HAP2-MEDIATED CELL FUSION IN A SEXUAL CILIATE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jennifer Fricke Pinello

January 2017

HAP2-MEDIATED CELL FUSION IN A SEXUAL CILIATE

Jennifer Fricke Pinello, Ph. D.
Cornell University 2017

A crucial step in the life cycle of eukaryotes is sexual reproduction. For productive sex to occur, the first cellular steps that must be taken are gamete recognition, adhesion, and membrane fusion. However, little is known about the proteins governing these initial events. Helpful in this regard, are studies on HAP2/GCS1, an ancient and highly conserved gamete membrane protein necessary for successful fertilization in a wide range of eukaryotic species. Efforts have already recognized the necessity of HAP2-mediated membrane fusion in the male gametes of multiple species, but they have come short of describing its mechanism of fusion, structure, or potential origins. Here, we investigated the expression, genetic necessity, and fusogenic capacity of HAP2 during the sexual cycle of the model single-celled ciliate species, *Tetrahymena thermophila*. Unlike in other species, we found that HAP2 is expressed in all seven mating types of *T. thermophila*, and fusion and fertility is only prevented when HAP2 is deleted from both cells of a mating pair. Along the way, we also functionally examined a hypothetical HAP2 accessory protein, ZFR1, probed the life cycle of a ciliated fish parasite for the presence of a conjugation stage, and accidentally found instances of unexpected parthenogenesis in *T. thermophila*. Perhaps most importantly, we developed a flow-cytometry-based assay for quantifying sexual cell fusion events in an *in vivo* model of eukaryotic fertilization, identified a predicted high-confidence structural homology between HAP2 and class II viral fusogens, and tested the functional extent of this homology through targeted deletions and biophysical tests of the HAP2 fusion loop. In addition to providing indications of a viral-like fusogenic mechanism for HAP2, these findings that show a structural and functional similarity between a viral and gamete fusogen harken back to early theories that a “selfish” genetic element promoting membrane fusion as a means for intergenomic transmission led to the origin of eukaryotic sex.

BIOGRAPHICAL SKETCH

The author of this work, Jennifer F. Pinello, is the oldest of three children and was born in the state of Vermont, United States of America on April 25, 1986 to parents John F. Fricke and Jacqueline M. Powers. She grew up in the small town of Northfield, Vermont where she attended Northfield High School and graduated as Salutatorian. While in High School, she attended a few courses at a local college, Norwich University. It was through a Principles of Biology Course at Norwich, that she discovered a fascination with the life sciences, and, upon later admission to the University of Vermont (UVM) on a full tuition Green and Gold Scholarship, she choose to major in Biological Sciences. The summer prior to the start of her freshman year, she had the opportunity to travel on a Freeman Foundation funded trip to China (which had been delayed from the year before due to the 2003 SARS epidemic). This experience inspired her to take Chinese language courses throughout college, study-abroad in Yunnan, China, and eventually declare a second major in Asian Studies. Infectious disease was a gradually developing interest at UVM, where a series of internships in the ecology of invasive species (first lamprey, than paper wasps) culminated in her undergraduate honors thesis work under Dr. Joseph J. Schall describing the population genetics of a malaria parasite's invasion of its lizard host. After graduating *cum laude* from UVM she moved with her future husband, Thomas Pinello, to Boston, MA for two years. There, she worked as a technician under Dr. Daniel Kavanagh at the Ragon Institute of MGH, MIT and Harvard, an AIDS research center, on the development novel vaccine technologies including characterizing murine immune responses to different formulations of an mRNA-loaded nanoparticle vaccine and assisting in a clinical dendritic-cell based therapeutic vaccine trial. In the fall of 2011 she was accepted to and began her graduate studies at Cornell University in the field of Immunology and Infectious Disease. In 2012, she married Thomas, was awarded an NSF Graduate Research Fellowship, and joined Dr. Theodore Clark's laboratory. She greatly enjoyed the many wonderful people she met and the dynamic environment of scientific discovery she experienced as a member of the Clark lab.

Dedicated to my parents, John and Jackie Fricke

ACKNOWLEDGMENTS

It is thanks to many family, friends, and mentors that I've been able to pursue this path. Thanks to my husband, Thomas Pinello, who has been a constant source of love and support through my entire time in graduate school. He was always there to offer a second opinion, watch presentations, and selflessly delayed his own career goals to allow me to pursue mine. I also owe a huge debt of gratitude to my parents, John and Jackie, and brothers, Tommy and Sean, My family always encouraged writing and public speaking practice, and taught me the importance of always trying my best. Thanks also for the support from my Fricke, Powers, and Pinello extended family members.

Thank you to my advisor Dr. Ted Clark, for believing in my ideas and giving me the freedom to pursue any research question. He found this worthwhile research project, taught me how to effectively speak and write scientifically, and always challenged me to see the "big picture". Other members of the Clark lab also helped foster my scientific development, especially Dr. Donna Cassidy-Hanley. She had the patience to train me in the finer details of ciliate genetics and I especially enjoyed discussions trouble-shooting PCR reactions or cloning issues with her and my long-term comrade in laboratory work, Daniel Kolbin. Thanks also to my other Clark lab coworkers; Kelsey Fryer, Gary Tan, Crystal Gergye, Karuna Katariwala, Morgan Kellibrew, Mozammal Hossain, Catherine Devine and visiting scholars Drs. Xueming Dan, Hong Zeng, and Liu Xiaodong.

Thanks also to my committee members Drs. Dave Holowka, Holger Sondermann, and Gary Whittaker. Dr. Holowka was always willing to help me through lending reagents or project discussions and helped review my manuscript. Dr. Sondermann, in addition to being my committee member was my professor in BIOAP6100, and taught me critical grant-writing skills and how to use PyMol. Dr. Whittaker, although a late addition to my committee, helped me immensely, generously allowing me to work in his lab for over three months to test HAP2 fusogenic-sufficiency and advising me on the expected characteristics of viral fusion proteins and their fusion peptides. I would also like to thank the professors I rotated with in my first year,

Drs. Eric Denkers and Brian Rudd. It was Dr. Denkers who primarily mentored me in NSF grant writing during my first year, and he was the first to suggest trying to CFSE label *Tetrahymena.* It was Dr. Rudd who taught me how to CFSE label T cells and analyze flow cytometry data. Thanks as well to the professors who I TA'd for Drs. Tim Huffaker, Maria Garcia-Garcia, Ken Kemphues, and Tony Bretscher. I would also like to thank many other Cornell professors, including Drs. Jim Casey, Maurine Linder, Linda Nicholson, Avery August, William Brown, and David Lin for their critical advice at different points in my graduate studies. In particular, thanks to Dr. Lin especially for all the work he has done to help BBS graduate students and initiating the Fall into Science Outreach program. Thanks as well to all the M & I Department members including Walter Iddings and Rod Getchell, all the students and post docs who were my friends; especially Kristeen Pareja, Hana Kim, Sachi Horibata, and Drs. Norah Smith, Sara Cohen, and Ruth Fahey, and my friends and mentors who offered advice or were collaborators in this project; Drs. Jean Millet, Alex Liqi Lai, Hung-lun Hsu, and Michele Bialecki.

Importantly, I would like to thank our other collaborators early in this project; Drs. Eric Cole, Mark Winey, Courtney Ozzello, Thomas Giddings Jr. and Jack Freed. I would also like to thank all those within the ciliate community that offered advice, especially Drs. Wei-Jen Chang, Aaron Turkewitz, Laura Landweber, Marcella Cervantes, and Ed Orias, and all those in membrane fusion research with whom I had helpful discussions, especially Drs. William Snell, Mark Johnson, Jennifer Forcina, Felix Rey, Juliette Fedry, and Thomas Krey. Thanks also to Dr. Garegin Papoian, Hao Wu, and Haiqing Zhao who offered advice on protein structure modeling.

Finally I would like to thank other key mentors I've had in my prior scientific training. My supervisor at the Ragon Institute, Dr. Daniel Kavanagh, taught me so much practical immunology and encouraged me to attend conferences and participate in scientific publishing. A big thanks to my undergraduate thesis advisor, Dr. Joseph Schall who gave me a passion for research & the ecology and evolution of parasites. Thanks also to my mentors Drs. Anne Vardo-Zalik, Ellen Marsden, Kurt Pickett, Jane Molofsky, my Chinese professors, Drs. Diana Sun & John Yin, and to my biology professors at Norwich; Dr. Lauren Howard and Virginia Kunkel.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter one

Introduction and literature review[1]

[1] In this Chapter, several figures were minimally adapted from previously published works and licensed for use here. These include: Figure 1.1, reprinted from Current Biology, 24, Eric Cole et al., Function of the Male-Gamete Specific Fusion Protein HAP2 in a Seven-Sexed Ciliate, 2168-2173, ©2014, with permission from Elsevier; Figure 1.2, reprinted from Trends in Cell Biology, 20, Wong, J. L. and Johnson, M. A., Is HAP2-GCS1 an ancestral gamete fusogen?, 134-141, ©2010, with permission from Elsevier; Figure 1.4A, adapted by permission from Macmillan Publishers Ltd: Nature Reviews, Microbiology, ©2006; and Figure 1.4B, reprinted from Cell, 157, Perez-Vargas, J. et al., Structural basis of eukaryotic cell-cell fusion, 407-419, ©2014, with permission from Elsevier.

A. Relevant Background

A1. Defining the nature and problem of eukaryotic sex

A common feature of eukaryotic life is the presence of sexual reproduction[1,2]. Despite the widespread nature of sex, it is surprising that there is not a consensus definition for this important biological act[3]. While many may define sexual reproduction as the combination of genetic material from two parents to form offspring[4], this definition is overly broad to define the modern biological imperatives of eukaryotic sex. Under this definition, the transfer of plasmid DNA during bacterial conjugation could even be considered a parasexual act[5]. Modern eukaryotic sex has advanced beyond the mere exchange of a few strands of DNA, and involves orchestrated genomic and cellular acts, precisely executed for the proper development of progeny cells. It is thus more appropriate to include in any modern **eukaryotic definition of sex, the presence of meiosis, cell-cell fusion, and nuclear fusion in generating a genetically distinct and epigenetically reprogrammed progeny**. Defining sex in this way then necessarily forces the definition of asexuality to encompass manifestations ranging from mitotic clonal cellular division, to versions of parthenogenesis that can take many forms (see Chapter 6), which occasionally incorporate some, but never all of these sexual characteristics.

The consequence of sexual reproduction on the course of eukaryotic evolution is so obvious it's often overlooked. The common transition of eukaryotic life between haploid and diploid life stages, is generally associated with phenotypic changes in an organism's outward appearance spurred either from development after gamete cell fusion to restore diploidy, or meiotic reductions to attain haploidy[6]. In different species, the multicellularity of haploid and diploid states differ, such that some organisms adopt haploid dominant stages as their multicellular or clonal stage (e.g. *Chlamydomonas, Schizosaccharomyces,* etc.), diploid dominant stages (Humans, brown algae, *Tetrahymena,* etc.) or "diplohaplontic" stages, where both diploid and haploid forms mitotically divide and often form multicellular bodies (e.g. some flowering plants, red alga *Mastocarpus*). Breeding restrictions among species, whether

geographical or otherwise[7], then compound this observable diversity, generating a variety of life persisting in accordance with both an individual organism's environmental success and the transmission capacity it's genetic information. All this complexity adds an intrinsic beauty, but also a certain degree of wonder to our perception of the natural world.

The ubiquity of eukaryotic sex, along with the presence of each of these highly-adapted alternative life cycles as a result of biparental sex presents a great paradox to evolutionary biology: Why has it been so beneficial for individuals to fuse with one another and combine genetic material rather than reproducing asexually [1,4,8–10]? While many theories have been proposed to address this question [1,11–15], supporting evidence is scant. In the work presented in this dissertation we fell inadvertently into this problem while studying the sexual cell-cell fusion events mediated by an ancient gamete fusion protein in the free-living single celled ciliate, *Tetrahymena thermophila*. In the first part of this introduction the important potential applications of basic knowledge gained from gamete interaction and fusion studies will be addressed, along with an introduction to *Tetrahymena thermophila's* life cycle. Previous research on the gamete fusion protein HAP2/GCS1 will then be reviewed followed by an introduction to membrane fusogens.

A2. Importance and applications of gamete fusion studies

A crucial step in the life cycle of eukaryotes is sexual reproduction. For sex to occur, the first cellular steps that must be taken are gamete recognition, adhesion, and membrane fusion. Although little is known about the proteins governing these events, initial research indicates that the proteins involved in gamete recognition and adhesion vary tremendously from species to species[16], while a protein called HAP2/GCS1 is remarkably conserved in the function of gamete membrane fusion throughout eukaryotic life[17,18]. Importantly, this suggests that many, if not all eukaryotes share a common mechanism of gamete fusion[19]. Thus, one could expect that more basic scientific knowledge on HAP2/GCS1 might lead to better treatments for infertility, better birth control methods, improved agricultural yields, or even enhanced methods

to block the sexual stages of certain parasites.

The causative agents of many major human and agricultural parasitic diseases including *Toxoplasma, Plasmodium, Trypanosoma, Theileria, Babesia, Neospora, Leishmania, Cryptosporidium*, *Eimeria etc.* have all been shown to contain HAP2/GCS1 orthologs[20–22]. Since these diseases cause significant health and economic burden worldwide, a high importance is placed on the development of vaccines and therapeutics against these agents[23,24]. Since a sexual stage is integral to the successful transmission of many of these parasites, vaccines that block HAP2-mediated fusion should block the spread of these devastating diseases. So far, *HAP2* has been shown necessary for the fertility of male gametocytes of the malaria parasites *Plasmodium berghei* and *Plasmodium falciparum*[25,26], and antibodies generated to target the *Plasmodium berghei* HAP2 ortholog have been shown to block the development of infective oocysts in the mosquito vector and block vector-borne transmission for this mouse malaria parasite[27,28]. Identification of the key epitopes important for neutralization of fusogenic activity is important to future development of vaccines utilizing HAP2 as a target.

In the same vein, HAP2/GCS1 was first identified as essential for the double fertilization events of sperm in angiosperms[29–31], and methods which control or enhance the reproduction of agricultural plant species could be useful for plant reproduction research and/or increasing crop production. It is also conceivable that companies specializing in the production of seed varieties would wish to protect their intellectual property or have new methods to prevent the unintended cross-fertilization of F1 hybrids. To this end, certain techniques using diphtheria-toxin A expressed from the *HAP2* locus of *Arabidopsis* sperm[32], or the heat inducible removal of the *HAP2/GCS1* allele for generation of homozygous mutants[33] have been developed.

Finally, although a *HAP2/GCS1* ortholog is not present in vertebrate species, basic research into the function of such a conserved gamete fusogen could lead to future insights into

how human gamete membrane fusion is accomplished. There are only a limited number of known protein structures used to mediate membrane fusion events, and it's possible structural similarities still exist between vertebrate gamete fusogens and HAP2, despite an underlying lack of sequence homology (see Chapter seven). Because the essential features of sex, namely, meiosis, cell-cell fusion, and nuclear fusion are conserved throughout many eukaryotes[1], the study of these events in simple model systems with easily inducible sexual cycles and access to gamete material is relatively straighforward[25]. Furthermore, since fertilization and gamete membrane fusion are intrinsically very difficult to study in many species, especially humans, the development of such model systems should promote rapid advances in our understanding of and ability to control these essential sexual events.

A3. Tetrahymena sexual conjugation

The model organism used in these studies was the single-celled, freshwater dwelling, ciliated protist *Tetrahymena thermophila.* This organism is genetically amenable and can easily be induced to sexually conjugate semi-synchronously in mass cultures, making it ideal for these studies. Briefly, sex in *Tetrahymena* is a process in which two complementary mating types fuse temporarily to mutually exchange haploid meiotic products across a specialized region of membrane called the conjugation junction, in essentially an act of double cross-fertilization. Interestingly, at this membrane interface, the mating cells form not one, but hundreds of fusion pores which do not expand, but are instead repaired at the conclusion of mating, to allow for the separation of the individual progeny cells. In this section, the events of this ciliate's life cycle and the cellular and nuclear maneuvers associates with its sexual conjugation stage are reviewed.

To understand sexual reproduction in *Tetrahymena,* it is first necessary to grasp some fundamental features of these exceptional creatures. Ciliated protozoa have a long and storied past as a cellular models dating back to the initial observations of Robert Hooke and Anton van Leeuwenhoek in the 17th century[34]. Among the nearly 7200 characterized ciliate species, is

Tetrahymena thermophila, a holotrichous ciliate with cilia that uniformly cover its whole body that it uses for both swimming and ingesting food[35]. *Tetrahymena* has been domesticated for axenic cultivation in the laboratory environment on bacterized or artificial media, but can also easily be found in the wild, growing in freshwater ponds and lakes during summer months [36,37]. The pyriform-shaped cells themselves are large, approximately 30-50μm in length, and have precisely positioned cellular structures including a posterior cytoproct, contractile vacuoles and an anterior mouth containing the cell's four eponymous ciliary membranelles[36].

One of the most salient features of all ciliates is nuclear dimorphism, which refers to the maintenance of separate germline and somatic genomes within a single cell. These genomes are present as two genetically, physically, and functionally distinct nuclei within the cell. The micronucleus (Mic), contains the transcriptionally silent germline genome, and the macronucleus (Mac) contains the transcriptionally active, somatic genome. The diploid Mic contains 5 pairs of chromosomes and is 220Mb in size, while the Mac is polyploid, contains approximately 200 linear minichromosomes amplified 45 fold, and is 105Mb[36]. At each round of sexual reproduction the zygotic Mic product is used as template for the generation of a new Mac and Mic in all progeny cells. The discrepancy in chromosome number between the two genomes arises from the developmental fragmentation of each zygotic Mic chromosome at sequence-specific chromosome breakage sites[36]. The discrepancy in size between the two genomes comes from the fact that during development, approximately 30% of the Mic genome is deleted in an RNAi-mediated epigenetic remodeling event removing an estimated 6000 transposable-element-rich sequences called "IESs" from the progeny Mac[38–41].

During asexual vegetative growth, *T. thermophila* expands rapidly and to high cellular densities (~1.5 x 10^6 cells/mL), with one division via binary fission every 2-3 hrs under optimal conditions. During asexual growth, the diploid Mic divides normally through mitosis and the Mac divides unevenly through a still incompletely understood process called amitosis[42]. Despite uncertainty about the mechanisms of amitotic division, the unequal dispersal of

minichromosomes to daughter cells, called phenotypic assortment, is a useful tool in selecting cell lines that are homozygous in their polyploid Mac for an introduced transgene encoding a trait of interest[43]. Phenotypic assortment and high ploidy levels in the Mac are also posited to be a natural means for asexually reproducing populations of *Tetrahymena* to remain highly-resistant to the expected deleterious accumulation of mutations known as Muller's ratchet[44].

There are three notable peculiarities to ciliate sexual conjugation that must be discussed to prevent any confusion before proceeding to the timeline of mating events. First, ciliates do not make individual gamete cells. Instead, they make haploid gametic-like pronuclei which are exchanged between cellular mating partners[45]. In comparison with other mating systems, *Tetrahymena* would be considered an example of both isogamy and hologamy, which means that individual cells housing these gametic nuclei are both morphologically indistinguishable from each other and their vegetatively reproducing counterparts, respectively. Second, unlike the majority of sexual life with its bipolar female and male mating types, *Tetrahymena thermophila* has seven different sexes[46]. In this system, any mating type (I-VII) may conjugate with any other mating types except its own. This system does allow for triplets of cells to adhere and form progeny when three different mating types of *Tetrahymena* are mixed, but in these cases the gametic pronuclei are not exchanged bidirectionally [35]. Third, mating *Tetrahymena* cell pairs do not permanently fuse together. Rather, sexual cell-cell fusion occurs in a transient and spatially limited way at the conjugation junction, a fenestrated zone of adhered membranes between the two mating partners containing 100-200 individual, non-expanding fusion pores that are 0.1 - 0.45 μm in diameter [47–50]. These pores allow the selective passage of protein, RNA, and migratory pronuclei during mating[51], but at the conclusion of mating these membrane pores are repaired, allowing what were originally the two parental cells to separate into two individual maturing progeny cells.

The timeline of sexual conjugation in ciliates is a 12 hr program (at 30^{o} C) divided into the distinct stages, with the initial stages classified mainly by changes in cell morphology, and

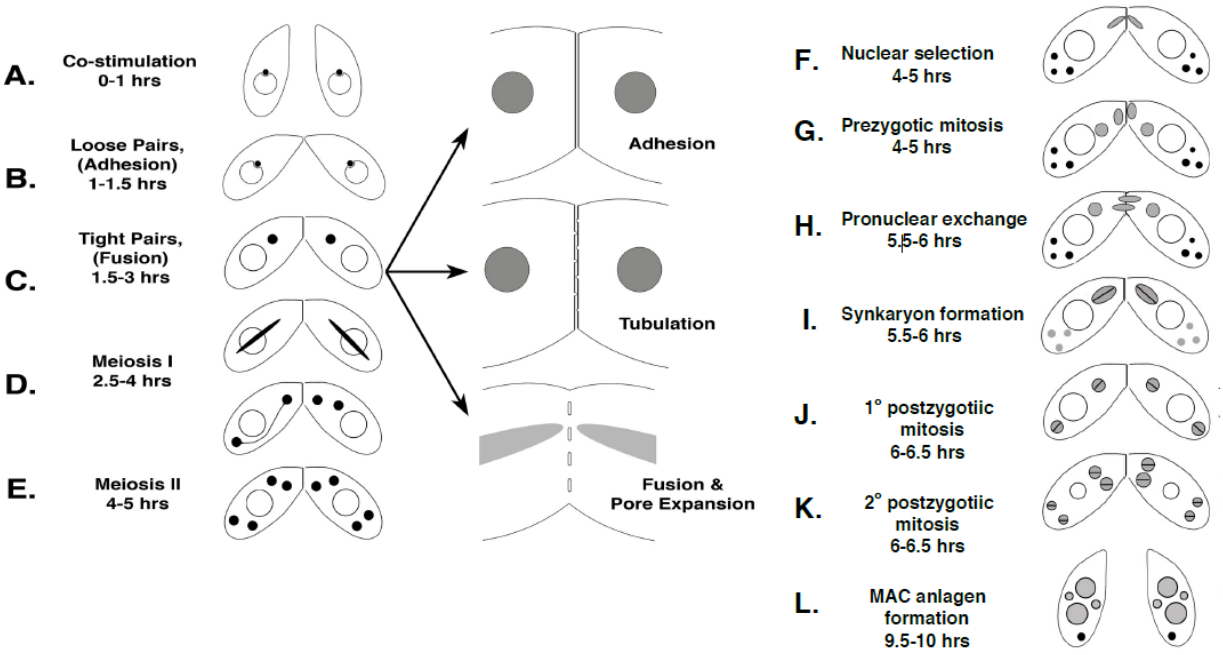
the later stages by changes in nuclear morphology and number (Figure 1.1). The first stage is called initiation and is induced by a minimum of two hours of nutritional starvation at 30^{o} C [52]. An important task of initiation is to allow cells at different stages of their cell cycle from their time in vegetative culture, to arrest at G1 stage[53]. Under prolonged periods of starvation, arrested cells become increasingly slender in shape and occasionally develop to a highly motile, fast-dispersal form[54]. Any nutrient sources reintroduced to the culture at this time will block the initiation process[48]. After initiation, mixed cells of different mating types require one hour of co-stimulation with their partner before forming pairs[47].

