EXPANSION OF THE *CHLOROVIRUS* GENUS BY STUDIES ON VIRUS NATURAL HISTORY AND *CHLORELLA* HOST METABOLISM

by

Cristian F. Quispe

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Biological Sciences
(Genetics, Cell, and Molecular Biology)

Under the Supervision of Professor James L. Van Etten

Lincoln, Nebraska
December, 2015

EXPANSION OF THE *CHLOROVIRUS* GENUS BY STUDIES ON VIRUS NATURAL HISTORY AND *CHLORELLA* HOST METABOLISM

Cristian F. Quispe, Ph.D.
University of Nebraska, 2015
Adviser: James L. Van Etten

Inland waters cover about 2.5 percent of our planet and harbor huge numbers of known and unknown microorganisms including viruses. Viruses likely play dynamic, albeit largely undocumented roles in regulating microbial communities and in recycling nutrients in the ecosystem. Phycodnaviruses are a genetically diverse, yet morphologically similar, group of large dsDNA-containing viruses (160- to 560-kb) that inhabit aquatic environments. Members of the genus *Chlorovirus* are common in freshwater. They replicate in eukaryotic, single-celled, chlorella-like green algae that normally exist as endosymbionts of protists in nature. Very little is known about the natural history of the chloroviruses and how they achieve high-titer and long-term persistence in nature. To study their natural history, we examined chloroviruses over a three-year period to determine their abundance, prevalence, and genetic diversity in a small lake in Nebraska (Chapter II). These studies indicated that the amount of infectious virus particles was seasonal and both host- and site-dependent. *Chlorovirus* populations persisted year-round, suggesting that the viruses are either very stable or that viral production occurs in an unknown natural host(s). During this study, a new viral group was discovered and characterized, expanding the *Chlorovirus* genus

(Chapter III). This group, designated as *Only Syngen viruses* (OSy), replicates in *Chlorella variabilis* (Syngen 2-3) cells. Furthermore, OSy viruses also have non-permissive features in two phylogenetically related *C. variabilis* sub-species and constitute the first report of a post-infection host mechanism that results in resistance against infection. In Chapter IV, five symbiotic-virus suceptible and four free-living *Chlorella* species were evaluated for their capabilities to assimilate nutrients. Hierarchical clustering reveals a clear distinction of both groups based on their assimilation of galactose, nitrate, asparagine, proline, and serine. Additionally, genomic and differential expression analyses of symbiotic algae confirm an abundance of amino acid transporter genes, some of which are constitutively expressed when the symbiotic algae either grow axenically or as an endosymbiont within their host. Such similarities indicate a parallel coevolution of shared metabolic pathways across multiple independent symbiotic events and suggest that physiological changes driving the *Chlorella* symbiotic phenotype also contribute to their natural fitness.

To the mysterious and invisible forces that have made everything work in the past, the present and the future…

ACKNOWLEDGMENTS

I would like to acknowledge the wonderful people with whom I have interacted through my life; because, directly or indirectly, they have helped me in shaping my personality. I do not have enough time and space to name everyone, but I want to include my family, friends and professors that have shared with me the much-needed support, friendship, advice and encouragement throughout these years. You have changed my life for the better, and all of my achievements are shared with you. Overall, these six years at the University of Nebraska-Lincoln (UNL) have been a great and invaluable experience, although tough at moments, I could never have done it without all of you.

I would like to express my utmost appreciation and thankfulness to my adviser and mentor, Dr. James L. Van Etten. His guidance and patience during these years have been invaluable. He has allowed me to pursue my own scientific interests while imparting his advice and wisdom. He has contributed to my professional and personal growth, for which I will always be indebted to him. Dr. Van Etten, thank you for letting me be a part of the fascinating world of giant viruses.

