



HAL
open science

Genetic diversity of *Echinococcus granulosus sensu stricto* infecting humans in western Algeria

Daouia Moussa, Kheira Senouci, Nori Midoun, Mohamed Lacheheb, Benali Tabeti, Noureddine Benmaarouf, Christophe Hennequin

► To cite this version:

Daouia Moussa, Kheira Senouci, Nori Midoun, Mohamed Lacheheb, Benali Tabeti, et al.. Genetic diversity of *Echinococcus granulosus sensu stricto* infecting humans in western Algeria. *Parasitology Research*, 2021, 120 (9), pp.3195-3202. 10.1007/s00436-021-07223-7. hal-03365912

HAL Id: hal-03365912

<https://hal.sorbonne-universite.fr/hal-03365912v1>

Submitted on 5 Oct 2021

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.

1 **Genetic diversity of *Echinococcus granulosus sensu stricto* infecting humans in western**
2 **Algeria**

3

4 Daouia Moussa¹, Kheira Senouci¹, Nori Midoun², Mohamed Lacheheb³, Benali Tabeti⁴,
5 Nouredine Benmaarouf⁴, Christophe Hennequin⁵.

6

7 1. Natural and Life Sciences Faculty, Department of Biology, University of Oran 1 Ahmed
8 Ben Bella, Oran, Algeria.

9 2. Department of Epidemiology and Preventive Medicine, University Hospital of Oran
10 (EHU), Oran, Algeria.

11 3. Department of Thoracic Surgery, University Hospital of Oran (EHU), Oran, Algeria.

12 4. Department of Hepatobiliary and Liver Transplant, University Hospital of Oran (EHU),
13 Oran, Algeria.

14 5. Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, CRSA, AP-HP, Hôpital
15 Saint-Antoine, Service de Parasitologie-Mycologie, F-75012 Paris, France

16

17

18 **Corresponding author:**

19 C. Hennequin, Service de Parasitologie-Mycologie, Hôpital St Antoine, 34 rue Crozatier,
20 75012 Paris, France

21 Phone + 33 1 4928 3412

22 Fax +33 1 4928 2251

23 Email: christophe.hennequin-sat@aphp.fr

24 **Abstract:**

25 Human cystic echinococcosis is a zoonosis due to the flat worm *Echinococcus granulosus sensu*
26 *lato*. The disease remains a major public health problem in Northern Africa. Molecular typing
27 enables a better understanding of the parasite circulation from animals to humans. In this study
28 we investigated the genotypic diversity of 46 *Echinococcus granulosus* isolates collected from
29 humans in the western part of Algeria by the mean of partial sequences of 4 mitochondrial loci,
30 namely *cox1a*, *cox1b*, *nd3* and *atp6*. Nucleotide polymorphism range from 0.6% (*nd3*) to 2.7%
31 (*cox1a*). Eight alleles had not been previously reported. Multilocus analysis showed that all the
32 isolates were from the *Echinococcus granulosus sensu stricto* (G1 genotype). Nineteen
33 different haplotypes made of the concatenation of 4 sequenced loci were observed, the most
34 common type clustering 13 isolates (36.1%). Twelve of these haplotypes had never been
35 described previously and fifteen (41.7%) haplotypes were represented by only one isolate.
36 Using sequences from this study and others retrieved from the Genbank database, any clustering
37 either according to the geographic origin within Algeria or according to the human or animal
38 origin of the isolates could be demonstrated supporting that genotype G1 population genetics
39 has been shaped by intensive animal breeding.

40

41

42

43

44

45 **Keywords:** *Echinococcus granulosus*; *cox1*; *nd3*; *atp6*; mitochondrion; genetic diversity;

46 Algeria.

47

48

49 The authors declare that no funding was obtained for this study

50 The authors declare to have nothing to disclose

51 All the data and remaining materials are available to the scientific community

52 DM collected the samples, performed all the experiments and writes a first draft of the paper.

53 KS supervise the biological diagnosis and review the paper. NM, ML, NB and BT took care

54 of the patients and review the paper.

55 CH supervised the experiments, analyzed the data and writes the paper

56

57

58

59

60 **Introduction**

61 Cystic echinococcosis (CE) is due to the development of the larval stage of the cestode
62 *Echinococcus granulosus sensu lato (s.l.)*. Human CE is a cosmopolitan zoonosis that
63 represents a major public health problem in endemic countries where collective prophylaxis
64 measures remain insufficient, notably in breeding areas such as North Africa (Dakkak 2010;
65 Eckert et al. 2001). This is the case of Algeria where, despite improved educational hygiene
66 and control over the slaughter of animals, the prevalence of CE remains high, with about 700
67 cases notified yearly (Dakkak 2010). This is due to the domestic life cycle of the parasite where
68 dogs are the predominant definitive host of the parasite and the main source of contamination
69 for humans, while intermediate hosts are represented by various species of ruminants. Previous
70 reports from Algeria have shown that 16 to 42% of stray-dogs harbor *E. granulosus s.l.* in their
71 intestine (Bentounsi et al. 2009), and that sheep, goat, cattle, horses, dromedaries, and wild
72 boars can act as intermediate hosts, sheep being by far the most common (Bardonnnet et al. 2003;
73 Eckert et al. 2001; Kouidri et al. 2014; Laatamna et al. 2019; Maillard et al. 2007; Zait et al.
74 2016).

75 A high degree of genetic diversity has been documented in *E. granulosus s.l.* that partly
76 correlates with intermediate host susceptibility (Alvarez Rojas et al. 2014; Romig et al. 2015).
77 According to the current nomenclature, *Echinococcus granulosus s.l.* is now considered a
78 species complex that includes *Echinococcus granulosus sensu stricto (s.s.)*, *Echinococcus*
79 *equinus*, *Echinococcus ortleppi*, *Echinococcus felidis* and *Echinococcus canadensis* (Alvarez
80 Rojas et al. 2014; Romig et al. 2015). *E. granulosus s.s.* is the predominant species in many
81 regions of the world, while *E. canadensis*, G6/G7 and, G8, G10 genotypes are dominant in
82 eastern and north of Europe, respectively (Alvarez Rojas et al. 2014). *E. granulosus s.s.* includes
83 2 genotypes denoted G1 and G3, G2 being now considered a microvariant of G3 (Kinkar et al.
84 2017). The G1 genotype is predominant in all intermediate hosts, including Humans and is

85 widely found in cattle in North Africa (Alvarez Rojas et al. 2014; Deplazes et al. 2017). G3,
86 initially identified from water buffalo, can also infect other intermediate hosts including
87 Humans (Romig et al. 2015).

