Title: High Genetic Diversities Between Isolates of the Fish Parasite *Cryptocaryon irritans* (Ciliophora) Suggest Multiple Cryptic Species

Hongshu Chi¹, Patricia Taik², Emily J. Foley², Alycia C. Racicot², Hilary M. Gray², Katherine E. Guzzetta², Hsin-Yun Lin³, Yen-Ling Song³, Che-Huang Tung⁴, Kosuke Zenke⁵, Tomoyoshi Yoshinaga⁵, Chao-Yin Cheng⁶, Wei-Jen Chang^{2,6#}, Hui Gong^{1,7#}

¹ Biotechnology Institute, Fujian Academy of Agricultural Sciences, Fuzhou, Fujian 350003, China

² Department of Biology, Hamilton College, Clinton, NY 13323, USA

³ Department of Life Science, National Taiwan University, Taipei 106, Taiwan

⁴ Department of Aquatic Biosciences, National Chyai University, Chyai City 60004, Taiwan

⁵ Department of Aquatic Bioscience, Graduate School of Agriculture, University of Tokyo, Tokyo 113-8657, Japan

⁶ Institute of Molecular Biology, Academia Sinica, Taipei 115, Taiwan

⁷ State Key Laboratory of Large Yellow Croaker Breeding, Ningde, Fujian 352103, China

Corresponding authors:

Wei-Jen Chang (wchang@hamilton.edu)
Tel: +1-315-8533126

Hui Gong (ghxfjm@163.com)
Tel: +86-591-87817514

Abstract

The ciliate protozoan *Cryptocaryon irritans* parasitizes marine fish and causes lethal white spot disease. Sporadic infections as well as large-scale outbreaks have been reported globally and the parasite's broad host range poses particular threat to the aquaculture and ornamental fish markets. In order to better understand *C. irritans*' population structure, we sequenced and compared mitochondrial *cox-1*, SSU rRNA, and ITS-1 sequences from 8 new isolates of *C. irritans* collected in China, Japan, and Taiwan. We detected two SSU rRNA haplotypes, which differ at three positions, separating the isolates into two main groups (I and II). *Cox-1* sequences also support the division into two groups, and the *cox-1* divergence between these two groups is unexpectedly high (9.28% for 1,582 nucleotide positions). The divergence is much greater than that detected in *Ichthyophthirius multifiliis*, the ciliate protozoan causing freshwater white spot disease in fish, where intraspecies divergence on *cox-1* sequence is only 1.95%. ITS-1 sequences derived from these eight isolates and from all other *C. irritans* isolates (deposited in the GenBank) not only support the two groups, but further suggest the presence of a third group with even greater sequence divergence. Finally, a small Ka/Ks ratio estimated from *cox-1* sequences suggests that this gene in *C. irritans* remains under strong purifying selection. Taken together, the *C. irritans* species may consist of many subspecies and/or syngens. Further work is needed to determine if there is reproductive isolation between the groups we have defined.

Keywords

Cryptic species; Tetrahymena; phylogeny; barcoding; variants; cytochrome c oxidase subunit 1; sexual reproduction

1. Introduction

The causative agent of marine fish white spot disease, *Cryptocaryon irritans*, is one of the most prominent threats to today's aquaculture and ornamental fish industries. For one, the ciliated protozoan parasite shows low host specificity: it can infect not only most marine fish (Coloni and Burgess, 1997), but also freshwater fish cultivated in artificial seawater (Burgess, et al, 1995). Further, *C. irritans* infections are highly contagious and generally lethal (Kaige and Miyazaki, 1985; Rasheed, 1989; Zuo, et al, 2012), and outbreaks occurring in high-density

mariculture result in substantial economic losses (Kaige and Miyazaki, 1985; Mo, et al, 2016; Hirazawa, et al, 2001b; Hirazawa, et al, 2001a).

C. irritans seems to be widespread; cases and outbreaks have been reported in South Africa, the USA, Japan, Korea, China, Taiwan, Thailand, Malaysia, Australia, and Israel (Kaige and Miyazaki, 1985; Vaughan and Christison, 2007; Diggles and Adlard, 1997; Yambot, et al, 2003; Jee, et al, 2000; Yoshinaga and Nakazoe, 1997; Diamant, et al, 1991; Sun, et al, 2006b; Hatanaka, et al, 2008; Leethochavalit, et al, 2011; Taniguchi, et al, 2011; Imajoh, et al, 2016). Based on sequences of the first internal transcribed spacer region (ITS-1), *C. irritans* isolates can be classified into at least two distinct groups (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006b; Leethochavalit, et al, 2011). The first group consists of isolates collected from geographically distant regions including Australia (Moreton Bay and Heron Island), Thailand, Malaysia, Taiwan (aquarium), Japan, the USA, and Israel. In contrast to the first group's wide distribution, isolates in the second group have only been reported in Japan, China, Taiwan (local isolates), and Australia (Moreton Bay) (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006b; Hatanaka, et al, 2008; Leethochavalit, et al, 2011; Taniguchi, et al, 2011; Imajoh, et al, 2016).