Co-stimulation is a cellular recognition event accompanied by morphological changes when two cells of opposite mating type are mixed. While contact-dependent interactions between cell surface bound pheromones called gamones are thought to trigger this event, there have been some reports that *Tetrahymena* may also secrete a substance to aid in this recognition event [48,55,56]. The protein and/or chemical identities of these surface-bound gamones and secreted factor(s) have yet to be elucidated. What is known is that co-stimulation is accompanied by increases in transcription, protein synthesis, and a wave of tyrosine phosphorylation that both activates cells and matures them for future adhesion events[57–59]. Co-stimulation can be halted by physical agitation (such as shaking) up to 90 min after mixing [48]. During co-stimulation, the cells undergo a subtle change in appearance, slightly rounding in shape and remodeling the zone of membrane anterior to their mouth apparatus to form a flat, deciliated site expressing a high density of concanavalin A (ConA) binding-proteins[48,60–62]. Because addition of ConA or tunicamycin (a glycosylation blocking drug) to the cell mixture at this point inhibits pair formation, it is thought that the glycosylated protein receptor(s) that bind to ConA are responsible for cellular adhesion events[61,63]. Current efforts are underway to identify these proteins (Cole, personal comm.). Together, these alterations to the anterior membrane - referred to as “tip transformation” - create the future membrane surface for adhesion between mating partners.

Cell-cell adhesion events start approximately 1 hr after the mixing opposite mating types at 30^{o} C. Mating pairs are at first weakly bound at their anterior tips and can fall apart easily with agitation or feeding, but then develop into firm pairs resistant to such perturbations around 2 hrs after mixing. The progression to firm pairs is expected to be aided by cell-cell fusion events tethering cells together at their conjugation junction[58,64]. The ultrastructure of the *Tetrahymena* conjugation junction has been characterized in detail by electron microscopic studies of pairs fixed at different time points during their fusion pore formation. At first, cellular membranes are closely apposed at a distance of ~50nm, with apposed bilayers having electron dense submembranous scaffolding that appears to fade slightly at distinct sites. Regions of reduced electron density then form membrane "dimples" or outpocketings from either partner bilayer into the intercellular cleft. The space between these outpocketings then fills with electron dense material before resolving into individual fusion pores[50,65]. The composition of phospholipids at the conjugation junction has also been shown to be different than in the rest of the cellular membrane, containing a high concentration of 2-aminoethyl phosphonolipid, a *Tetrahymena*-specific cone-shaped lipid with properties similar to phospholipid ethanolamine that enables high-curvature of membrane structures, such as fusion pores [66].

During the period of adhesion and cytoplasmic continuity, the series of germline nuclear events that accompany ciliate sex begin and are outlined in detail in Figure 1[64]. These nuclear events last from 1.5 – 10 hr after mixing of opposite mating types and include all events from meiosis through the development of the zygotic Mac and Mic within progeny cells.

Figure 1.1 Diagram of the conjugation pathway in *Tetrahymena thermophila*. Panels A-E illustrate the early stages of pre-pairing and pair formation. **(A)** When starved cells of complementary mating types are mixed (time = zero hrs; T = 30^{o} C), a period of co-stimulation leads to modification of the anterior end of the cell, creating an adhesion plaque. **(B)** Between 1 and 1.5 hrs, cells begin to form loose pairs. **(C)** Between 1.5 and 3 hrs (depending on the density and synchrony of the culture), loose pairs transform into "tight pairs" due, at least in part, to the formation of membrane protrusions into the junction cleft that ultimately fuse with the plasma membrane of the mating partner, leading to the formation of several hundred intercellular pores. **(D)** At the onset of Meiosis I, pores are already well established, and undergoing a process of pore expansion from 2.5-4 hrs. **(E)** Pore expansion continues until at least 4.5 hours by which time Meiosis II is underway. This fenestrated junction is further modified into a "curtain" of membrane tubules, as the pore-expansion fronts meet and collide. Panels F-L illustrate the events leading up to and following exchange of migratory pronuclei. **(F)** After Meiosis II, one haploid product (gray) is selected and becomes anchored at the nuclear exchange junction. **(G)** This product undergoes mitosis and one pronucleus from each partner docks at the well-perforated exchange junction, while the other resides some distance away (pronuclei are depicted in gray). Meanwhile, the other 3 haploid meiotic products begin to degenerate (shown in black at cell posterior). **(H)** Migratory pronuclei are then exchanged through the now fenestrated membrane junction. **(I)** As migratory pronuclei fuse with the stationary pronuclei on each side of the junction to form the zygotic synkarya they both adopt elongate shapes. **(J)** The first postzygotic mitosis creates two diploid nuclear products in each cell. **(K)** These divide again forming four diploid nuclear products. In both (J) and (K) the nuclei are depicted with a line bisecting them to emphasize their diploid, recombinant nature. The anterior-most nuclei undergo endo-reduplication of their DNA thereby becoming polyploid macronuclei while the posterior two nuclei remain diploid as micronuclei. **(L)** Finally, the cells separate as the parental macronucleus moves to the posterior and is resorbed. This figure was Reprinted from Current Biology, 24, Eric Cole et al., Function of the Male-Gamete Specific Fusion Protein HAP2 in a Seven-Sexed Ciliate, 2168-2173, Copyright (2014), with permission from Elsevier [64].


A.
Co-stimulation
0-1 hrs
B.
Loose Pairs,
(Adhesion)
1-1.5 hrs
C.
Tight Pairs,
(Fusion)
1.5-3 hrs
D.
Meiosis I
2.5-4 hrs
E.
Meiosis II
4-5 hrs
Adhesion
Tubulation
Fusion &
Pore Expansion
F.
Nuclear selection
4-5 hrs
G.
Prezygotic mitosis
4-5 hrs
H.
Pronuclear exchange
5.5-6 hrs
I.
Synkaryon formation
5.5-6 hrs
J.
1° postzygotiic
mitosis
6-6.5 hrs
K.
2° postzygotiic
mitosis
6-6.5 hrs
L.
MAC anlagen
formation
9.5-10 hrs

All nuclear events happen in each of the two mating partners contemporaneously. First, the germline Mic of each parental cell undergoes two rounds of meiosis creating four haploid nuclear products. One of these products is "selected" and divides mitotically to generate a stationary and migratory pronuclei, each of which are destined to be involved in the fertilization event. The migratory pronucleus of each mating partner is then propelled across the conjugation junction by a microtubular basket and fuses with the stationary pronucleus of its partner cell[67]. This fertilization nucleus then undergoes two post-zygotic divisions, with the specific cytological arrangement of the four products of division establishing their individual developmental fates. The anterior most nuclei in each cell are destined to become the new macronuclei, and start the process of endoreduplication and epigenetic remodeling. One of the posterior most nuclei is destined to become the future germline Mic, while the other one is eliminated. During this time, the old parental Mac is also destroyed. When cells separate (~10-12 hr after mixing) they are called "exconjugants" and do not divide again until they are refed[68]. When refed, the first division of the exconjugants gives one Mac to each of the daughter cells (along with one Mic produced by mitosis of the single remaining Mic) to yield the four "karyonoids" which transition into clonal, vegetative growth. It is within the initial fission of these karyonoids that mating type is established[69], but "young" cells require a period of clonal immaturity (~40-100 fissions) before they may again sexually conjugate[44]. Thus, each individual parental *Tetrahymena* cell is capable of all the tasks required to generate a future offspring during sexual reproduction.

B. The primordial gamete fusion protein HAP2/GCS1

B1. Identification and characterization.

The protein HAP2/GCS1 (*Hapless 2 / Generative Cell Specific 1*) has now been identified in a wide range of taxa extending from single-celled protozoa to flowering plants and arthropods as a male gamete fusion factor necessary for fertilization (Figure 1.2) [19,22]. It was first discovered in the model plant, *Arabidopsis thaliana,* by Dr. Mark Johnson in 2004 in a

forward genetic screen for haploid-disrupting mutants causing defects in pollen tube growth, guidance, or pollen grain development [29]. Meanwhile, in an independent study, Dr. Toshiyuki Mori noticed the same gene due to it's high expression levels in late stage pollen development of the Trumpet Lily, *Lilium longiflorum,* and named the gene and its corresponding protein product GCS1[30]. In a functional study of this protein in *Arabidopsis thaliana,* Mori and co-workers found that male HAP2/GCS1 mutants were completely infertile, and they were also the first to recognize the presence of the gene in many taxa (including *Plasmodium*) and suggest it functioned in either adhesion or fusion[30]. Important work that really solidified the potential role of this protein as a membrane fusogen came from Dr. William Snell's lab where it was shown that adhesion of (+) and (–) gametes of the biflagellated single celled algae, *Chlamydomonas reinhardtii,* was unaffected by HAP2 deletion[25]. A HAP2-induced post-adhesion block to fertilization strongly implied a role for the protein specifically in membrane fusion. This conclusion was supported in more recent studies with other species including plants[30,31], green algae[25,70], cnidarians[71,72], insects[73], amoebae[74], and protozoa (including *Plasmodium*)[25,26,64] showing that HAP2/GCS1 is exclusively expressed in male gamete cells and necessary for male gamete fertility. Much of this subsequent research was focused on identifying key domains in HAP2/GCS1 necessary for protein function. In brief, these studies found that deletions or substitutions of regions within the protein's amino terminus disrupted function, and that certain poly-basic and potentially palmitoylated cysteine residues in the cytosolic domains are important as well (see Chapters 3 & 5)[20,75,76]. Significant interest was also engendered with regard to HAP2's potential use as a transmission-blocking malarial vaccine candidate[26,28,77].

B2. Regulation.

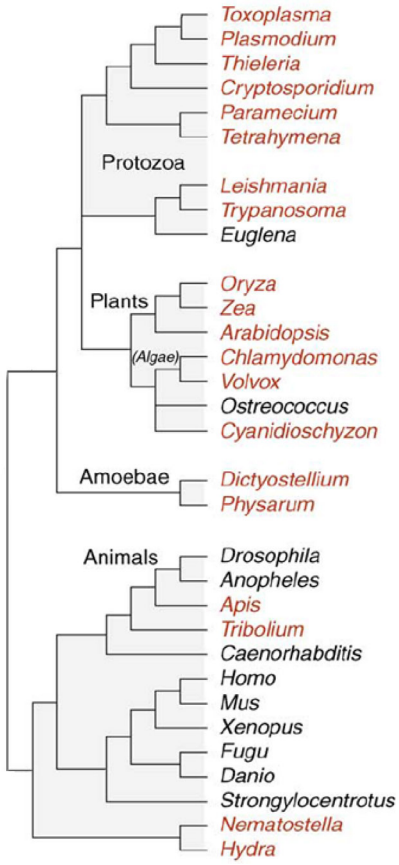
It has been noted that regulation of HAP2/GCS1 is tightly controlled at both the level of mRNA transcription and protein expression in mating cells[78]. In many species, transcription is limited to the sexual life cycle stage and specifically to male gametes[30,47]. In some species,

Figure 1.2 HAP2 protein domains and distribution of HAP2 orthologs. (A) A model of the typical arrangement of protein domains in a standard HAP2 ortholog. The 742 amino acid long *Tetrahymena thermophila* HAP2 protein product is shown for this example with the signal peptide (SP) shown in red, the HAP2 domain in orange, the transmembrane domain (TM) in yellow, and the poly-basic domain (+) in green. **(B)** A phylogenetic tree representation of the diverse genera containing a HAP2 ortholog (red) and genera that do not have a HAP2 ortholog (black). Panel (B) was reprinted from Trends in Cell Biology, 20, Wong, J. L. and Johnson, M. A., Is HAP2-GCS1 an ancestral gamete fusogen?, 134-141, Copyright (2010), with permission from Elsevier[19].

A


SP
HAP2
TM
+
1
19
281
329
545
596
742

B


Toxoplasma
Plasmodium
Thieleria
Cryptosporidium
Paramecium
Tetrahymena
Protozoa
Leishmania
Trypanosoma
Euglena
Oryza
Zea
Plants
Arabidopsis
(Algae)
Chlamydomonas
Volvox
Ostreococcus
Cyanidioschyzon
Amoebae
Dictyostellium
Physarum
Animals
Drosophila
Anopheles
Apis
Tribolium
Caenorhabditis
Homo
Mus
Xenopus
Fugu
Danio
Strongylocentrotus
Nematostella
Hydra

additional levels of protein control have also been observed. For example, in *Arabidopsis thaliana,* the HAP2 protein only translocates to the plasma membrane after the sperm cell comes within vicinity of the egg, specifically the egg cell-secreted protein EC1 [79]. The mechanisms of this specific membrane translocation event are still unknown. In *Chlamydomonas reinhardtii*, the HAP2 protein is rapidly degraded on the cell membrane after gamete fusion [78,80]. Lastly, while the volvocine alga, *Gonium pectorale*, expresses both HAP2/GCS1 transcript and protein in female gametes, the protein is specifically degraded in females and never reaches the plasma membrane[70]. These additional layers of HAP2/GCS1 regulation are not so surprising, as a misappropriately expressed gamete fusogen could result in polyspermy, negatively impacting an organism's fecundity[78].

B3. Features of a bona-fide fusogen.

What evidence is necessary to show that HAP2/GCS1 is a bona-fide fusogen? According to a review of viral and developmental fusogens published by Sapir et al. in 2008, a candidate protein is only elevated to the status of bona-fide fusogen upon fulfilling three criteria: expression and localization at the fusion site, necessity for fusion, and sufficiency to induce the fusion of liposomes or heterologous cells[81,82]. From this definition, the only bona-fide developmental fusogens are syncytin, a captive retroviral envelope protein involved in placenta formation [83–85], the Fusion Failure (FF) family of fusogens, AFF-1 and EFF-1[86–88], which are responsible for syncytial cell fusion events during *C. elegans* development, and myoblast fusion proteins responsible for muscle cell development. Results are perplexing for the myoblast fusion events necessary for muscle cell formation, as it appears that two proteins have fulfilled these requirements, the FGFRL1 receptor[89–91] and Myomaker[92,93]. Additional experiments and/or comparisons of the structures of these proteins to known viral and developmental fusogens could potentially resolve which of these proteins is actually responsible for myoblast fusion[94–96]. With regards to HAP2/GCS1, the first two out of these three requirements have been satisfied for its classification as a bona fide fusogen.

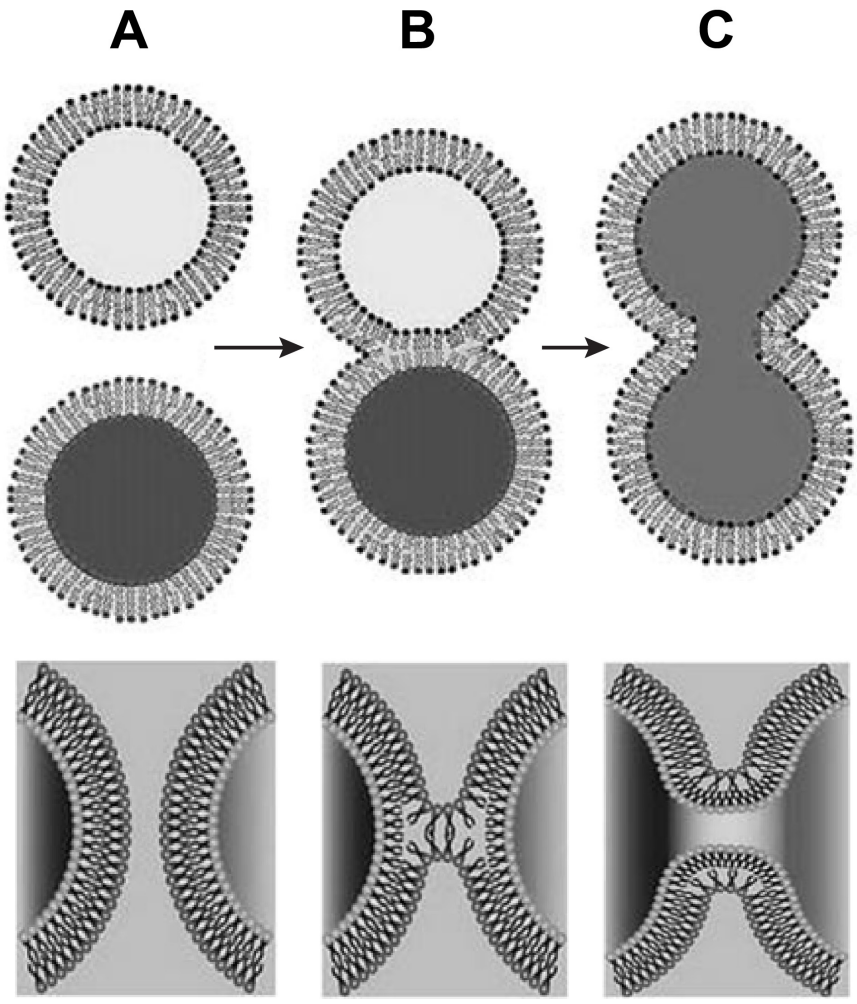
C. Membrane Fusogens

Membrane fusion is the process of generating a single continuous membrane from two independent lipid bilayers. Fusion events generally occur under the direction of specific proteins and through the formation of a hemi-fusion intermediate membrane structure, in which first outer leaflets of the bilayer fuse, followed by inner leaflets (Figure 1.3) [81,97]. In this section, I briefly review a selection of the different membrane fusion events existing in cells and viruses and the proteins that help orchestrate them. A particular emphasis is placed on class II fusion protein structure and function, with coverage of the triggers, fusion peptides, conformational changes and two proposed models of mechanistic action.

C1. Intracellular membrane fusion events

Intracellular vesicle trafficking*,* comprising the entire eukaryotic exocytic pathway, allows the separation and delivery of protein content within and outside the cell [98]. These are among the best-studied membrane fusion events, and the proteins that control them are thought to be nearly as ancient as eukaryotic cells themselves [99]. The docking and merger of vesicles to target membranes within cells is controlled by a complex of molecular components known as SNAREs (*S*oluble *N*-ethylmaleimide *A*ttachment protein *R*eceptors)[100,101]. These transmembrane proteins mediate membrane fusion through the heterotypic zippering interactions of coiled-coil domains present on both vesicle and target membranes in a calcium-dependent fashion[101,102]. After fusion, SNAREs are unzipped and recycled for a new round of fusion by *N*-ethylmaleimide *S*ensitive *F*usion protein (NSF) and *S*oluble *N*SF *A*ttachment *P*rotein (SNAP). For SNAREs, only after the interacting partners of a myc-tagged NSF[103] were identified were researchers able to reconstruct the minimal necessary components sufficient to fuse membranes in alternate systems. Although only known to carry out their biological functions intracellularly, minimal reconstitutions of essential SNARE components assembled on the outer leaflets of cellular plasma membranes have been shown to be sufficient for ectoplasmic cell-cell fusion events[104]. Interestingly it has been recently proposed that

Figure 1.3 Fundamental steps in membrane fusion. A cartoon diagramming the key steps in vesicle fusion is shown (top) and zoomed in images of the membrane dynamics at each step (below). **(A)** Membrane contact occurs between the apposed bilayers of two separate membrane-enclosed compartments. **(B)** The outer leaflet of the phospholipid bilayer mixes with the cognate outer leaflet of the apposed bilayer to form what is called the "hemi-fusion intermediate" structure. **(C)** The inner leaflet of the phospholipid bilayer then mixes allowing for fusion pore formation and content mixing between the two vesicular compartments. This figure was adapted from a similar figure originally published under a Creative Commons Attribution License, Copyright 2011 Sutapa Mondal Roy and Munna Sarkar (PMID: 21660306) [105].


A
B
C

SNARE protein structure is reminiscent of Class I viral fusion proteins, suggesting a potential relationship between these two proteins[106–108].

Mitofusion events unite two individual mitochondria within the cell. Mitochondria, thought to be the evolutionary derivatives of an endosymbiotic prokaryote, are the double membrane-bound organelles responsible for energy generation in eukaryotic cells. However, while free-living bacterial cells do not have the ability to completely fuse to one another (likely due to their peptidoglycan walls), mitochondria do so regularly. On the basis of their fusion protein structures, the mechanism of mitofusion is less related to SNAREs[109], and more similar to those of the atlastin GTPases which mediate the fusions of endoplasmic reticulum membranes[110–112]. Fzo1, Ugo1 and Mgm1 are all yeast proteins which have been identified as important for promoting the outer-membrane fusion events of mitochondria [113]. They are expected to do so specifically in a GTP- and pH-dependent manner through a homotypic trans-interaction of antiparallel coiled-coiled repeats of Fzo1 (mammalian homolog is Mitofusin Mfn1)[114]. It is not clear, however, whether Fzo1 only functions in docking of individual mitochondria or both docking and fusion[112]. Relevant to this discussion, other mitochondrial fusion mutant phenotypes were attributed to defects in placental trophoblast giant cells[115] (when Mfn2 was deleted in mice), and male sterility in *Drosophila melanogaster* Fzo1 mutants due to mitochondria deficiencies during spermatogenesis[116].

Nuclear envelope fusion during karyogamy, or the formation of a zygotic nucleus from two haploid pronuclei, is another example of a fusion event between two double membrane-bound organelles within the cell. In yeast karyogamy, SNAREs and Prm3 mediate outer membrane fusion, but recently an ancient conserved membrane protein called Kar5 (a.k.a., brambleberry) has drawn interest as potentially being involved in both outer and inner membrane fusion as well [117–119]. A common theme in karyogamy in both *Tetrahymena* and yeast is that microtubular networks are key to initially draw the two nuclei destined for fusion together [120–122]. In karyogamy, it is possible that these microtubular dynamics may even

functionally replace the necessity for a specific “recognition” event between heterologous nuclei prior to membrane adhesion.

C2. Classes of viral fusogens

In contrast to the complexes of proteins controlling intracellular membrane fusion, many viruses use only a single protein to accomplish the task[123]. The diversity of proteins enveloped viruses use to initiate membrane fusion in their entry of host cells during infection are classified into three structural categories described below. For each of these categories a short description and a few examples are given. Regardless of the variety existing in the structure, triggers, and fusion peptide adopted by each viral system, a common theme in these protein’s mechanisms is the unilateral interaction of pre-fusion trimers with target bilayers and the conformational transitions of these oligomers to a post-fusion “trimer-of-hairpins” [124].

Class I fusion protein monomers are characterized by two α-helices and an amphiphilic fusion peptide[81]. In their pre-fusion trimeric form, the six-helical bundle centrally buries the amino terminal, fusion peptide-containing, α-helical coiled coil of the three monomers, wrapping them in an external layer of their membrane-proximal helices[124]. When activated, conformational changes extend the exposed fusion peptide in the amino terminal domain towards the target membrane, forming a rod-like structure[81]. After peptide anchorage in the target bilayer, this amino terminal domain folds back towards its partner helices, drawing the two membranes together and leading to fusion as it forms a post-fusion six-helical bundle. Fusion triggers for these proteins are generally either low pH, receptor-binding, or a combination of the two[124]. Some classical examples of viruses containing Class I fusion proteins are Human Immunodeficiency Virus (HIV), Simian virus, Ebola virus, Influenza virus, and the Severe Acute Respiratory Syndrome (SARS) virus.