88 Despite the importance of this public health problem in Algeria, epidemiological studies, and
89 particularly molecular epidemiological studies, focused on *E. granulosus s.l.* from this country
90 are still scarce. They conclude on the predominance of the *E. granulosus s.s.* G1 genotype as
91 the main source of human infections (Bardonnet et al. 2003; Maillard et al. 2007; Zait et al.
92 2016). Some of these surveys may have limited conclusion notably due to length of the
93 nucleotide sequence analyzed. However, such studies are important for different purposes.
94 Even considering that the genotype is not strictly restricted to one intermediate host, the
95 identification of the predominant species in a region would help health authorities to drive their
96 prevention measures. Results of such studies can also give light on the circulation of the parasite
97 between different regions, and to analyze the different hosts involved, including humans.

98 To increase our knowledge on the genetic diversity of *E. granulosus s.l.* circulating in Algeria,
99 we initiated a genotyping study based on a 4 partial mitochondrial loci system focused on
100 human isolates from the Western part of Algeria, where sheep breeding is mostly intensive.

101

102 **Materials and methods**

103 **Biological samples**

104 Fluid from of 46 hydatid cysts (19 pulmonary and 27 hepatic) obtained after surgical resection
105 from 45 patients admitted at the Hospital of Oran (EHU) were available for the genotyping
106 study (Online Resource 1). Patients were from different localities mainly in the north-west of
107 Algeria (Figure S1). Fluids were first stored in 0.9% sodium chloride solution to avoid
108 protoscolex desiccation. They were then centrifuged at 50xg for 5min and the pellet containing
109 protoscolex and/or the germinal layer were stored at -20°C in 90% ethanol for further analysis.

110 **Molecular analysis**

111 *DNA extraction*

112 DNA was extracted from hydatid tissue samples (protoscolex and/or germinal layer) using a
113 mechanical disruption (MagnaLyser, Roche) followed by column-based extraction (QIAamp
114 DNA Blood kit, Qiagen) according to the manufacturer's instructions.

115 *Amplification and direct sequencing*

116 Four mitochondrial fragments were selected for genotyping (Figure S2). They target the
117 cytochrome c oxidase subunit 1 (2 fragments referred as *cox1a* and *cox1b*), the dehydrogenase
118 subunit 3 (*nd3*) and the ATPase subunit 6 (*atp6*). Briefly, amplification was performed with a
119 mixture containing 0.3 U Taq polymerase (Taq Gold, Invitrogen), 1µM forward and reverse
120 primers and 5 µl DNA in a final volume of 25 µl. Primers have been previously described and
121 their sequences are available in Online Resource 2 (Bowles et al. 1992; Lavikainen et al. 2006).
122 For all the targets, amplification was achieved according to the following thermal program: 10
123 min at 94 °C, followed by 35 cycles of denaturation at 94°C; hybridization at 50°C and
124 elongation at 72°C, each of these steps lasted 30 sec; then a final extension at 72 °C during 10
125 min. Amplification was checked by running a 1% agarose gel electrophoresis.

126 PCR products were then purified (QIAquick PCR Purification kit) and subjected to direct
127 sequencing with the BigDye Terminator v3.1 reagent (Invitrogen), according to the
128 manufacturer's protocol and using the primers used for amplification. The corresponding
129 chromatograms were obtained using the 3500xl Dx Genetic Analyzer automatic sequencer.

130 **Phylogenetic analysis**

131 The sense and anti-sense chromatograms were aligned and manually edited using the BioEdit
132 software version 7.2.6. Consensus sequences were then compared by the mean of the Blastn
133 program available on the National Center for Biotechnology Information website
134 (<https://blast.ncbi.nlm.nih.gov/>). Single locus multiple alignments were performed including
135 the reference mitochondrion genome (Genbank accession number AB786664) published by
136 Nakao et al. (Nakao et al. 2013) and using the MegaX (vs 10.1.8) software with the Clustal
137 algorithm.

138 Phylogenetic analysis was achieved using these sequences and sequences available in Genbank
139 and corresponding to Algerian *E. granulosus s.s.* isolates (Bardonnnet et al. 2003; Kinkar et al.
140 2018; Laatamna et al. 2019; Maillard et al. 2007; Zait et al. 2016). We performed both an
141 analysis using the *cox1* locus because a higher number of sequences are available in Genbank
142 with 100% query cover and another one using the concatenated sequences of the 4 loci further
143 in called haplotypes, obtained in our study. The phylogenetic relationship between sequences
144 identified in our series and sequences retrieved in Genbank was illustrated by the construction
145 of minimum spanning trees using an optimal eBurst algorithm (Francisco et al. 2009). Graphical
146 representation was done using the Phyloviz online program (<http://www.phyloviz.net/>). In this
147 graphical representation, isolates with the same allelic profile fall in the same circle, the size of
148 which is proportional to the number of isolates with that particular profile.

149

150 **Ethics statement**

151 The protocol has been approved by the Ethics Committee of the Hospital of Oran (ref. CA
152 O2/2020). Patients were informed during their hospital stay that data from their medical charts
153 could be used in a medical study that would be possibly published in a scientific journal. At the
154 beginning of the study, each specimen was given a code and all the experiments were performed
155 without knowledge of the patients' information, so that anonymization did not allow their
156 identification.

157 **Results**

158 Amplification failed for the 4 genetic targets for two specimens. Positive amplification and
159 sequencing were obtained for 37, 42, 43 and 43 isolates for *atp6*, *nd3*, *cox1a*, *cox1b*, respectively
160 making complete haplotype available for 36 isolates.

161 Single locus analysis revealed 8, 3, 10, 6 alleles for locus 1 to 4, respectively (Table 1). The level
162 of diversity with highest for *cox1a* locus (2.7%) and lowest for *nd3* locus (0.6%). Eight of these
163 alleles have not been reported previously and have been deposited in GenBank (from
164 MW417352 to MW417359 and MW446898-MW446899).

165 Phylogenetic analysis of concatenated sequences from our study and those found in GenBank
166 showed that all our isolates were of *E. granulosus s.s.* G1 genotype (data not shown). Multilocus
167 analysis on the 36 isolates with a complete haplotype allowed the differentiation of
168 19 haplotypes named Hap01 to Hap19 (Table 2). Hap01 was the most common type isolates
169 (36.1%). Twelve of these haplotypes had never been reported previously and fifteen (41.7%)
170 haplotypes were represented by only one isolate. Notably 56% of the pulmonary isolates had
171 unique haplotype versus 20% for hepatic isolates.