I am further grateful to all the professors in the Nebraska Center for Virology, especially Drs. Clinton Jones, Matthew Wiebe, Asit Pattnaikt and Qingsheng Li, for imparting their knowledge, which has helped improved my experiences at UNL. I would also like to thank the members of my supervisory committee: Drs. Kenneth Nickerson, Amit Mitra, Zhang Lwen, and Wayne Riekhof. Thank you for your comments and collaborative effort to improve the work presented in this dissertation. Additionally, I would like to thank Dr. James Steadman at the Department of Plant Pathology for connecting me with the opportunity to begin my journey at UNL.

A special thanks to all current and past members and students in the Van Etten Lab; it has been a great joy to share these years with you. I was very lucky to work with such a unique and supportive group of people. A special mention

goes to Anya Seng, Olivia Sonderman and Michelle McQuinn for their constant support and understanding, especially in difficult moments during my Master and Ph.D. programs. Also, I extend my highest gratitude to Dr. Les Lane for believing in my “new prospective” to study *Chlorovirus*. Additionally, thanks to James Gurnon, who taught me all necessary protocols and tips to work with the viral-host system.

In a wider context, none of this work would have been possible without the support of my family. All your love, care and wisdom significantly contributed to the final completion of this work. I am very grateful to my mother, Ligia, who raised me to pursue dreams, value knowledge and honor education. I also want to offer my appreciation to all of my friends in the US, who have made my experiences here so unique and unforgettable. Additional appreciation goes to my friends in Ecuador and around the world, who are the best friends anyone could ask for; thanks for staying close to me all these years.

This work was supported by the UNL School of Biological Sciences Teaching Assistantships including Special Funds, the Nebraska Center for Virology, the Department of Plant Pathology, the undergraduates funded by the UCARE and ARD scholarships, the NSF-EPSCoR, Stanley Medical Research Institute, and the COBRE grants.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

Literature Review

Three-year Survey of Abundance, Prevalence and Genetic Diversity of Chlorovirus Populations in a Small Urban Lake

Characterization of a New Chlorovirus Type with Permissive and Non-Permissive Features on Phylogenetically Related Algae Strains

Comparative Genomics, Transcriptomics and Metabolism Distinguish Symbiotic from Free-living Chlorella

TABLES

Literature Review

Three-year Survey of Abundance, Prevalence and Genetic Diversity of Chlorovirus Populations in a Small Urban Lake

Characterization of a New Chlorovirus Type with Permissive and Non-Permissive Features on Phylogenetically Related Alga

Comparative Genomics, Transcriptomics and Metabolism Distinguish Symbiotic from Free-living Chlorella

CHAPTER I

INTRODUCTION

Literature Review

Viruses: cosmopolitan, abundant and important entities in nature

Viruses are ubiquitous members of the biosphere as they are found in essentially every ecosystem on the planet (Short, 2012). For example, analyses of aquatic environmental samples indicate that high concentrations of viruses (10^5 to 10^9 particles/ml) that infect microorganisms, primarily bacteria, are present in marine and inland waters (e.g., Lim et al., 2013; Rodriguez-Brito et al., 2010; Short, 2012; Yau et al., 2011). The virus number typically exceeds that of cellular organisms by at least an order of magnitude; thus, the number of different viruses within a community is huge. Their functions of predation and gene transfer make viruses key drivers in the dynamics of microbial ecosystems (Mokili et al., 2012; Suttle, 2007). Furthermore, viruses play important roles in the global biogeochemical cycling of carbon and nutrients (Bratbak et al., 1990; Rohwer & Thurber, 2009).

Studies from diverse biomes show that different environments possess distinct viral community structures. Even small and individual ecosystems, such as human feces, contain around 1,000 viral genotypes, whereas viral communities in seawater, although they are more diverse, contain around 5,000 genotypes (Breitbart et al., 2002; Breitbart et al., 2003). In both environments, the predominant viral type accounted for at least 1% of the total population. In contrast, samples collected from near-shore marine sediments were highly diverse, hosting between 10,000 and one million viral genotypes, with the most

abundant type representing less than 0.01% of the community. Thus, our biosphere is abundant with genetic information, which mainly is of viral origin and we have not yet been able to assign a role, function or evolutionary origin (Cortez et al., 2009).