Influenza virus hemagglutinin (HA) is perhaps the best understood class I fusion protein. It is a type I transmembrane protein whose disulfide-bonded proteolytic cleavage products, HA1 and HA2, mediate both viral adhesion and fusion, respectively[81]. The proteolytic cleavage of

HA also results in the N' terminal positioning of the its fusion peptide. First, HA1 binds sialic acid on the host-cell receptor surface, and the virion is endocytosed[81]. Exposure to the low pH of the endosome activates HA through protonation of both subunits, allowing the decoupling of HA2 from HA1 for ensuing rearrangements forming the extended rod structure and hooking fusion peptides into the bilayer[124]. At this stage, it is expected that clusters of trimeric HA units form prior to fold-back to ensure sufficient fusogenic force for membrane merger[125].

Class II fusion protein monomers have an elongated three-domain architecture composed almost entirely from β-sheets and an internal hydrophobic fusion loop. Although class II fusion proteins (CII) lack underlying conservation at the amino acid level, the extreme similarities observed in the tertiary structures of these fusogens suggest that they were derived from a common ancestor [86,95,106]. In viral CII proteins, before virion assembly they are transiently associated with a protein chaperone (prM for Tick Borne Encephalitis Virus and p62 for SFV) that functions in their proper folding and transport, and is cleaved off by furin before virion release[95]. In their pre-fusion form, ectodomains lie flat along the viral membrane in a herringbone pattern of antiparallel dimers, with the C' terminal transmembrane proximal domains I and III of one monomer embracing the fusion loop-containing domain II of the other[126,127]. Interestingly, the arrangement of dimers on the virion membrane in certain arboviruses was discovered to be less tightly packed when the virus is passing through its human versus mosquito host. These changes are reflective of the temperature differences between the two host environments, and explain the accessibility of certain viral epitopes (that were originally thought to be buried) to the mammalian immune system[128].

The internal fusion loop itself (~16 amino acids long) is sandwiched between beta-strands 'c' and 'd' at the very apical end of domain II. Although the canonical viral fusion loop sequences of each viral family are highly conserved, the observed mutation frequencies of these loop sequences in some viral isolates are unexpectedly high (~27% non-synonymous mutation rate), suggesting some degree of flexibility in the overall amino acid composition of the

peptide, especially under conditions of superinfection [129]. In general, fusion peptides are known to form independently folded domains in the membrane upon insertion that alter the structure of the lipid bilayer to be more highly-ordered, in an enthalpy-driven “non-classical” hydrophobic effect[130–132]. With respect to CII fusion loops, the overall amino acid content, specifically of hydrophobic, flexible glycines and/or large bulky residues, centrally located within the loop [133,134], along with peptide clustering (likely involving secondary structural changes of the peptide in the bilayer evironment) [135] are both important for membrane interaction. At least from recent data (see Chapter three), however, it appears that the secondary structure changes of CII fusion peptides just before or during insertion in the outer leaflet of the bilayer are more essential for their membrane interactions than their overall amino acid composition. The degree to which this substantial peptide-induced membrane stress may contribute to viral fusion events (in addition to the fusion peptide simply serving as an anchor for fusion protein conformational changes) has not been well established.

In the **classical model** of viral CII fusion (Figure 1.4A), upon activation, the dimers rearrange to a prefusion homotrimer through a metastable dimer intermediate, peeling the oligomerizing ectodomains up off the viral surface and towards the apposing membrane, and finally stabilizing into a trimer only once the fusion loops have inserted into the target bilayer[95]. From here, the C-terminal end of molecule (domains I & III) folds back towards the trimer’s core axis (which is formed in large part by the homotypic interactions of domain II from the individual monomers). This advances domain III into a position much closer to domain II’s fusion loop and brings the target membranes together for fusion. Larger complexes or networks formed by lateral interactions of these CII trimers are thought to further facilitate fusion[95].

In some examples of CII viral entry, adhesion to the target membrane is carried out by a separate receptor-binding glycoprotein (alphaviruses), while in others the CII fusion protein itself functions in both receptor-binding and fusion (flaviviruses). In the latter case, due to the high number of cell lines permissive to infection, the receptor is assumed to be a molecule

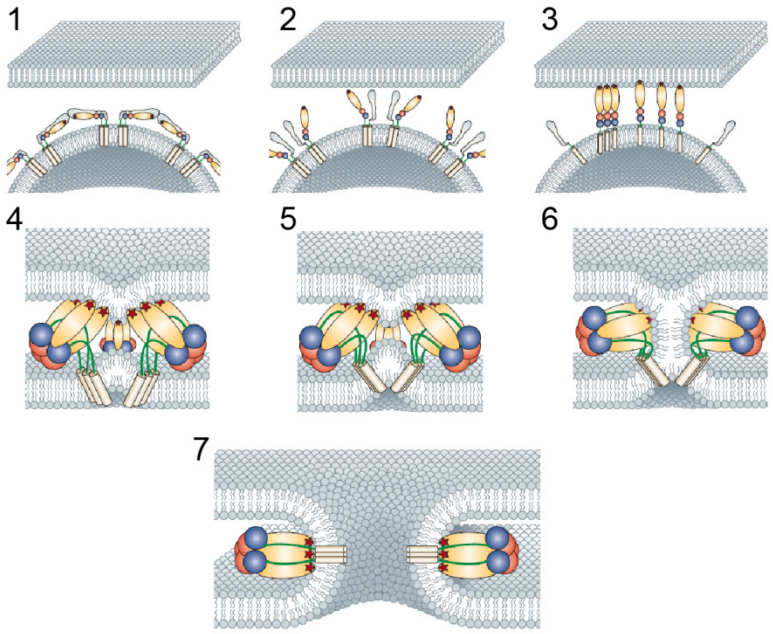
ubiquitously expressed on the surface of many different cell types or multiple receptors may exist [136]. Some known receptors of the Dengue virus class II fusogens are heparan sulfate, Heat Shock protein 90 and 70, neolactotetraosylceramide, CD14, BiP, mannose receptor, and DCSIGN [136]. In these cases receptor binding is mediated by domain III of the fusogen. After adhesion to the cell surface, flavivirus virions are generally taken into the cell via clathrin-mediated endocytosis and fuse in the late endosome[137] to release their +ssRNA viral genome into the host cell cytosol. Formation of new virions occurs through budding into the endoplasmic reticulum (flaviviruses), or at the plasma membrane (alphaviruses)[95].

The recognized fusion trigger for CII viral fusion is low (~6) pH [123], however recent studies have found that calcium is a requirement for at least one example [138,139], while still other proteins were shown to depend on target bilayer compositions containing higher concentrations of cholesterol or negatively-charged, cone-shaped lipids (like PE) [95,137]. Membrane compositions in general, are already known to have profound effects on the likelihood of progression through fusion intermediate structures, so these lipid requirements are expected to pose more of a restriction to fusion pore formation rather than function as a trigger[140,141].

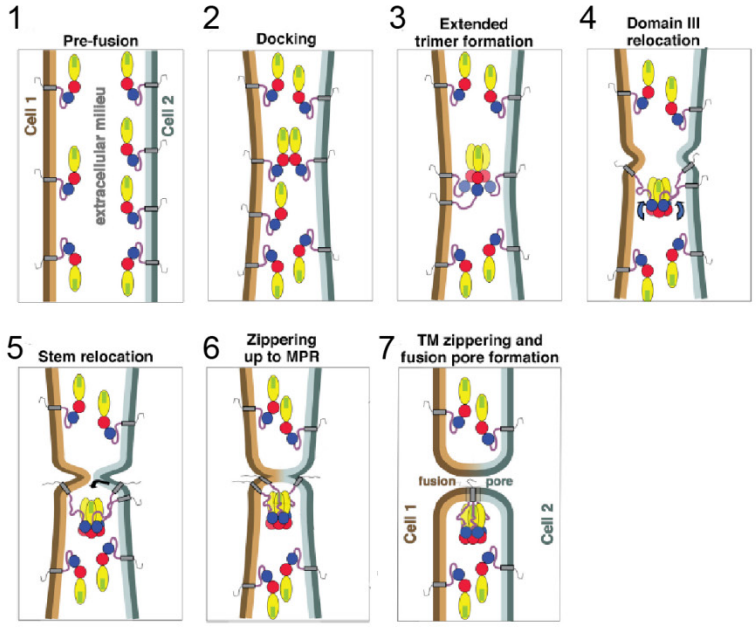
Classical examples of class II fusion proteins are found in the *Flaviviridae* and *Togaviridae* families, which contain members such as Chikungunya virus, Yellow fever virus, Dengue Virus, Zika virus, and Tick-borne encephalitis virus[95]. Interestingly, CII fusion proteins are also now known to mediate certain developmental cell-cell fusion events (see discussion below and Chapter 3). While some of these viruses have vaccines in development, the only two successful vaccines that have been developed so far for CII-fusion protein-containing viruses are for Japanese Encephalitis Virus and Yellow fever virus[142]. It is thought that neutralizing antibody responses against these CII-fusogen envelope proteins are key correlates of protection from these infections. Other therapeutic interventions, such as fusion-inhibiting peptides analogous to the envelope stem, are being developed in parallel[143].

Figure 1.4 Proposed models of Class II Membrane Fusion. (A) The classical model of viral class II membrane fusion, adapted by permission from Macmillan Publishers Ltd: Nature Reviews, Microbiology [95], copyright 2006. (B) The developmental model of viral class II membrane fusion as reprinted from Cell, 157, Perez-Vargas, J. *et al.*, Structural basis of eukaryotic cell-cell fusion, 407-419, Copyright (2014), with permission from Elsevier [144]. The three domains of the CII fusion protein are labeled as follows: domain I = red, domain II= yellow, domain III = blue. Steps in the predicted fusion process are numbered 1-7. In step **1**, antiparallel dimers lay on the virion surface in A or upright on the both membranes in B. **2.** An activating trigger in A causes a conformational change that shifts the dimers up off the virion's surface and exposes their fusion loop, while in B pre-fusion monomers cluster together forming trans-interactions between monomers on opposed membranes. **3.** In A, insertion of the fusion loops into the target bilayer allows for a stable pre-fusion homotrimer to form while in B, addition of a third monomer to docked monomers creates an extended intermediate structure without the necessity of a fusion loop interaction. **4.** In both cases, domain III relocates to a new, more forwardly-located binding site on the trimer between domains I and II, drawing the two associated membranes together and distorting them through the cooperative actions of the trimer (and fusion loop in A). **5.** The "stem" (sequence connecting domain III to the transmembrane domain) relocates to a position tightly against the trimer axis, bringing transmembrane domains closer to one another, further distorting the membrane. In B, the stem is brought into a position to interact with the tip of domain II in the trimers. **6.** As all three stems reposition to their final destinations, the two membranes advance to a hemifusion intermediate stage. **7.** As the transmembrane segments of each dimer come into alignment together, a complete fusion pore is formed.

A


1
2
3
4
5
6
7

B


1
Pre-fusion
Cell 1
extracellular milieu
Cell 2
2
Docking
3
Extended trimer formation
4
Domain III relocation
5
Stem relocation
6
Zippering up to MPR
7
TM zippering and fusion pore formation
fusion pore
Cell 1
Cell 2

Class III fusion protein monomers combine structural traits seen in both CI and CII fusion proteins. These structures for prefusion trimers containing both α-helices and β-sheets that possibly end up in post-fusion six helix bundles[81]. They also have short, bipartite and hydrophobic fusion loops that are exposed all the time (not just upon activation) [81,124]. One distinctive feature of CIII fusion proteins is that their low pH or receptor triggered conformational changes appear to be reversible, unlike those of CI and CII proteins[124]. Example viruses containing these fusion proteins are Rabies Virus, Herpes Simplex Virus, and Vesicular Stomatitis Virus.

C3. Developmental cell-cell fusion

Cell-cell fusion is known to be critical for many important biological processes ranging from gamete cell-cell fusion in sex, to nonsexual cellular fusions generating giant multinucleate cells as precursors for the development of an array of different tissues such as bone, placenta, eye lens, and muscle[82,145]. The diversity of the uses for cell-cell fusion in different organisms abounds, with fusions in blastomeres of sponges (Porifera), self-fusions in the hypha of *Neurospora,* and over 300 somatic cell-cell fusions during formation of various tissues and organ systems in the nematode worm *C. elegans* [94,145]. A role for cellular fusion has even been suggested in the pathology of cancer[146]. In this emerging field, however, many of the fusion proteins controlling these events are still unknown, and with them the array of potential fusion mechanisms that might help us better understand the merger of membranes[141]. While identification of these proteins is key, some cell-cell fusogens are already starting to be understood.

Among these, probably the best characterized are the somatic cell fusions instigated by the fusion failure (FF) proteins in the nematode worm *Caenorhabditis elegans* [86,145]. These proteins are responsible for the generation of muscles in the pharynx and epithelia, as well as the uterus, nervous system, vulva, hymen, and excretory system of the developing worm[145,147,148]. The fusion proteins involved are called Epithelial cell Fusion Failure (EFF-

1) and Anchor cell Fusion Failure (AFF-1), and constitute the first known family of developmental fusogens [87]. The FF proteins are model fusogens, necessary for cell fusion during worm development, and resulting in syncytia formation when they are heterologously expressed in mammalian cells. As in the worm, expression of these fusion proteins in cultured cells was necessary on both target and donor membranes for fusion to occur, and mediated fusion of transfected BHK-21 cells[149], as well as of EFF-1 carrying pseudotyped viral particles to an EFF-1 expressing target membrane[88]. The frequency and degree to which EFF-1 heterologous cell-cell fusion occurs can be increased via localized actin polymerization at the plasma membrane [150].

Finally, the EFF-1 protein was the first cell-cell fusogen for which a structure was solved [144], leading to the startling discovery that it highly resembled a class II viral fusion protein. The highly similar structure, along with the fact that CII envelope proteins are found in some retroviruses of *C. elegans* is more than a bit suggestive as to the origins of these proteins, but a clear relationship has yet to be established [151]. The EFF-1 structure, however, differed from classical CII proteins in a number of respects, such as the lack of a fusion loop. Instead EFF-1 had a very short acidic patch of amino acids in the protein region where the loop should have been. It was speculated that these acidic residues may possibly destabilize the outer leaflet of the target bilayer, but were not expected to directly insert into the membrane. In addition, in contrast with the membrane parallel pre-fusion dimers of CII viral fusogens, the EFF-1 protein was surprisingly found to be positioned upright and monomeric on the cell surface[152].

Given the necessity of the EFF-1 molecule on both donor and target membranes for fusion events to occur, a **developmental model** for CII fusion by EFF-1 has been proposed (Figure 1.4B)[144]. In this model, the major conformational changes (the “fold-back”) within the CII structure are the same, but the formation of the prefusion trimer differs. The prefusion trimer is expected to form through a homotypic trans-interaction between EFF-1 molecules from opposite membranes to form an intermembrane prefusion trimer containing transmembrane

anchors in both bilayers [144]. The fusion loops of this trimer are not expected to insert into the membrane, although it is possible that domain II would still be orientated towards one of the two bilayers if the “acidic patch” really does play a role. Regardless, the trans-orientation of the trimer would already have anchors in both membranes rendering any insertion of a fusion loop unnecessary. Thus, it is purely the conformational changes resulting in the post-fusion trimer of hairpins that is thought to be the sole force drawing the two apposed membranes close enough together to initiate fusion. While this is an alluring mechanism, no biochemical data yet exists to show that EFF-1 monomers from two different cells can functionally interact in a trans-oligomerization event. In addition to this, the environmental trigger of FF-mediated fusion events is still unknown.

In the realm of gamete-cell fusion, while HAP2/GCS1 is known to be important in many eukaryotes, HAP2 orthologs have not been found in fungi or vertebrates[18]. While the tetraspanin Prm1p[153–155] is implicated in yeast gamete cell fusion, the identity of the gamete fusogen in vertebrates is still unknown, but several candidate proteins have been suggested [156,157]. Of these, the proteins already known to cause blocks to fertility in mice are the tetraspanin CD9 on egg cells[158–160], ADAMs[161], Izumo of sperm[162,163], and Izumo’s binding partner on the egg, Juno [164]. With all of these (including Prm1p in yeast), it has been difficult to delineate a precise function in fusion, and it is very possible that in many, if not all of these cases, the protein’s function is in gamete membrane adhesion rather than fusion.

D. Objectives and Organization of the Dissertation

The original goal of this dissertation was to elucidate the functional role of HAP2/GCS1 in the ciliate *Tetrahymena thermophila,* in order to more broadly understand the potential fusogenic activity of this ancient gamete fusion protein. At the time we started this work, it was still very unclear as to whether HAP2/GCS1 was the instigator of gamete membrane fusion, or maybe just a necessary member of a functional complex of proteins accomplishing this task. The work was prompted by a discussion between Dr. Timothy Springer (Harvard) and my

advisor, Dr. Ted Clark just before I started my rotation in the lab in my first year. From that conversation, my rotation project in the lab became using *Tetrahymena* as a protein expression platform to generate large amounts of the HAP2 protein for a potential structure. When these efforts proved ineffective, we focused our attention on the interesting questions arising in the biology HAP2/GCS1 function during ciliate mating. We realized that *Tetrahymena*'s easily induced, synchronous cell cycle and facile modes of genetic transformation and progeny analysis made it ideal for *in vivo* function studies of this key developmental cell-cell fusion event. Furthermore, the eccentricities intrinsic to this model organism's sexual habits added an extra dimension to the project's potential impact.

The bulk of my dissertation research is presented in the first two data chapters. In Chapter 2, the function of HAP2/GCS1 in *Tetrahymena* is described. Chapter 3 focuses on HAP2 protein-domain function, and is based primarily on structural homology predictions that led to the identification of topological similarities between HAP2 orthologs and class II viral fusion proteins, and the presence of a functional fusion loop. Chapters 4-6 describe the results of side projects done in tandem with the studies shown in Chapters 2 and 3. The results of the later chapters are in line with the overall theme of HAP2-mediated sexual cell fusion, but are likely of higher interest to a more targeted audience working in ciliate biology. To summarize the work, in Chapter 7 I provide a short synopsis of the highlights of my dissertation as well as potential directions for future research in this field.

REFERENCES

1. Speijer, D., Lukeš, J., and Eliáš, M. (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc. Natl. Acad. Sci. *112*, 8827–8834.

2. Ramesh, M.A., Malik, S.-B., and Logsdon Jr., J.M. (2005). A Phylogenomic Inventory of Meiotic Genes: Evidence for Sex in Giardia and an Early Eukaryotic Origin of Meiosis. Curr. Biol. *15*, 185–191.

3. Lehtonen, J., and Kokko, H. (2014). Sex. Curr. Biol. *24*, R305–R306.

4. Barton, N.H., and Charlesworth, B. (1998). Why Sex and Recombination? Science *281*, 1986–1990.

5. Dougherty, E.G. (1955). Comparative Evolution and the Origin of Sexuality. Syst. Biol. *4*, 145–169.

6. Otto, S.P., and Gerstein, A.C. (2008). The evolution of haploidy and diploidy. Curr. Biol. *18*, R1121–R1124.

7. Vieira, A., and Miller, D.J. (2006). Gamete interaction: is it species-specific? Mol. Reprod. Dev. *73*, 1422–1429.

8. Smith, J.M. (1978). The Evolution of Sex (CUP Archive).

9. Otto, S.P., and Gerstein, A.C. (2006). Why have sex? The population genetics of sex and recombination. Biochem. Soc. Trans. *34*, 519–522.

10. Schön, I., Martens, K., and Dijk, P. eds. (2009). Lost Sex (Dordrecht: Springer Netherlands) Available at: http://link.springer.com/10.1007/978-90-481-2770-2 [Accessed November 8, 2016].

11. Hickey, D.A. (1982). Selfish DNA: A Sexually-Transmitted Nuclear Parasite. Genetics *101*, 519–531.

12. Bell, P.J.L. (2006). Sex and the eukaryotic cell cycle is consistent with a viral ancestry for the eukaryotic nucleus. J. Theor. Biol. *243*, 54–63.

13. Brockhurst, M.A. (2011). Sex, Death, and the Red Queen. Science *333*, 166–167.

14. Ghiselin, M.T. (1974). The Economy of Nature and the Evolution of Sex (University of California Press).

15. Muller, H.J. (1964). THE RELATION OF RECOMBINATION TO MUTATIONAL ADVANCE. Mutat. Res. *106*, 2–9.

16. Swanson, W.J., and Vacquier, V.D. (2002). The rapid evolution of reproductive proteins. Nat. Rev. Genet. *3*, 137–144.

17. Dresselhaus, T., and Snell, W.J. (2014). Fertilization: A Sticky Sperm Protein in Plants. Curr. Biol. *24*, R164–R166.

18. Goodman, C.D., and McFadden, G.I. (2008). Gamete fusion: key protein identified. Curr. Biol. CB *18*, R571-3.

19. Wong, J.L., and Johnson, M.A. (2010). Is HAP2-GCS1 an ancestral gamete fusogen? Trends Cell Biol. *20*, 134–141.

20. Liu, Y., Pei, J., Grishin, N., and Snell, W.J. (2015). The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in Chlamydomonas and regulates the fusion reaction. Development *142*, 962–971.

21. Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., *et al.* (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*, D279–D285.

22. Pfam: Family: HAP2-GCS1 (PF10699) Available at: http://pfam.xfam.org/family/PF10699#tabview=tab7 [Accessed November 27, 2016].

23. Wellems, T.E., Hayton, K., and Fairhurst, R.M. (2009). The impact of malaria parasitism: from corpuscles to communities. J. Clin. Invest. *119*, 2496–2505.

24. Hotez, P.J., Alvarado, M., Basáñez, M.-G., Bolliger, I., Bourne, R., Boussinesq, M., Brooker, S.J., Brown, A.S., Buckle, G., Budke, C.M., *et al.* (2014). The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases. PLoS Negl. Trop. Dis. *8*, e2865.

25. Liu, Y., Tewari, R., Ning, J., Blagborough, A.M., Garbom, S., Pei, J., Grishin, N.V., Steele, R.E., Sinden, R.E., Snell, W.J., *et al.* (2008). The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in Chlamydomonas and Plasmodium gametes. Genes Dev. *22*, 1051–1068.

26. Hirai, M., Arai, M., Mori, T., Miyagishima, S.Y., Kawai, S., Kita, K., Kuroiwa, T., Terenius, O., and Matsuoka, H. (2008). Male fertility of malaria parasites is determined by GCS1, a plant-type reproduction factor. Curr. Biol. CB *18*, 607–613.

27. Blagborough, A.M., and Sinden, R.E. (2009). Plasmodium berghei HAP2 induces strong malaria transmission-blocking immunity in vivo and in vitro. Vaccine *27*, 5187–5194.

28. Hirai, M., and Mori, T. (2010). Fertilization is a novel attacking site for the transmission blocking of malaria parasites. Acta Trop. *114*, 157–161.

29. Johnson, M.A., Besser, K. von, Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z., and Preuss, D. (2004). Arabidopsis hapless mutations define essential gametophytic functions. Genetics *168*, 971–982.

30. Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). Generative Cell Specific 1 is essential for angiosperm fertilization. Nat. Cell Biol. *8*, 64–71.

31. Besser, K. von, Frank, A.C., Johnson, M.A., and Preuss, D. (2006). Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development *133*, 4761–4769.