172 The first phylogenetic tree was based on the *cox1b* locus taking the advantage of 92 homolog
173 sequences in GenBank (Kinkar et al. 2018; Laamanna et al. 2019; Zait et al. 2016). It shows that
174 36 (83.7%) of our isolates harbor the predominant sequence (*cox1b_01*), similar to the reference
175 mitochondrial genome. Sequences from all the isolates but 2 were also direct satellites from
176 this central clade with links corresponding to a single nucleotide mutation. No segregation
177 according to the host nor to the geographic origin could be demonstrated (Fig 1a, Fig 1b). When
178 using complete concatenated haplotypes using the sequences of four loci tested in our study,
179 Hap01 appeared as the founder, giving birth to closely related haplotypes either tested in this
180 study or retrieved in Genbank. However, Hap10, which originated indirectly from Hap01 (2
181 nucleotide mutations), forms a second cluster distantly with further diversification (Figure 2).

182 All the sequences were obtained from cysts collected from humans and no apparent geographic
183 segregation trend was observed between these haplotypes.

184

185 **Discussion**

186 Because the life cycle of *E. granulosus s.l.* includes domestic animals, mainly dogs and
187 livestock, eradication of CE can be reached by control measures that could be set up from data
188 obtained in epidemiological surveys. There are several studies focused on the genetic
189 diversity of *E. granulosus s.l.* conducted in the Maghreb but those conducted in Algeria are
190 more rare (Bardonnet et al. 2003; Boufana et al. 2014; Kinkar et al. 2018; Laatamna et al.
191 2019; M'Rad et al. 2020; Oudni-M'rad et al. 2016; Zait et al. 2016).

192 In our study, we were able to genotype 44 *E. granulosus s.l.* isolates collected from humans
193 living in western Algeria. Previous studies have shown that whatever the country or the host
194 mitochondrial targets offer a high power of discrimination in regards to nuclear targets
195 (Laurimae et al. 2016; Lavikainen et al. 2006; Maillard et al. 2007). We used a multilocus
196 sequence typing scheme based on 4 fragments encoding mitochondrial genes with 1,626
197 nucleotides analyzed. This allowed the description of 19 haplotypes of which 12 had never been
198 reported, with 13 alleles newly described in Algeria of which 8 had not been previously
199 deposited in GenBank.

200 *E. granulosus s.s.* G1 genotype was the unique genotype identified for the all isolates tested in
201 this study. This is in accordance with previous studies conducted in Algeria that found a
202 frequency of 90.7 to 100% of this genotype in CE cases (Bardonnet et al. 2003; Laatamna et al.
203 2019; Zait et al. 2016). This also agrees with the worldwide predominance of *E. granulosus s.s.*
204 that would cause 88.44% of the human CE cases (Alvarez Rojas et al. 2014). Surveys conducted
205 in Tunisia and Morocco, two neighboring countries, also confirm that G1 genotype is
206 predominant as causing human CE in North Africa (M'Rad et al. 2005; Oudni-M'rad et al. 2016;
207 Tahiri et al. 2019). However, while not detected in our study, it should be noted that both *E.*
208 *granulosus s.s.* (G3) and *E. canadensis* (G6) have already been detected in Algeria (Maillard et
209 al. 2007; Zait et al. 2016), the latter being mostly found in southern parts of Algeria where

210 camelid breeding is common. These results are in accordance with the predominance of the G1
211 genotype in North Africa, warranting further studies to better understand the circulation of this
212 species between the various hosts

213 Molecular epidemiology surveys having analyzed the genetic diversity of *E. granulosus s.l.*
214 diversity in Algeria remain rather rare. Zait et al. using a two-loci molecular system (*cox1b* and
215 *nad1*) identified twenty different haplotypes among the 70 isolates tested, mostly originating
216 from the Eastern part of Algeria, with a predominant one aggregating 71.6% of the isolates
217 being identical to the reference mitogenome (Zait et al. 2016). More recently, Laatamna et al,
218 used the full-length sequence of the *cox1* gene to demonstrate a particularly high genetic
219 diversity (73 different haplotypes among 125 isolates tested) in *E. granulosus* from livestock
220 animals from the Djelfa province(Laatamna et al. 2019). However, they also concluded on the
221 identity between the predominant haplotype and the reference mitogenome.

222 To enhance our knowledge on the genetic diversity of *E. granulosus s.s.* in Algeria, we
223 compared our data with previously published sequences. Comparison was limited due to
224 different selections of genetic targets and the length of sequences analyzed. Nevertheless, we
225 were able to compared the haplotype of 92 isolates collected from different regions of Algeria
226 using a part of the *cox1*gene (locus *cox1b* of our study). Indeed, partial or complete sequence
227 of the *cox1* gene has been the most commonly used target for *E. granulosus s.l.* genotyping
228 studies. Using this locus, we confirm the predominance of the *cox1a_01*allele, synonymous to
229 the EG01 genotype, previously shown to cause most of human and animal CE cases in Algeria
230 (Laatamna et al. 2019; Zait et al. 2016). No segregation between isolates collected from humans
231 or animals was noted supporting the integration of both hosts in the life cycle of the parasite.
232 Similarly, Algerian isolates of the G1 genotype irrespective of their geographic origin, shared
233 a predominant haplotype or closely-related haplotypes, supporting the clonal expansion of the
234 parasite throughout the country without host barrier, with further diversification (satellite

235 haplotypes with limited genetic variation). This is in accordance with the mode of reproduction
236 of the parasite that undergoes a sexual reproduction from a single worm thanks to a
237 hermaphrodite genital apparatus.

238 However, using the 4 loci scheme, we were able to detect a substantially higher level of genetic
239 diversity within the G1 genotype than using a single locus. All the haplotypes detected in our
240 study were genetically close to but different from the G1 reference mitogenome available in
241 GenBank (AB786664) that has been isolated from a man in China (Nakao et al. 2013). As
242 previous studies conducted in other parts of Algeria, we showed that the Hap01 haplotype is
243 predominant in western Algeria. This synonymous nucleotide mutation, in a locus rarely
244 included in the typing method, has already been described in Italian, Australian, Tunisian and
245 Algerian isolates (Kinkar et al. 2019; Kinkar et al. 2018; Sgroi et al. 2019). The Hap01
246 haplotype belongs to a clonal complex that includes a low level of diversity with satellite
247 haplotypes (9 haplotypes diverge only by one single nucleotide mutation). However, the second
248 most frequent haplotype Hap02, that derives from Hap01, seems to form a new complex that
249 gives birth to new closely related haplotypes. Whether this is due to the natural evolution of the
250 species or the introduction of either cattle or sheep, or dogs from other regions should be further
251 evaluated. The contamination of humans in other regions can also be hypothesized.