"Virus" misleading us for decades

Virology officially started in 1898 (Beijerinck & Johnson, 1964), and for decades, viruses were defined by what they were not: very small entities (ultra filterable microbes) not visible under the microscope and not culturable in the absence of a host. Viruses were first considered as possible intermediate forms between mineral and true cellular life (Witzany, 2012). Additionally, at the end of the 20th century, the first viruses known to the public were those causing malignant phenotypes in clinical or agricultural organisms such as yellow fever in humans, mosaic disease in tobacco, and foot-and-mouth disease in livestock. Not surprisingly, virus is the Latin term for "poison, venom, or slimy fluid," which reflects its common strategy to survive (Witzany, 2012). Although at the time, the depiction of viruses as malicious killers was very appealing, new information about viruses emphasizes the important roles they possess, not only in the evolution of all life, but also as symbionts or co-evolutionary partners of host organisms (Witzany, 2012). Thus, with the advances in sequencing and technology of modern science, we are just now able to rediscover and understand what "viruses" really encompass.

Viruses are unlimited sources for diversity in their composition and shape

Viral genomes are the major source of genetic information in the biosphere, as they are known to evolve rapidly (Koonin et al., 2015). Despite being the most diverse biological entities, viruses are also the least characterized microbe in terms of their genetic, taxonomic, and functional diversity; for example, they often contain unique genes for which no homologues exist (Witzany, 2012).

There are countless unique genes in viruses with the potential to have unique and completely unexpected functions; in such a diverse pool, genes can produce structurally and functionally conserved proteins that have no apparent cellular ancestors. For example, novel proteins are able to generate unlimited structures, as evidenced by the various shapes seen in prokaryotic viruses: lemon-shaped viruses, tulip-shaped viruses, bottle-shaped viruses, stick-shaped viruses with hooks and pleomorphic-viruses along with others with globular, icosahedral and filamentous shapes (Pietila et al., 2014) (Figure 1).

Viruses also lack a universal gene (Rohwer & Edwards, 2002) such as the ribosomal RNA genes that are used to assess microbial diversity. Some genes, however, are conserved within particular taxonomic groups, as evidenced in the sequenced genomes of viral isolates; thus, their sequences are similar enough at the nucleotide level to facilitate taxonomic identifications (Mokili et al., 2012). Taxonomically, viruses are also classified by the nature of their nucleic acids following the Baltimore classification (Baltimore, 1971).

