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

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

Daouia Moussa[1], Kheira Senouci[1], Nori Midoun[2], Mohamed Lacheheb[3], Benali Tabeti[4], Noureddine Benmaarouf[4], Christophe Hennequin[5].

1. Natural and Life Sciences Faculty, Department of Biology, University of Oran 1 Ahmed Ben Bella, Oran, Algeria.
2. Department of Epidemiology and Preventive Medicine, University Hospital of Oran (EHU), Oran, Algeria.
3. Department of Thoracic Surgery, University Hospital of Oran (EHU), Oran, Algeria.
4. Department of Hepatobiliary and Liver Transplant, University Hospital of Oran (EHU), Oran, Algeria.
5. Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, CRSA, AP-HP, Hôpital Saint-Antoine, Service de Parasitologie-Mycologie, F-75012 Paris, France

Corresponding author:

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

Phone + 33 1 4928 3412

Fax +33 1 4928 2251

Email: christophe.hennequin-sat@aphp.fr

Abstract:

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

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

The authors declare that no funding was obtained for this study

The authors declare to have nothing to disclose

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

DM collected the samples, performed all the experiments and writes a first draft of the paper. KS supervise the biological diagnosis and review the paper. NM, ML, NB and BT took care of the patients and review the paper.

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

Introduction

Cystic echinococcosis (CE) is due to the development of the larval stage of the cestode *Echinococcus granulosus sensu lato* (*s.l.*). Human CE is a cosmopolitan zoonosis that represents a major public health problem in endemic countries where collective prophylaxis measures remain insufficient, notably in breeding areas such as North Africa (Dakkak 2010; Eckert et al. 2001). This is the case of Algeria where, despite improved educational hygiene and control over the slaughter of animals, the prevalence of CE remains high, with about 700 cases notified yearly (Dakkak 2010). This is due to the domestic life cycle of the parasite where dogs are the predominant definitive host of the parasite and the main source of contamination for humans, while intermediate hosts are represented by various species of ruminants. Previous reports from Algeria have shown that16 to 42% of stray-dogs harbor *E. granulosus s.l.* in their intestine (Bentounsi et al. 2009), and that sheep, goat, cattle, horses, dromedaries, and wild boars can act as intermediate hosts, sheep being by far the most common (Bardonnet et al. 2003; Eckert et al. 2001; Kouidri et al. 2014; Laatamna et al. 2019; Maillard et al. 2007; Zait et al. 2016).

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

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

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

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

Materials and methods

Biological samples

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

Molecular analysis

DNA extraction

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

Amplification and direct sequencing

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

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

Phylogenetic analysis

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

Phylogenetic analysis was achieved using these sequences and sequences available in Genbank and corresponding to Algerian *E. granulosus s.s.* isolates (Bardonnet et al. 2003; Kinkar et al. 2018; Laatamna et al. 2019; Maillard et al. 2007; Zait et al. 2016). We performed both an analysis using the cox1blocus because a higher number of sequences are available in Genbank with 100% query cover and another one using the concatenated sequences of the 4 loci further in called haplotypes, obtained in our study. The phylogenetic relationship between sequences identified in our series and sequences retrieved in Genbank was illustrated by the construction of minimum spanning trees using an optimal eBurst algorithm (Francisco et al. 2009). Graphical representation was done using the Phyloviz online program (http://www.phyloviz.net/).In this graphical representation, isolates with the same allelic profile fall in the same circle, the size of which is proportional to the number of isolates with that particular profile.

Ethics statement

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

Results

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

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

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

The first phylogenetic tree was based on the *cox1b* locus taking the advantage of 92homolog sequences in GenBank (Kinkar et al. 2018; Laatamna et al. 2019; Zait et al. 2016). It shows that 36 (83.7%) of our isolates harbor the predominant sequence (*cox1b_01*), similar to the reference mitochondrial genome. Sequences from all the isolates but 2 were also direct satellites from this central clade with links corresponding to a single nucleotide mutation. No segregation according to the host nor to the geographic origin could be demonstrated (Fig1a, Fig 1b). When using complete concatenated haplotypes using the sequences of four loci tested in our study, Hap01 appeared as the founder, giving birth to closely related haplotypes either tested in this study or retrieved in Genbank. However, Hap10, which originated indirectly from Hap01 (2 nucleotide mutations), forms a second cluster distantly with further diversification (Figure 2).