252 Interestingly, this haplotype has already been described from Kenchela, located in eastern
253 Algeria. Similarly, Hap04 was shared by two isolates from Boumerdes and Laghouat, 2 cities
254 distant from western Algeria. More recently, whole mitochondrial genome sequence using
255 primer walking has been proposed for *E. granulosus s.l.* typing and offers a better
256 discriminatory power and capability to analyze phylogenetic relationship). So, Kinkar et al.
257 showed that Algerian isolates of *E. granulosus s.l.* are split into different geographical clusters
258 that also include European and North African isolates (Kinkar et al. 2018). This and results
259 from other studies (Boufana et al. 2014; Hassan et al. 2017; Kinkar et al. 2018), notably showing

260 the low fixation indices between human and animal isolates, strongly suggest that genotype G1
261 population genetics has been shaped by intensive animal trade.

262 While the results of the present study are based on a relatively low number of isolates, they
263 highlight the particularly high genetic diversity of *E. granulosus s.s.* in Algeria, with some
264 specific features. To improve our knowledge and better understand how the parasite spread into
265 this country and elsewhere, a consensus typing method should be defined allowing to aggregate
266 together the data from offering the opportunity to conduct robust phylogeography analysis. This
267 would be further applied to type both human and animal *E. granulosus s.s.* isolates

268

269

270 **Acknowledgements**

271 This study was supported by the Parasitology laboratory of the University of Oran1 Ahmed

272 Ben Bella, the Department of Parasitology at the Hospital of Oran (EHU).

273 The authors would like to thank Chafika Tachema for referring the specimens.

274 The authors also thank the Joint Laboratory of Genetics and Molecular Biology and Sandra

275 Vellaissamy from the Parasitology Department of St Antoine Hospital for excellent technical

276 assistance.

277

278 **Conflict of interest**

279 The authors state that they have no conflict of interest.

280

281 **References**

- 282 Alvarez Rojas, C. A., Romig, T., Lightowlers, M. W. (2014). *Echinococcus granulosus sensu*
283 *lato* genotypes infecting humans--review of current knowledge. Int J Parasitol 44 (1),
284 9-18. doi:10.1016/j.ijpara.2013.08.008
- 285 Bardonnet, K., Piarroux, R., Dia, L., Scheegans, F., Beurdeley, A., Godot V., Vuitton DA
286 (2002) Combined eco-epidemiological and molecular biology approaches to assess
287 *Echinococcus granulosus* transmission to human in Mauritania: occurrence of the camel
288 strain and human cystic echinococcosis. Trans R Soc Trop Med Hyg 96:383–386
- 289 Bardonnet K, et al. (2003) Cystic echinococcosis in Algeria: cattle act as reservoirs of a sheep
290 strain and may contribute to human contamination. Veterinary parasitology 116(1):35-
291 44 doi:10.1016/s0304-4017(03)00255-3
- 292 Alvarez Rojas CA, Romig T, Lightowlers MW (2014) *Echinococcus granulosus sensu lato*
293 genotypes infecting humans--review of current knowledge. Int J Parasitol 44(1):9-18
294 doi:10.1016/j.ijpara.2013.08.008
- 295 Bardonnet K, et al. (2003) Cystic echinococcosis in Algeria: cattle act as reservoirs of a sheep
296 strain and may contribute to human contamination. Vet Parasitol 116(1):35-44
297 doi:10.1016/s0304-4017(03)00255-3
- 298 Bentounsi B, Meradi S, Ayachi A, Cabaret J (2009) Cestodes of Untreated Large Stray Dog
299 Populations in Algeria: A Reservoir for Herbivore and Human Parasitic Diseases. The
300 Open Veterinary Science Journal 3:64-7
- 301 Boufana B, et al. (2014) Genetic variability and haplotypes of *Echinococcus* isolates from
302 Tunisia. Trans R Soc Trop Med Hyg 108(11):706-14 doi:10.1093/trstmh/tru138
- 303 Bowles J, Blair D, McManus DP (1992) Genetic variants within the genus *Echinococcus*
304 identified by mitochondrial DNA sequencing. Mol Biochem Parasitol 54(2):165-73
305 doi:10.1016/0166-6851(92)90109-w
- 306 Dakkak A (2010) Echinococcosis/hydatidosis: a severe threat in Mediterranean countries. Vet
307 Parasitol 174(1-2):2-11 doi:10.1016/j.vetpar.2010.08.009
- 308 Deplazes P, et al. (2017) Global Distribution of Alveolar and Cystic Echinococcosis. Adv
309 Parasitol 95:315-493 doi:10.1016/bs.apar.2016.11.001
- 310 Eckert J, Gemmell M, Meslin Fo-X, Pawlowski Z, Organization WH (2001) WHO/OIE
311 manual on echinococcosis in humans and animals: a public health problem of global
312 concern.

313 Francisco AP, Bugalho M, Ramirez M, Carrico JA (2009) Global optimal eBURST analysis
314 of multilocus typing data using a graphic matroid approach. BMC Bioinformatics
315 10:152 doi:10.1186/1471-2105-10-152

316 Hassan ZI, et al. (2017) Two haplotype clusters of *Echinococcus granulosus sensu stricto* in
317 northern Iraq (Kurdistan region) support the hypothesis of a parasite cradle in the
318 Middle East. Acta Trop 172:201-207 doi:10.1016/j.actatropica.2017.04.028

319 Kinkar L, et al. (2019) Long-read sequencing reveals a 4.4 kb tandem repeat region in the
320 mitogenome of *Echinococcus granulosus (sensu stricto)* genotype G1. Parasit Vectors
321 12(1):238 doi:10.1186/s13071-019-3492-x

322 Kinkar L, et al. (2018) Global phylogeography and genetic diversity of the zoonotic
323 tapeworm *Echinococcus granulosus sensu stricto* genotype G1. Int J Parasitol 48(9-
324 10):729-742 doi:10.1016/j.ijpara.2018.03.006

325 Kinkar L, et al. (2017) New mitogenome and nuclear evidence on the phylogeny and
326 taxonomy of the highly zoonotic tapeworm *Echinococcus granulosus sensu stricto*.
327 Infect Genet Evol 52:52-58 doi:10.1016/j.meegid.2017.04.023

