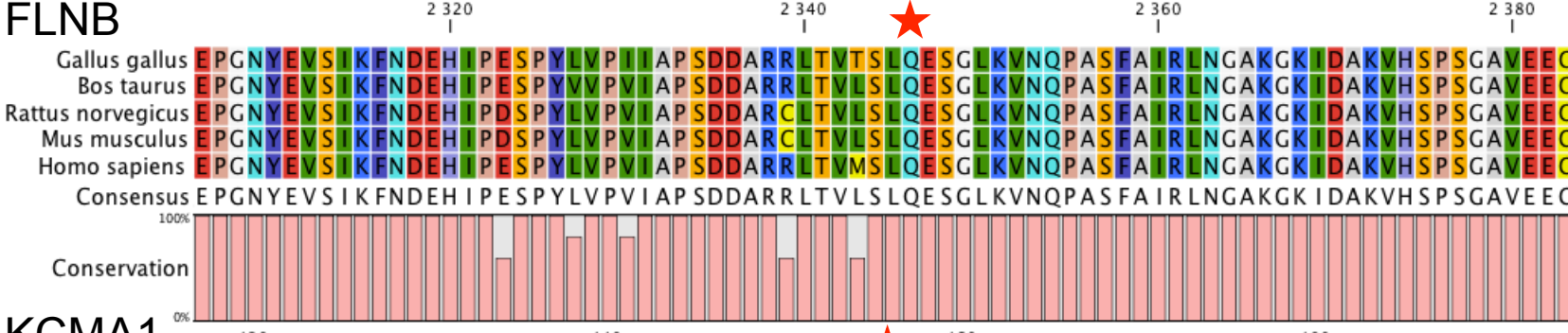
COG3



CES1



FLNB



KCMA1