32. Frank, A.C., and Johnson, M.A. (2009). Expressing the Diphtheria Toxin A Subunit from the HAP2(GCS1) Promoter Blocks Sperm Maturation and Produces Single Sperm-Like Cells Capable of Fertilization. PLANT Physiol. *151*, 1390–1400.

33. Nagahara, S., Takeuchi, H., and Higashiyama, T. (2015). Generation of a homozygous fertilization-defective gcs1 mutant by heat-inducible removal of a rescue gene. Plant Reprod. *28*, 33–46.

34. Nanney, D.L., and Simon, E.M. (1999). Chapter 1 Laboratory and Evolutionary History of Tetrahymena thermophila. In Methods in Cell Biology, D. J. A. and J. D. Forney, ed. (Academic Press), pp. 3–25. Available at: http://www.sciencedirect.com/science/article/pii/S0091679X08615277 [Accessed November 28, 2016].

35. Nanney, D.L. (1980). Experimental Ciliatology: An Introduction to Genetic and Developmental Analysis in Ciliates (John Wiley & Sons).

36. Ruehle, M.D., Orias, E., and Pearson, C.G. (2016). Tetrahymena as a Unicellular Model Eukaryote: Genetic and Genomic Tools. Genetics *203*, 649–665.

37. Doerder, F.P., Arslanyolu, M., Saad, Y., Kaczmarek, M., Mendoza, M., and Mita, B. (1996). Ecological Genetics of Tetrahymena thermophila: Mating Types, i-Antigens, Multiple Alleles and Epistasis. J. Eukaryot. Microbiol. *43*, 95–100.

38. Mochizuki, K. (2012). Developmentally Programmed, RNA-Directed Genome Rearrangement in Tetrahymena. Dev. Growth Differ. *54*, 108–119.

39. Schoeberl, U.E., and Mochizuki, K. (2011). Keeping the soma free of transposons: programmed DNA elimination in ciliates. J. Biol. Chem. *286*, 37045–37052.

40. Fillingham, J.S., Thing, T.A., Vythilingum, N., Keuroghlian, A., Bruno, D., Golding, G.B., and Pearlman, R.E. (2004). A Non-Long Terminal Repeat Retrotransposon Family Is Restricted to the Germ Line Micronucleus of the Ciliated Protozoan Tetrahymena thermophila. Eukaryot. Cell *3*, 157–169.

41. Chalker, D.L., Meyer, E., and Mochizuki, K. (2013). Epigenetics of Ciliates. Cold Spring Harb. Perspect. Biol. *5*, a017764–a017764.

42. Cervantes, M.D., Coyne, R.S., Xi, X., and Yao, M.-C. (2006). The Condensin Complex Is Essential for Amitotic Segregation of Bulk Chromosomes, but Not Nucleoli, in the Ciliate Tetrahymena thermophila. Mol. Cell. Biol. *26*, 4690–4700.

43. Turkewitz, A.P., Orias, E., and Kapler, G. (2002). Functional genomics: the coming of age for Tetrahymena thermophila. Trends Genet. *18*, 35–40.

44. Zufall, R.A. (2016). Mating Systems and Reproductive Strategies in Tetrahymena. In Biocommunication of Ciliates, G. Witzany and M. Nowacki, eds. (Springer International Publishing), pp. 195–220. Available at: http://link.springer.com/chapter/10.1007/978-3-319-32211-7_12 [Accessed September 12, 2016].

45. Orias, E. (2014). Membrane Fusion: HAP2 Protein on a Short Leash. Curr. Biol. *24*, R831–R833.

46. Umen, J.G. (2013). Genetics: Swinging Ciliates' Seven Sexes. Curr. Biol. *23*, R475–R477.

47. Cole, E.S. (2016). Cell-Cell Interactions Leading to Establishment of a Mating Junction in Tetrahymena and Paramecium, Two "Contact-Mediated" Mating Systems. In Biocommunication of Ciliates, G. Witzany and M. Nowacki, eds. (Springer International Publishing), pp. 195–220. Available at: http://link.springer.com/chapter/10.1007/978-3-319-32211-7_12 [Accessed September 12, 2016].

48. Cole, E.S. (2006). The Tetrahymena Conjugation Junction. In Cell-Cell Channels, F. Baluska, D. Volkmann, and P. W. Barlow, eds. (New York, NY: Springer New York), pp. 39–62. Available at: http://dx.doi.org/10.1007/978-0-387-46957-7_3.

49. Wolfe, J. (1982). The conjugation junction of Tetrahymena: Its structure and development. J. Morphol. *172*, 159–178.

50. Wolfe, J. (1985). Cytoskeletal reorganization and plasma membrane fusion in conjugating Tetrahymena. J. Cell Sci. *73*, 69–85.

51. McDonald, B.B. (1966). The exchange of RNA and protein during conjugation in Tetrahymena. J. Protozool. *13*, 277–285.

52. Cole, E.S. (2000). The Tetrahymena Conjugation Junction. Madame Curie Biosci. Database. Available at: http://www.ncbi.nlm.nih.gov/books/NBK6002/.

53. Wolfe, J. (1976). G1 arrest and the division/conjugation decision in Tetrahymena. Dev. Biol. *54*, 116–126.

54. Nelsen, E.M. (1978). Transformation in Tetrahymena thermophila. Development of an inducible phenotype. Dev. Biol. *66*, 17–31.

55. Adair, W.S., Barker, R., Jr, R.S.T., and Wolfe, J. (1978). Demonstration of a cell-free factor involved in cell interactions during mating in Tetrahymena. Nature *274*, 54–55.

56. Takahashi, M. (1973). Does fluid have any function for mating in Tetrahymena? Sci. Rep Tohoku Univ *36*, 223–229.

57. Xiong, J., Lu, Y., Feng, J., Yuan, D., Tian, M., Chang, Y., Fu, C., Wang, G., Zeng, H., and Miao, W. (2013). Tetrahymena functional genomics database (TetraFGD): an integrated resource for Tetrahymena functional genomics. Database J. Biol. Databases Curation *2013*, bat008.

58. Allewell, N.M., Oles, J., and Wolfe, J. (1976). A physicochemical analysis of conjugation in Tetrahymena pyriformis. Exp. Cell Res. *97*, 394–405.

59. Kume, A. (2012). Tyrosine phosphorylation in Tetrahymena thermophila during intercellular communication. Available at: http://yorkspace.library.yorku.ca/xmlui/handle/10315/31920 [Accessed November 28, 2016].

60. Ofer, L., Levkovitz, H., and Loyter, A. (1976). Conjugation in Tetrahymena pyriformis. The effect of polylysine, concanavalin A, and bivalent metals on the conjugation process. J. Cell Biol. *70*, 287–293.

61. Frisch, A., and Loyter, A. (1977). Inhibition of conjugation in Tetrahymena pyriformis by ConA. Exp. Cell Res. *110*, 337–346.

62. Fujishima, M., Tsuda, M., Mikami, Y., and Shinoda, K. (1993). Costimulation-induced rounding in Tetrahymena thermophila: early cell shape transformation induced by sexual cell-to-cell collisions between complementary mating types. Dev. Biol. *155*, 198–205.

63. Pagliaro, L., and Wolfe, J. (1987). Concanavalin A binding induces association of possible mating-type receptors with the cytoskeleton in Tetrahymena. Exp. Cell Res. *168*, 138–152.

64. Cole, E.S., Cassidy-Hanley, D., Pinello, J.F., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Jr, T.G., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. Curr. Biol. CB *24*, 2168–2173.

65. Orias, E. (1986). Ciliate Conjugation. In The Molecular Biology of Ciliated Protozoa, Gall, J.G., ed. (New York, NY: Academic Press), pp. 45–94.

66. Ostrowski, S.G., Bell, C.T.V., Winograd, N., and Ewing, A.G. (2004). Mass spectrometric imaging of highly curved membranes during Tetrahymena mating. Science *305*, 71–73.

67. Orias, J.D., Hamilton, E.P., and Orias, E. (1983). A microtubule meshwork associated with gametic pronucleus transfer across a cell-cell junction. Science *222*, 181–184.

68. Yin, L., Gater, S.T., and Karrer, K.M. (2010). A Developmentally Regulated Gene, ASI2, Is Required for Endocycling in the Macronuclear Anlagen of Tetrahymena. Eukaryot. Cell *9*, 1343–1353.

69. Cervantes, M.D., Hamilton, E.P., Xiong, J., Lawson, M.J., Yuan, D., Hadjithomas, M., Miao, W., and Orias, E. (2013). Selecting one of several mating types through gene segment joining and deletion in Tetrahymena thermophila. PLoS Biol. *11*, e1001518.

70. Kawai-Toyooka, H., Mori, T., Hamaji, T., Suzuki, M., Olson, B.J.S.C., Uemura, T., Ueda, T., Nakano, A., Toyoda, A., Fujiyama, A., *et al.* (2014). Sex-Specific Posttranslational Regulation of the Gamete Fusogen GCS1 in the Isogamous Volvocine Alga Gonium pectorale. Eukaryot. Cell *13*, 648–656.

71. Steele, R.E., and Dana, C.E. (2009). Evolutionary history of the HAP2/GCS1 gene and

sexual reproduction in metazoans. PloS One *4*, e7680.

72. Ebchuqin, E., Yokota, N., Yamada, L., Yasuoka, Y., Akasaka, M., Arakawa, M., Deguchi, R., Mori, T., and Sawada, H. (2014). Evidence for participation of GCS1 in fertilization of the starlet sea anemone Nematostella vectensis: Implication of a common mechanism of sperm–egg fusion in plants and animals. Biochem. Biophys. Res. Commun. *451*, 522–528.

73. Garcia, V.E. (2012). A Generative Cell Specific 1 Ortholog in Drosophila melanogaster. Available at: https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/20270/Garcia_washington_0250O_10191.pdf?sequence=1 [Accessed January 19, 2016].

74. Okamoto, M., Yamada, L., Fujisaki, Y., Bloomfield, G., Yoshida, K., Kuwayama, H., Sawada, H., Mori, T., and Urushihara, H. (2016). Two HAP2-GCS1 homologs responsible for gamete interactions in the cellular slime mold with multiple mating types: Implication for common mechanisms of sexual reproduction shared by plants and protozoa and for male-female differentiation. Dev. Biol. *415*, 6–13.

75. Wong, J.L., Leydon, A.R., and Johnson, M.A. (2010). HAP2(GCS1)-dependent gamete fusion requires a positively charged carboxy-terminal domain. PLoS Genet. *6*, e1000882.

76. Mori, T., Hirai, M., Kuroiwa, T., and Miyagishima, S.Y. (2010). The functional domain of GCS1-based gamete fusion resides in the amino terminus in plant and parasite species. PloS One *5*, e15957.

77. Snell, W.J., Liu, Y., Billker, O., Sinden, R.E., and Tewari, R. (2011). Parasite Vaccine. Available at: http://www.google.com/patents/US20110091526 [Accessed November 28, 2016].

78. Johnson, M.A. (2010). Fertilization: monogamy by mutually assured destruction. Curr. Biol. CB *20*, R571-3.

79. Sprunck, S., Rademacher, S., Vogler, F., Gheyselinck, J., Grossniklaus, U., and Dresselhaus, T. (2012). Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. Science *338*, 1093–1097.

80. Liu, Y., Misamore, M.J., and Snell, W.J. (2010). Membrane fusion triggers rapid degradation of two gamete-specific, fusion-essential proteins in a membrane block to polygamy in Chlamydomonas. Dev. Camb. Engl. *137*, 1473–1481.

81. Sapir, A., Avinoam, O., Podbilewicz, B., and Chernomordik, L.V. (2008). Viral and developmental cell fusion mechanisms: conservation and divergence. Dev. Cell *14*, 11–21.

82. Oren-Suissa, M., and Podbilewicz, B. (2007). Cell fusion during development. Trends Cell Biol. *17*, 537–546.

83. Blond, J.-L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., and Cosset, F.-L. (2000). An Envelope Glycoprotein of the Human Endogenous Retrovirus HERV-W Is Expressed in the Human Placenta and Fuses Cells Expressing the Type D Mammalian Retrovirus Receptor. J. Virol. *74*, 3321–3329.

84. Cornelis, G., Heidmann, O., Degrelle, S.A., Vernochet, C., Lavialle, C., Letzelter, C., Bernard-Stoecklin, S., Hassanin, A., Mulot, B., Guillomot, M., *et al.* (2013). Captured retroviral envelope syncytin gene associated with the unique placental structure of higher ruminants. Proc. Natl. Acad. Sci. *110*, E828–E837.

85. Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., *et al.* (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature *403*, 785–789.

86. Mohler, W.A., Shemer, G., Campo, J.J. del, Valansi, C., Opoku-Serebuoh, E., Scranton, V., Assaf, N., White, J.G., and Podbilewicz, B. (2002). The type I membrane protein EFF-1 is essential for developmental cell fusion. Dev. Cell *2*, 355–362.

87. White, J.M. (2007). The first family of cell-cell fusion. Dev. Cell *12*, 667–668.

88. Avinoam, O., Fridman, K., Valansi, C., Abutbul, I., Zeev-Ben-Mordehai, T., Maurer, U.E., Sapir, A., Danino, D., Grunewald, K., White, J.M., *et al.* (2011). Conserved eukaryotic fusogens can fuse viral envelopes to cells. Science *332*, 589–592.

89. Steinberg, F., Gerber, S.D., Rieckmann, T., and Trueb, B. (2010). Rapid Fusion and Syncytium Formation of Heterologous Cells upon Expression of the FGFRL1 Receptor. J. Biol. Chem. *285*, 37704–37715.

90. Zhuang, L., Pandey, A.V., Villiger, P.M., and Trueb, B. (2015). Cell-cell fusion induced by the Ig3 domain of receptor FGFRL1 in CHO cells. Biochim. Biophys. Acta *1853*, 2273–2285.

91. Bluteau, G., Zhuang, L., Amann, R., and Trueb, B. (2014). Targeted Disruption of the Intracellular Domain of Receptor FgfrL1 in Mice. PLOS ONE *9*, e105210.

92. Millay, D.P., O'Rourke, J.R., Sutherland, L.B., Bezprozvannaya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2013). Myomaker: A membrane activator of myoblast fusion and muscle formation. Nature *499*, 301–305.

93. Millay, D.P., Gamage, D.G., Quinn, M.E., Min, Y.-L., Mitani, Y., Bassel-Duby, R., and Olson, E.N. (2016). Structure–function analysis of myomaker domains required for myoblast fusion. Proc. Natl. Acad. Sci. *113*, 2116–2121.

94. Aguilar, P.S., Baylies, M.K., Fleissner, A., Helming, L., Inoue, N., Podbilewicz, B., Wang, H., and Wong, M. (2013). Genetic basis of cell-cell fusion mechanisms. Trends Genet. TIG *29*, 427–437.

95. Kielian, M., and Rey, F.A. (2006). Virus membrane-fusion proteins: more than one way to make a hairpin. Nat. Rev. *4*, 67–76.

96. Abmayr, S.M., and Pavlath, G.K. (2012). Myoblast fusion: lessons from flies and mice. Development *139*, 641–656.

97. Jahn, R., Lang, T., and Südhof, T.C. (2003). Membrane fusion. Cell *112*, 519–533.

98. Monck, J.R., and Fernandez, J.M. (1996). The fusion pore and mechanisms of biological membrane fusion. Curr. Opin. Cell Biol. *8*, 524–533.

99. Dacks, J.B., and Field, M.C. (2007). Evolution of the eukaryotic membrane-trafficking system: origin, tempo and mode. J. Cell Sci. *120*, 2977–2985.

100. Plattner, H. (2010). Membrane trafficking in protozoa SNARE proteins, H+-ATPase, actin, and other key players in ciliates. Int. Rev. Cell Mol. Biol. *280*, 79–184.

101. Jahn, R., and Scheller, R.H. (2006). SNAREs — engines for membrane fusion. Nat. Rev. Mol. Cell Biol. *7*, 631–643.

102. Söllner, T., Whiteheart, S.W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J.E. (1993). SNAP receptors implicated in vesicle targeting and fusion. Nature *362*, 318–324.

103. Wilson, D.W., Whiteheart, S.W., Wiedmann, M., Brunner, M., and Rothman, J.E. (1992). A multisubunit particle implicated in membrane fusion. J. Cell Biol. *117*, 531–538.

104. Hu, C., Ahmed, M., Melia, T.J., Sollner, T.H., Mayer, T., and Rothman, J.E. (2003). Fusion of cells by flipped SNAREs. Science *300*, 1745–1749.

105. Mondal Roy, S., and Sarkar, M. (2011). Membrane Fusion Induced by Small Molecules and Ions. J. Lipids *2011*, e528784.

106. Modis, Y. (2014). Relating structure to evolution in class II viral membrane fusion proteins. Curr. Opin. Virol. *5*, 34–41.

107. Donald, J.E., Zhang, Y., Fiorin, G., Carnevale, V., Slochower, D.R., Gai, F., Klein, M.L., and DeGrado, W.F. (2011). Transmembrane orientation and possible role of the fusogenic peptide from parainfluenza virus 5 (PIV5) in promoting fusion. Proc. Natl. Acad. Sci. U. S. A. *108*, 3958–3963.

108. Skehel, J.J., and Wiley, D.C. (1998). Coiled coils in both intracellular vesicle and viral membrane fusion. Cell *95*, 871–874.

109. Wickner, W., and Schekman, R. (2008). Membrane fusion. Nat. Struct. Mol. Biol. *15*, 658–664.

110. Bian, X., Klemm, R.W., Liu, T.Y., Zhang, M., Sun, S., Sui, X., Liu, X., Rapoport, T.A., and Hu, J. (2011). Structures of the atlastin GTPase provide insight into homotypic fusion of endoplasmic reticulum membranes. Proc. Natl. Acad. Sci. U. S. A. *108*, 3976–3981.

111. Byrnes, L.J., and Sondermann, H. (2011). Structural basis for the nucleotide-dependent dimerization of the large G protein atlastin-1/SPG3A. Proc. Natl. Acad. Sci. U. S. A. *108*, 2216–2221.

112. Koshiba, T., Detmer, S.A., Kaiser, J.T., Chen, H., McCaffery, J.M., and Chan, D.C. (2004). Structural Basis of Mitochondrial Tethering by Mitofusin Complexes. Science *305*, 858–862.

113. Okamoto, K., and Shaw, J.M. (2005). Mitochondrial morphology and dynamics in yeast and multicellular eukaryotes. Annu. Rev. Genet. *39*, 503–536.

114. Meeusen, S., McCaffery, J.M., and Nunnari, J. (2004). Mitochondrial Fusion Intermediates Revealed in Vitro. Science *305*, 1747–1752.

115. Chen, H., Detmer, S.A., Ewald, A.J., Griffin, E.E., Fraser, S.E., and Chan, D.C. (2003). Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. J. Cell Biol. *160*, 189–200.

116. Hales, K.G., and Fuller, M.T. (1997). Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase. Cell *90*, 121–129.

117. Abrams, E.W., Zhang, H., Marlow, F.L., Kapp, L., Lu, S., and Mullins, M.C. (2012). Dynamic Assembly of Brambleberry Mediates Nuclear Envelope Fusion during Early

Development. Cell *150*, 521–532.

118. Ning, J., Otto, T.D., Pfander, C., Schwach, F., Brochet, M., Bushell, E., Goulding, D., Sanders, M., Lefebvre, P.A., Pei, J., *et al.* (2013). Comparative genomics in Chlamydomonas and Plasmodium identifies an ancient nuclear envelope protein family essential for sexual reproduction in protists, fungi, plants, and vertebrates. Genes Dev. *27*, 1198–1215.

119. Rogers, J.V., and Rose, M.D. (2015). Kar5p Is Required for Multiple Functions in Both Inner and Outer Nuclear Envelope Fusion in Saccharomyces cerevisiae. G3 GenesGenomesGenetics *5*, 111–121.

120. Melloy, P., Shen, S., White, E., McIntosh, J.R., and Rose, M.D. (2007). Nuclear fusion during yeast mating occurs by a three-step pathway. J. Cell Biol. *179*, 659–670.

121. Gibeaux, R., and Knop, M. (2013). When yeast cells meet, karyogamy!: An example of nuclear migration slowly resolved. Nucleus *4*, 182–188.

122. Hamilton, E.P., Suhr-Jessen, P.B., and Orias, E. (1988). Pronuclear fusion failure: an alternate conjugational pathway in Tetrahymena thermophila, induced by vinblastine. Genetics *118*, 627–636.

123. Stiasny, K., Fritz, R., Pangerl, K., and Heinz, F.X. (2011). Molecular mechanisms of flavivirus membrane fusion. Amino Acids *41*, 1159–1163.

124. White, J.M., Delos, S.E., Brecher, M., and Schornberg, K. (2008). Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. Crit. Rev. Biochem. Mol. Biol. *43*, 189–219.

125. Hamilton, B.S., Whittaker, G.R., and Daniel, S. (2012). Influenza virus-mediated membrane fusion: determinants of hemagglutinin fusogenic activity and experimental approaches for assessing virus fusion. Viruses *4*, 1144–1168.

126. Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C., and Harrison, S.C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. Nature *375*, 291–298.

127. Kuhn, R.J., Zhang, W., Rossmann, M.G., Pletnev, S.V., Corver, J., Lenches, E., Jones, C.T., Mukhopadhyay, S., Chipman, P.R., Strauss, E.G., *et al.* (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell *108*, 717–725.

128. Rey, F.A. (2013). Dengue virus: two hosts, two structures. Nature *497*, 443–444.

129. Seligman, S.J. (2008). Constancy and diversity in the flavivirus fusion peptide. Virol. J. *5*, 27.

130. Tamm, L.K., Han, X., Li, Y., and Lai, A.L. (2002). Structure and function of membrane fusion peptides. Biopolymers *66*, 249–260.

131. Ge, M., and Freed, J.H. (2009). Fusion Peptide from Influenza Hemagglutinin Increases Membrane Surface Order: An Electron-Spin Resonance Study. Biophys. J. *96*, 4925–4934.

132. Martin, I., and Ruysschaert, J.M. (2000). Common properties of fusion peptides from diverse systems. Biosci. Rep. *20*, 483–500.

133. Mendes, Y.S., Alves, N.S., Souza, T.L.F., Jr, I.P.S., Bianconi, M.L., Bernardi, R.C., Pascutti, P.G., Silva, J.L., Gomes, A.M.O., and Oliveira, A.C. (2012). The Structural Dynamics of the Flavivirus Fusion Peptide–Membrane Interaction. PLOS ONE *7*, e47596.

134. Melo, M.N., Sousa, F.J.R., Carneiro, F.A., Castanho, M.A.R.B., Valente, A.P., Almeida, F.C.L., Da Poian, A.T., and Mohana-Borges, R. (2009). Interaction of the Dengue Virus Fusion Peptide with Membranes Assessed by NMR: The Essential Role of the Envelope Protein Trp101 for Membrane Fusion. J. Mol. Biol. *392*, 736–746.

135. Stauffer, F., Melo, M.N., Carneiro, F.A., Sousa, F.J.R., Juliano, M.A., Juliano, L., Mohana-Borges, R., Poian, A.T.D., and Castanho, M.A.R.B. (2008). Interaction between dengue virus fusion peptide and lipid bilayers depends on peptide clustering. Mol. Membr. Biol. *25*, 128–138.