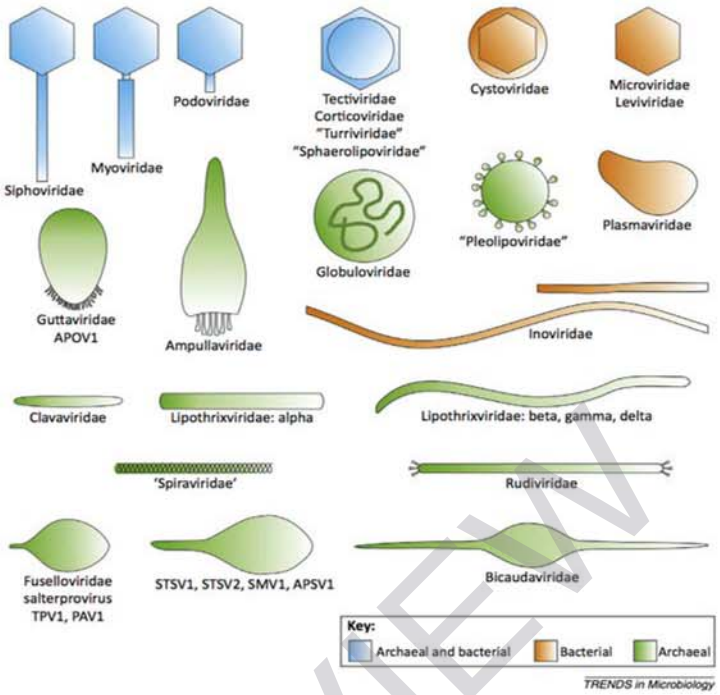

Figure 1. Diversity of virion morphotypes of prokaryotic viral structures generated by novel proteins. Virions are not drawn to scale. Abbreviations: APOV1, Aeropyrum pernix ovoid virus 1; APSV1, Aeropyrum pernix spindle-shaped virus 1; PAV1, Pyrococcus abyssi virus 1; SMV1, Sulfolobus monocaudavirus 1; STSV1, Sulfolobus tengchongensis spindle-shaped virus 1; STSV2, Sulfolobus tengchongensis spindle-shaped virus 2; TPV1, Thermococcus prieurii virus 1. Illustration taken from Pietila et al. 2014.

This classification system organizes viruses into one of seven groups depending on a combination of their genetic material (DNA or RNA), strandedness (single-stranded or double-stranded), sense (positive or negative), and replication approach.

Order "Megavirales" a breakthrough in the virology world

Generally, viral genomes are small compared to those of cellular organisms. In recent years however, the discovery of several groups of giant viruses has

dramatically changed this paradigm (Abergel et al., 2015; Koonin et al., 2015; Van Etten & Dunigan, 2012). Currently, viral genome sizes range from about 2 kilobases (kb) to more than 2.5 megabases (Mb). This expansion blurs the differences between cells and viruses in terms of genome size and complexity. The genomes of some giant viruses are even larger than numerous bacteria, archaea, and a few eukaryotic organisms (Koonin et al., 2015; Koonin et al., 2015a).

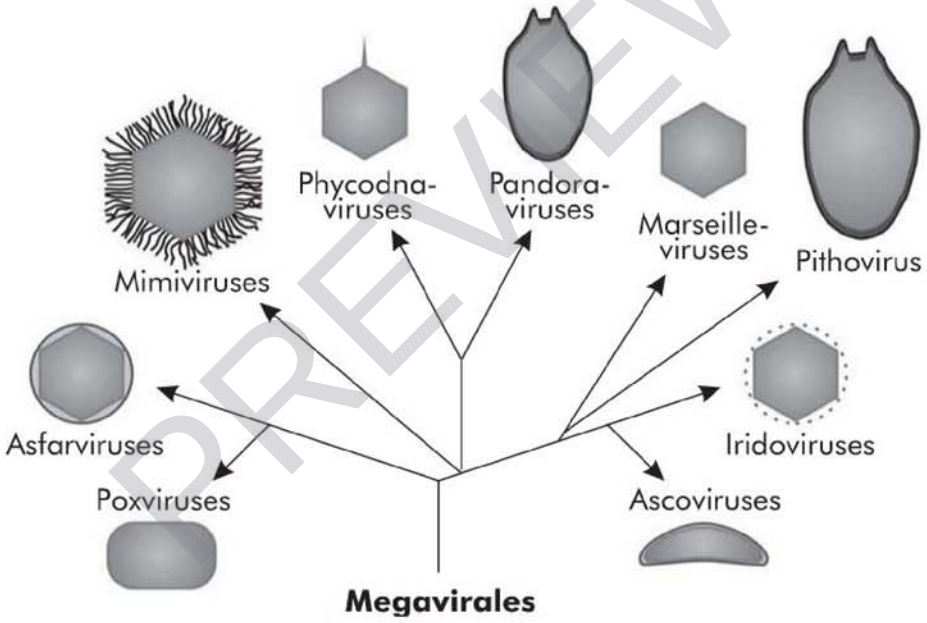

Figure 2. Nine viral families within the Megavirales order. Illustration taken from Koonin, Dolja and Krupovic 2015 with modifications.

