International Journal for Parasitology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Genetic differentiation of the G6/7 cluster of *Echinococcus canadensis* based on mitochondrial marker genes

Francis Addy [a,*], Marion Wassermann [a], Dorothy Kagendo [b], Dennis Ebi [a], Eberhard Zeyhle [b], Ibrahim E. Elmahdi [c], Gerald Umhang [d], Adriano Casulli [e], Majid F. Harandi [f], Ortwin Aschenborn [a], Peter Kern [g], Ute Mackenstedt [a], Thomas Romig [a]

[a] *University of Hohenheim, Parasitology Unit 220B, Emil-Wolff-Str. 34 70599 Stuttgart, Germany*
[b] *Meru University of Science and Technology, School of Health Sciences, P.O. Box 972, 60200 Meru Kenya*
[c] *University of Gezira, Faculty of Health and Environmental Sciences, P.O. Box 20, Wad Medani, Sudan*
[d] *Anses LRFSN, Wildlife Surveillance and Eco-epidemiology unit, Technopôle Agricole et Vétérinaire, B.P. 40009, 54220 Malzéville, France*
[e] *WHO Collaborating Centre for the Epidemiology, Detection and Control of Cystic and Alveolar Echinococcosis, Istituto Superiore di Sanità, viale Regina Elena 299, 00161 Rome, Italy*
[f] *Research Center for Hydatid Disease in Iran, Kerman University of Medical Sciences, Kerman, Iran*
[g] *Ulm University Hospitals, Comprehensive Infectious Diseases Center, Albert-Einstein-Allee 23, D-89081 Ulm, Germany*

ARTICLE INFO

Article history:
Received 31 March 2017
Received in revised form 1 June 2017
Accepted 7 June 2017
Available online xxxx

Keywords:
Echinococcus canadensis
Microdiversity
Haplotypes
Zoonosis
Cystic echinococcosis

ABSTRACT

Among the genotype/species causing cystic echinococcosis , the taxonomic status of *Echinococcus canadensis* is only partially resolved. Within *E. canadensis*, four genotypes (G6, G7, G8 and G10) have been described based on short mitochondrial sequences, of which G6 and G7 (the 'camel' and the 'pig' strain, respectively) are closely related and variously regarded as microvariants of a single strain G6/7. Globally, this G6/7 cluster is the second most important agent of human cystic echinococcosis and is the predominant *Echinococcus* taxon in large parts of sub-Saharan Africa. To add data on the internal structure and the geographical distribution of this cluster, we analysed diversity and population structure of 296 isolates of *E. canadensis* from sub-Saharan Africa, the Middle East and Europe using the complete mitochondrial cytochrome c oxidase subunit 1 (*cox1*) (1,608 bp) and NADH dehydrogenase subunit 1 (*nad1*) (894 bp) gene sequences. Polymorphism of the mtDNA loci gave 51 (*cox1*), 33 (*nad1*) and 73 (*cox1-nad1* concatenated) haplotypes. African and Middle Eastern isolates mainly grouped in a star-like structure around a predominant haplotype, while the European isolates produced more diversified networks. Although the *cox1* diagnostic sequence for G6 is frequent in the African/Middle Eastern sub-cluster, and that for G7 is common in the European isolates, numerous intermediate variants prevent a clear distinction into 'G6' or 'G7', and the entire taxon is best treated as a common haplotype cluster G6/7. In contrast, wildlife isolates from the northern hemisphere showed that the G6/7 cluster is clearly distinct from both G8 and G10, and isolates of the latter genotypes were remarkably distant from each other. It is clear from the present study based on mitochondrial data that G6/7 is a coherent genotypic entity within *E. canadensis* that retains substantial intraspecific variance, and sub-populations share common ancestral polymorphisms and haplotypes. This study provides the basis for wider biogeographic comparison and population genetics studies of this taxon.