136. Smit, J.M., Moesker, B., Rodenhuis-Zybert, I., and Wilschut, J. (2011). Flavivirus cell entry and membrane fusion. Viruses *3*, 160–171.

137. Zaitseva, E., Yang, S.T., Melikov, K., Pourmal, S., and Chernomordik, L.V. (2010). Dengue virus ensures its fusion in late endosomes using compartment-specific lipids. PLoS Pathog. *6*, e1001131.

138. Dubé, M., Rey, F.A., and Kielian, M. (2014). Rubella virus: first calcium-requiring viral fusion protein. PLoS Pathog. *10*, e1004530.

139. DuBois, R.M., Vaney, M.-C., Tortorici, M.A., Kurdi, R.A., Barba-Spaeth, G., Krey, T., and Rey, F.A. (2013). Functional and evolutionary insight from the crystal structure of rubella virus protein E1. Nature *493*, 552–556.

140. Chernomordik, L.V., and Kozlov, M.M. (2003). Protein-lipid interplay in fusion and fission of

biological membranes. Annu. Rev. Biochem. *72*, 175–207.

141. Chernomordik, L.V., and Kozlov, M.M. (2008). Mechanics of membrane fusion. Nat. Struct. Mol. Biol. *15*, 675–683.

142. Pulendran, B. (2009). Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology. Nat. Rev. Immunol. *9*, 741–747.

143. Koehler, J.W., Smith, J.M., Ripoll, D.R., Spik, K.W., Taylor, S.L., Badger, C.V., Grant, R.J., Ogg, M.M., Wallqvist, A., Guttieri, M.C., *et al.* (2013). A Fusion-Inhibiting Peptide against Rift Valley Fever Virus Inhibits Multiple, Diverse Viruses. PLoS Negl. Trop. Dis. *7*. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3772029/ [Accessed November 30, 2016].

144. Perez-Vargas, J., Krey, T., Valansi, C., Avinoam, O., Haouz, A., Jamin, M., Raveh-Barak, H., Podbilewicz, B., and Rey, F.A. (2014). Structural basis of eukaryotic cell-cell fusion. Cell *157*, 407–419.

145. Oren-Suissa, M., and Podbilewicz, B. (2010). Evolution of programmed cell fusion: common mechanisms and distinct functions. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *239*, 1515–1528.

146. Powell, A.E., Anderson, E.C., Davies, P.S., Silk, A.D., Pelz, C., Impey, S., and Wong, M.H. (2011). Fusion between Intestinal epithelial cells and macrophages in a cancer context results in nuclear reprogramming. Cancer Res. *71*, 1497–1505.

147. Oren-Suissa, M., Hall, D.H., Treinin, M., Shemer, G., and Podbilewicz, B. (2010). The fusogen EFF-1 controls sculpting of mechanosensory dendrites. Science *328*, 1285–1288.

148. Sapir, A., Choi, J., Leikina, E., Avinoam, O., Valansi, C., Chernomordik, L.V., Newman, A.P., and Podbilewicz, B. (2007). AFF-1, a FOS-1-regulated fusogen, mediates fusion of the anchor cell in C. elegans. Dev. Cell *12*, 683–698.

149. Podbilewicz, B., Leikina, E., Sapir, A., Valansi, C., Suissa, M., Shemer, G., and Chernomordik, L.V. (2006). The C. elegans developmental fusogen EFF-1 mediates homotypic fusion in heterologous cells and in vivo. Dev. Cell *11*, 471–481.

150. Shilagardi, K., Li, S., Luo, F., Marikar, F., Duan, R., Jin, P., Kim, J.H., Murnen, K., and Chen, E.H. (2013). Actin-propelled Invasive Membrane Protrusions Promote Fusogenic Protein Engagement During Cell-Cell Fusion. Science *340*, 359–363.

151. Frame, I.G., Cutfield, J.F., and Poulter, R.T. (2001). New BEL-like LTR-retrotransposons in Fugu rubripes, Caenorhabditis elegans, and Drosophila melanogaster. Gene *263*, 219–230.

152. Zeev-Ben-Mordehai, T., Vasishtan, D., Siebert, C.A., and Grunewald, K. (2014). The full-length cell-cell fusogen EFF-1 is monomeric and upright on the membrane. Nat. Commun. *5*, 3912.

153. Olmo, V.N., and Grote, E. (2010). Prm1 functions as a disulfide-linked complex in yeast mating. J. Biol. Chem. *285*, 2274–2283.

154. Engel, A., Aguilar, P.S., and Walter, P. (2010). The yeast cell fusion protein Prm1p requires covalent dimerization to promote membrane fusion. PloS One *5*, e10593.

155. Heiman, M.G., and Walter, P. (2000). Prm1p, a pheromone-regulated multispanning membrane protein, facilitates plasma membrane fusion during yeast mating. J. Cell Biol. *151*, 719–730.

156. Primakoff, P., and Myles, D.G. (2007). Cell-cell membrane fusion during mammalian fertilization. FEBS Lett. *581*, 2174–2180.

157. Rubinstein, E., Ziyyat, A., Wolf, J.-P., Le Naour, F., and Boucheix, C. (2006). The molecular players of sperm–egg fusion in mammals. Semin. Cell Dev. Biol. *17*, 254–263.

158. Chalbi, M., Barraud-Lange, V., Ravaux, B., Howan, K., Rodriguez, N., Soule, P., Ndzoudi, A., Boucheix, C., Rubinstein, E., Wolf, J.P., *et al.* (2014). Binding of sperm protein Izumo1 and its egg receptor Juno drives Cd9 accumulation in the intercellular contact area prior to fusion during mammalian fertilization. Development *141*, 3732–3739.

159. Kaji, K., Oda, S., Shikano, T., Ohnuki, T., Uematsu, Y., Sakagami, J., Tada, N., Miyazaki, S., and Kudo, A. (2000). The gamete fusion process is defective in eggs of Cd9-deficient mice. Nat. Genet. *24*, 279–282.

160. Le Naour, F., Rubinstein, E., Jasmin, C., Prenant, M., and Boucheix, C. (2000). Severely reduced female fertility in CD9-deficient mice. Science *287*, 319–321.

161. Yuan, R., Primakoff, P., and Myles, D.G. (1997). A role for the disintegrin domain of cyritestin, a sperm surface protein belonging to the ADAM family, in mouse sperm-egg plasma membrane adhesion and fusion. J. Cell Biol. *137*, 105–112.

162. Inoue, N., Ikawa, M., Isotani, A., and Okabe, M. (2005). The immunoglobulin superfamily

protein Izumo is required for sperm to fuse with eggs. Nature *434*, 234–238.

163. Kato, K., Satouh, Y., Nishimasu, H., Kurabayashi, A., Morita, J., Fujihara, Y., Oji, A., Ishitani, R., Ikawa, M., and Nureki, O. (2016). Structural and functional insights into IZUMO1 recognition by JUNO in mammalian fertilization. Nat. Commun. *7*, 12198.

164. Bianchi, E., Doe, B., Goulding, D., and Wright, G.J. (2014). Juno is the egg Izumo receptor and is essential for mammalian fertilization. Nature *508*, 483–487.

Chapter two

Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate[1]

[1] Reprinted from Current Biology, 24, Eric S. Cole, Donna Cassidy-Hanley, **Jennifer Fricke Pinello**, Hong Zeng, Marion Hsueh, Daniel Kolbin, Courtney Ozzello, Thomas Giddings, Jr., Mark Winey, and Theodore G. Clark, Function of the Male-Gamete Specific Fusion Protein HAP2 in a Seven-Sexed Ciliate, 2168-2173, ©2014, with permission from Elsevier, *http://dx.doi.org/10.1016/j.cub.2014.07.064*.

Summary

HAP2, a male-gamete-specific protein conserved across vast evolutionary distances, has garnered considerable attention as a potential membrane fusogen required for fertilization in taxa ranging from protozoa and green algae to flowering plants and invertebrate animals[1–6]. However, its presence in *Tetrahymena thermophila*, a ciliated protozoan with seven sexes or mating types that bypasses the production of male gametes, raises interesting questions regarding the evolutionary origins of gamete-specific functions in sexually dimorphic species. Here we show that HAP2 is expressed in all seven mating types of *T. thermophila* and that fertility is only blocked when the gene is deleted from both cells of a mating pair. HAP2 deletion strains of complementary mating types can recognize one another and form pairs; however, pair stability is compromised and membrane pore formation at the nuclear exchange junction is blocked. The absence of pore formation is consistent with previous studies suggesting a role for HAP2 in gamete fusion in other systems. We propose a model in which each of the several hundred membrane pores established at the conjugation junction of mating *Tetrahymena* represents the equivalent of a male/female interface, and that pore formation is driven on both sides of the junction by the presence of HAP2. Such a model supports the idea that many of the disparate functions of sperm and egg were shared by the ‘‘isogametes’’ of early eukaryotes and became partitioned to either male or female sex cells later in evolution.

Results

Genetic evidence for HAP2 function in multiple mating types

Tetrahymena thermophila has seven sexes or mating types, any of which can mate with the other six, but not with itself[7,8]. Cells become competent to mate following nutritional starvation. When mixed together, starved cells of complementary mating types undergo a period of morphological transformation, producing a region of smooth, deciliated membrane at their anterior ends where they subsequently pair[9–11]. Approximately 1 hr after mixing at 30ºC, cells begin to form loose associations that are easily disrupted. By 2 hr, these loose associations give way to tight pairing, and a series of meiotic events ensues in which two haploid pronuclei, one stationary and one migratory, are generated in each cell of the mating pair (diagrammed in Figure 1.1). The transformation of a loose adhesion into a mature, mechanically robust mating junction coincides with the appearance of 0.1–0.2 μm diameter intercellular pores that form as a result of several hundred independent membrane fusion events. Over time, these pores expand up to ten times their initial diameter, allowing the exchange of migratory pronuclei[11,12]. Following exchange, migratory and stationary pronuclei fuse, membrane integrity is restored, and cells disengage to complete the program of conjugal development.

The macronuclear genome sequence of *T. thermophila* predicts a single homolog encoding HAP2, a male-gamete-specific protein thought to play a role in sperm/egg fusion in a wide range of species[1–6]. To confirm this, we obtained a full-length cDNA and deduced amino acid sequence for the protein (GenBank accession number KJ629172). Importantly, analysis of the Tetrahymena Functional Genomics Database indicated that the *HAP2* gene is highly upregulated when cells of complementary mating types are mixed[13]. With the idea that HAP2 is required for membrane pore formation in conjugating *Tetrahymena*, we initiated experiments to examine the effect of *HAP2* deletion on mating success as defined by the completion of all the normal events of conjugation including cross-fertilization and new macronucleus development. Initially, we found that deletion of the HAP2 gene from the vegetative macronucleus of *T. thermophila* mating type VII (strain ΔHAP2-428) had no effect on progeny

development when crossed with wild-type cells of mating type II (SB1969; data not shown). One interpretation of this result was that HAP2 is dispensable for mating in *Tetrahymena*. Alternatively, HAP2 expression could be mating-type specific (and confined to mating types other than VII), or ubiquitous in all mating types but required on only one side of a pair to allow membrane fusion and pronuclear exchange to occur. To explore this further, we performed additional crosses using *HAP2* deletion strains (Δ*HAP2*) carrying different drug-selectable markers in their germline micronuclei. Reciprocal matings were conducted between {wild-type x Δ*HAP2*} cells, {Δ*HAP2* x Δ*HAP2*} cells, and control {wild-type x wild-type} cells of different mating types, with the parents of each cross containing either cycloheximide (Cy) or 6-methyl purine (6-MP) resistance markers in their micronuclei. In such crosses, fertilization success was scored as the percentage of mating pairs that gave rise to progeny resistant to both drugs (see Figure S-2.3). The results are shown in Table 2.1. Whereas reciprocal crosses between knockout cell lines and cells carrying the wild-type *HAP2* gene were clearly fertile (giving rise to double drug resistant progeny), deletion of *HAP2* from both mating types completely blocked the appearance of cross-fertilized progeny. To test the penetrance of the infertility phenotype, we performed mass matings followed by sequential drug challenge. From mating cultures of {Δ*HAP2* x Δ*HAP2*} cell lines involving over 10^6 mating pairs, no double drug-resistant progeny were produced. Insertion of either the wild-type gene or an HA-tagged version of the full-length *HAP2* cDNA into the endogenous *HAP2* locus of both cells of a {Δ*HAP2* x Δ*HAP2*} cross restored fertility, demonstrating the necessity of HAP2 in cross-fertilization and mating success (Table 2.1).

Table 2.1 Effects of *HAP2* deletion on mating success. Reciprocal crosses (x) between wild type and *HAP2* knockout (Δ*HAP2*) strains were carried out in various combinations. Δ*HAP2*[a] and Δ*HAP2*[b] are, respectively CU428.2 and CU427.4 cell lines in which the *HAP2* gene has been completely replaced by a neomycin resistance cassette. For each cross, pairs were individually isolated into drops of media and allowed to establish "synclones" (mixed clonal descendants from both exconjugants), which were characterized phenotypically as explained in detail in Figure S-2.3. Three types of progeny synclones were phenotypically distinguished: those that successfully completed all steps of conjugation and developed a new macronucleus (cross-fertilizers; *R/R*), those that were blocked in gamete pronucleus exchange but proceeded normally otherwise (self-fertilizers; *R/S* or *S/R*), and those that failed to complete development of new macronuclei and retained their parental macronucleus (back-outs; *S/S*). % survival is the percent of isolated pairs that generated living cultures. The percentages of the three classes of progeny are shown in the remaining columns. The data presented for each cross are the pooled results from two to four independent pair isolation experiments. The actual number of pairs representing a particular outcome out of the total number of surviving pairs is shown in parentheses. Among progeny that developed a new macronucleus, the fraction of true progeny (i.e., cross-fertilizers) was directly correlated with the number of wild-type *HAP2* parents in the cross (2, 1, or 0), and the differences were highly statistically significant ($C^2_{6, n=1918} = 781.7$, $p<0.0001$). Rescue strains are cell lines in which *HAP2* gene knock-outs were replaced in the macronucleus with either the wild-type copy or cDNA version of the gene, in both parents of the cross (see Methods and Figure S-2.6). Because of the drug phenotype of the rescue strains, "self-fertilizers" and "back-outs" were not distinguishable from one another (indicated as NA, not applicable)

Mating	**% Survival**	**% "Cross-fertilizers"** ***R/R***	**% "Self-fertilizers"** ***R/S* or *S/R***	**% "Back-outs"** ***S/S***
CU427.4 x CU428.2	96 (404/423)	90 (364/404)	2 (9/404)	8 (31/404)
ΔHAP2[a] x CU427.4	91 (514/564)	17 (86/514)	12 (63/514)	71 (365/514)
ΔHAP2[b] x CU428.2	91 (774/846)	28 (217/774)	7 (55/774)	65 (502/774)
ΔHAP2[a] x *ΔHAP2*[b]	96 (226/235)	0 (0/226)	4 (9/226)	96 (217/226)
Genomic Rescue Cross	86 (284/329)	27 (77/284)	NA	NA
HAP2-HA cDNA Rescue Cross	66 (217/329)	25 (54/217)	NA	NA

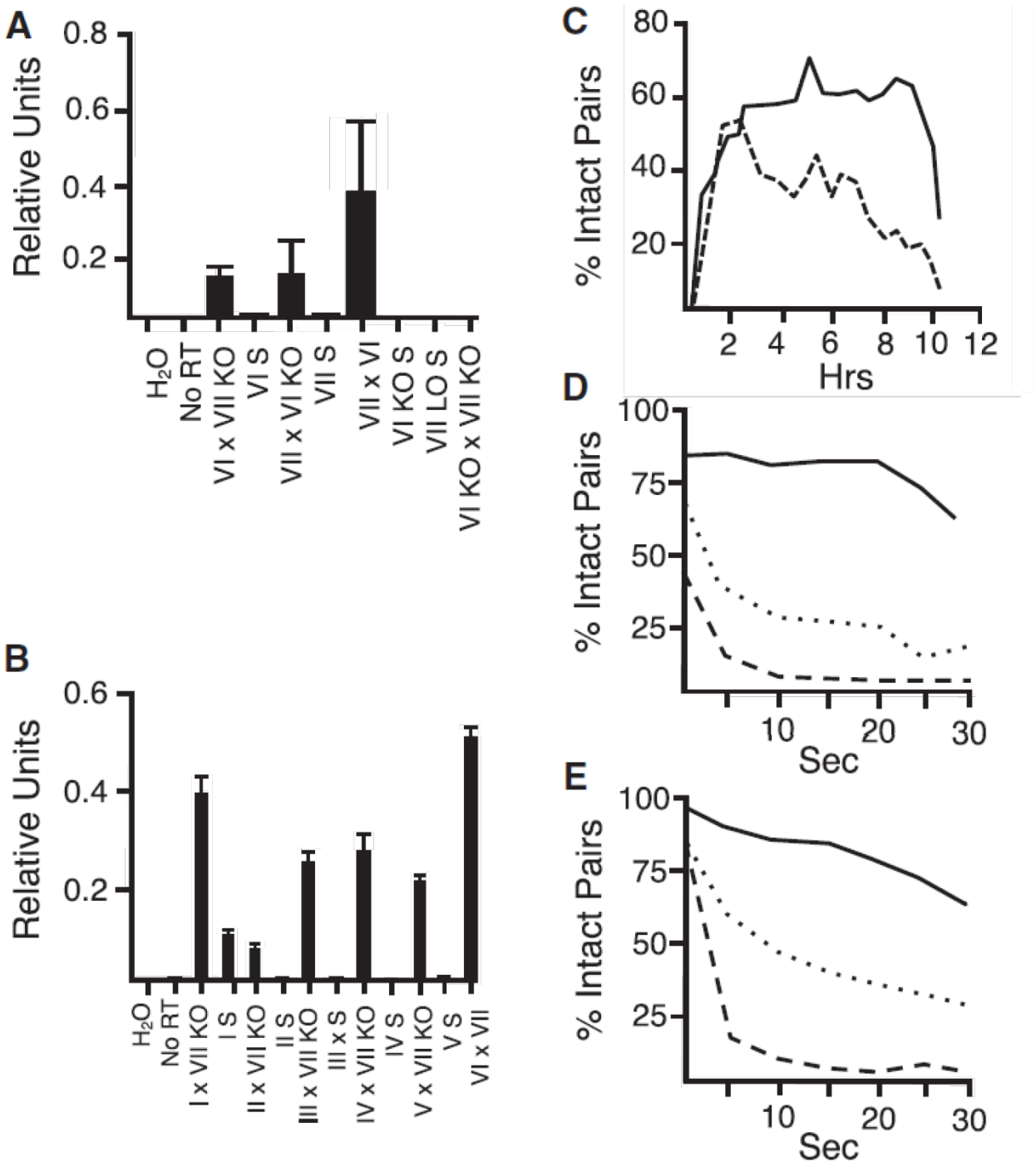
Expression of HAP2 in all seven mating types of *T. thermophila*

Taken together, the results described above indicate that HAP2 is functional in both members of a mating pair regardless of mating type. To test this further, we conducted quantitative real-time PCR experiments to determine *HAP2* mRNA levels, and we found that the gene was indeed expressed in reciprocal {wild-type x Δ*HAP2*} crosses of mating types VI and VII, regardless of which mating type carried the wild-type allele. In such crosses, *HAP2* transcript levels were roughly half of those in control {wild-type x wild-type} crosses (Figure 2.2A). Furthermore, in crosses between the Δ*HAP2* knockout in mating type VII and wild-type cells from all other mating types (i.e., I–V), *HAP2* mRNA was always detected, verifying that expression is not confined to any specific mating type. All seven sexes of *T. thermophila* express the *HAP2* gene during early conjugal development (Figure 2.2B).

Requirement for HAP2 in pair stability

At the light-microscopy level, deletion of *HAP2* had little or no effect on the ability of complementary mating types to recognize one another and form pairs. Nevertheless, when left undisturbed, {Δ*HAP2* x Δ*HAP2*} pairs came apart more rapidly than {wild-type x wild-type} pairs (Figure 2.2C). To quantify pair stability, we exposed mating cells to varying periods of mechanical agitation and counted the ratio of intact pairs to single cells in the culture. Figure 2.2D shows that wild-type pairs were relatively insensitive to vortexing, while {Δ*HAP2* x Δ*HAP2*} partners were highly sensitive, with over 90% of pairs separating after 5 s of agitation. We tested pairs from both early (2 hr) and mid-stage matings (4 hr) and obtained similar results, indicating that the loss of stability of {Δ*HAP2* x Δ*HAP2*} mating pairs at 2 hr was not simply the result of delayed development. Interestingly, {wild-type x Δ*HAP2*} partners showed an intermediate level of sensitivity to vortexing (Figure 2.2E), which correlates with *HAP2* expression at the transcript level.

Figure 2.2 *HAP2* transcript levels and pair stability in wild type and *ΔHAP2* crosses. **(A)** *HAP2* mRNA transcript levels in control and reciprocal {wild type x *ΔHAP2*} and {*ΔHAP2* x *ΔHAP2*} crosses as determined by quantitative reverse transcriptase PCR. {wild type x *ΔHAP2*} crosses showed roughly half the level of expression seen in wild type control matings, while { *ΔHAP2 x ΔHAP2} crosses showed no detectable HAP2* expression. Knockout (KO) strains used in this study were ΔHAP2-428 (mating type VII; Methods) and ΔHAP2-427 (mating type VI; Methods). Error bars indicate the average expression levels +/- SEM (n=3). **(B)** Relative *HAP2* transcript levels in crosses between mating type VII knockout strain ΔHAP2-428 and wild type strains of mating types 1-V (strains SB3539, B2086.2, SB281, CU438.1, and C3 368.1, respectively; Methods). *HAP2* mRNA expression in a wild type cross between mating types VII and VI (CU428.2 and CU427.4) is shown at the far right. Error bars represent mean and SEM for technical triplicates of cDNA samples from either mating (x) or starvation (S) cultures. **(C)** Wild type cells (CU428.2 and CU427.4) or *ΔHAP2* strains (ΔHAP2-428 and ΔHAP2-427) were starved, mixed together at time zero, and measured for pair frequency at 30 min intervals thereafter. **(D and E)** Cells were mixed and then physically disrupted using a Vortex-Genie at (D) 2-2.5 hr or (E) 4-4.5 hr after mixing. Cells were carefully withdrawn from culture vessels and gently placed on microscope slides before visualization under bright field. For each time point, 100 subjects were counted three separate times, and the percent of intact pair was averaged for each time point. Each experiment took 30 min to conduct. A diagram depicting the cellular events associated with conjugation of wild type cells at different time points is shown in Figure S-2.3.