The order "Megavirales" unites diverse families of giant viruses. The genome size of members of this group ranges from 100-kb to 2.5-Mb. Viruses in this order are believe to have a monophyletic lineage based on evolutionary genomic analysis, and they include nine large dsDNA virus families: Phycodnaviridae,

Poxviridae, Asfarviridae, Iridoviridae, Ascoviridae, Mimiviridae, Marseillevirus, Pandoravirus, and Pithovirus (Chen & Suttle, 1996; Koonin et al., 2015) (Figure 2).

Collectively these viruses are designated as nucleo-cytoplasmic large dsDNA viruses (NCLDV). They infect animals and diverse unicellular eukaryotes, and they replicate either exclusively in the cytoplasm of the host cells or possess both cytoplasmic and nuclear stages in their life cycle (Koonin et al., 2015; Van Etten et al., 2010). Intriguingly, NCLDVs have not yet been reported in any higher plants.

Phycodnaviruses are cosmopolitan in marine and freshwater environments

The Phycodnaviridae family represents icosahedral dsDNA viruses that infect marine and freshwater eukaryotic algae. Phycodnaviruses are key elements in aquatic ecosystems with important roles in the regulation of algal microbial habitats such as communities of red and brown algae (Coll et al., 2010; Kaiser, 2000; Short, 2012).

The family Phycodnaviridae is divided into six genera based on their host range: *Chlorovirus, Coccolithovirus, Phaeovirus, Prasinovirus, Prymnesiovirus, and Raphidovirus*. These divisions are supported not only by phylogenetic analysis, but also by sequence identity and structural conformation of their major capsid proteins. Their genomes range in size from 100- to 550-kb (Larsen et al., 2008; Van Etten et al., 2002).

Genus Chlorovirus

Members of the genus *Chlorovirus*, are ubiquitous in nature and have been isolated from inland waters collected throughout the world (Yamada, et al., 2006) including North and South America, Europe, Asia and Australia (Cho et al., 2002; Short & Short, 2009; Van Etten et al.,1985a; Van Etten, et al.,1985b; Zhang et al., 1988) (Figure 3). Chloroviruses infect certain unicellular, eukaryotic, ex-symbiotic chlorella-like green algae, often referred to as zoochlorellae (Meints, et al.,1984; Reisser et al.,1991).

Typically, chlorovirus titers in native waters fluctuate between 1-100 plaque-forming units (PFU) per ml; however, titers as high as 100,000 PFU/ml of indigenous water have been observed. Titers fluctuate with the seasons, with the highest titers occurring in the spring; however, the mechanism(s) in nature that allows long-term chlorovirus persistence and distribution in freshwater is still unknown (Cho, et al.,2002; Reisser, et al.,1988; Van Etten, 1995; Yamada, et al.,1991; Yamada, et al.,1993; Zhang, et al.,1988).

Common endosymbiosis between zoochlorellae and protist species

Green algae are some of the most abundant and ancient organisms on the planet. They have emerged as significant contributors in global energy and biogeochemical recycling (Grossman, 2005). Algae form a group of diverse photosynthetic organisms, ranging from multicellular to single-celled genera such as *Chlorella* (Proschold et al., 2011). Members in the genus *Chlorella* (Phylum

Chlorophyta) are small (2 to 10 μm in diameter), coccoid, nonmotile, unicellular green algae with a rigid cell wall and a single chloroplast, that exist as one of the most widely distributed algae in freshwater throughout the world. They reproduce by mitotic division in a simple developmental cell cycle. Vegetative cells increase in size and divide into two, four, eight, or more progeny depending on the species and environmental conditions. The progeny is then released by rupture or enzymatic digestion of the parental walls (Shihira & Krauss, 1965; Van Etten & Meints, 1999). Although most *Chlorella* species are free-living, *C. variabilis* is a species that exists as an endosymbiont of ciliated protozoan *P. bursaria* in nature (Table 1). They are often referred to as ex-symbiotic chlorellae or zoochlorellae (Proschold et al., 2011, Jolley & Smith, 1978, Siegel, 1960) (Figure 3).