While ITS sequences are useful in classifying *C. irritans* isolates, they are short, and it can be difficult to make intra- and interspecies comparisons. The other frequently used phylogenetic marker, the mitochondrial cytochrome c oxidase subunit 1 (*cox-1*) sequence, has not been used in resolving intra- and interspecies phylogenetic relationships in *C. irritans* largely because the sequence was unavailable in this species before this study. *Cox-1* is a well-conserved protein coding gene, making it easy to align among orthologs in different species. Its DNA sequence is also longer (full length ~ 2,500 nt), and can potentially provide more evolutionary signals: both GenBank and the Consortium for the Barcode of Life use *cox-1* sequences as their default DNA barcoding sequences. Recent successes in using *cox-1* sequences to classify isolates of another ciliate parasite, *Ichthyophthirius multifiliis*, also suggest a similar approach may be applicable to help characterize isolates of *C. irritans* at the molecular level (MacColl, et al, 2015).

In this study we obtained *C. irritans cox-1* sequence from transcriptome sequencing using Illumina NGS technology. We then sequenced and compared *cox-1*, small ribosomal RNA, and ITS-1 sequences from 8 new isolates of *C. irritans* sampled from China, Taiwan, and Japan.

All three molecular markers support that *C. irritans* consists of two distinct populations, and ITS-1 sequences suggest the presence of a third group. As expected, *cox-1* sequences help resolve phylogenetic relationships among *C. irritans* isolates with similar ITS-1 sequences. We also detected greater than expected molecular divergences among *C. irritans* isolates.

2. Material and Methods

2.1 C. irritans DNA isolation

Tomonts and theronts of *C. irritans* were collected following established protocols (Yambot and Song, 2006; Yoshinaga, et al, 1994). *C. irritans* DNA was extracted by using DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA) following manufacturer's protocols.

2.2 *C. irritans* mitochondrial *cox-1* sequence

Since the mitochondrial *cox-1* sequence of *C. irritans* was not available before this study, we conducted a small scale transcriptome sequencing on the total RNA extracted from the CN_ND_2014 sample. Six cDNA libraries containing ~ 300 bp inserts made from poly-A selected RNA were constructed, and sequences of 412,760,652 paired-end, high-quality 101 bp reads were determined using an Illumina HighSeq sequencer (Genewiz, South Plainfield, NJ). Short read sequences were then assembled into contigs by using Trinity with default settings (Grabherr, et al, 2011), and *Ichthyophthirius multifiliis cox-1* sequence (GenBank: YP_004841742) was used to blast search for the best match in these contigs. The best matching contig consists of 2,950 bp, and an ORF can be putatively translated into a protein sequence belonging to the heme_Cu_oxidase I superfamily using the mitochondrial genetic code. Results derived from preliminary phylogenetic analyses further suggest that this sequence belongs to a ciliate species (data not shown). Based on these findings, we deduce that the contig contains *C. irritans cox-1* sequence (GenBank: KY427005), and primers were designed for subsequent PCR amplifications.

2.3 PCR, cloning, and sequencing

PCR was conducted following protocols published elsewhere (MacColl, et al, 2015). Briefly, each PCR mixture contained 1X GoTaq (Promega, Madison, WI), 0.2 uM of each primer, and 5 - 20 ng of DNA. Typical cycling conditions were: initial denaturation at 98 °C for 2 min, followed by 35 cycles of 95 °C 20 s, 50 °C (or another specified annealing temperature) 30 s, 72 °C 1 min (2 min for small rDNA), and a final extension step at 72 °C for 5 min. PCR products were either cloned into pGEM-T Easy vector (Promega) and sequenced (Genewiz), or directly sequenced after column purification (Zymo Research, Irvine, CA). Primer sequences are provided in Table S1.