1. Introduction

Cystic echinococcosis (CE) is a neglected zoonotic disease of global significance. It is well known that CE is caused by diverse aetiological agents belonging to the genus *Echinococcus*, some of which have been given species status in recent revisions (Nakao et al., 2013; Romig et al., 2015). Among these, *Echinococcus canadensis* is genetically the most variable species, containing various 'strains' that are geographically and epidemiologically separated. Initially described as G6 (camel strain), G7 (pig strain), G8 ('American' cervid strain), and G10 ('Fennoscandian' cervid strain), they form a monophyletic cluster based on mitochondrial genomes and nuclear marker genes (Nakao et al., 2013). An additional strain G9, described from internal transcribed spacer region 1 (ITS1)-RFLP patterns, is now thought to represent a microvariant of G7, although this is unconfirmed (Scott et al., 1997; Romig et al.,

* Corresponding author. Fax: +49 711 459 22274.
E-mail addresses: francisaddy02@gmail.com, francis.addy@uni-hohenheim.de (F. Addy).

http://dx.doi.org/10.1016/j.ijpara.2017.06.003
0020-7519/

Please cite this article in press as: Addy, F., et al. Genetic differentiation of the G6/7 cluster of *Echinococcus canadensis* based on mitochondrial marker genes. Int. J. Parasitol. (2017), http://dx.doi.org/10.1016/j.ijpara.2017.06.003

2015). While G6 and G7 show only minor differences between each other in the (whole) mitochondrial genome and are often treated as microvariants, G8 and G10 are more distantly related and a taxonomic split of *E. canadensis* into three species has been proposed (Lymbery et al., 2015a). The validity of this proposal has been challenged, however, mainly for lack of nuclear genomic sequence data (Nakao et al., 2015), so that the taxonomy of *E. canadensis* remains to be resolved. Ecologically, the genotypes G8 and G10 are predominantly transmitted in wildlife cycles between wolves and ungulates of the deer family (Cervidae) in the northern hemisphere, while the globally distributed genotypes G6 and G7 are known from domestic lifecycles that involve dogs as definitive hosts, and pigs, camels and some ruminants (mainly goats) as intermediate hosts. This distinction into 'sylvatic' and 'domestic' strains is blurred, however, because the cervid strains can also be transmitted in domestic or semi-domestic lifecycles that involve dogs and domesticated reindeer (Oksanen and Lavikainen, 2015), and there are numerous examples of wild host animal involvement in the transmission of G6 and G7 genotypes (Romig et al., 2017).

The genetic variability within the G6/7 cluster is not well known. The camel and pig strains were initially distinguished from other CE agents due to the morphology of the adult worm, the shorter time worms need to reach gravidity in dogs, and on their lifecycles which are based on camels (in northern and eastern Africa, the Middle East and central Asia) and domestic pigs (e.g. in central and eastern Europe) (Eckert et al., 1989, 1993). Genetic characterization as G6 (camel strain) and G7 (pig strain) was initially based on partial sequences of the mitochondrial cytochrome c oxidase subunit 1 (*cox1*) (366 bp) and NADH dehydrogenase subunit 1 (*nad1*) (471 bp) genes (Bowles et al., 1992; Bowles and McManus, 1993). Sequences of G6 (from three isolates from east African camels and a goat) and G7 (from two isolates from Polish pigs) were clearly different from other agents of CE, but diverged from each other only by one (*cox1*) and three (*nad1*) nucleotide substitutions. This close affinity between G6 and G7 was later confirmed by analysis of longer sequences including the entire mitochondrial genomes (Nakao et al., 2013). Despite this close affinity, numerous isolates were subsequently differentiated into G6 and G7 in molecular studies from various parts of the world. This differentiation is typically based on the short sequences mentioned above or on comparison with various 'reference' sequences available in GenBank (e.g. AB208063 and AB235847). To date, no molecular study is available that provides a rational basis to retain G6 and G7 as discrete mitochondrial genotypes by assessing the genetic variability of appropriate numbers of geographically spaced isolates. This has practical relevance, because *E. canadensis* G6/7 is the second most frequent cause of human CE worldwide with 11% of 1,661 genetically characterised cyst isolates from humans (Alvarez Rojas et al., 2014), and knowledge on the pathways of transmission of this parasite is important for prevention and control. Genotypes G6 and G7 are often assumed to be epidemiologically separated by different species of intermediate hosts. It is unclear, however, if the two genotypes reflect different host preferences (G6 for camel, G7 for pigs), or if they are microvariants that reflect different geographical dispersal. Genotype G6 was mainly detected in the Middle East and adjacent regions where camels are important as livestock, while G7 is typical for central and eastern Europe with its tradition of pig farming. In the Americas, G6 has been identified from goats and G7 from pigs (Soriano et al., 2010), but this may reflect the region of livestock origin rather than host preferences. Of particular interest is the situation in sub-Saharan Africa, where G6 is known from a dog–camel lifecycle in vast regions in the northern and northeastern parts of the continent, but also from further south, where camels are not present and other livestock species (particularly goats) act as intermediate hosts (Addy et al., 2012), and where wild mammals can be heavily involved in the transmission (O. Aschenborn, unpublished data).