A
Relative Units
0.8
0.6
0.4
0.2
H2O
No RT
VI x VII KO
VI S
VII x VI KO
VII S
VII x VI
VI KO S
VII LO S
VI KO x VII KO
B
Relative Units
0.6
0.4
0.2
H2O
No RT
I x VII KO
I S
II x VII KO
II S
III x VII KO
III x S
IV x VII KO
IV S
V x VII KO
V S
VI x VII
C
% Intact Pairs
80
60
40
20
2
4
6
8
10
12
Hrs
D
% Intact Pairs
100
75
50
25
10
20
30
Sec
E
% Intact Pairs
100
75
50
25
10
20
30
Sec

HAP2 is required for membrane pore formation

Bilateral exchange of gametic pronuclei in *T. thermophila* has been linked to the formation of intercellular pores at the nuclear exchange junction[12,14]. Using transmission electron microscopy (TEM), we found that HAP2 was necessary for the formation of intercellular pores at this site, and consequently exchange of gametic pronuclei between mating cells. As shown in Figure 2.3A and 2.3E, respectively, {wild-type x wildtype} and {Δ*HAP2* x Δ*HAP2*} mating pairs exhibited smooth adhesion zones between closely apposed plasma membranes, along with membrane outpocketings extending from both partners into the intercellular cleft between cells, beginning with the loose-pairing interval at 1–1.5 hr post mixing. In the wild-type pairs, these membrane protrusions have been shown to fuse with the plasma membrane of partnering cells (E.S.C., unpublished data), and approximately 2 hr after cell mixing, pores could be observed at numerous loci within the adhesion zone (Figure 2.3A–D). By contrast, pores were not observed in {Δ*HAP2* x Δ*HAP2*} pairs examined at 2 hr (Figure 2.3E–H), or at later time points even when meiotic nuclear configurations were observed at the mating junction (Figure 2.3F). The presence of large, seemingly hypertrophied membrane vesicles within the junctional lumens of {Δ*HAP2* x Δ*HAP2*} pairs were likely a consequence of the inability of membrane protrusions to fuse and form pores (Figure 2.3G). A proposed model for the events at the mating junction of {Δ*HAP2* x Δ*HAP2*} and {wild-type x wild-type} crosses is shown in Figure S-2.4.

HAP2 localizes to the nuclear exchange junction

Consistent with a role for HAP2 in membrane pore formation, a GFP-tagged version of the HAP2 protein localized predominantly to the interface between conjugating cells, decorating the entire mating junction, excluding established pores (Figure 2.4A–C). Immunoelectron microscopy of freeze-substituted material showed some cytoplasmic labeling of the GFP-tagged protein (consistent with a site of synthesis within the endomembrane system), as well as localization to the extracellular space between fusion pores at the conjugation junction (Figure S-2.5). Because the chimeric GFP-tagged fusion protein was overexpressed in wild-type cells in the presence of the native HAP2 protein, we conducted additional localization studies using

Δ*HAP2* deletion strains complemented with a C-terminal, HA-tagged version of the full-length *HAP2* cDNA in both strains crossed. As shown in Figure 2.4E and F, the epitope-tagged version of HAP2 localized to the conjugation junction in a pattern almost identical to that of the HAP2::GFP fusion protein.

Discussion

The experiments described here show that HAP2 is expressed in all mating types of *T. thermophila* and functions in both cells of a mating pair to allow fertilization to occur. Interestingly, cells of complementary mating types can pair in the complete absence of HAP2, indicating that the protein acts downstream of mating type recognition and cell adhesion, which are likely controlled by multiple factors, beginning with cell-to-cell contact[15,16] and interaction of products of the mating type alleles[8]. Nevertheless, the fact that HAP2 both localizes to the conjugation junction and is required for membrane pore formation provides strong support for previous arguments that HAP2 acts as a gamete fusogen either by itself or in conjunction with other proteins[1–6]. Finally, we observed that the physical interaction of mating cells was stabilized by HAP2. Rather than a programmed event, premature dissociation of {Δ*HAP2* x Δ*HAP2*} knockout pairs was likely the result of an inability to form pores and highlights a novel role for membrane fusion in pair stabilization that may have functional consequences in species other than *Tetrahymena* with highly motile gametes.

It is worth noting that the frequencies of progeny development in {wild-type x Δ*HAP2*} crosses and {Δ*HAP2* x Δ*HAP2*} genomic rescue crosses (in which the *HAP2* knockout construct was partially replaced in both parental strains with the native *HAP2* gene) were less than that in control {wildtype x wild-type} matings (Table 2.1). In each case, this may have been due to a dosage effect, since in the {wild-type x Δ*HAP2*} crosses *HAP2* was expressed in only one cell of a mating pair, and in the genomic rescue crosses we were unable to restore the gene to its full 45N copy number in the macronuclei of the rescued Δ*HAP2* knockout cell lines for technical reasons. Although other possibilities exist, a reduction in the number of membrane pores, a

Figure 2.3 Ultrastructure of the nuclear exchange junction. (A-D) TEM images of wild-type nuclear exchange junctions. (E-H) TEM images of {*ΔHAP2* x *ΔHAP2*} mating junctions. **(A)** Wild-type junction showing a membrane tubule (black arrow) protruding into the extracellular space 2 hr in to mating. White arrow indicates adjacent pore. Scale bar represents 100 nm. **(B)** Low-magnification view of a wild-type junction 4 hr into mating, showing multiple, complete junction pores (region bracketed by white arrows). Black arrows indicate meiotic nucleus in the extended "crescent" (or prophase I) configuration. This illustrates that in matings of wild-types cells, pores are complete by the onset of meiosis I. Scale bar represents 500 nm. **(C)** Wild-type junction (2 hr into mating) showing a typical junction pore (arrow). Scale bar represents 100 nm. **(D)** Image of wild-type (4 hr) junction showing multiple complete fusion pores. (Scale bar represents 500 nm). **(E)** {*ΔHAP2* x *ΔHAP2*} junction at 2 hr showing protrusions of plasma membrane into extracellular space (arrows). Scale bar represents 100 nm. **(F)** Low-magnification image of a {*ΔHAP2* x *ΔHAP2*} pair during late meiosis I (~4 hr) showing the complete junction devoid of pores. Lower arrows indicate condensed nuclear chromatin. The upper horizontal arrow indicates the "neck" of the intranuclear meiotic spindle full of microtubules. This pair was at late anaphase or early telophase of meiosis I. The vertical arrow at the top indicates the exchange junction, devoid of pores. **(G)** {*ΔHAP2* x *ΔHAP2*} pair at 2 hr showing a region of the junction with swollen extracellular vesicles collecting in the junction cleft. Note also the absence of pores. **(H)** Complete {*ΔHAP2* x *ΔHAP2*} exchange junction at 2 hr showing complete absence of fusion pores. Scale bar represents 500 nm. A diagram of membrane events at the junctions of wild-type and *ΔHAP2* crosses is shown in Figure S-2.4.

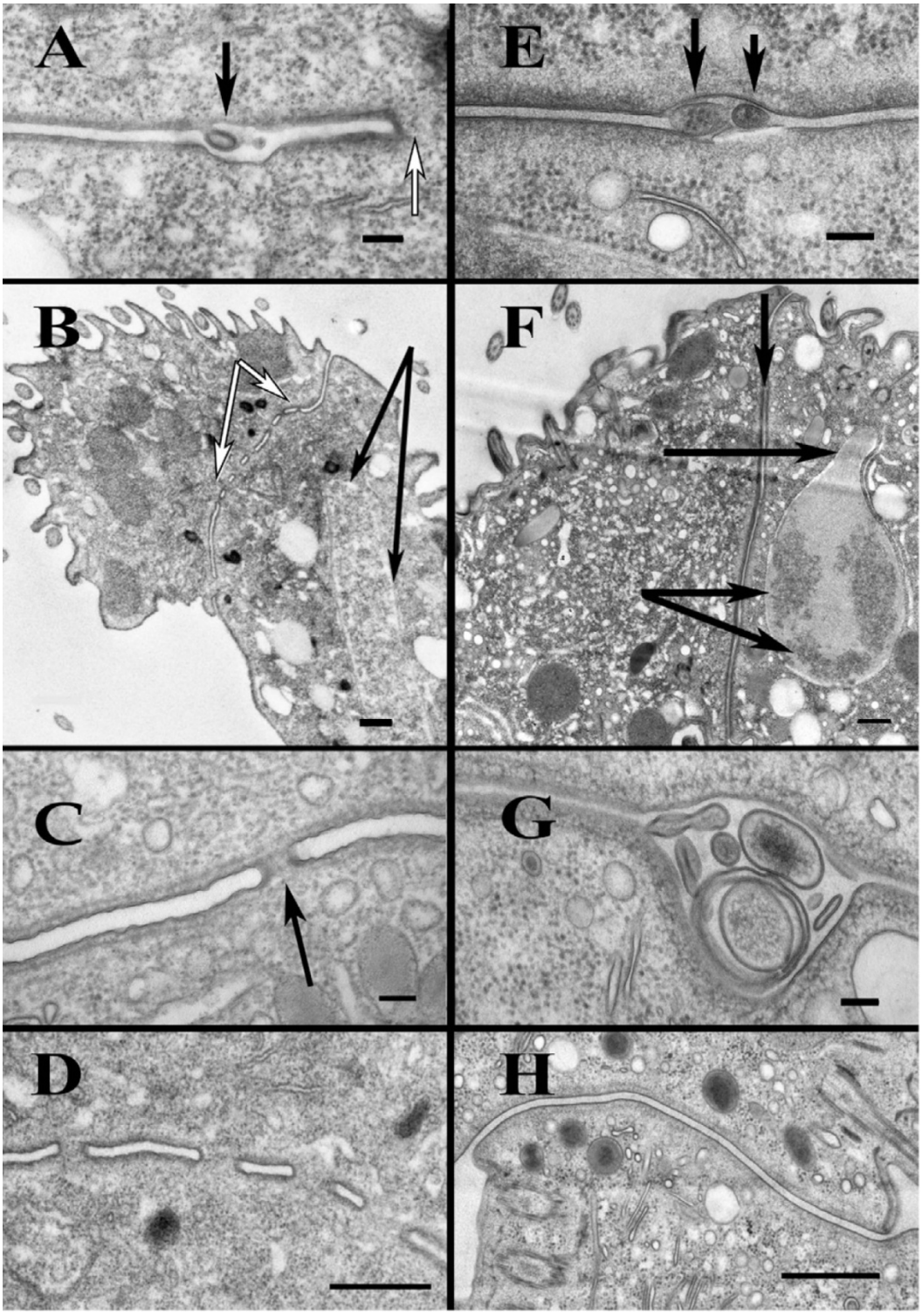
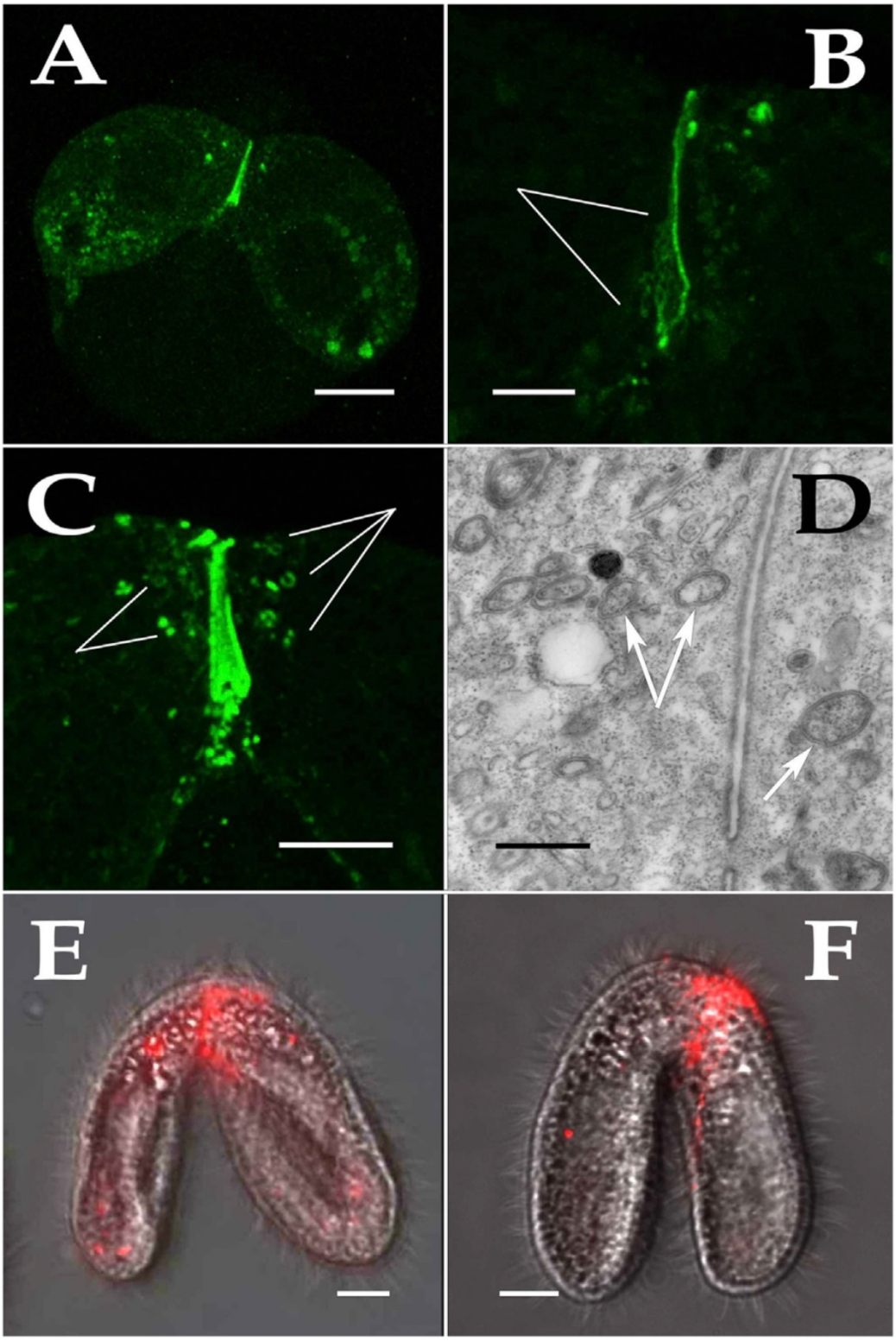

A
E
B
F
C
G
D
H

Figure 2.4 Localization of HAP2 in mating cells. (A-D) A chimeric *HAP2::GFP* fusion construct under the control of a cadmium-inducible promoter was introduced into the macronucleus of *T. thermophila* mating type IV cells (strain CU522). Transformants were starved, induced with $CdCl_2$, and mixed with mating type II cells under starvation conditions. At varying times thereafter, live mating cells were visualized using fluorescence optics. **(A)** A mating pair 3 hr after mixing, with strong GFP fluorescence over the mating junction. Scale bar represents 20 μm. **(B)** A higher magnification image of the mating junction from a later pair (3.5 hr) shown on edge. White lines bracket a region perforated by pores. Scale bar represents 5 μm. **(C)** Mating junction at 4 hr with punctate staining (semicircular and circular profiles) on either side of the margins of the junction (white lines). This punctate staining pattern resembles autophagosomes in the same region of the mating junction shown in (D). Scale bar represents 10 μm. **(D)** Transmission electron micrograph through a region near the mating junction (1.5 hr into mating) with autophagic vacuoles (arrows) shown in cross-section. Scale bar represents 500 nm. Immunogold labeling of GFP-tagged HAP2 at the ultrastructural level is shown in Figure S-2.5. **(E and F)** A C-terminal HA-tagged version of the full-length *HAP2* cDNA was introduced into the endogenous *HAP2* locus of *ΔHAP2* knockout strains (mating types VI and VII) by homologous recombination. Cells were mated, fixed and labeled with mouse anti-HA antibodies and secondary rhodamine red conjugated goat anti-mouse IgG as described in the text. Two different mating pairs are shown in (E) and (F). Scale bars represent 10 μm.


A
B
C
D
E
F

decrease in pair stability, or both could account for the reduction in progeny development seen in these matings.

Previous studies have shown that HAP2 function is restricted to male gametes in sexually dimorphic, anisogamous species (i.e., organisms whose gametes are dissimilar in form and function)[1–5]. Nevertheless, isogamous organisms almost certainly arose before anisogamous species[17,18], and phenotypic traits that are now fixed in either sperm or egg (such as HAP2 activity) may well have been shared by the gametes of ancestral, isogamous life forms. In this regard, *Tetrahymena*, although producing neither male nor female gametes, may offer a glimpse into how the problem of gamete fusion was initially solved in the earliest eukaryotes, namely, through the formation of one or more pores (fusion events) initiated independently by both cells of a mating pair, each exhibiting either ‘‘male’’ or ‘‘female’’ character at given points of membrane contact at their interface. In this regard, HAP2 expression patterns have been examined in three isogamous species other than *Tetrahymena*: the slime mold *Physarum polycephalum*[2] and two algal species; *Chlamydomonas reinhardtii*[4] and *Gonium pectorale*[19]. Consistent with what we describe here, HAP2 was reported to be expressed in the two mating types of *Physarum*, and in both plus and minus mating types of *Chlamydomonas* and *Gonium*[4,19]. While these observations suggest that HAP2 expression is not restricted to particular mating types in isogamous species, deletion of the *HAP2* gene from the *minus* (‘‘male’’) but not the *plus* (‘‘female’’) mating type of *C. reinhardtii* blocks fertilization[4], and while HAP2 is made, the protein is rapidly degraded in *plus* gametes of *G. pectorale* following gamete activation[19]. On the one hand, the apparent vestigial nature of HAP2 expression in the *plus* mating types of these species (along with the molecular and ultrastructural differences in their gametes[19–22]) may indicate they are in transition from isogamous to anisogamous forms[23,24]. At the same time, *Tetrahymena* is itself anomalous, in that it bypasses the production of ‘‘male’’ and ‘‘female’’ gametes altogether and instead produces stationary (‘‘female’’) and migratory (‘‘male’’) pronuclei that are exchanged between complementary mating types. Thus, the requirement for a male-gamete-specific fusogen, HAP2, in all mating

types of *T. thermophila* might be an adaptation to its nuclear exchange behavior rather than a characteristic of isogamous life forms in general.

Although it remains to be determined whether *T. thermophila* represents the exception rather than the rule in terms of HAP2 expression in isogamous species, an analysis of membrane dynamics in mating cells indicates that membrane fusion events, which are typically initiated by male gametes, are driven on both sides of the conjugation junction in mating *Tetrahymena*. This begins with the formation of membrane protrusions, or tubules of about the size of the singular mating structure in *minus* gametes of *C. reinhardtii* (i.e.,~50 nm diameter), which extend from the plasma membranes of each partner into the junction cleft. As shown here, in the absence of HAP2 these membrane events appeared to initiate normally, with protrusions extending from both cells into the junction cleft. However, no pores formed, and cytoplasmic continuity was never established in the critical developmental interval immediately following cell adhesion and the formation of membrane protrusions. This, and the fact that HAP2 appeared targeted to a region of specialized membrane (the conjugation junction) where pore formation occurs, argues strongly for a role for HAP2 in membrane fusion.

From a practical standpoint, the ability to induce synchronous mass mating in cells that can be cultured to high density on a large scale, along with methods for isolating the conjugation junction itself[25], makes *Tetrahymena* a potentially powerful system in which to examine HAP2 function at a biochemical level. It should also be noted that a recent screen of proteins upregulated during conjugation in *T. thermophila* led to the identification of a gene encoding a predicted zinc finger domain-containing protein (*ZFR1*; TTHERM_01285910), which localizes to the conjugation junction of mating cells and, as with HAP2, appears necessary for normal fertility and pair stability[26]. The possibility that these and other proteins act in concert to drive membrane fusion, and perhaps pore expansion, clearly bears further investigation. Finally, although these studies provide strong evidence of a role for HAP2 in membrane fusion, *Tetrahymena* and other ciliate species appear unique in their ability to limit cell-cell fusion and to reverse the establishment of cytoplasmic continuity between cells during the mating process.

Materials and methods

***Tetrahymena* strains and culture conditions.** *Tetrahymena* strains were obtained from the Tetrahymena Stock Center, Cornell University. These and all new strains constructed in this study are described in detail in Table S-2.1: *Tetrahymena thermophila* Strains. Cells were grown at 30°C in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 33.3μM $FeCl_3$) on a platform shaker at ~ 100 rpm. For mating, log phase cells were starved at ~2 ×10^5 cells/ml for 12-18 hrs in 10 mM Tris buffer (pH 7.5), and mixed in equal concentration at 30°C. For somatic transformation, target cells were grown to mid-to-late log stage (~6 × 10^5 – 1 × 10^6 cells/ml) in NEFF medium, and starved overnight in 10 mM Tris buffer (pH 7.5) at ~2 × 10^5 cells/ml prior to biolistic transformation[27,28]. Following transformation, cells were transferred to NEFF medium containing 250μg/ml Penicillin G, 250 μg/ml Streptomycin, and 1.25μg/ml Fungizone, and grown at 30°C.

***Tetrahymena* strain construction.**

Δ***HAP2 deletion strains*:** The endogenous *HAP2* gene was knocked out in macronuclei of *T. thermophila* strains CU428.2, CU427.4, and B2086.2 by homologous recombination using a Neo4 drug resistance cassette[29]. Knockout construct Neo4 HAP2 KO was made by separately amplifying 5' and 3' flanking regions of the HAP2 gene by PCR using *T. thermophila* genomic DNA as the template, and primer pairs HAP2_5' flankfor/HAP2_5' flankrevXHOI and HAP2_3'forSACI/HAP2_3' flankrev, respectively (design of DNA constructs for strain development along with PCR primer sequences are shown in Figure S-2.6: Construct Design; and, Table S-2.2: PCR Primers, respectively). The Neo resistance cassette was amplified from the Neo4 vector[29] using primer pair Neo4for2XhoI/Neo4cyrevSacI. The three PCR products were cleaved with the appropriate restriction enzymes (either *Xho*I, *Sac*I or both), gel purified and sequentially ligated in reactions using T4 DNA ligase (New England Biolabs) at 16ºC overnight. The final ligation product was gel purified, blunt-end cloned into PCR4BLUNTTOPO and sequenced. The resulting transformation vector was linearized and transformed without

further purification into the somatic macronuclei via biolistic bombardment[27,30]. Δ*HAP2* clones were selected initially in NEFF medium containing 70 µg/ml paromomycin, and pushed to complete macronuclear gene replacement by incrementally increasing the paromomycin concentration to 800 µg/ml. Complete replacement of the endogenous gene by the knockout construct was verified in each case by PCR using genomic DNA from transformed clones and primers that spanned the insertion site (HAP25' flankfor and HAP23' flankrev). Strains that showed only PCR fragments with sizes expected for the knockout construct (ΔHAP2-428, ΔHAP2-427) were grown for several more transfers in 800 µg/ml paromomycin, and then frozen in liquid nitrogen[31].

GFP-tagged HAP2 expressing strains For GFP-expressing cell lines, *Bcl*I and *Mlu*I sites were added to the 5' and 3' ends, respectively, of full-length *HAP2* cDNA by PCR, and the resulting product cloned into corresponding restriction sites in the plasmid vector, pMTT1-NRK2-GFP [31], to create pMTT1-HAP2-GFP (panel D, Figure S-2.6: Construct Design). Vector DNA was then introduced into modified *T. thermophila* strain, CU522 (mating type IV) via biolistic bombardment [29] leading to insertion of the expression cassette (including the cadmium-inducible MTT1 promoter) into the β-tubulin-1 locus via homologous recombination[32]. Positive transformants (designated HAP2GFP522) were selected by growth in 20 µM taxol.