2.4 Sequence and phylogenetic analyses

Sequences were manipulated using BioEdit software (Hall, 1999). GenBank accession numbers of newly reported sequences (KY426994 - KY427004, KY440336-KY440343, and KY486177-KY486184) and those published elsewhere are provided in Table S2. Bayesian

analyses were conducted using MrBayes (Ronquist, et al, 2012) with the following settings: GTR+I+G model, two independent chains, 100,000 MCMC steps, and a burnin of 2,500. By the end of MCMC the standard deviation of split frequencies reached 0.0052. Maximum likelihood (ML) tree analyses were conducted using PhyML (Guindon and Gascuel, 2003) with 100 bootstrapping replicates. The best model GTR+I+G (HKY+I for ITS-1) was predetermined by jModelTest (Darriba, et al, 2012). For maximum parsimony (MP), PAUP was used with 10,000 bootstrapping replicates. (Wilgenbusch and Swofford, 2003)

3. Results

3.1 *C. irritans* samples

Eight *C. irritans* samples were collected from infected fish found in local outbreaks in China, fish retailers at two local fishing ports in Taiwan, and a pet shop in Japan between 2012 and 2015 (Table 1). All samples are labeled by two letter abbreviations of the countries, followed by the abbreviations of cities where isolates were obtained, and then the years when samples were collected. For example, JP_TY_2014 represents an isolate collected in Tokyo, Japan, in 2014; CN_CL_2012 indicates a sample collected from Changle (Fujian Province), China, in 2012.

3.2 *C. irritans* *cox-1* sequences

The *cox-1* ORF recovered from the transcriptome assembly is 1,896 bp in length, and is likely to miss a short segment at its 5' end, based on preliminary blast comparisons to known *cox-1* sequences. SmartBlast results show that this sequence is most similar to *cox-1* sequences of *Tetrahymena* spp., but clearly indicate that this sequence does not belong to a *Tetrahymena* sp. (data not shown). Results derived from our phylogenetic analyses place this sequence and sequences from other *C. irritans* isolates published in this study into a clade next to the Oligohymenophorea clade (Supplementary Fig. S1), which is in agreement with the phylogenetic position of *C. irritans* deduced from 18S SSU rRNA sequences (Diggles and Adlard, 1995). We were therefore confident that this is the (partial) *cox-1* sequence of *C. irritans* and designed primers to amplify a 1,582 bp region.

All 8 *C. irritans* DNA yielded ~ 1,600 bp PCR products; the PCR products were cloned, sequenced, and aligned for phylogenetic analyses. In isolates CN_CL_2012, CN_ND_2013, TW_WL_2015, and JP_TY_2014 we recovered two haplotypes, indicated by a 1 or 2 at the ends of the isolate names, e.g. CN_CL_2012_1 and CN_CL_2012_2; for JP_TY_2014 we use 81 and 86. There are 14, 20, and 11 nt differences between the two *cox-1* haplotypes identified in CN_CL_2012, CN_ND_2013, and TW_WL_2015, respectively. The two JP_TY *cox-1* haplotypes differ by only four nucleotides.

3.3 Phylogenetic analyses of *C. irritans* *cox-1* sequences

Phylogenetic analyses with maximum parsimony, maximum likelihood, and Bayesian methods show that the *cox-1* haplotypes identified in *C. irritans* isolated from China and Taiwan can be classified into several subgroups (Group I, Fig. 1). With strong bootstrap supports, we assign CN_FD_2012, TW_WL_2015_2, CN_CL_2012_2, CN_ND_2013_2, and CN_ND_2014 to a subgroup Ib, which is in contrast to the other haplotypes recovered from the CN_CL_2012, TW_WL_2015, and CN_ND_2013 samples. We arbitrarily assign the later haplotypes, namely CN_CL_2012_1, TW_WL_2015_1, and CN_ND_2013_1, to a subgroup Ia, which together with subgroup Ib constitutes a well-supported clade (Fig. 1). Our results suggest that in Southeast China and Taiwan there are at least two distinct subgroups of *C. irritans* involved in outbreaks. For instance, the 2013 outbreak in Ningde (Fujian Province, China, CN_ND_2013) could be attributed to the spread of both subgroup Ia (CN_ND_2013_1) and Ib (CN_ND_2013_2). A sample collected from northern Taiwan in 2015, TW_WL_2015, shows a similar pattern: we isolated both subgroup Ia (TW_WL_2015_1) and subgroup group Ib (TW_WL_2015_2) from the same infected fish.

TW_KL_2013 and TW_KL_2014 contain identical *cox-1* sequences and were collected from the same fishing port and from the same species of host fish (*Acanthopagrus schlegelii*) in two consecutive years (Table 1). These two isolates, while genetically closer to groups Ia and Ib than to group II (see below), are unique by themselves (2.83% difference between TW_KL_2014 and CN_FD_2012, Fig. 1). Because *A. schlegelii* came from net-cage fish farming, our results indicate that an isolated *C. irritans* population may have persisted in that farming area for years.