328 Kouidri M, Benchaib Khoudja F, Boulkaboul A, Ammar Selles SM (2014) Comparative
329 Study on Fertility, Viability and Biochemical Profiles of Hydatid Cyst Fluid from the
330 Lungs and Liver of Sheep in Algeria. World Applied Sciences Journal 31(10): 1699-
331 703

332 Laatamna AE, et al. (2019) Frequency and genetic diversity of *Echinococcus granulosus*
333 *sensu stricto* in sheep and cattle from the steppe region of Djelfa, Algeria. Parasitol
334 Res 118(1):89-96 doi:10.1007/s00436-018-6118-x

335 Laurimae T, et al. (2016) Genetic diversity and phylogeography of highly zoonotic
336 *Echinococcus granulosus* genotype G1 in the Americas (Argentina, Brazil, Chile and
337 Mexico) based on 8279bp of mtDNA. Infect Genet Evol 45:290-296
338 doi:10.1016/j.meegid.2016.09.015

339 Lavikainen A, Lehtinen MJ, Laaksonen S, Agren E, Oksanen A, Meri S (2006) Molecular
340 characterization of *Echinococcus* isolates of cervid origin from Finland and Sweden.
341 Parasitology 133(Pt 5):565-70 doi:10.1017/S0031182006000667

342 M'Rad S, et al. (2005) Molecular evidence of ovine (G1) and camel (G6) strains of
343 *Echinococcus granulosus* in Tunisia and putative role of cattle in human
344 contamination. Vet Parasitol 129(3-4):267-72 doi:10.1016/j.vetpar.2005.02.006

345 M'Rad S, et al. (2020) Microsatellite Investigations of Multiple *Echinococcus Granulosus*
346 *Sensu Stricto* Cysts in Single Hosts Reveal Different Patterns of Infection Events
347 between Livestock and Humans. *Pathogens* 9(6) doi:10.3390/pathogens9060444

348 Maillard S, et al. (2007) Taxonomic position and geographical distribution of the common
349 sheep G1 and camel G6 strains of *Echinococcus granulosus* in three African countries.
350 *Parasitol Res* 100(3):495-503 doi:10.1007/s00436-006-0286-9

351 Nakao M, et al. (2013) Mitochondrial phylogeny of the genus *Echinococcus* (Cestoda:
352 *Taeniidae*) with emphasis on relationships among *Echinococcus canadensis*
353 genotypes. *Parasitology* 140(13):1625-36 doi:10.1017/S0031182013000565

354 Oudni-M'rad M, et al. (2016) First molecular evidence of the simultaneous human infection
355 with two species of *Echinococcus granulosus sensu lato*: *Echinococcus granulosus*
356 *sensu stricto* and *Echinococcus canadensis*. *Parasitol Res* 115(3):1065-9
357 doi:10.1007/s00436-015-4836-x

358 Romig T, Ebi D, Wassermann M (2015) Taxonomy and molecular epidemiology of
359 *Echinococcus granulosus sensu lato*. *Vet Parasitol* 213(3-4):76-84
360 doi:10.1016/j.vetpar.2015.07.035

361 Sgroi G, et al. (2019) Cystic echinococcosis in wild boars (*Sus scrofa*) from southern Italy:
362 Epidemiological survey and molecular characterization. *Int J Parasitol Parasites Wildl*
363 9:305-311 doi:10.1016/j.ijppaw.2019.04.013

364 Tahiri S, et al. (2019) Genotyping of human *Echinococcus granulosus* cyst in Morocco. *J*
365 *Parasit Dis* 43(4):560-565 doi:10.1007/s12639-019-01127-1

366 Zait H, et al. (2016) Molecular characterization of *Echinococcus granulosus sensu stricto* and
367 *Echinococcus canadensis* in humans and livestock from Algeria. *Parasitol Res*
368 115(6):2423-31 doi:10.1007/s00436-016-4994-5

369

370

371
372

	Number of bases sequenced	Number of variable sites (%)	Alleles defined	Alleles not previously described	
				In Algeria	In the world
<i>atp6</i>	586	7 (1.2)	8	4	2
<i>nd3</i>	330	2 (0.6)	3	2	1
<i>cox1a</i>	371	10 (2.7)	10	6	5
<i>cox1b</i>	339	6 (2.6)	6	1	0
Total	1626	25 (1.5)	22		

373

374

375 Table 1: Sequence variability of 4 mitochondria loci observed in this study

376

377

	<i>atp6</i>																
	5802	5832	5841	5858	5861	5932	6060	6108	6126	6143	6145	6224	6286	6302	6329	6330	
Reference	T	G	C	C	T	T	C	T	A	T	G	T	A	C	A	T	
<i>apt6_01</i>									G								
<i>apt6_02</i>			T						G								
<i>apt6_03</i>					C				G								
<i>apt6_04</i>									G					T			
<i>apt6_05</i>			T	T					G								
<i>apt6_06</i>			T						G				T				
<i>apt6_07</i>		A							G								
<i>apt6_08</i>	C								G								
<i>apt6_09*</i>							T		G								
<i>apt6_10*</i>									G						G		
<i>apt6_11*</i>			T						G	C							
<i>apt6_12*</i>						C			G								
<i>apt6_13*</i>								C	G		A	C				C	
	<i>nd3</i>																
	8681	8726	8738	8820	8828	8842	8843										
Reference	C	T	T	C	T	A	C										
<i>nd3_01</i>																	
<i>nd3_02</i>				T													
<i>nd3_03</i>						T											
<i>nd3_04*</i>	T																
<i>nd3_05*</i>			C														
<i>nd3_06*</i>		C			C		T										

		<i>cox1a</i>														
		9019	9038	9042	9052	9054	9072	9104	9132	9134	9137	9148	9169	9170	9207	9242
Reference		G	A	C	T	A	G	C	A	A	A	T	T	G	A	C
<i>cox1a_01</i>																
<i>cox1a_02</i>															G	
<i>cox1a_03</i>									G							
<i>cox1a_04</i>														A		
<i>cox1a_05</i>										G						
<i>cox1a_06</i>				T						G						
<i>cox1a_07</i>													C			
<i>cox1a_08</i>												C				
<i>cox1a_09</i>																T
<i>cox1a_10</i>								T								
<i>cox1a_11*</i>			G		C	G										
<i>cox1a_12*</i>							A									
<i>cox1a_13*</i>		A														
		<i>cox1b</i>														
		9822	9830	9853	9863	9870	9908	10054	10073							
Reference		A	G	C	C	T	C	T	A							
<i>cox1b_01</i>																
<i>cox1b_02</i>				T												
<i>cox1b_03</i>					T											
<i>cox1b_04</i>						T		G								
<i>cox1b_05</i>						C										
<i>cox1b_06</i>		G														
<i>cox1b_07*</i>			A		T			C								

