Infection, Genetics and Evolution 45 (2016) 256–261

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Research paper

Development of a multilocus sequence typing tool for high-resolution subtyping and genetic structure characterization of *Cryptosporidium ubiquitum*

Ying Tang [a,b,1], Na Li [c,1], Mingxin Song [b], Dawn M. Roellig [a], Yaoyu Feng [c,*], Lihua Xiao [a,**]

[a] Division of Foodborne, Waterborne and Environmental Diseases, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30329, USA
[b] College of Veterinary Medicine, Northeast Agricultural University, Harbin, Heilongjiang 150030, China
[c] School of Resources and Environmental Engineering, East China University of Science and Technology, Shanghai 200237, China

ARTICLE INFO

Article history:
Received 30 July 2016
Received in revised form 8 September 2016
Accepted 10 September 2016
Available online 12 September 2016

Keywords:
Cryptosporidium ubiquitum
Multilocus sequence typing
Zoonotic transmission
Host adaptation

ABSTRACT

Cryptosporidium ubiquitum is an emerging zoonotic pathogen in humans. Recently, a subtyping tool targeting the 60-kDa glycoprotein (gp60) gene was developed for *C. ubiquitum*, and identified six subtype families (XIIa-XIIf). In this study, we selected five genetic loci known to be polymorphic in *C. hominis* and *C. parvum* for the development of a multilocus subtyping tool for *C. ubiquitum*, including CP47 (cgd6_1590), MSC6-5 (cgd6_4290), cgd6_60, cgd2_3690, and cgd4_370. PCR primers for these targets were designed based on whole genome sequence data from *C. ubiquitum*. DNA sequence analyses of 24 *C. ubiquitum* specimens showed the presence of 18, 1, 5, 4, and 5 subtypes at the CP47, MSC6-5, cgd6_60, cgd2_3690, and cgd4_370 loci, respectively. Altogether, 18 multilocus sequence typing (MLST) subtypes were detected among the 19 specimens successfully sequenced at all polymorphic loci. Phylogenetic analyses of the MLST data indicated that the rodent subtype families of XIIe and XIIf were highly divergent from others, and the ruminant XIIa subtype family formed a monophyletic group genetically distant from other rodent subtype families XIIb, XIIc, and XIId. The latter showed no consistent grouping of specimens and formed one large cluster in phylogenetic analysis of concatenated multilocus sequences. This was supported by results of STRUCTURE and F_{ST} analyses, which further suggested that XIIa originated from one common ancestor whereas XIIb, XIIc, and XIId contained mixed ancestral types, reflecting a close relatedness of the three subtype families and the likely occurrence of genetic recombination among them. Thus, an MLST tool was developed for high-resolution subtyping of *C. ubiquitum* and results of preliminary characterizations of specimens from humans and animals supported the conclusion on the existence of ruminant and rodent-adapted *C. ubiquitum* groups.

1. Introduction

Cryptosporidium ubiquitum, previously known as the *Cryptosporidium* cervine genotype, infects a broad range of host species (Fayer et al., 2010). It has been found in human cases worldwide, especially in industrialized countries (Blackburn et al., 2006; Chalmers et al., 2009; Cieloszyk et al., 2012; Davies et al., 2009; Elwin et al., 2012; Feltus et al., 2006; Molloy et al., 2010; Ong et al., 2002; Trotz-Williams et al., 2006). In addition to causing human disease, *C. ubiquitum* is detected in a variety of animals, especially ruminants and rodents (Diaz et al., 2015; Fayer et al., 2010; Feng et al., 2007; Li et al., 2016; Mirhashemi et al., 2016). Thus, *C. ubiquitum* has emerged as an important zoonotic species in humans. Nevertheless, little is known on the transmission route of *C. ubiquitum* in humans and animals and the significance of zoonotic infection in its epidemiology.

To characterize the transmission of human-pathogenic *Cryptosporidium* spp., various molecular diagnostic tools have been developed. For *C. ubiquitum*, a subtyping tool targeting the 60-kDa glycoprotein (gp60) gene has identified six subtype families (XIIa-XIIf) (Li et al., 2014). These subtype families differ in host range and geographic distribution. However, this observation is yet to be supported by multilocus sequence typing (MLST). Thus far, MLST tools are available for five *Cryptosporidium* species, including *C. hominis*, *C. parvum*, *C. meleagridis*, *C. muris*, and *C. andersoni*, and have been shown to be useful in genetic

* Correspondence to: Y. Feng, School of Resources and Environmental Engineering, East China University of Science and Technology, Shanghai 200237, China.
** Correspondence to: L. Xiao, Division of Foodborne, Waterborne and Environmental Diseases, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30329, USA.
E-mail addresses: yyfeng@ecust.edu.cn (Y. Feng), lxiao@cdc.gov (L. Xiao).
[1] Contributed equally to this work.

http://dx.doi.org/10.1016/j.meegid.2016.09.011
1567-1348/

characterizations of these species (Cama et al., 2006; Feng et al., 2011; Gatei et al., 2007; Gatei et al., 2006; Wang et al., 2014).

In this study, we selected five genetic loci known to be polymorphic in *C. hominis* and *C. parvum*, developed an MLST tool for high-resolution subtyping of *C. ubiquitum* from humans and animals.