3.4 The Japanese *C. irritans* isolate has a distinct *cox-1* haplotype

The genetic distance between the Japanese isolate and the others from China and Taiwan is unusually long, and thereby we classify the Japanese isolate into a distinct group II. Between CN_FD_2012 and JP_TY_81 we detected 147 differences among 1,582 positions (9.29%). This genetic distance is slightly longer than that of *Tetrahymena australis* and *Tetrahymena canadensis*, and shorter than that of *Tetrahymena limacis* and *Tetrahymena*

borealis (Fig. 1). The 9.29% intraspecies divergence is also much higher than those reported in most *Tetrahymena* spp., where averaged intraspecies divergences were smaller than 1% on 689 positions of *cox-1* sequences (Chantangsi, et al, 2007). The averaged interspecies divergence within the *Tetrahymena* genus was found to be 10.47% on *cox-1* sequences (Chantangsi, et al, 2007). Finally, the ratio between the nonsynonymous and synonymous substitution rates (Ka/Ks) calculated from CN_FD_2012 and JP_TY_81 using the software KaKs_Calculator (Zhang, et al, 2006) with the LWL method of model averaging (Li, et al, 1985) was 0.039, suggesting that the *cox-1* gene in *C. irritans* remains under strong purifying selection.

3.5 Genetic divergences of SSU rRNA sequences

Because of the greater than expected intraspecies divergences found on the *cox-1* locus between the group I and II *C. irritans* isolates, we further ask whether a similar degree of divergence could be detected in the SSU rRNA sequence. We found three allelic sites among 1,720 positions, or 0.17%, in the SSU rRNA sequence between group I and II *C. irritans* isolates (GenBank: KY440336-43). In comparisons, the averaged interspecies divergence found in SSU rRNA sequences within the *Tetrahymena* genus was 1.39% from 1,639 positions (Chantangsi, et al, 2007). However, the degree of divergences varies: there is only a one site difference between *T. limacis* and *T. borealis*. In contrast, there are 22 sites different between *T. canadensis* and *T. australis*.

3.6 Genetic divergences of ITS-1 sequences

Because the ITS-1 sequence has been the genetic marker used to classify *C. irritans* isolates (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006b), it is worth comparing phylogenetic relationships derived from—and genetic divergences detected—for both ITS-1 and *cox-1* sequences. In general, compared to *cox-1* sequences, we detected much lower genetic divergences in ITS-1 sequences. In Group Ia and Ib isolates (CN_CL, CN_FD, CN_ND, and TW_WL), instead of multiple *cox-1* haplotypes we recovered from these isolates, we detected only one allele of ITS-1 sequence among all of them. The other two Group I isolates (TW_KL) share identical ITS-1 sequences, which are only one nucleotide apart from those of other Group I isolates (CN_CL, CN_FD, CN_ND, and TW_WL) among 305 positions.

This low genetic divergence presents a sharp contrast to the 2.83% variance observed for the *cox-1* locus (see above). In fact, ITS-1 sequences derived from all China and Taiwan isolates described in this study, and from some isolates collected from China, Taiwan, Japan, and Australia in other studies, differ from one another by no more than 3 nucleotides on this locus (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006b; Taniguchi, et al, 2011; Imajoh, et al, 2016). We thus used the sequence derived from the CN_ND_2013 isolate (GenBank: KY486179) to represent Group I and those closely related isolates for further phylogenetic analyses (Group I*, Fig. 2). The ITS-1 sequence derived from the JP_TY_2014 isolate (Group II) is distinct from the Group I ITS-1 sequence, and differs from a previously published sequence derived from an Israeli isolate by one nucleotide (GenBank: AY029269, Fig. 2)(Diggles and Adlard, 1997).