381
382

383

384 Table 2: Polymorphic sites of the 4mitochondrial loci used in this study.

385 Polymorphism in comparison with the reference mitochondrial genome (genotype G1, GenBank accession number AB786664) are indicated in

386 **Bold**

387 *Denotes sequences retrieved in GenBank

388

389

Haplotype	Number of				
	haplotype	<i>atp6</i>	<i>nad3</i>	<i>cox1a</i>	<i>cox1b</i>
HAP01	13	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP02	1	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_04</i>
HAP03	1	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_02</i>	<i>cox1b_01</i>
HAP04	2	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_03</i>	<i>cox1b_01</i>
HAP05	1	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_06</i>	<i>cox1b_01</i>
HAP06	1	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_08</i>	<i>cox1b_02</i>
HAP07	2	<i>atp6_01</i>	<i>nd3_02</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP08	1	<i>atp6_01</i>	<i>nd3_03</i>	<i>cox1a_01</i>	<i>cox1b_05</i>
HAP09	1	<i>atp6_02</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP10	1	<i>atp6_02</i>	<i>nd3_01</i>	<i>cox1a_10</i>	<i>cox1b_06</i>
HAP11	4	<i>atp6_02</i>	<i>nd3_01</i>	<i>cox1a_02</i>	<i>cox1b_01</i>
HAP12	1	<i>atp6_02</i>	<i>nd3_01</i>	<i>cox1a_02</i>	<i>cox1b_02</i>
HAP13	1	<i>atp6_03</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP14	1	<i>atp6_03</i>	<i>nd3_01</i>	<i>cox1a_05</i>	<i>cox1b_01</i>
HAP15	1	<i>atp6_04</i>	<i>nd3_01</i>	<i>cox1a_07</i>	<i>cox1b_01</i>
HAP16	1	<i>atp6_05</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP17	1	<i>atp6_06</i>	<i>nd3_01</i>	<i>cox1a_02</i>	<i>cox1b_01</i>
HAP18	1	<i>atp6_07</i>	<i>nd3_01</i>	<i>cox1a_09</i>	<i>cox1b_02</i>
HAP19	1	<i>atp6_08</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP20*	1	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_11</i>	<i>cox1b_01</i>
HAP21*	1	<i>atp6_01</i>	<i>nd3_04</i>	<i>cox1a_01</i>	<i>cox1b_01</i>

HAP22*	1	<i>atp6_03</i>	<i>nd3_05</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP23*	1	<i>atp6_09</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_03</i>
HAP24*	1	<i>atp6_10</i>	<i>nd3_01</i>	<i>cox1a_01</i>	<i>cox1b_01</i>
HAP25*	1	<i>atp6_11</i>	<i>nd3_01</i>	<i>cox1a_02</i>	<i>cox1b_01</i>
HAP26*	1	<i>atp6_12</i>	<i>nd3_04</i>	<i>cox1a_09</i>	<i>cox1b_01</i>
HAP27*	1	<i>atp6_01</i>	<i>nd3_01</i>	<i>cox1a_12</i>	<i>cox1b_01</i>
HAP28*	1	<i>atp6_13</i>	<i>nd3_06</i>	<i>cox1a_13</i>	<i>cox1b_07</i>

391

392 Table 3: Allelic profiles of each haplotype

393 *denotes haplotypes found in Genbank

394

395

396 Fig. 1. Minimum spanning tree of the *cox1b* locus drawn using the sequences obtained in this
397 study and 98 sequences from *E. granulosus* s.s. isolated from Algeria retrieved from GenBank.

398 1a: classification according to the host either human or animal; 1b: classification according to
399 the geographical origin

400 Circles represent sequence type; the distance between circles represents the number of
401 mutations; circle diameter represents number of specimens with the corresponding sequence
402 type.

403 Accession numbers of sequences from GenBank used in this figure: AF408686.1,
404 KR349027.1, KR381826.1, KU925429.1, MG672128.1, MG672283.1, MG672284.1,
405 MG672285.1, MG672287.1, MG672288.1, MG672289.1, MG672291.1, MG672292.1,
406 MG672293.1, MG808283.1, MG808285.1, MG808287.1, MG808290.1, MG808292.1,
407 MG808293.1, MG808294.1, MG808295.1, MG808296.1, MG808298.1, MG808302.1,
408 MG808303.1, MG808307.1, MG808308.1, MG808309.1, MG808313.1, MG808315.1,
409 MG808316.1, MG808317.1, MG808318.1, MG808324.1, MG808327.1, MG808328.1,
410 MG808330.1, MG808331.1, MG808334.1, MG808336.1, MG808337.1, MG808338.1,
411 MG808339.1, MG808340.1, MG808343.1, MG808344.1, MG808345.1, MG808346.1,
412 MG808348.1, MG808349.1, KR349028.1, MG672290.1, MG808286.1, MG808305.1,
413 MG808321.1, MG808322.1, MG808347.1, KR349030.1, MG672286.1, MG808304.1,
414 MG808326.1, KT316341.1, MG808342.1, MG808341.1, KR349034.1, MG808335.1,
415 MG808329.1, MG808320.1, MG808297.1, MG808319.1, MG808310.1, MG808288.1,
416 MG808301.1, MG808284.1, MG808299.1, MG808291.1, MG808289.1, KR349029.1,
417 KR349033.1, MG808333.1, MG808314.1, MG808332.1, MG808325.1, MG808312.1,
418 MG808311.1, MG808306.1, MG808282.1, KR349032.1, KR349031.1, MG808300.1,
419 MG808323.1

420

421 Fig. 2. Minimum spanning tree of the concatenated haplotypes (*atp6*, *nd3*, *cox1a*, *cox1b*) from

422 this study (dark blue) and 12 sequences from *E. granulosus* s.s. isolated from Algeria

423 retrieved from GenBank (clear blue).