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

Discussion

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

In our study, we were able to genotype 44 *E. granulosus s.l.* isolates collected from humans living in western Algeria. Previous studies have shown that whatever the country or the host mitochondrial targets offer a high power of discrimination in regards to nuclear targets (Laurimae et al. 2016; Lavikainen et al. 2006; Maillard et al. 2007). We used a multilocus sequence typing scheme based on 4 fragments encoding mitochondrial genes with 1,626 nucleotides analyzed. This allowed the description of 19 haplotypes of which12 had never been reported, with 13 alleles newly described in Algeria of which 8 had not been previously deposited in GenBank.

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

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

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

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

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

However, using the 4 loci scheme, we were able to detect a substantially higher level of genetic diversity within the G1 genotype than using a single locus. All the haplotypes detected in our study were genetically close to but different from the G1reference mitogenome available in GenBank (AB786664) that has been isolated from a man in China (Nakao et al. 2013). As previous studies conducted in other parts of Algeria, we showed that the Hap01 haplotype is predominant in western Algeria. This synonymous nucleotide mutation, in a locus rarely included in the typing method, has already been described in Italian, Australian, Tunisian and Algerian isolates (Kinkar et al. 2019; Kinkar et al. 2018; Sgroi et al. 2019). The Hap01 haplotype belongs to a clonal complex that includes a low level of diversity with satellite haplotypes (9 haplotypes diverge only by one single nucleotide mutation). However, the second most frequent haplotype Hap02, that derives from Hap01, seems to form a new complex that gives birth to new closely related haplotypes. Whether this is due to the natural evolution of the species or the introduction of either cattle or sheep, or dogs from other regions should be further evaluated. The contamination of humans in other regions can also be hypothesized. Interestingly, this haplotype has already been described from Kenchela, located in eastern Algeria. Similarly, Hap04 was shared by two isolates from Boumerdes and Laghouat, 2 cities distant from western Algeria. More recently, whole mitochondrial genome sequence using primer walking has been proposed for *E. granulosus s.l.* typing and offers a better discriminatory power and capability to analyze phylogenetic relationship). So, Kinkar et al. showed that Algerian isolates of *E. granulosus s.l.* are split into different geographical clusters that also include European and North African isolates (Kinkar et al. 2018). This and results from other studies (Boufana et al. 2014; Hassan et al. 2017; Kinkar et al. 2018), notably showing

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

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

Acknowledgements

This study was supported by the Parasitology laboratory of the University of Oran1 Ahmed Ben Bella, the Department of Parasitology at the Hospital of Oran (EHU).

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

The authors also thank the Joint Laboratory of Genetics and Molecular Biology and Sandra Vellaissamy from the Parasitology Department of St Antoine Hospital for excellent technical assistance.

Conflict of interest

The authors state that they have no conflict of interest.

References

Alvarez Rojas, C. A., Romig, T., Lightowlers, M. W. (2014). *Echinococcus granulosus sensu lato* genotypes infecting humans--review of current knowledge. Int J Parasitol 44 (1), 9-18. doi:10.1016/j.ijpara.2013.08.008

Bardonnet, K., Piarroux, R., Dia, L., Scheegans, F., Beurdeley, A., Godot V., Vuitton DA (2002) Combined eco-epidemiological and molecular biology approaches to assess *Echinococcus granulosus* transmission to human in Mauritania: occurrence of the camel strain and human cystic echinococcosis. Trans R Soc Trop Med Hyg 96:383–386

Bardonnet K, et al. (2003) Cystic echinococcosis in Algeria: cattle act as reservoirs of a sheep strain and may contribute to human contamination. Veterinary parasitology 116(1):35-44 doi:10.1016/s0304-4017(03)00255-3

Alvarez Rojas CA, Romig T, Lightowlers MW (2014) Echinococcus granulosus sensu lato genotypes infecting humans--review of current knowledge. Int J Parasitol 44(1):9-18 doi:10.1016/j.ijpara.2013.08.008