Phylogenies inferred from the ITS-1 sequences reveal one well supported clade consisting of *C. irritans* isolates collected from the USA, Australia, Israel, South Africa, Japan, Malaysia, and an aquarium in Taiwan (Diggles and Adlard, 1997; Yambot, et al, 2003; Hatanaka, et al, 2008)(Fig. 2). Our JP_TY_2014 is also a member of this clade, suggesting that Group II defined by us may well be expanded to represent the whole monophyletic clade (Fig. 2). The lineage Group I* represents all isolates collected from China, most of the isolates from Taiwan and Japan, and one from Australia as mentioned above. The tree also reveals yet another independent group consisting of isolates collected from Kaoshiung City and Penghu Island, Taiwan (AF490383 and AF490382, respectively, Fig. 2)(Yambot, et al, 2003). Unfortunately we do not have *cox-1* sequences from these isolates. However, given that *cox-1* sequences tend to accumulate more mutations than ITS-1 (e.g. 2.83% divergence vs. 0.3% between CN_FD and TW_WL; 9.29% vs. 3.27% between CN_FD and JP_TY), *cox-1* sequences from the Kaoshiung and Penghu isolates are expected to display even greater divergences (ITS-1 divergence is 6.56% between the Kaoshiung isolate (AF490383) and CN_FD (KY486177)).

4. Discussion

In this study we show that *cox-1* sequences can be used to effectively distinguish 8 *C. irritans* isolates. Similar approaches have been applied to and yield successful outcomes in detecting both intra- and interspecies variation in *Tetrahymena* spp. (Chantangsi, et al, 2007; Chantangsi and Lynn, 2008), and intraspecies variations in another ciliate parasite, *Ichthyophthirius multifiliis* (MacColl, et al, 2015). Based on *cox-1* sequences *C. irritans* can be divided into two distinct groups, I and II, with 9.29% sequence divergence between these two groups, which is far greater than some interspecies divergences in *Tetrahymena* spp. (Fig. 1, and (Chantangsi, et al, 2007)). This greater-than-expected divergence is much higher than that detected in the protozoan fish parasite *I. multifiliis*. There, 18 geographically distinct *I. multifiliis* isolates were classified into three groups (Zaila, et al, 2017), but the largest intraspecies divergence was only 1.95% across 1,493 *cox-1* nucleotide positions (between KT783594 and KJ690555).

We also find two SSU rDNA haplotypes differing by three nucleotides in *C. irritans*. In *I. multifiliis* only one SSU rDNA haplotype was detected at this locus among 9 isolates (MacColl, et al, 2015). The two SSU rDNA haplotypes detected in *C. irritans* also distinguish Group I isolates from Group II. Compared to *Tetrahymena* spp., the three-nucleotides difference is smaller than the averaged interspecies divergence found on the SSU rDNA locus (~23 positions, 1.39% from 1,639 positions), but is greater than that of some species pairs (e.g. one nucleotide difference between *T. limacis* and *T. borealis*). (Chantangsi, et al, 2007).

The high intraspecies divergence of the *cox-1* locus between the two main groups, and the group-distinguishing SSU rDNA haplotypes of *C. irritans*, suggest the presence of cryptic species (syngens). Furthermore, ITS-1 sequences not only support the division of two main groups, but indicate the presence of a third group (Fig. 2). High sequence divergence of ITS-1 (> 10%) was also noted between isolates in Group II and III (Fig. 2) (Diggles and Adlard, 1997; Yambot, et al, 2003). Because *cox-1* sequences display higher divergences than ITS-1 sequences in isolate pairs (i.e. JP_TY vs. CN_FD, see Results), we expect more mitochondrial haplotypes present in Group II isolates, and even higher divergences in Group III isolates in *cox-1* sequences. It would be interesting to know whether isolates from different groups are sexually isolated, and thus biological species. However, the mating process in *C. irritans*, if any, is completely unknown to date.

Given that the *cox-1* Ka/Ks ratio for *C. irritans* is low, the high divergences do not seem to be contributed by relaxed selection; high mutation rates, large effective population size, and/or long evolutionary time more likely account for the observed high genetic divergences. Clearly, more sequencing data from different *C. irritans* groups will be needed to further elucidate the evolutionary history of this parasite.

Within Group I, variations in *cox-1* sequences, but not ITS-1, allow us to detect co-infections of different *C. irritans* subgroups with different mitochondrial *cox-1* haplotypes in some local outbreaks (Ia and Ib, Fig. 1). The Ia and Ib subgroups thus seem to be the dominant strains in Southeast China and Taiwan. Because both subgroups are present in both China and Taiwan, the Taiwan Strait does not seem to effectively block cross-strait transmission, which may be mediated by currents, human transportations, and/or bird migration. In addition to subgroups Ia and Ib, a separately evolved population was also discovered. Isolates TW_KL_2013 and TW_KL_2014 were collected from the same species of fish host provided by the same fishing port vendor in two consecutive years. These two share identical *cox-1* sequences, but their sequences are distinct from those of other group I isolates. Given that the host fish *A. schlegelii* is now raised in cage fish farms, this distinct *C. irritans* population might have persisted in that fish farming area for years. Noteworthy, the divergence of *cox-1* sequences between TW_KL_2014 and CN_FD_2012 (group Ia) was found to be 2.83%. This is still much greater than the largest intraspecies divergence found in *I. multifiliis* (1.95%).