424 Accession numbers of sequences from GenBank used in this figure: MG672288 MG672291

425 MG672284 MG672290 MG672128 MG672283 MG672285 MG672286 MG672287

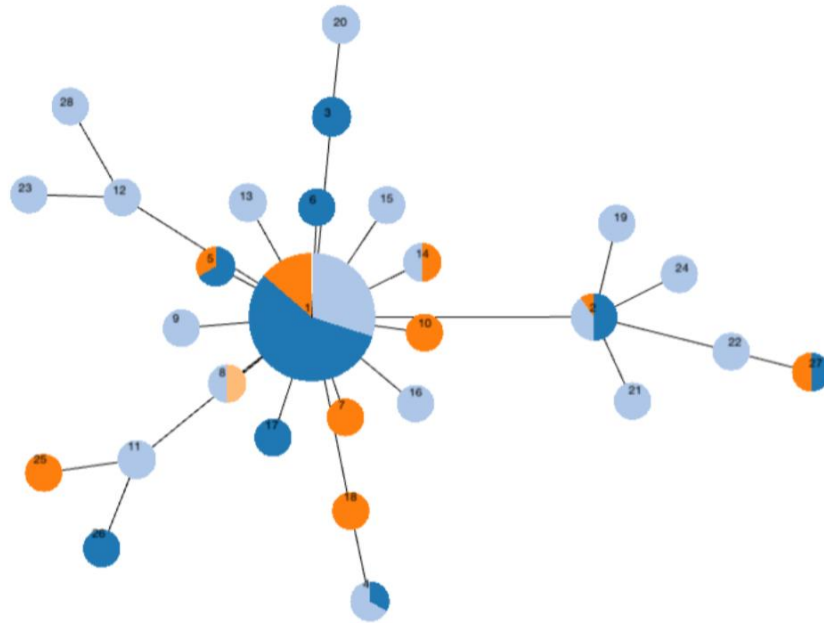
426 MG672289 MG672292 MG672293 MG682544

427

428

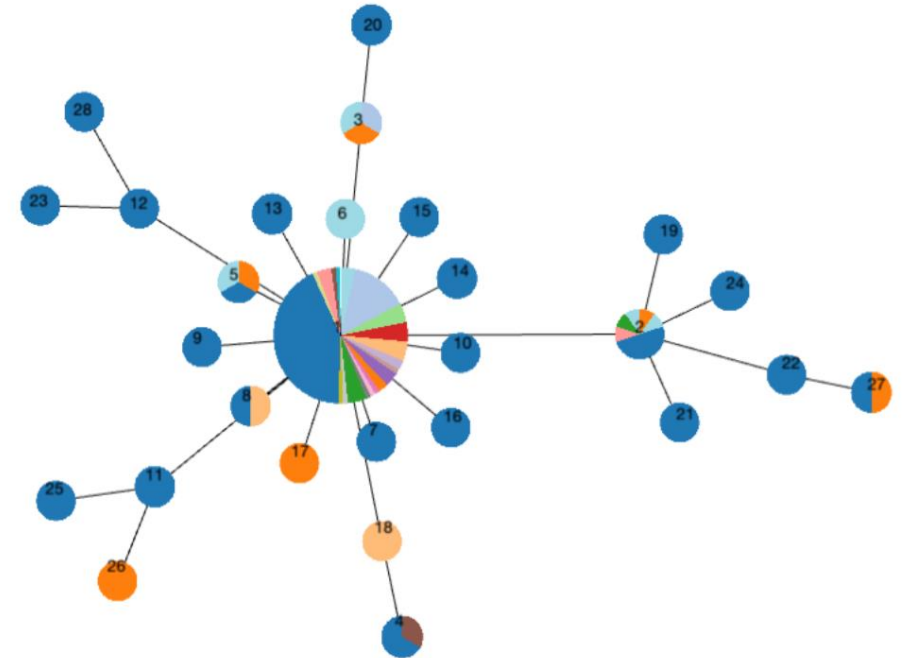
429

430



- Homo sapiens
- Sheep
- Cattle
- Goat

Figure 1a



- | | |
|------------------|--------------|
| ■ Djelfa | ■ Mostaganem |
| ■ Oran | ■ Tlemcen |
| ■ Algiers | ■ Ouargla |
| ■ Tiaret | ■ Khenchela |
| ■ Ain Defla | ■ Bouira |
| ■ Relizane | ■ Laghouat |
| ■ Mascara | ■ Bourmerdes |
| ■ Sidi Bel Abbas | ■ Adrar |
| ■ Ain Temouchent | ■ Naama |
| ■ Chlef | ■ Others |