2. Materials and methods

2.1. Specimens

DNA extractions from 24 *C. ubiquitum* specimens were used in this study. The specimens were from humans, one Verreaux's sifaka (*Propithecus verreauxi coquereli*), and various species of ruminants and rodents in the United States, Spain, the Slovak Republic, South Africa, China, and Nepal (Table 1). The specimens were assigned to XIIa-XIIf subtype families (Table 1) by gp60 sequence analysis in a previous study (Li et al., 2014).

2.2. MLST markers

For multilocus subtyping of *C. ubiquitum*, we selected five genetic loci based on their polymorphic nature in *C. parvum* and *C. hominis*. Three polymorphic markers of cgd6_60, CP47 (cgd6_1590) and MSC6-5 (cdg6_4290) were previously used in several MLST and population genetic analyses of *C. hominis* and *C. parvum* (Feng et al., 2014; Feng et al., 2013; Li et al., 2013), and two other loci of cgd2_3690 and cgd4_370 were found to be polymorphic in a recent comparison of *C. parvum* and *C. hominis* genomes (Guo et al., 2015). Their orthologs in *C. ubiquitum* were identified by alignment of whole genome sequence of *C. ubiquitum* obtained in a previous study (Li et al., 2014) with the reference sequence from the *C. parvum* IOWA genome (Abrahamsen et al., 2004). Primers for nested PCR were designed based on semi-conserved sequence of each locus with expected PCR products ranging from 619 bp to 1055 bp (Table 2).

2.3. MLST PCR

Nested PCR was used for the amplification of MLST markers. For each nested PCR, the total volume of PCR was 50 μl, which consisted of 1 μl of DNA (primary PCR) or 2 μl of the primary PCR product (secondary PCR), primers at a concentration of 0.25 μM (primary PCR using F1 and R1) or 0.5 μM (secondary PCR using F2 and R2), 200 mM deoxyribonuleotide triphosphate mix (Promega, Madison, WI), 3 mM $MgCl_2$ (Promega), 1 × GeneAmp PCR buffer (Applied Biosystems, Foster City, CA), and 1.25 U of Taq DNA polymerase (Promega). The primary PCR reactions also contained 400 ng/μl of non-acetylated bovine serum albumin (Sigma, Louis, MO). PCR amplification consisted of an initial denaturation at 94 °C for 5 min; 35 cycles of 94 °C for 45 s, the specified annealing temperature for each primer set (Table 2) for 45 s, and 72 °C for 1 min; and a final 7-min extension at 72 °C. Two negative controls (reagent-grade water) for primary PCR and secondary PCR were used in each PCR run. The secondary PCR products were detected by agarose gel electrophoresis and ethidium bromide staining. Two PCR replicates were used to analyze each DNA extraction.

2.4. DNA sequencing

The secondary PCR products of the expected size were sequenced in both directions using an ABI 3130 Genetic Analyzer (Applied Biosystems, Foster city, CA). A second PCR product was sequenced if the initial sequence was unreadable. The sequences obtained were assembled using ChromasPro version 1.5 (http://www.technelysium.com.au/ChromasPro.html), edited using BioEdit version 7.1 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html), and aligned using ClustalX version 1.81 (http://www.clustal.org/).

2.5. Data analysis

To assess the phylogenetic relationships among various gp60 subtype families of *C. ubiquitum*, neighbor-joining trees of nucleotide

Table 1
Specimens of *Cryptosporidium ubiquitum* used in the study and their subtype designations at the five selected loci.

Specimen	Host	Source location	gp60 subtype family[a]	MLST subtype				
				CP47	MSC6-5	cgd6_60	cgd2_3690	cgd4_370
38664	Sheep	China	XIIa	4	1	1	1	3
38665	Sheep	China	XIIa	4	+	1	1	3
32563	Impala	South Africa	XIIa	4	1	1	1	4
32560	Buffalo	South Africa	XIIa	5	1	1	1	4
30238	Swamp deer	Nepal	XIIa	2	−	1	1	4
30241	Swamp deer	Nepal	XIIa	3	−	1	1	4
31570	Sheep	Spain	XIIa	1	1	1	1	4
34141	Human	USA	XIIb	7	1	2	2	1
33456	Verreaux's sifaka	USA	XIIb	6	1	2	2	1
18868	Chipmunk	USA	XIIb	18	−	−	−	−
37192	Human	USA	XIIb	6	1	2	3	5
37202	Human	USA	XIIb	8	−	2	3	5
17574	Human	USA	XIIc	9	1	2	2	1
36638	Porcupine	USA	XIIc	10	1	2	2	1
37211	Human	USA	XIIc	9	+	2	3	5
17577	Human	USA	XIIc	11	1	2	2	1
29361	Human	USA	XIId	13	1	3	2	1
37204	Human	USA	XIId	14	1	2	2	1
14223	Beaver	USA	XIId	15	−	−	2	1
18370	Eastern gray squirrel	USA	XIId	−	−	−	2	2
14886	Eastern chipmunk	USA	XIId	12	1	2	2	1
37193	Human	USA	XIId	13	1	2	3	5
37646	Field mouse	Slovak Republic	XIIe	16	+	4	4	+
37644	Field mouse	Slovak Republic	XIIf	17	−	5	+	+

+, PCR positive but produced noisy signals at sequencing; −, PCR negative.
[a] The gp60 subtype family data of these specimens are obtained from a previous report (Li et al., 2014).