***Genomic HAP2 and HA-tagged HAP2 cDNA rescue strains*:** A genomic *HAP2* rescue construct was created by amplifying the genomic *HAP2* gene and adjacent flanking regions, and inserting the pHrpl29-B cycloheximide resistance selection cassette (J. Bowen, personal communication) within the flanking region, ~ 400 bp downstream of the 3' end of the *HAP2* gene (panel B, Figure S-2.6: Construct Design). Three separate PCR products were generated using primers shown in Table S-2.2: PCR Primers, and the resulting products cleaved with the appropriate restriction enzymes (either *Xho*I, *Sac*I or both), gel purified and sequentially ligated in reactions using T4 DNA ligase as above. The final ligation product was gel purified and blunt-end cloned into plasmid, PCR4BLUNT-TOPO. A similar HA-tagged HAP2 cDNA rescue

construct was created by introducing *Bam*HI and *Kpn*I sites into the genomic rescue vector 5' and 3', respectively, of the genomic HAP2 gene, and placing the same restriction sites at the 5' and 3' ends of the HA-tagged HAP2 cDNA (panel C, Figure S-2.6: Construct Design), using primers listed in Table S-2.2: PCR Primers. Following PCR, both vector and cDNA insert were cleaved with *Bam*HI and *Kpn*I, gel purified, and ligated using T4 DNA ligase. Both the genomic and cDNA vectors were transformed into *E. coli* and frozen at -80°C. For use in *Tetrahymena* transformation, each vector was linearized and transformed without further purification into the somatic macronuclei of Δ*HAP2* deletion strains of mating type VII (CU428.2) and mating type VI (CU427.4). Separate transformations for each vector and strain combination were carried out via biolistic bombardment as above. *HAP2* rescue clones (HAP2genomicRes428, HAP2genomicRes427, HAP2cDNARes428, HAP2cDNARes427 (Table S-2.1: *Tetrahymena thermophila* strains) were selected in NEFF medium containing 25 μg/ml cycloheximide.

Quantitative PCR. Internal primers for quantitation of *Tetrahymena HAP2* gene expression (viz. TthermPair2for and TthermPair2rev (Table S-2.2: PCR Primers) were designed using PrimerExpress3 software (Life Technologies). PCR reactions were carried out in triplicate in SyberSelect MasterMix (Invitrogen) in 10μl reactions using an ABI Viia7 Real-Time thermocycler (Life Technologies) and first-strand cDNA as the template. To generate a quantitation standard, synthetic *HAP2* sense RNA was made by *in vitro* transcription using SP6 RNA polymerase to drive expression from the full-length *T. thermophila HAP2* cDNA in the pCR4Blunt-TOPO vector according to the manufacturer's instructions (Promega). Plasmid DNA was linearized before use in *in vitro* transcription reactions. SP6 generated RNA was quantified with a QuaWell UV spectrophotometer and stored in aliquots at -80°C. Total RNA from samples of mating cells, along the *in vitro* transcribed standard, were then used to prepare template cDNA using the Invitrogen SuperScript III First Strand Synthesis system for RT-PCR as described above. Real-time PCR was carried out with dilutions of cDNA from the *in vitro* synthesized RNA to generate a standard curve. The amount of *HAP2* transcript in experimental samples was finally

determined by comparison of C_ts generated from experimental samples relative to those of the standard curve. Reactions were run according to FastCycling Mode parameters; 1 cycle of 50°C for 2min, 95°C for 2min, followed by 40 cycles of 95°C for 1sec, 59°C for 30sec. The percent efficiency ranged from 94.55 – 95.67%, while the slope of the trend line was -3.46. A dissociation curve was performed immediately following the conclusion of the run, and did not contain any off-target amplicons.

Pair disruption assays. Mating pairs were physically disrupted using a Vortex GenieTM set at setting 4/5. Three mL samples of mating cells (at 5-7×10^5 cells/mL) were transferred to a 15 mL centrifuge tube. The tube was lowered onto the vibrating Vortex GenieTM platform as it was running, until contact was made and a vigorous swirl induced. Then at 5 second intervals, 11 µL samples were withdrawn with a P20 and laid on a clean microscope slide. A coverslip was lowered gently so that the edge of the coverslip was in contact with the slide (at an angle) and the drop made gentle contact with the coverslip before it was released (to minimize extra agitation). 100 “subjects” were counted for each time point (scored as single cells or pairs). The number of pairs/total “subjects” multiplied by 100 yielded % pair data. The same procedures were also used to measure pair stability over time in the absence of mechanical disruption as described in the text.

High-pressure freeze-substitution for TEM and immuno-gold TEM localization. Cultures with mating pairs of *Tetrahymena* were prepared for transmission electron microscopy or immuno-electron microscopy as described in Meehl et al.[33] and Giddings et al.[34]. Briefly, cells were harvested by centrifugation into a cryoprotectant solution consisting of 15% dextran (9-11KD, Sigma) and 5% BSA in 10 mM Tris and frozen in a Wohlwend Compact 02 High Pressure Freezer. Matings involving Δ*HAP2* mutants were centrifuged gently using a handcrank centrifuge to avoid pair disruption. For analysis of ultrastructure in wild type and mutant mating pairs, samples were freeze-substituted in 2% osmium tetroxide, 0.1% uranyl acetate in acetone and embedded in Spurr’s epoxy resin. Serial thin sections were imaged in a Philips CM100

TEM. For immunolocalization of the chimeric HAP2:GFP gene product, samples were freeze substituted in 0.25% glutaraldehyde, 0.1% uranyl acetate in acetone and embedded in Lowicryl HM20. Serial thin sections were immuno-labeled with a rabbit anti-GFP polyclonal antibody (kind gift from Chad Pearson, University of Colorado-Denver) followed by goat-anti-rabbit antibody conjugated to 15nm colloidal gold (Ted Pella, Inc.).

Imaging of GFP and HA-tagged *HAP2* transformants. For localization of HAP2 during mating, a cell line carrying the *HAP2::GFP* fusion gene (HAP2GFP522) was mated to a wild type partner (CU428.2). The engineered cell line was induced to begin expressing the HAP2::GFP fusion protein by applying CdCl2 at 0.1 µg/mL during the last 2 hrs of starvation. Mixing of the GFP-strain and a wild type partner at a ratio of 1:1 (vol:vol) resulted in a final concentration of 0.05 µg/mL CdCl2 during co-stimulation and pairing. Mating pairs in which one partner was expressing the GFP-fusion protein were then imaged using an Olympus BX-50 fluorescence microscope equipped with a Fluoview scanning laser confocal imaging system. Conventional fluorescence images were captured using a DP-72 digital camera.

To localize HA-tagged HAP2 in mating *ΔHAP2* rescue strains, mating pairs were isolated at 2.5-3 hrs after mixing and washed in 20mM HEPES before fixation for 1 hr in IC Fixation Buffer (eBioscience). Subsequent washes and antibody incubations were carried out in eBioscience 1x Permeabilization Buffer. Pairs were labeled with primary mouse antibody against the HA-tag (Covance) at dilutions of 1/200-1/500, followed by incubation with a secondary rhodamine red goat anti-mouse IgG at 1/300 dilution (Invitrogen). Pairs were imaged using a Zeiss Axio Imager M1 microscope equipped with an AxioCamMR3 camera. Composite images were assembled using AxioVision Software (AxioVs40 V 4.6.1.0).

Analysis of cross-fertilization via drug selection. Mating success was assayed by constructing cell lines of different mating types deleted for the *HAP2* gene within their expressed macronuclei, and carrying different drug-resistance markers in their germline micronuclei. The ΔHAP2-427 cell line carried a dominant cycloheximide resistance allele (Cyr) in its

micronucleus, while the other, ΔHAP2-428, carried a dominant 6-methyl-purine resistance allele (Mpr) in its micronucleus (see Figure S-2.3). Neither cell line carried a drug resistance marker in its macronucleus making the parental lines sensitive to both drugs. Assays were performed on cells derived from either single isolated pairs or mass cultures as follows. Six hours after mixing, mating cultures were diluted in growth medium (this discourages “late” pairing partners and causes dissociation of pairs very early in the mating program). Starting at 7 hours, (0.5-1 hr after feeding) persistent pairs were isolated by pipette into solitary culture. Typically, at least 96 pairs were isolated from each mating condition and their progeny were grown in 96-well microtiter plates after initial expansion in hanging drops. These “synclones” (mixed progeny expanded from both of the resulting exconjugants) were then replicated into plates containing either 15μg/ml 6-methyl purine or 25 μg/ml cycloheximide. After 4 days in drug, plates were scored for survival, and replicated to the reciprocal drug. For matings that produced <1/96 double-drug resistant progeny, mass drug selections were performed to detect the existence of rare cross fertilization progeny. In such experiments, an entire mating culture (>1×10^6 pairs), were diluted into growth medium, and subjected to sequential drug challenge *en masse*. HAP2 rescue strains were made by insertion of either the wild type *HAP2* gene, or an HA-tagged version of the full-length HAP2 cDNA into the Δ*HAP2* deletion strains, ΔHAP2-427 and ΔHAP2-428 (see panels B and C, Figure S-2.6: Construct Design). Rescue of mating success was determined as above by mating HAP2 rescue strains to one another.

REFERENCES

1. Johnson, M.A., Besser, K. von, Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z., and Preuss, D. (2004). Arabidopsis hapless mutations define essential gametophytic functions. Genetics *168*, 971–982.

2. Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). Generative Cell Specific 1 is essential for angiosperm fertilization. Nat. Cell Biol. *8*, 64–71.

3. Besser, K. von, Frank, A.C., Johnson, M.A., and Preuss, D. (2006). Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development *133*, 4761–4769.

4. Liu, Y., Tewari, R., Ning, J., Blagborough, A.M., Garbom, S., Pei, J., Grishin, N.V., Steele, R.E., Sinden, R.E., Snell, W.J., *et al.* (2008). The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in Chlamydomonas and Plasmodium gametes. Genes Dev. *22*, 1051–1068.

5. Steele, R.E., and Dana, C.E. (2009). Evolutionary history of the HAP2/GCS1 gene and sexual reproduction in metazoans. PloS One *4*, e7680.

6. Wong, J.L., and Johnson, M.A. (2010). Is HAP2-GCS1 an ancestral gamete fusogen? Trends Cell Biol. *20*, 134–141.

7. Nanney, D.L. (1953). Nucleo-Cytoplasmic Interaction during Conjugation in Tetrahymena. Biol. Bull. *105*, 133–148.

8. Cervantes, M.D., Hamilton, E.P., Xiong, J., Lawson, M.J., Yuan, D., Hadjithomas, M., Miao, W., and Orias, E. (2013). Selecting one of several mating types through gene segment joining and deletion in Tetrahymena thermophila. PLoS Biol. *11*, e1001518.

9. Suganuma, Y., Shimode, C., and Yamamoto, H. (1984). Conjugation in Tetrahymena: formation of a special junction area for conjugation during the co-stimulation period. J. Electron Microsc. (Tokyo) *33*, 10–18.

10. Wolfe, J. (1985). Cytoskeletal reorganization and plasma membrane fusion in conjugating Tetrahymena. J. Cell Sci. *73*, 69–85.

11. Orias, E. (1986). Ciliate Conjugation. In The Molecular Biology of Ciliated Protozoa, Gall, J.G., ed. (New York, NY: Academic Press), pp. 45–94.

12. Cole, E.S. (2006). The Tetrahymena Conjugation Junction. In Cell-Cell Channels, F. Baluska, D. Volkmann, and P. W. Barlow, eds. (New York, NY: Springer New York), pp. 39–62. Available at: http://dx.doi.org/10.1007/978-0-387-46957-7_3.

13. Xiong, J., Lu, Y., Feng, J., Yuan, D., Tian, M., Chang, Y., Fu, C., Wang, G., Zeng, H., and Miao, W. (2013). Tetrahymena functional genomics database (TetraFGD): an integrated resource for Tetrahymena functional genomics. Database J. Biol. Databases Curation *2013*, bat008.

14. Wolfe, J. (1982). The conjugation junction of Tetrahymena: Its structure and development. J. Morphol. *172*, 159–178.

15. Bruns, P.J., and Palestine, R.F. (1975). Costimulation in Tetrahymena pyriformis: A developmental interaction between specially prepared cells. Dev. Biol. *42*, 75–83.

16. Finley, M.J., and Bruns, P.J. (1980). Costimulation in Tetrahymena II. A nonspecific response to heterotypic cell-cell interactions. Dev. Biol. *79*, 81–94.

17. Parker, G.A., Baker, R.R., and Smith, V.G.F. (1972). The Origin and Evolution of Gamete Dimorphism and the Male-Female Phenomenon. J Theor Biol *36*, 529–553.

18. Smith, J.M. (1978). The Evolution of Sex (CUP Archive).

19. Kawai-Toyooka, H., Mori, T., Hamaji, T., Suzuki, M., Olson, B.J.S.C., Uemura, T., Ueda, T., Nakano, A., Toyoda, A., Fujiyama, A., *et al.* (2014). Sex-Specific Posttranslational Regulation of the Gamete Fusogen GCS1 in the Isogamous Volvocine Alga Gonium pectorale. Eukaryot. Cell *13*, 648–656.

20. Gametic differentiation in Chlamydomonas reinhardtii. III. Cell wall lysis and microfilament-associated mating structure activation in wild- type and mutant strains (1975). J. Cell Biol. *67*, 623–637.

21. Lin, H., and Goodenough, U.W. (2007). Gametogenesis in the Chlamydomonas reinhardtii minus Mating Type Is Controlled by Two Genes, MID and MTD1. Genetics *176*, 913–925.

22. Ning, J., Otto, T.D., Pfander, C., Schwach, F., Brochet, M., Bushell, E., Goulding, D., Sanders, M., Lefebvre, P.A., Pei, J., *et al.* (2013). Comparative genomics in Chlamydomonas and Plasmodium identifies an ancient nuclear envelope protein family essential for sexual reproduction in protists, fungi, plants, and vertebrates. Genes Dev. *27*, 1198–1215.

23. Umen, J.G. (2011). Evolution of sex and mating loci: An expanded view from Volvocine algae. Curr. Opin. Microbiol. *14*, 634–641.

24. Hiraide, R., Kawai-Toyooka, H., Hamaji, T., Matsuzaki, R., Kawafune, K., Abe, J., Sekimoto, H., Umen, J., and Nozaki, H. (2013). The Evolution of Male–Female Sexual Dimorphism Predates the Gender-Based Divergence of the Mating Locus Gene MAT3/RB. Mol. Biol. Evol. *30*, 1038–1040.

25. Cole, E.S., Anderson, P.C., Fulton, R.B., Majerus, M.E., Rooney, M.G., Savage, J.M., Chalker, D., Honts, J., Welch, M.E., Wentland, A.L., *et al.* (2008). A Proteomics Approach to Cloning Fenestrin from the Nuclear Exchange Junction of Tetrahymena. J. Eukaryot. Microbiol. *55*, 245–256.

26. Xu, J., Tian, H., Wang, W., and Liang, A. (2012). The zinc finger protein Zfr1p is localized specifically to conjugation junction and required for sexual development in Tetrahymena thermophila. PloS One *7*, e52799.

27. Bruns, P.J., and Cassidy-Hanley, D. (2000). Biolistic transformation of macro- and micronuclei. Methods Cell Biol. *62*, 501–512.

28. Seashell Technology - Carrier Particle Protocols for Plasmid DNA Available at: http://www.seashelltech.com/protocols.shtml [Accessed October 24, 2016].

29. Mochizuki, K. (2008). High efficiency transformation of Tetrahymena using a codon-optimized neomycin resistance gene. Gene *425*, 79–83.

30. Cassidy-Hanley, D., Bowen, J., Lee, J.H., Cole, E., VerPlank, L.A., Gaertig, J., Gorovsky, M.A., and Bruns, P.J. (1997). Germline and Somatic Transformation of Mating Tetrahymena thermophila by Particle Bombardment. Genetics *146*, 135–147.

31. Bruns, P.J., Smith, H.R., and Cassidy-Hanley, D. (2000). Long-term storage. Methods Cell Biol. *62*, 213–218.

32. Gaertig, J., Thatcher, T.H., Gu, L., and Gorovsky, M.A. (1994). Electroporation-mediated replacement of a positively and negatively selectable beta-tubulin gene in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. *91*, 4549–4553.

33. Meehl, J., Giddings, T., Jr., and Winey, M. (2010). High Pressure Freezing, Electron Microscopy, and Immuno-Electron Microscopy of Tetrahymena thermophila Basal Bodies. In Cytoskeleton Methods and Protocols Methods in Molecular Biology., R. H. Gavin, ed. (Humana Press), pp. 227–241. Available at: http://dx.doi.org/10.1007/978-1-60761-376-3_12 [Accessed October 24, 2016].

34. Giddings Jr., T.H., Meehl, J.B., Pearson, C.G., and Winey, M. (2010). Chapter 6 - Electron Tomography and Immuno-labeling of Tetrahymena thermophila Basal Bodies. In Methods in Cell Biology Electron Microscopy of Model Systems., T. Müller-Reichert, ed. (Academic Press), pp. 117–141. Available at: http://www.sciencedirect.com/science/article/pii/S0091679X10960068 [Accessed October 24, 2016].

Supplemental figures

Table S-2.1 *Tetrahymena thermophila* strains.

Strain[b]	Genotype (micronucleus)	Genotype; Phenotype (macronucleus)
CU428.2	*CHX1/CHX1*; *mpr1-1/mpr1-1*	*MPR1, CHX1;* mp-s, cy-s, VII
CU427.4	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1, CHX1;* mp-s, cy-s, VI
B2086.2	*CHX1/CHX1; MPR1/MPR1*	*MPR1, CHX1;* mp-s, cy-s, II
CU522	*mpr1-1/mpr1-1; btu1-1::btu1-1M350K/btu1-1::btu1-1M350K*	*mpr1-1, btu1-1::btu1-1M350K;* mp-r, ory-r, tax-s, IV
SB3539	*chx1[C3]-1/chx[C3]-1*	*CHX1[C3];* cy-s, I
CU438.1	*pmr1-1/pmr1-1*	*PMR1;* pm-s, IV
SB281	*exoB2/exoB2; gal1-1/gal1-1*	*exoB2, gal1-1;* exo-, gal-r, III
C3 368.1	*CHX1/CHX1[C3]; MPR1/MPR1[C3]*	*MPR1[C3], CHX1[C3]*; mp-s, cy-s, V
^HAP2Δ428	*CHX1/CHX1; mpr1-1/mpr1-1*	*MPR1, CHX1, hap2-1[Δ::neo4];* mp-s, cy-s, pm-r, VII
^HAP2Δ427	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1, CHX1, hap2-1[Δ::neo4];* mp-s, cy-s, pm-r, VI
^HAP2ΔB2086	*CHX1/CHX1; MPR1/MPR1*	*MPR1, CHX1, hap2-1[Δ::neo4];* mp-s, cy-s, pm-r, II
^HAP2GFP522	*mpr1-1/mpr1-1; btu1-1::btu1-1M350K/btu1-1::btu1-1M350K*	*mpr1-1, btu1-1::btu1-1M350K/btu1-12::hap2-2[GFP$_c$];* mp-r, ory-s, tax-r, IV
^HAP2genomicRes428	*CHX1/CHX1; mpr1-1/mpr1-1*	*MPR1, CHX1, hap2-1[Δ::neo4/hap2-3(3'cy2)];* mp-s, pm-r, cy-r, VII
^HAP2genomicRes427	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1, CHX1, hap2-1[Δ::neo4/hap2-3(3'cy2)];* mp-s, pm-r, cy-r, VI
^HAP2cDNARes428	*CHX1/CHX1; mpr1-1/mpr1-1*	*MPR1, CHX1, hap2 1[Δ::neo4/hap2cDNA (3'cy2)];* mp-s, pm-r, cy-r, VII
^HAP2cDNARes427	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1, CHX1, hap2 1[Δ::neo4/hap2cDNA (3'cy2)];* mp-s, pm-r, cy-r, VI

[b] CU428.2 and CU427.4 are functional heterokaryons, phenotypically sensitive to 6-methylpurine and cycloheximide respectively in the macronucleus, but homozygous for resistance to 6-methylpurine (CU428.2) or cycloheximide (CU427.4) in the micronucleus. B2086.2 is a wild type strain genotypically and phenotypically sensitive to both drugs. CU522 carries a mutation in the β-tubulin 1 gene conferring resistance to oryzalin and sensitivity to taxol. C3 368.1 and SB3539 are C3 strains of *T. thermophila*. All other strains listed are B strain *T. thermophila*. Strains marked with ^ are transformant strains created during the course of this project.

Table S-2.2 PCR Primers.

Primer Name[c]	Sequence, 5' → 3'	Primer Use
HAP2 5'flankfor	GTTATTTTCAGCATCTTCTTTCATTTG	HAP2Δ
HAP2 5'flankrevXHOI	TAAGTActcgagGCTAGAAGTAATATACTCACACCTGTTC	HAP2Δ
HAP2 3'flankforSacI	CTTATGgagctcTCTAAAAGTTGTTTCAACATCTCC	HAP2Δ
HAP2 3' flankrev	ATCTCTTCTGATCATAGAGCACC	HAP2Δ
Neo4for2XhoI	GACTTActcgagAATAAGGGTTTTGAATAACTCCT	Hap2Δ
Neo4cyrevSacI	ATTCTAgagctcTGCATTTTTCCAGTAAAAATTTGA	Hap2Δ, Rescue Hap2
HAP25'FLANKFor	GTTATTTTCAGCATCTTCTTTCATTTG	Rescue HAP2
Hap2Int3'FlankArevXhoI	AATGTActcgagGGCAACCTTTGTTAAAAACC	Rescue HAP2
HrpL29BforXhoI	AGCTCActcgagATATCTTCAAAGTATGGATTAATTATTTCA	Rescue HAP2
Hap2for	ATGAAATTTTTGGCTTTTGGATTGATTTATTTTC	Full HAP2 cDNA
GSP1rev	TCATTCAATTAGTAGATAGAGAGGAGATGTTGA	Full HAP2 cDNA
3'GFPHap2BclIFor	GATTACtgatcaATGAAATTTTTGGCTTTTGGATTGATTTATTTTC	HAP2GFP
3'GFPHap2MluIrev	TATACGacgcgtTTCAATTAGTAGATAGAGAGGAGATGTTG	HAP2GFP
Oligod(G)Anchorprimer	CATAGAGCTCGGATCCGGGIIGGGIIGGGIIGGGIIGDN	5'RACE
TtHAP2GSP2Rev	CACGATCTATCATAACTGCAAC	5'RACE
TtHAP2GSP3Rev	AAGCCTAATACAAATTCAGCAC	5'RACE
Hap2 3'FlankBforSacI	TACAGAgagctcCCATAAAAAAAGATAATTTTTAAGTATG	Genomic Rescue
HAP2 3'FlankRev	ATCTCTTCTGATCATAGAGCACC	Genomic Rescue
BamHIHAP2For	GATTACggatccATGAAATTTTTGGCTTTTGG	cDNA Rescue
TtFLHAP2HAKpnIR	TAGTACggtaccTCAAGCATAATCAGGAACATCATAAGGATATTCAATTAGTAGATAGAGAGGAGATGTTG	cDNA Rescue
BamHI5'FlankRev	GATTACggatccTTTTTATATTATGAATGTGTTAATATAATTTTTATTC	cDNA Rescue
KpnI flank Afor	GATTACggtaccATGAAAAATATTTATTTATAATTTTTAATATTTTTG	cDNA Rescue
TthermPair2for	TTCCAGCTACAGATCCAAGAGTTCT	qPCR primer
TthermPair2rev	CAGCGTAAACATGGTTTTGTCAA	qPCR primer
TtHAP2Sp6RNAfor	GATTTAGGTGACACTATAGAAATGAAATTTTTGGCTTTTGGATTG	Sp6 HAP2 cDNA
TtHAP2Sp6rev	TCATTCAATTAGTAGATAGAGAGGAGATG	Sp6 HAP2 cDNA

[c] All primers used in this project were supplied by Sigma-Aldrich, stored as 100 μM stock solutions, and diluted to 10μM working solutions prior to use in PCR reactions. Tms were determined using Modified Breslauer's thermodynamics, dH and dS parameters as recommended by the manufacturer. Added restriction sites are shown in lower case letters.