Studies on the intraspecific variability of *E. canadensis* G6/7 are only available from geographically restricted regions such as Mongolia (Ito et al., 2013, 2014) and Russia (Konyaev et al., 2013). To characterise the genetic variability within *E. canadensis* G6/G7 and identify the correlation of haplotype frequencies with host species and geographical distribution, we analysed close to 300 isolates from eastern and southern Africa, the Middle East, and Europe.

2. Materials and methods

2.1. Echinococcus isolates

For G6/7 genotypes, a total of 296 cysts or worm isolates were collected during surveys or by opportunistic sampling in Kenya ($n = 131$, from camels (95), cattle (18), goats (11), dogs (2), and human patients (5)), Sudan ($n = 56$, from camels (50), cattle (2) and human patients (4)), Namibia ($n = 8$, from *Oryx gazella*), Mauritania ($n = 7$, from camels), Ethiopia ($n = 1$, from cattle), Corsica (France) ($n = 48$, from pigs), Serbia ($n = 13$, from pigs (12) and wild boar (1)), Slovakia ($n = 8$, from pigs), Hungary ($n = 5$, from pigs), Romania ($n = 3$, from cattle (1) and wild boars (2)), Iran ($n = 12$, from camels (9) and human patients (3)) and Armenia ($n = 4$, from pigs). Isolates used here include the germinal layer or protoscolices of individual cysts, and adult worms (see Supplementary Table S1 for detailed descriptions of isolates, also available at doi:10.17632/9sxv4mz4j3.1). All isolates were stored in 70% ethanol until use, except the Mauritanian isolates that were sampled using FTA® cards (VWR International, France) and stored dry as described earlier (Boué et al., 2017).

For G8 and G10 genotypes, DNA aliquots of the following isolates were used; G8 from Russia ($n = 3$, from moose (*Alces alces*) (2) and wolf (*Canis lupus*) (1) – AB777910) and USA ($n = 1$, from moose – AB235848) and G10 from Finland ($n = 1$, from reindeer (*Rangifer tarandus*) – AB745463)

2.2. Mitochondrial cox1 and nad1 DNA amplification and sequencing

DNA was obtained for PCR by lysing single protoscolices or adult worms in 10 µl of 0.02 M NaOH at 95 °C for 10 min (Nakao et al., 2003) or by lysing small pieces of cyst tissue in 50–100 µl of 0.02 M NaOH at 99 °C for 30 min and centrifuging at 8,000*g* for 1 min (Addy et al., 2017). Where these methods gave inadequate DNA for PCR, genomic DNA was extracted from cyst tissues using the High Pure PCR template Preparation Kit (Roche Diagnostics GmbH, Germany). DNA was extracted from approximately 1 cm^2 of the FTA® cards used in taking the Mauritania samples, and from cyst tissue of the Corsican (French) samples using the iPrep ChargeSwitch gDNA Tissue Kit, (Invitrogen, France).

The complete mitochondrial *cox1* (1608 bp) and *nad1* (894 bp) genes were obtained by nested PCR. The *cox1* gene was either amplified in one piece, including the complete gene, or in two overlapping fragments. In either case, primary PCR was performed in a 50 µl reaction mixture containing 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 20 pmol of each external primer, 0.2 mM dNTPs, 1.25 U of Ampli-*Taq* Polymerase (Applied Biosystems, Germany) and 1 µl of crude lysate/DNA, and the reaction cycled 25 times under the following conditions: denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s and elongation at 72 °C for 90 s. In the nested PCR, the internal primers were added, the $MgCl_2$ concentration was adjusted to 1.375 mM and 0.5 µl of primary PCR product used as template, and the reaction cycled 35 times under same thermal conditions as the primary PCR. For

Please cite this article in press as: Addy, F., et al. Genetic differentiation of the G6/7 cluster of *Echinococcus canadensis* based on mitochondrial marker genes. Int. J. Parasitol. (2017), http://dx.doi.org/10.1016/j.ijpara.2017.06.003