Chlorella variabilis Algal strain	*P. bursaria* strain	*P. bursaria* collection site
SAG 211-6		USA
ATCC 50258/CCAP 211/84 (NC64A)		North Carolina, USA
ATCC 30562 (Syngen 2-3)		Ohio, USA
N-1-A		USA
NIES-2541 (OK1-ZK)	OK1	Aichi, Japan
So13-ZK	So13	Nagano, Japan
NIES-2540 (F36-ZK)	F36	(cross breed, Japan-Japan)
KM2-ZK/pbKM2	KM2	Shimane, Japan
Dd1-ZK	Dd1	Ibaraki, Japan
Bnd1-ZK	Bnd1	Hiroshima, Japan
HB2-2-1	HB2-2	Hiroshima, Japan
shiP-7-A4	shiP-7	Miyazaki, Japan
takaP-3-A2	takaP-3	Oita, Japan
(uncultured)†	Cs2	Shanghai, China
(uncultured)†	MRBG1	Melbourne, Australia

Table 1. List of names and collection site of the *Chlorella variabilis* algae strains isolated from their respective *P. bursaria* hosts. Table taken from Fujishima et al., 2010 with modifications.

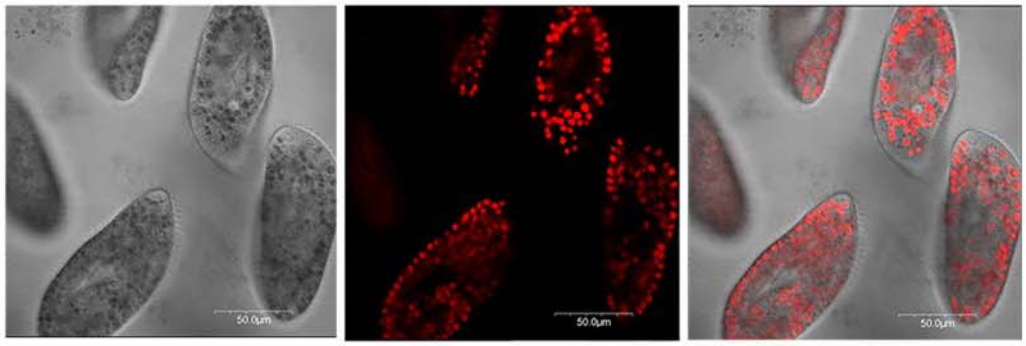

Figure 3. *Paramecium bursaria* shown in white light, ultraviolet light, and merged lights highlighting the red-fluorescing chlorophyll of the green algae housed within the symbiont.

C. variabilis is extremely variable in response and sensitive to small differences in culture conditions, thus its name *variabilis* (Shihira & Krauss, 1965).

In *P. bursaria*, hereditarily intracellular zoochlorellae inhabit the gastrodermal symbiosomes (perialgal vacuoles) of the protist and transfer an important amount of their photosynthetically fixed carbon (e.g. maltose, fructose) and amino acids to the non-photosynthetic partner (Cernichiari et al., 1969; Fujishima, 2010; Karakashian, 1975; Matzke, et al.,1990; Ziesenisz, et al.,1981) (Figure 3). Additionally, ex-symbiotic algae produce three times more oxygen than their free-living counterparts at low light intensity rates (Cronkite & van den Brink, 1981). Probably, high rate oxygen release in low light intensities is a special adaptive feature stemming from endosymbiotic interactions. Some individuals also differ in their uptake of nutrients. It has been suggested that ex-symbionts posses an efficient system to import and metabolize many organic nitrogen sources, while they can not utilize inorganic components, such as nitrate or nitrite, as their only