Figure 1b

431
432
433
434

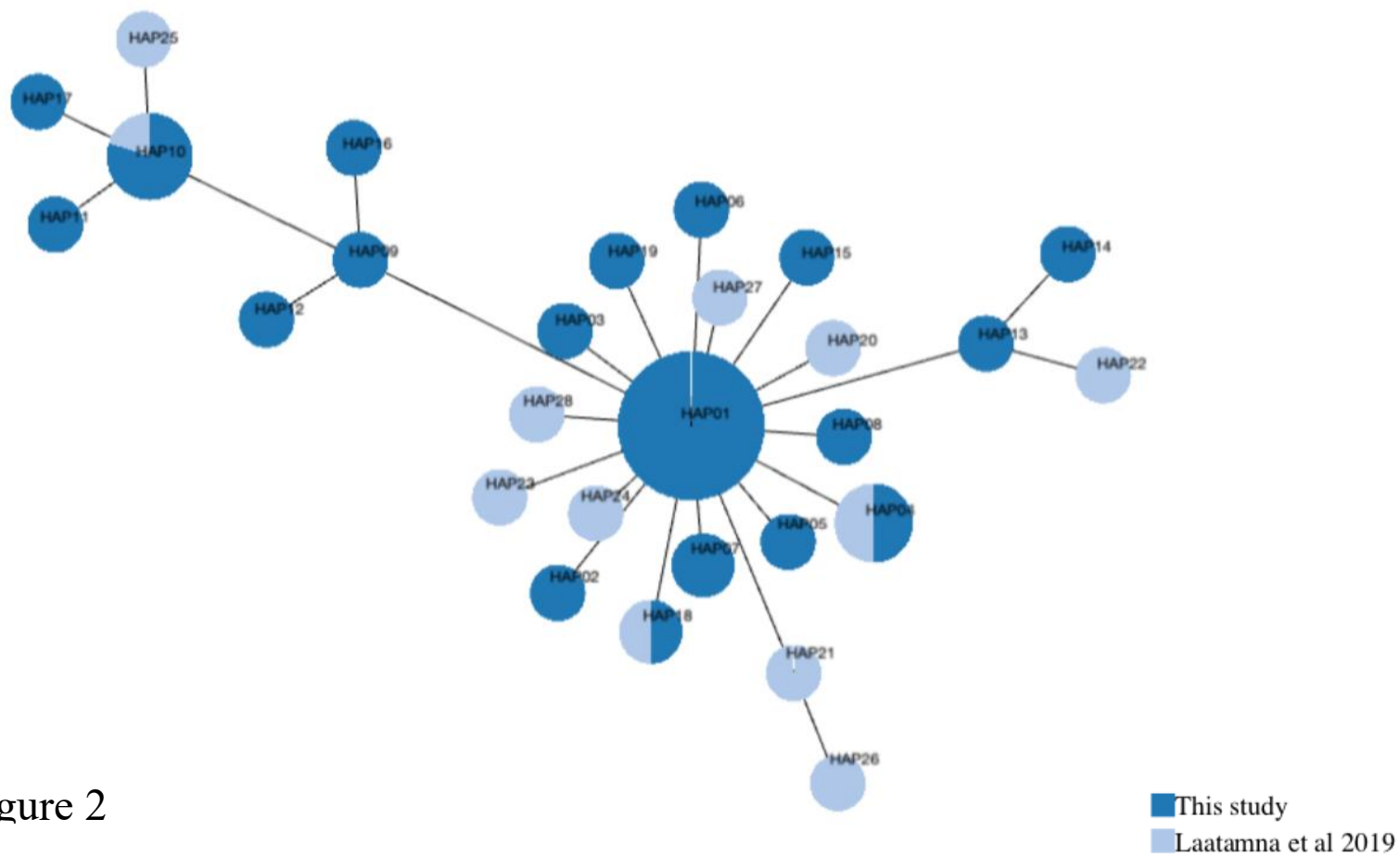


Figure 2

435

436 **Supplementary data**

437 Table S1: Demographic characteristics of *E. granulosus s.s.* analyzed in this study

438 Table S2: Primers used for amplification and sequencing of *E. granulosus s.l.* mitochondrial
439 targets.

440

441 Figure S1: Algerian localities of origin of the *E. granulosus s.s.* tested in this study (red dots)
442 and from previous studies conducted in Algeria (Blue dots)

443 Figure S2: Genetic map showing the 4 loci used in this study and their respective position on
444 the mitochondrial genome of *E. granulosus s.s.* Nucleotide positions refer to the reference
445 genome Genbank accession number AB786664)

446

447

Isolates	Age	Anatomic site of isolation	Living place
F1	41	Liver	Oran
F2	38	Liver	Oran
F3	35	Liver	Oran
F4	29	Liver	Oran
F6	60	Liver	Oran
F7	40	Liver	Relizane
F8	31	Liver	Illizi
F9	60	Liver	Mascara
F10	47	Liver	Tiaret
F11	27	Liver	Chlef
F12	22	Liver	Tlemcen
F13	68	Liver	Ain Temouchent
F14	25	Liver	Oran
F20	64	Liver	Ain Temouchent
F21	64	Liver	Oran
F22	43	Liver	Mascara
F23	54	Liver	Oran
F24	56	Liver	Oran
F25	32	Liver	Relizane
F26	29	Liver	Tiaret
F27	29	Liver	Tiaret
F27 ¹	29	Liver	Tiaret
F28	31	Liver	Ain Temouchent
F29	39	Liver	Mostaganem
F30	34	Liver	Oran
F31	60	Liver	Oran
P1	52	Lungs	Mascara
P2	32	Lungs	Chlef
P3	38	Lungs	Adrar
P4	35	Lungs	Sidi Bel Abbes
P5	59	Lungs	Tiaret

P7	64	Lungs	Mostaganem
P8	37	Lungs	Sidi Bel Abbes
P10	16	Lungs	Sidi Bel Abbes
P11	20	Lungs	Relizane
P12	40	Lungs	Ain Defla
P14	63	Lungs	Sidi Bel Abbes
P15	31	Lungs	Saida
P16	19	Lungs	El Bayadh
P17	23	Lungs	Tissemsilt
P18	37	Lungs	Oran
P22	77	Lungs	Mascara
P23	21	Lungs	Relizane
P24	17	Lungs	Naama

448

449 Table S1: Demographic characteristics of *E. granulosus s.s.* analyzed in this study

450 ¹ Denote isolation from same patient

451

452
453
454
455

Gene target	Primers	Primers sequence	Position¹	Fragment size	Previously described in
<i>atp6</i>	F	5'-TCAATTTGAAGCGTTGGAGATAACTT	5661-5686	770	Lavikainen <i>et al.</i> , 2006
	R	5'-GAAGGAACAATTGCCAACCC	6411-6430		Lavikainen <i>et al.</i> , 2006
<i>nd3</i>	F	5'-TTGGGTATCCTTGGTCTCGT	8407-8426	550	Lavikainen <i>et al.</i> , 2006
	R	5'-ATGACACAAAATTATTAGCAGTA	8934-8956		Lavikainen <i>et al.</i> , 2006
<i>cox1a</i>	F	5'-TTTTTTGGGCATCCTGAGGTTTAT	8878-8901	444	Bowles <i>et al.</i> , 1992
	R	5'-TAAAGAAAGAACATAATGAAAATG	9327-9348		Bowles <i>et al.</i> , 1992
<i>cox1b</i>	F	5'-TTGTTGAATTGTTTAGTGGGTATG	9762-9785	471	Lavikainen <i>et al.</i> , 2006
	R	5'-GCGGTAAATTCAAATCAGACAA	10182-10205		Lavikainen <i>et al.</i> , 2006

456

457 Table S2: Primers used for amplification and sequencing of *E. granulosus s.l.* mitochondrial targets.

458

459 ¹Position on the complete reference mitochondrial genome (Genbank accession number AB786664)

460

461

This study

- 1 Wilaya d'Adrar
- 2 Wilaya d'Aïn Defla
- 3 Wilaya d'Aïn Témouchent
- 4 Wilaya de Chlef
- 5 Wilaya d'El Bayadh
- 6 Wilaya de Mascara
- 7 Wilaya de Mostaganem
- 8 Wilaya de Naâma
- 9 Wilaya d'Oran
- 10 Wilaya de Relizane
- 11 Wilaya de Saïda
- 12 Wilaya de Sidi Bel Abbès
- 13 Wilaya de Tiaret
- 14 Wilaya de Tissemsilt
- 15 Wilaya de Tlemcen

Previous studies

- 1 Wilaya de Blida
- 2 Wilaya de Bouira
- 3 Wilaya de Bourmerdès
- 4 Wilaya de Khenchela
- 5 Wilaya de Laghouat
- 6 Wilaya d'Ouargla
- 7 Wilaya de Tipaza
- 8 Wilaya de Tiaret
- 9 Wilaya d'Aïn Defla
- 10 Alger
- 11 Wilaya de Djelfa



Figure S1: Algerian localities of origin of the *E. granulosus* strains tested in this study (red dots) and from previous studies conducted in Algeria (Blue dots)

Figure S2