Based on our findings and early findings derived from ITS-1 sequences (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006a), the geographical distribution of Group I *C. irritans* may be limited to China, Taiwan, Japan, and Australia (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006b; Taniguchi, et al, 2011; Imajoh, et al, 2016). This is in contrast to the broader spread of Group II *C. irritans*, which has been found in Israel (Mediterranean Sea), South Africa, Japan, Malaysia, Thailand, Australia (Diggles and Adlard, 1997; Yambot, et al, 2003; Hatanaka, et al, 2008; Leethochavalit, et al, 2011). *C. irritans* isolates collected from an aquarium in Taiwan (Pingtung) (Yambot, et al, 2003), a pet store in Japan (Tokyo, this study) and the USA (Georgia) (Diggles and Adlard, 1997), also belong to Group II. While it seems that Group I and Group II show distinct geographical distribution patterns, the sampling sizes remain too low to derive conclusive correlations, if any, between

groups/subgroups and localities. More sampling around the globe will be needed in order to visualize the distributions of these main groups.

Finally, among all the countries reporting *C. irritans* infections, Australia is the only one where both Groups I and II *C. irritans* were known to be present in wild fish populations: Group I *C. irritans* was found on infected fish from Moreton Bay (Lab C strain in (Diggles and Adlard, 1997)), and Group II from Heron Island (Diggles and Adlard, 1997), which is 700 km away from Moreton Bay. The significance of this observation is yet to be determined. Now, with the option to classify *C. irritans* isolates using *cox-1* sequences, it will be possible to trace migration/transportation and better understand the epidemiology and evolutionary history of this parasite.

Acknowledgements

The authors would like to thank Drs. Mitsuru Eguchi, Akimasa Hatanaka, and Masayuki Imajoh for providing information regarding *C. irritans* isolates; Dr. Georgi Marinov for sharing *Paramecium cox-1* sequences; Dr. Meng-Chao Yao for providing assistances to facilitate the project; Dr. Thomas Doak for critical comments on the manuscript.

Funding Resources

This research was supported by Key Projects of Fujian Province for Agriculture (2013NZ0002); the Science and Technology Program from Fujian Province (No. 2016R1018-4); National Science Foundation (MRI-0959297); Research Corporation Cottrell College Award (20976); the Wertimer Endowed Chair Fund to WC; and Hamilton College Summer Research Grants.

Figure Legend

Figure 1.

Phylogenetic tree inferred from bayesian analyses. Mitochondrial *cox-1* sequences derived from *C. irritans* isolates, and from species in Oligohymenophorea: *Paramecium* spp., *Tetrahymena* spp., and *Ichthyophthirius multifiliis* were aligned and subjected to maximum parsimony, maximum likelihood, and Bayesian phylogenetic analyses. Nodes with bootstrap and posterior probability values higher than 70 are denoted with solid circles. Vertical lines indicate groupings and subgroupings in *C. irritans* isolates. Tree was drawn to scale by using iTOL (Letunic, et al, 2016).

Figure 2.

Maximum likelihood tree inferred from ITS-1 sequences of *C. irritans* isolates. Nodes with bootstrap values higher than 70 from both maximum likelihood (HKY+I) and maximum parsimony analyses are denoted with solid circles. Group I* represents closely related ITS-1 sequences derived from isolates collected in Southeast China, Taiwan, Japan, and Australia (GenBank nos: AB608054, AF490381, AY029273, DQ270008-13, HQ148713-4, JN636814, KC357673, KU761582, LC128576, KY486177-80, KY486182-84, Table S2) (Diggles and Adlard, 1997; Yambot, et al, 2003; Sun, et al, 2006b; Taniguchi, et al, 2011; Imajoh, et al, 2016). Group II was first designated to JP_TY_2014 isolate and is expanded to represent the well supported monophyletic clade consisting of isolates collected worldwide (Diggles and Adlard, 1997; Yambot, et al, 2003; Hatanaka, et al, 2008). Isolates with identical sequences are separated by semicolons. Group III is restricted to the two distinct isolates from Taiwan (Yambot, et al, 2003). Tree was drawn to scale by using iTOL (Letunic, et al, 2016).