Bardonnet K, et al. (2003) Cystic echinococcosis in Algeria: cattle act as reservoirs of a sheep strain and may contribute to human contamination. Vet Parasitol 116(1):35-44 doi:10.1016/s0304-4017(03)00255-3

Bentounsi B, Meradi S, Ayachi A, Cabaret J (2009) Cestodes of Untreated Large Stray Dog Populations in Algeria: A Reservoir for Herbivore and Human Parasitic Diseases. The Open Veterinary Science Journal 3:64-7

Boufana B, et al. (2014) Genetic variability and haplotypes of Echinococcus isolates from Tunisia. Trans R Soc Trop Med Hyg 108(11):706-14 doi:10.1093/trstmh/tru138

Bowles J, Blair D, McManus DP (1992) Genetic variants within the genus Echinococcus identified by mitochondrial DNA sequencing. Mol Biochem Parasitol 54(2):165-73 doi:10.1016/0166-6851(92)90109-w

Dakkak A (2010) Echinococcosis/hydatidosis: a severe threat in Mediterranean countries. Vet Parasitol 174(1-2):2-11 doi:10.1016/j.vetpar.2010.08.009

Deplazes P, et al. (2017) Global Distribution of Alveolar and Cystic Echinococcosis. Adv Parasitol 95:315-493 doi:10.1016/bs.apar.2016.11.001

Eckert J, Gemmell M, Meslin Fo-X, Pawlowski Z, Organization WH (2001) WHO/OIE manual on echinococcosis in humans and animals: a public health problem of global 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

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

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

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

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

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

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

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

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

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

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

	atp6															
	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**
apt6_01									G							
apt6_02			T						G							
apt6_03					C				G							
apt6_04									G					T		
apt6_05			T	T					G							
apt6_06			T						G				T			
apt6_07		A							G							
apt6_08	C								G							
*apt6_09**							T		G							
*apt6_10**									G						G	
*apt6_11**			T						G	C						
*apt6_12**						C			G							
*apt6_13**								C	G		A	C				C

	nd3						
	8681	8726	8738	8820	8828	8842	8843
Reference	**C**	**T**	**T**	**C**	**T**	**A**	**C**
nd3_01							
nd3_02				T			
nd3_03						T	
*nd3_04**	T						
*nd3_05**			C				
*nd3_06**		C			C		T

	cox1a														
	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**
cox1a_01															
cox1a_02														G	
cox1a_03								G							
cox1a_04													A		
cox1a_05										G					
cox1a_06			T						G						
cox1a_07												C			
cox1a_08											C				
cox1a_09															T
cox1a_10							T								
*cox1a_11**		G		C	G										
*cox1a_12**						A									
*cox1a_13**	A														

	cox1b							
	9822	9830	9853	9863	9870	9908	10054	10073
Reference	**A**	**G**	**C**	**C**	**T**	**C**	**T**	**A**
cox1b_01								
cox1b_02			T					
cox1b_03				T				
cox1b_04						T		G
cox1b_05					C			
cox1b_06	G							
*cox1b_07**		A		T			C	

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

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

*Denotes sequences retrieved in GenBank

390

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

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

391

392 Table 3: Allelic profiles of each haplotype

393 *denotes haplotypes found in Genbank

394

Fig. 1. Minimum spanning tree of the *cox1b* locus drawn using the sequences obtained in this study and 98sequences from *E. granulosus s.s.* isolated from Algeria retrieved from GenBank. 1a: classification according to the host either human or animal; 1b: classification according to the geographical origin

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

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

Fig. 2. Minimum spanning tree of the concatenated haplotypes (*atp6, nd3, cox1a, cox1b*) from this study (dark blue) and 12 sequences from *E. granulosus s.s.* isolated from Algeria retrieved from GenBank (clear blue).