References

- Burgess, P.J. and Matthews, R.A., 1995. Fish host range of seven isolates of *Cryptocaryon irritans* (Ciliophora). *J. Fish Biol.* 46, 727-729.
- Chantangsi, C. and Lynn, D.H., 2008. Phylogenetic relationships within the genus *Tetrahymena* inferred from the cytochrome c oxidase subunit 1 and the small subunit ribosomal RNA genes. *Mol. Phylogenet. Evol.* 49, 979-987.
- Chantangsi, C., Lynn, D.H., Brandl, M.T., Cole, J.C., Hetrick, N., and Ikonomi, P., 2007. Barcoding ciliates: a comprehensive study of 75 isolates of the genus *Tetrahymena*. *Int. J. Syst. Evol. Microbiol.* 57, 2412-2425.
- Colorni, A. and Burgess, P., 1997. *Cryptocaryon irritans* Brown 1951, the cause of white spot disease in marine fish: an update. *Aquarium Sci. Conserv.* 1, 217-238.
- Darriba, D., Taboada, G.L., Doallo, R., and Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772.

Diamant, A., Issar, G., Colorni, A., and Paperna, I., 1991. A pathogenic Cryptocaryon-like ciliate from the Mediterranean Sea. *Bull. Eur. Ass. Fish Pathol.* 11, 122-124.

Diggles, B.K. and Adlard, R.D., 1997. Intraspecific variation in *Cryptocaryon irritans*. *J. Eukaryot. Microbiol.* 44, 25-32.

Diggles, B.K. and Adlard, R.D., 1995. Taxonomic affinities of *Cryptocaryon irritans* and *Ichthyophthirius multifiliis* inferred from ribosomal RNA sequence data. *Dis. Aquat. Org.* 22, 39-43.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644-652.

Guindon, S. and Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696-704.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95-98.

Hatanaka, A., Umeda, N., and Hirazawa, N., 2008. Molecular cloning of a putative agglutination/immobilization antigen located on the surface of a novel agglutination/immobilization serotype of *Cryptocaryon irritans*. *Parasitology* 135, 1043-1052.

Hirazawa, N., Oshima, S.-I., Hara, T., Mitsuboshi, T., and Hata, K., 2001a. Antiparasitic effect of medium-chain fatty acids against the ciliate cryptocaryon irritans infestation in the red sea bream *pagrus major*. *Aquaculture* 198, 219-228.

Hirazawa, N., Oshima, S.-I., and Hata, K., 2001b. In vitro assessment of the antiparasitic effect of caprylic acid against several fish parasites. *Aquaculture* 200, 251-258.

Imajoh, M., Morimitu, K., Sukeda, M., Umezaki, T., Monno, S., Goda, H., Kubo, E., and Oshima, S., 2016. TaqMan Real-Time PCR Detection and Phylogenetic Analysis of *Cryptocaryon irritans* in Nomi Bay, Kochi, Japan. *Fish Pathol.* 51, 103-111.

Jee, B.-Y, Ki, H.K., Park, S.-I, and Kim, Y.-C, 2000. A new strain of *Cryptocaryon irritans* from the cultured olive flounder *Paralichthys olivaceus*. *Dis. Aquat. Org.* 43, 211-215.

Kaige, N. and Miyazaki, T., 1985. A Histopathological Study of White Spot Disease in Japanese Flounder. *Fish Pathol.* 20, 61-64.

Leethochavalit, S., Watanachote, J., and Rittirut, N., 2011. Characterization of *Cryptocaryon* sp. isolated from marine fish in THAILAND and in vitro treatment.

Letunic, I. and Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242-5.

Li, W.H., Wu, C.I., and Luo, C.C., 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* 2, 150-174.

MacColl, E., Therkelsen, M.D., Sherpa, T., Ellerbrock, H., Johnston, L.A., Jariwala, R.H., Chang, W., Gurtowski, J., Schatz, M.C., Mozammel Hossain, M., et al, 2015. Molecular genetic diversity and characterization of conjugation genes in the fish parasite *Ichthyophthirius multifiliis*. *Mol. Phylogenet. Evol.* 86, 1-7.

Mo, Z.Q., Li, Y.W., Wang, H.Q., Wang, J.L., Ni, L.Y., Yang, M., Lao, G.F., Luo, X.C., Li, A.X., and Dan, X.M., 2016. Comparative transcriptional profile of the fish parasite *Cryptocaryon irritans*. *Parasit. Vectors* 9, 630.

Rasheed, V.M., 1989. Diseases of Cultured Brown-Spotted Grouper *Epinephelus tauvina* and Silvery Black Porgy *Acanthopagrus cuvieri* in Kuwait. *J. Aquat. Anim. Health* 1, 102-107.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539-542.