Accession numbers of sequences from GenBank used in this figure: MG672288 MG672291 MG672284 MG672290 MG672128 MG672283 MG672285 MG672286 MG672287 MG672289 MG672292 MG672293 MG682544

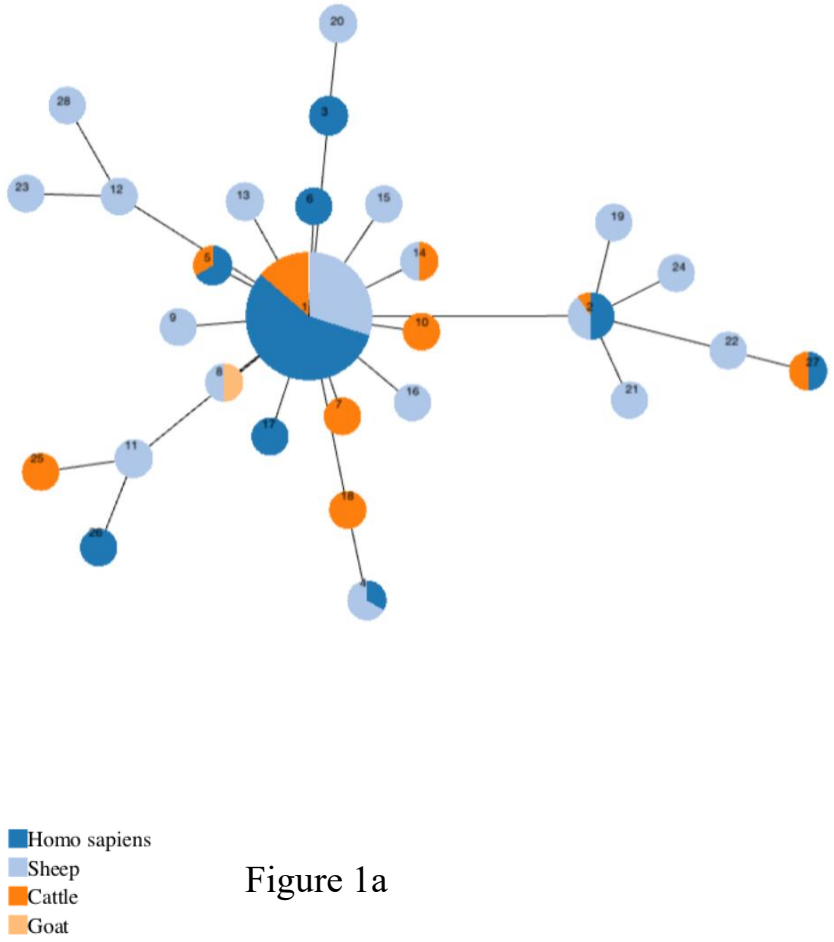


Figure 1a

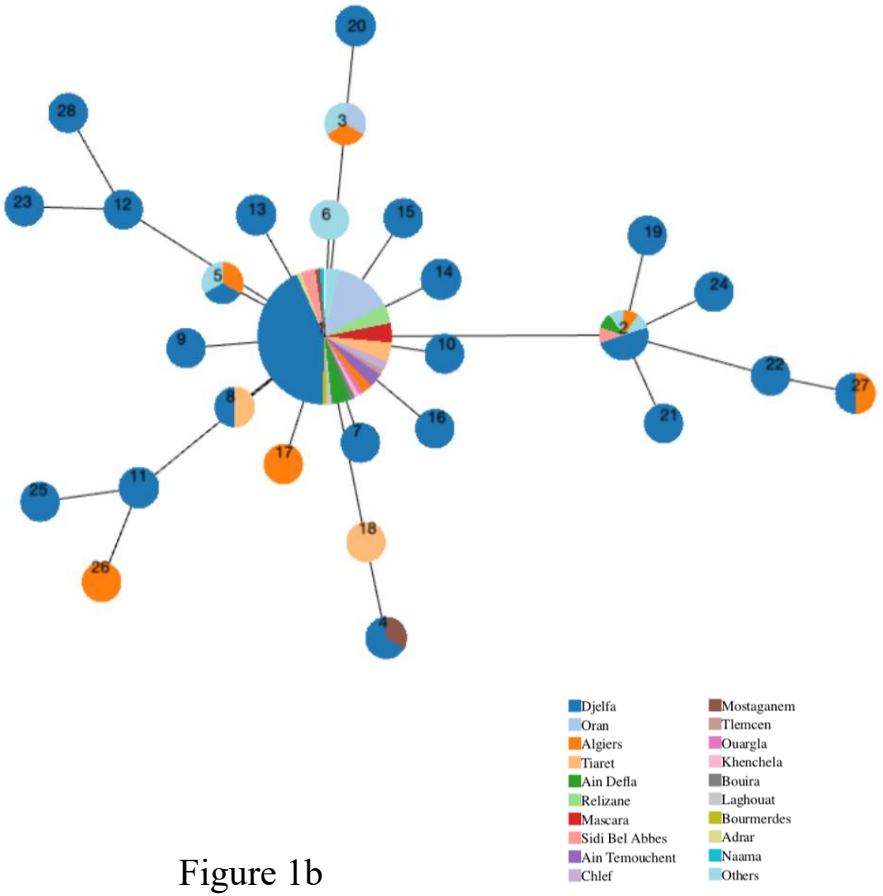


Figure 1b

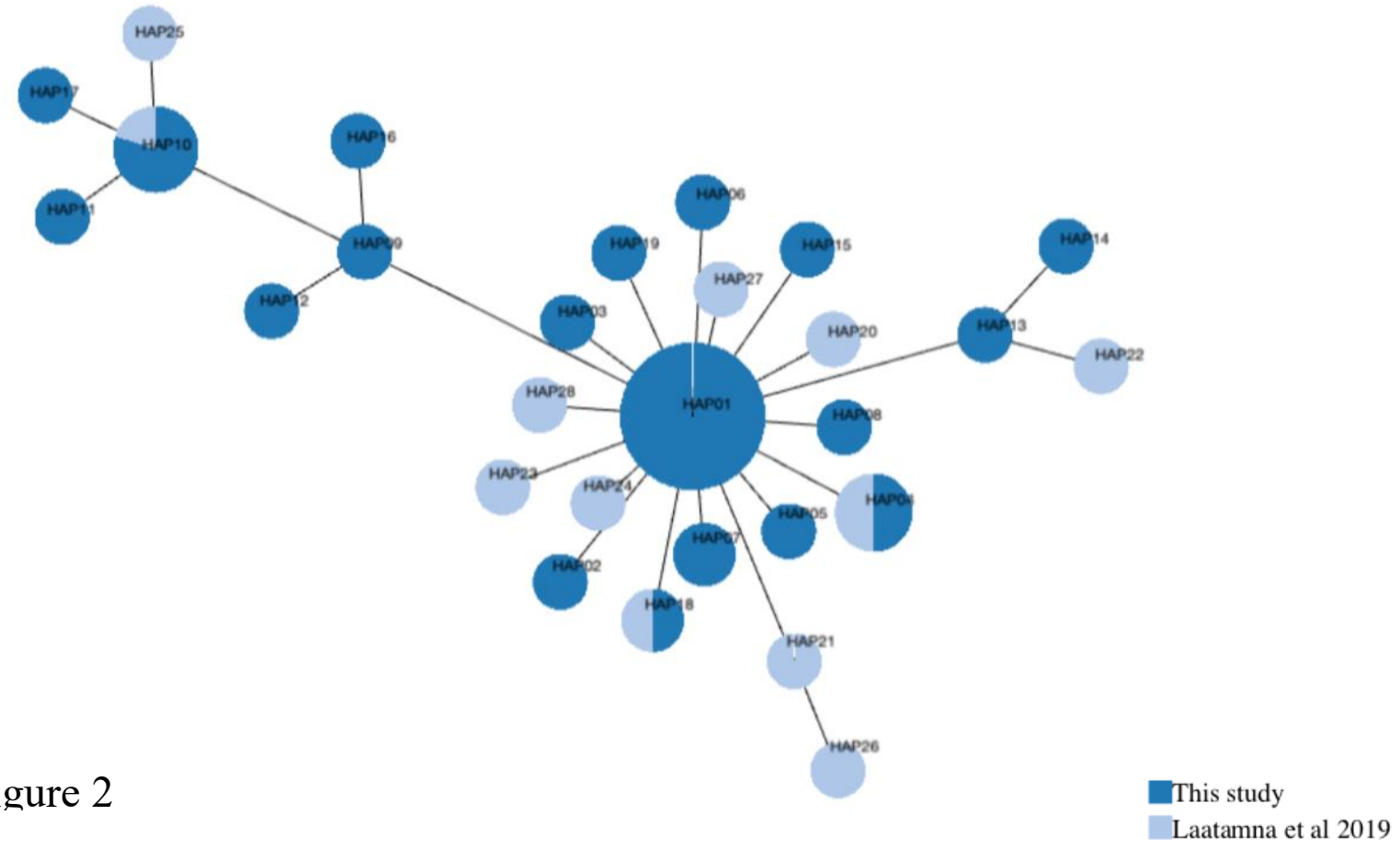

HAP25
HAP17
HAP10
HAP16
HAP11
HAP09
HAP12
HAP06
HAP19
HAP15
HAP14
HAP27
HAP03
HAP20
HAP13
HAP22
HAP28
HAP01
HAP08
HAP23
HAP24
HAP04
HAP05
HAP07
HAP02
HAP18
HAP21
HAP26
This study
Laatamna et al 2019

Figure 2

Supplementary data

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

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

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

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

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

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

[1] Denote isolation from same patient

Gene target	Primers	Primers sequence	Position[1]	Fragment size	Previously described in
atp6	F	5'-TCAATTTGAAGCGTTGGAGATAACTT	5661-5686	770	Lavikainen *et al.*, 2006
	R	5'-GAAGGAACAATTGCCAACCC	6411-6430		Lavikainen *et al.*, 2006
nd3	F	5'-TTGGGTATCCTTGGTCTCGT	8407-8426	550	Lavikainen *et al.*, 2006
	R	5'-ATGACACAAAATTATTAGCAGTA	8934-8956		Lavikainen *et al.*, 2006
cox1a	F	5'-TTTTTTGGGCATCCTGAGGTTTAT	8878-8901	444	Bowles *et al.*, 1992
	R	5'-TAAAGAAAGAACATAATGAAAATG	9327-9348		Bowles *et al.*, 1992
cox1b	F	5'-TTGTTGAATTGTTTAGTGGGTATG	9762-9785	471	Lavikainen *et al.*, 2006
	R	5'-GCGGTAAATTCAAATCAGACAA	10182-10205		Lavikainen *et al.*, 2006

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

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

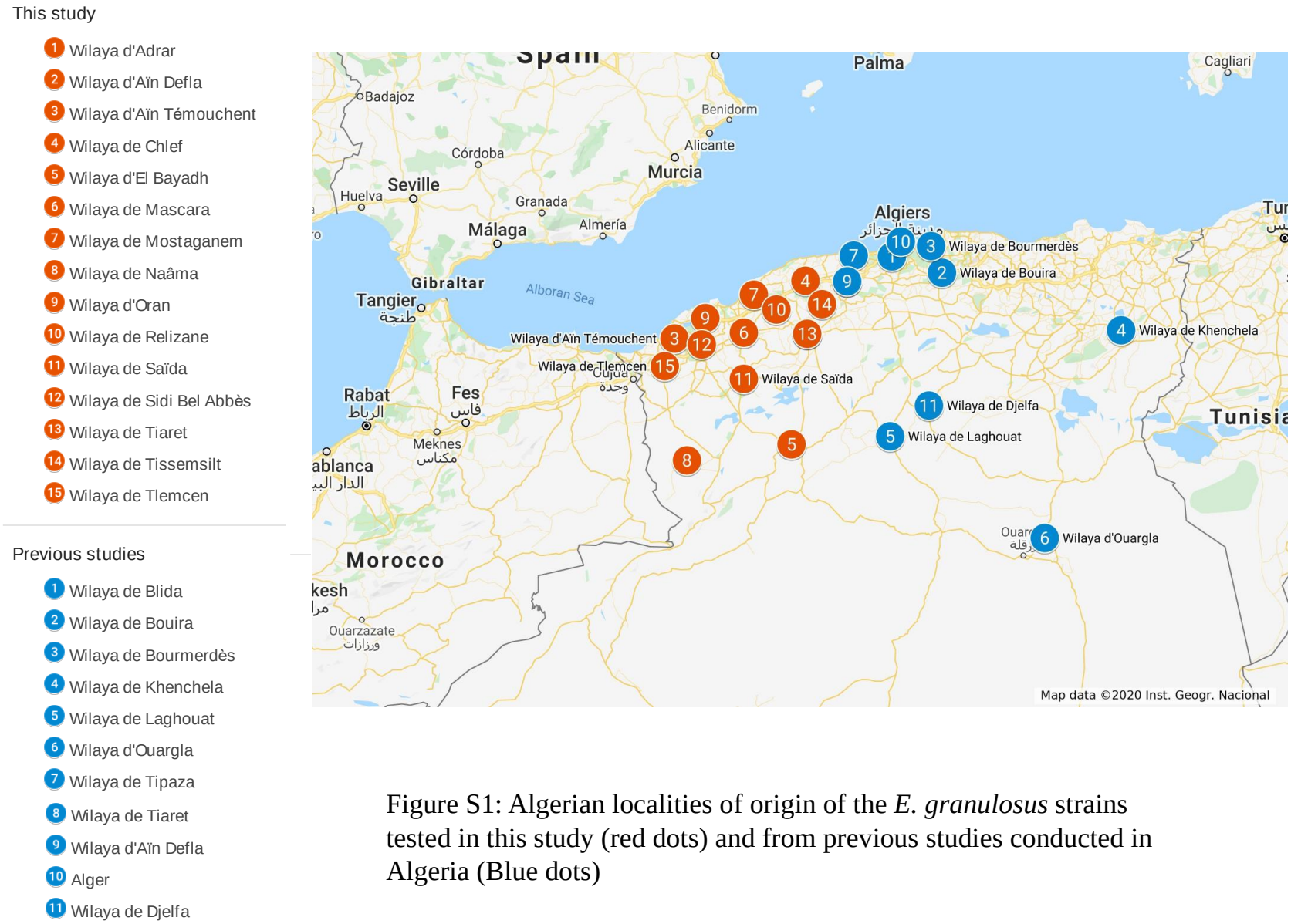
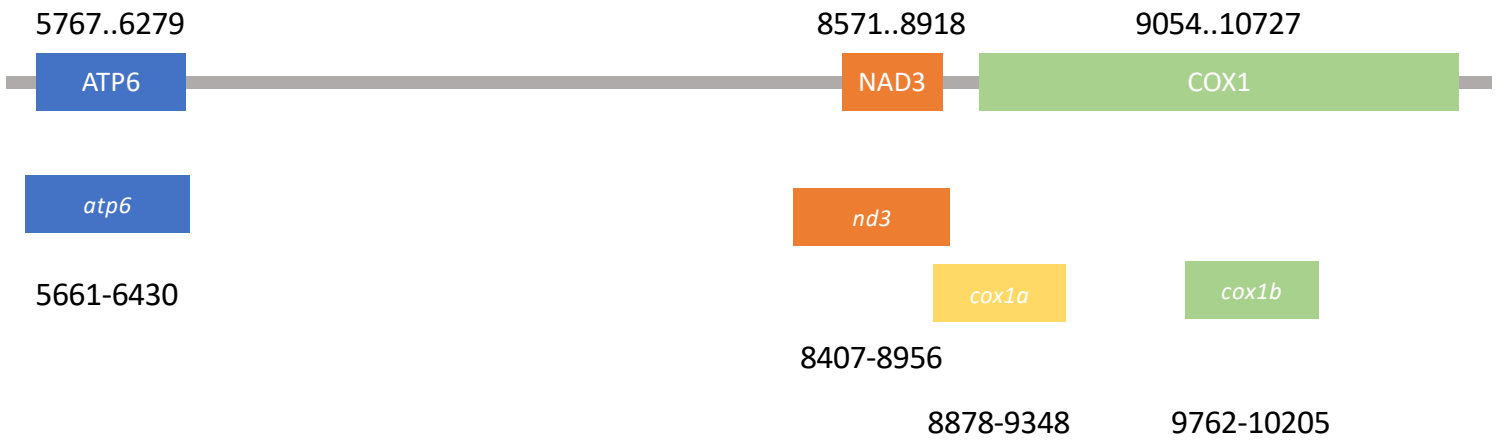


Figure S1: Algerian localities of origin of the *E. granulosus* strains tested in this study (red dots) and from previous studies conducted in Algeria (Blue dots)

462

Figure S2



463