Sun, H.Y., Zhu, X.Q., Xie, M.Q., Wu, X.Y., Li, A.X., Lin, R.Q., and Song, H.Q., 2006a. Characterization of *Cryptocaryon irritans* isolates from marine fishes in Mainland China by ITS ribosomal DNA sequences. *Parasitol. Res.* 99, 160-166.

Sun, H.Y., Zhu, X.Q., Xie, M.Q., Wu, X.Y., Li, A.X., Lin, R.Q., and Song, H.Q., 2006b. Characterization of *Cryptocaryon irritans* isolates from marine fishes in Mainland China by ITS ribosomal DNA sequences. *Parasitol. Res.* 99, 160-166.

Taniguchi, A., Onishi, H., and Eguchi, M., 2011. Quantitative PCR assay for the detection of the parasitic ciliate *Cryptocaryon irritans*. *Fisheries Science* 77, 607-613.

Vaughan, D. and Christison, K., 2007. *Cryptocaryon irritans*: food-fish for thought. *Parassitologia* 49, 131.

Wilgenbusch, J.C. and Swofford, D., 2003. Inferring evolutionary trees with PAUP*. *Curr. Protoc. Bioinformatics* Chapter 6, Unit 6.4.

Yambot, A.V. and Song, Y.-L., 2006. Immunization of grouper, *Epinephelus coioides*, confers protection against a protozoan parasite, *Cryptocaryon irritans*. *Aquaculture* 260, 1-9.

Yambot, A.V., Song, Y.-L., and Sung, H.-H., 2003. Characterization of *Cryptocaryon irritans*, a parasite isolated from marine fishes in Taiwan. *Dis. Aquat. Org.* 54, 147-156.

Yoshinaga, T. and Dickerson, H.W., 1994. Laboratory propagation of *Cryptocaryon irritans* on a saltwater-adapted *Poecilia* hybrid, the black molly. *J. Aquat. Anim. Health* 6, 197-201.

Yoshinaga, T. and Nakazoe, J.-I., 1997. Acquired protection and production of immobilization antibody against *Cryptocaryon irritans* (Ciliophora, Hymenostomatida) in *Mummichog* (*Fundulus heteroclitus*). *Fish Pathol.* 32, 229-230.

Zaila, K., Doak, T., Ellerbrock, H., Tung, C., Martins, M., Kolbin, D., Yao, M., Cassidy-Hanley, D., Clark, T., and Chang, W., 2017. Diversity and universality of endosymbiotic *Rickettsia* in the fish parasite *Ichthyophthirius multifiliis*. *Frontiers in Microbiology* 8, 189.

Zhang, Z., Li, J., Zhao, X.Q., Wang, J., Wong, G.K., and Yu, J., 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics* 4, 259-263.

Zuo, R., Ai, Q., Mai, K., Xu, W., Wang, J., Xu, H., Liufu, Z., and Zhang, Y., 2012. Effects of dietary n-3 highly unsaturated fatty acids on growth, nonspecific immunity, expression of some

immune related genes and disease resistance of large yellow croaker (*Larimichthys crocea*) following natural infestation of parasites (*Cryptocaryon irritans*). *Fish and Shellfish Immunology* 32, 249-258.

Table 1. Information of eight samples of *C. irritans*

Samples	Locations	Dates	Hosts	Sources	Stages
CN_FD_2012	Fuding, Fujian, China	2012	<i>Epinephelus awoara</i>	Fish farm	Tomonts
CN_CL_2012	Changle, Fujian, China	2012	<i>Paralichthys olivaceus</i>	Fish farm	Tomonts
CN_ND_2013	Ningde, Fujian, China	2013	<i>Larimichthys crocea</i>	Fish farm	Tomonts
CN_ND_2014	Ningde, Fujian, China	2014	<i>Larimichthys crocea</i>	Fish farm	Tomonts
TW_KL_2013	Keelung, Taiwan	2013	<i>Acanthopagrus schlegelii</i>	Fishing port	Tomonts [#]
TW_KL_2014	Keelung, Taiwan	2014	<i>Acanthopagrus schlegelii</i>	Fishing port	Theronts [#]
TW_WL_2015	Wanli, Taiwan	2015	<i>Plectorhinchus gibbosus</i>	Fishing port	Theronts [#]
JP_TY_2014	Tokyo, Japan	2014	<i>Paracanthurus hepatus</i>	Pet shop	Tomonts [#]

[#] From a single host fish.

Figure 1

Figure 2

Figure Abstract