.

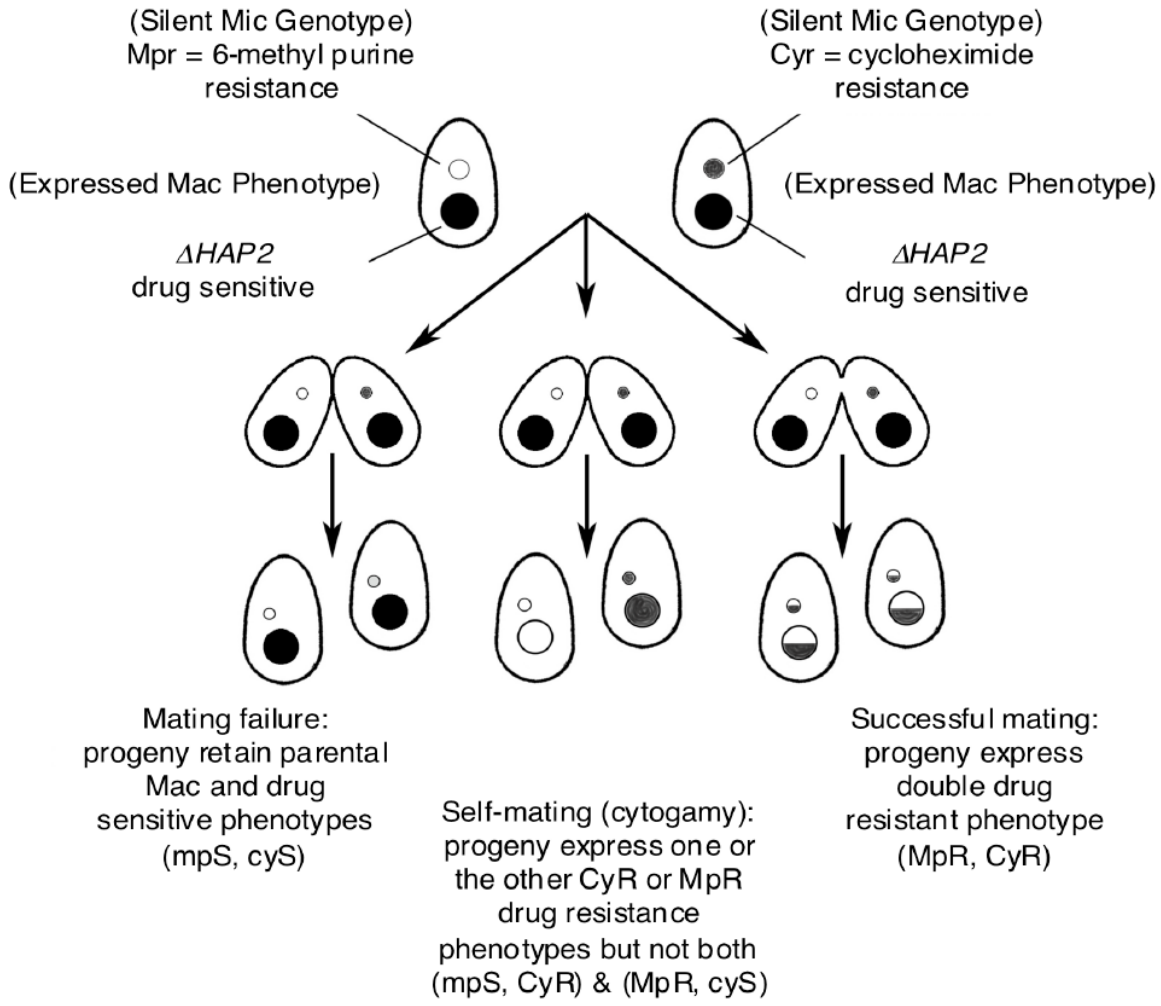


Figure S-2.3 Mating assay for fertilization success. Cells were constructed with drug-sensitive alleles in the macronuclei (Mac), and drug resistance markers in their transcriptionally silent, germline micronuclei (Mic). The Macs of each partner were engineered to express either the Δ*HAP2* deletion or the *HAP2* wild type allele. In this example, the cell on the left has a 6-methylpurine (6-MP) drug-resistance allele (Mpr) encrypted in its Mic, and is homozygous for the Δ*HAP2* knockout allele in its drug-sensitive Mac. Progeny will only express 6-MP resistance if the Mic genome comes into expression through completion of conjugal development. The cell on the right has a cycloheximide (CHX) drug resistance allele (Cyr) "encrypted" in its Mic, and is homozygous for the Δ*HAP2* knockout allele in its drug-sensitive Mac. Below, the three possible mating outcomes are illustrated. In the first outcome (left-most), mating is aborted, and pairs retain their drug sensitive parental Macs. They, and their clonal offspring, exhibit sensitivity to both CHX and 6-MP. In the second potential outcome (middle pathway), gametic pronuclei fail to be exchanged, yet each side completes development (undergoing some form of infrequently occuring cytogamy or self-mating). This allows each partner to express its own Mic's drug-resistance allele, but not the mating partner's (since nuclei were never exchanged). One partner will be resistant to CHX but not to 6-MP. The other partner will be resistant to 6-MP but not to CHX. In the third potential outcome (right-most pathway), cross-fertilization occurs producing progeny expressing both drug-resistance alleles. Cross-fertilization was never observed in matings between Δ*HAP2* deletion strains.

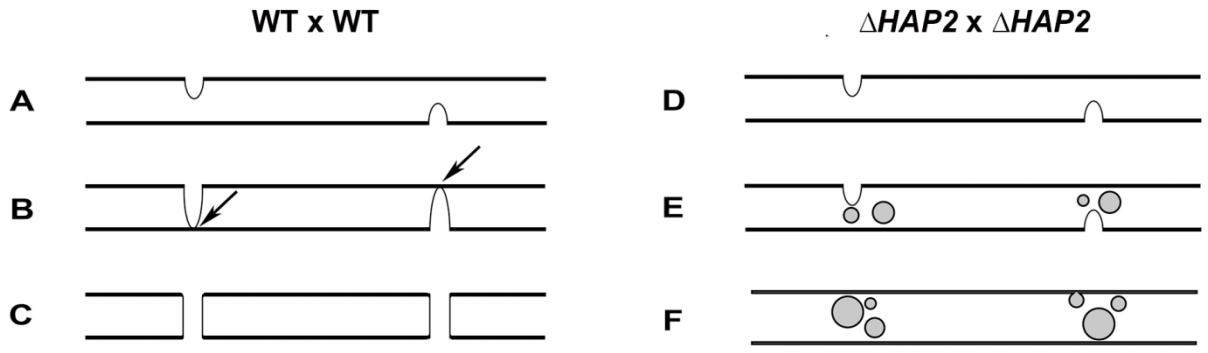


Figure S-2.4 Model of membrane events in *ΔHAP2* deletion (left) and wild type (WT, right) matings. Panels A-C illustrate how we envision pore formation happens in a wild type pair. **(A)** Membrane protrusions extend outward from the plasma membranes (solid parallel lines) of both partners and into the extracellular space of the junction cleft. **(B)** Membrane protrusions make contact (arrows) with plasma membrane of the mating partner. **(C)** Membrane fusion takes place creating cytoplasmic bridges ("pores") between mating partners. Panels (D-F) illustrate how we envision membrane dynamics go astray in {Δ*HAP2* x Δ*HAP2*} matings. **(D)** Protrusions extend from each mating partner into the extracellular space of the exchange junction. **(E)** Membrane fusion fails, resulting in the formation of tubular, multilammellar extracellular vesicles. **(F)** Without establishment of intercellular pores, vesicles expand and accumulate in the extracellular space of the junction, then, theoretically disperse into the media surrounding the paired cells.

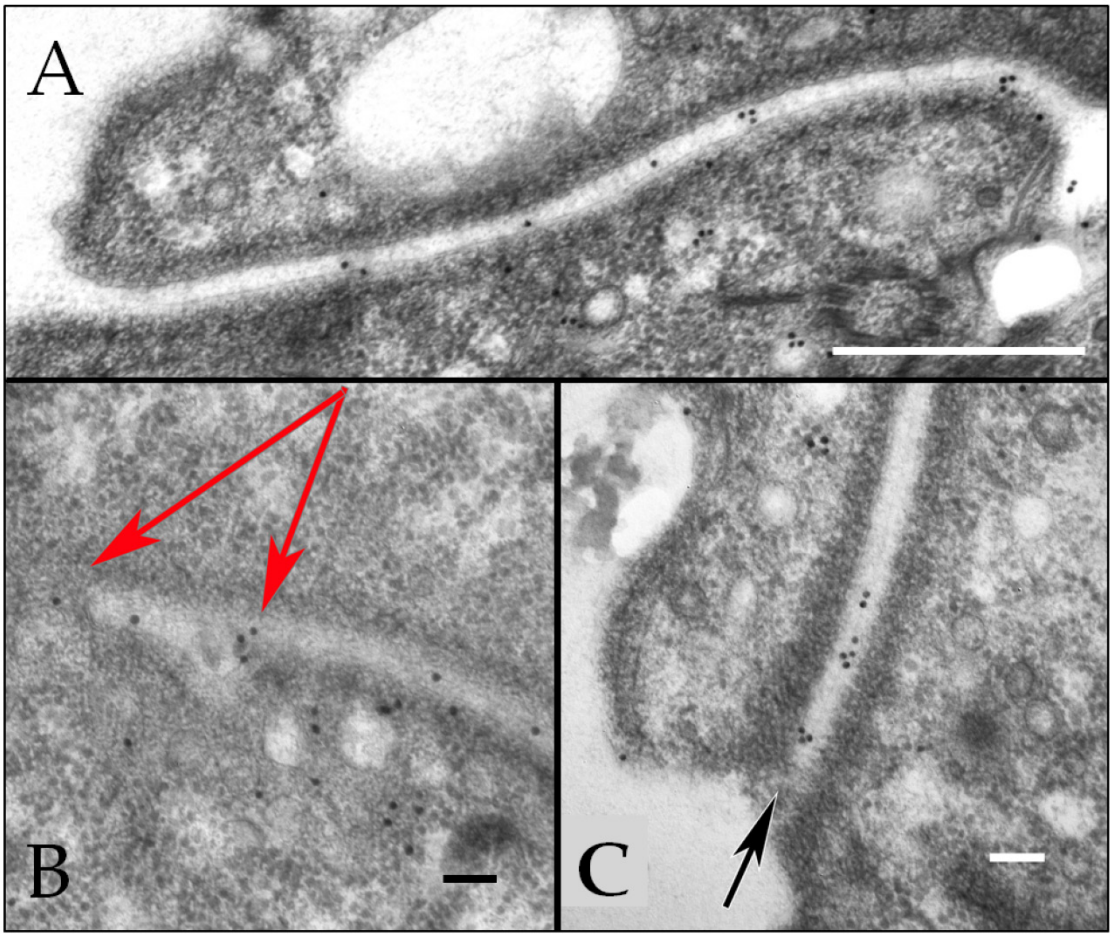


Figure S-2.5 Immunogold localization of GFP-tagged HAP2 at the mating junction in *Tetrahymena thermophila*. Cultures of wild type *T. thermophila* cells (mating type II) were crossed with CU522 (mating IV) cells that had been transformed with a chimeric *HAP2::GFP* construct. After 2 hrs, mating cells were fixed and processed for immunogold labeling as described in the text. **(A)** A low-magnification TEM image with immunogold directed against GFP indicating presence of the tagged protein in the extracellular space between paired cells (Scale bar = 500 nm). **(B)** TEM image capturing a region of the exchange junction in which a cytoplasmic channel (upper red arrow) has already been established through membrane fusion. The lower red arrow indicates membrane protuberances extending into the extracellular space of the exchange junction and decorated with anti-GFP gold-conjugated antibodies. (Scale bar = 100 nm). **(C)** A higher magnification image with immunogold labeling of the HAP2::GFP protein. Black arrow indicates the junctional cleft between mating cells. (Scale bar = 100 nm).

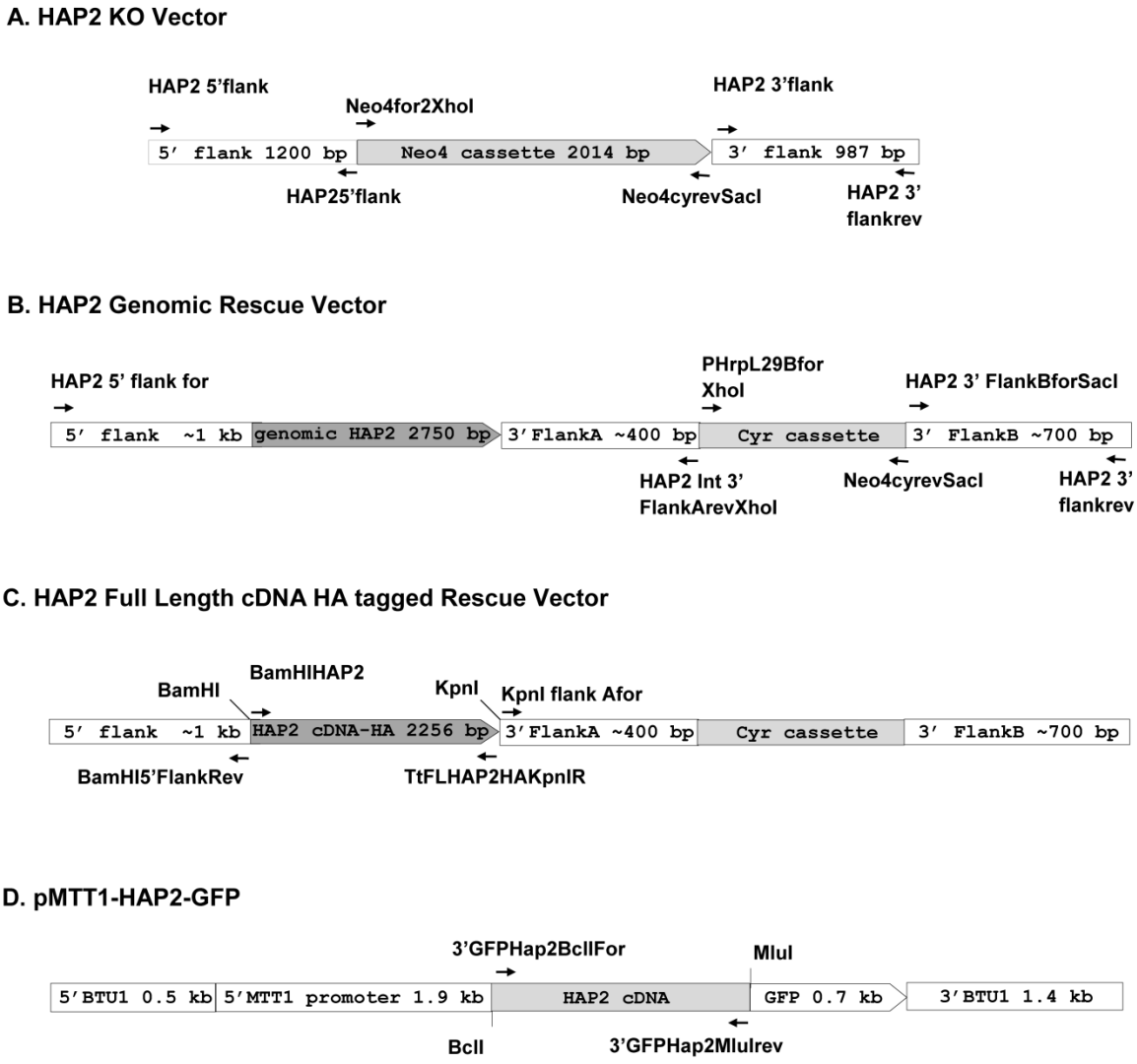


Figure S-2.6 Construct design for *HAP2* transformation vectors. (A) *HAP2* knockout (KO) vector for replacement of the endogenous *HAP2* gene with a functionally independent Neo4 selection cassette. *HAP2* 5' and 3' flanking regions and the Neo4 resistance cassette were amplified using the primers shown. Restriction sites were included in primers as indicated to facilitate vector construction. **(B)** *HAP2* genomic rescue vector used to replace the Neo4 cassette in previously constructed knockout strains with the genomic version of HAP2. Primers used to amplify genomic *HAP2* and adjacent flanking regions, and to amplify a cycloheximide resistance selection cassette for insertion within the *HAP2* 3' flank are shown. Restriction sites included in the primers are indicated. **(C)** *HAP2* HA-tagged cDNA rescue vector used to replace the Neo4 cassette in previously constructed knockout strains with a full length cDNA version of *HAP2* fused to an influenza hemagglutinin (HA) epitope tag. Primers used to amplify the tagged cDNA gene and to add *Bam*HI and *Kpn*I restriction sites to both insert and vector are shown. **(D)** *HAP2* 3' GFP (Green Fluorescent Protein) vector. *Bcl*I amd *Mlu*I restriction sites were added to the 5' and 3' ends of the *HAP2* cDNA gene respectively using the primers shown, and the amplified product was inserted into *Bcl*I and *Mlu*I sites of pMTT1-NRK2-GFP to create MTT1-HAP2-GFP. ß-tubulin-1 flanking regions target the construct to the somatic ß-tubulin-1 locus.

Chapter three

A link between virus and gamete fusion proteins: Implications for the origin of eukaryotic sex[a]

[a] Paper submitted to Current Biology on November 18, 2016 by **Jennifer Fricke Pinello**, Alex L. Lai, Jean Kaoru Millet, Donna Cassidy-Hanley, Jack H. Freed, and Theodore G. Clark

Summary

The conserved transmembrane protein, HAP2/GCS1, has been linked to fertility in a wide range of taxa and is hypothesized to be an ancient gamete fusogen. Using template-based homology modeling, we now show that the ectodomain of HAP2 orthologs from *Tetrahymena thermophila* and other species adopt a protein fold remarkably similar to the Dengue Virus E glycoprotein and related class II viral fusogens. Consistent with this finding, mutations to a region in and around a predicted “fusion loop” in the *Tetrahymena* protein abrogate membrane pore formation in mating cells, and a synthetic peptide corresponding to the loop interacts directly with model membranes in biophysical assays. These results raise interesting questions regarding the evolutionary relationships of class II membrane fusogens, and harken back to a long-held argument that eukaryotic sex arose as the by-product of selection for the horizontal transfer of a “selfish” genetic element from cell-to-cell via membrane fusion.

Introduction

Although sperm-egg fusion is a critical step in sexual reproduction, remarkably little is known about the molecular details of the process. Nevertheless, discovery of the conserved transmembrane protein, HAP2/GCS1[1, 2], has brought renewed focus to the problem and raised the intriguing possibility that HAP2 is an ancestral gamete fusogen dating to the last common ancestor of all eukaryotes[3].

HAP2 has been linked to fertility in a broad range of taxa extending from protists to flowering plants and insects[4–7], with gene knockout studies in unicellular eukaryotes, including *Chlamydomonas reinhardtii*, *Plasmodium berghei* and *Tetrahymena thermophila*[8–11] offering perhaps the strongest evidence that the protein plays a direct role in membrane fusion. In *C. reinhardtii* and *P. berghei*, disruption of the *HAP2* gene had no effect on adhesion of male and female gametes but completely blocked membrane fusion and zygote formation post-pairing[9]. Similarly, complementary mating types of *T. thermophila* were able to recognize and adhere to one another but failed to fuse in the absence of HAP2 expression[8]. Importantly, in *Chlamydomonas* and *Tetrahymena*, tagged versions of the protein localized to regions of the plasma membrane where fusion is initiated. HAP2 activity appears to be restricted to sperm, or the functional equivalent of male gametes in sexually dimorphic species[3, 6]. *T. thermophila*, on the other hand, has seven sexes or mating types and no true male gametes. Recent studies in this species have shown that the protein is made in all seven mating types, and a complete block to fertility occurs only when the *HAP2* gene is deleted from both cells of a mating pair[8].

While these and other studies[5, 7, 12, 13] clearly suggest a role for HAP2 in membrane fusion, primary sequence comparisons between HAP2 homologs and known membrane fusogens have shown no obvious similarities[3]. Indeed, our current understanding of HAP2 structure-function relationships is based almost entirely on the effects of targeted mutations to the extracellular and cytosolic regions of the protein on fertilization success. In short, these studies suggest that species-restricted functions of HAP2 reside within extracellular region, with

numerous residues/motifs including the conserved HAP2-GCS1 domain having a role in fertilization[14–16]. By contrast, some studies suggest that the cytosolic region is almost entirely dispensable for activity[16], while others point to positively charged residues as well as multi-cysteine motifs that may be palmitoylated as being important[14, 15].

While the use of engineered mutations has been informative, assays for HAP2 function have focused largely on blocks to fertility rather than membrane fusion per se. Such assays are indirect and often time-consuming, as are more direct assays for membrane fusion involving transmission electron microscopy. To address these issues in *Tetrahymena*, we sought to develop a high-throughput flow cytometry-based assay that would use exchange of fluorescently labeled proteins across the conjugation junction as a rapid and direct way to measure membrane pore formation in populations of synchronously mating cells. Concomitantly, we used template-based structural homology modeling to uncover conformational domains within HAP2 that are important for protein function. As shown here, predicted structures for the *T. thermophila* HAP2 ectodomain bear a striking resemblance to class II viral fusogens, in particular the Dengue Virus E glycoprotein. These homology models of HAP2 include a predicted fusion loop, a region in the viral proteins that inserts into target membranes and plays a key role in the fusion process. Deletion of the predicted HAP2 fusion loop or residues thought to stabilize the loop were found to block pore formation in mating *Tetrahymena*, and a synthetic peptide corresponding to this region was found to interact directly with model membranes in a variety of biophysical assays.

During the course of this work we became aware of successful efforts to crystalize and generate a high-resolution structure for *Chlamydomonas* HAP2, along with functional data that are consistent with the results reported here (Fedry et al. submitted). Together, these studies argue strongly that HAP2 mediates zygote formation through a membrane fusion mechanism analogous to that used by Dengue and related viruses to enter host cells.

Results

Transfer of labeled cytosolic proteins between mating cells

In *T. thermophila*, HAP2 is localized to the conjugation junction, a specialized region of membrane between the two mating cells where hundreds of fusion pores form[17]. These pores allow the exchange of haploid pronuclei as well as limited amounts of protein and RNA between mating partners. To determine whether the exchange of labeled proteins could be used to assay for membrane fusion events (pore formation) during sexual conjugation, live cells of complementary mating types were starved to render them mating competent, and then separately labeled either green or red with the amine reactive dyes, carboxyfluorescein diacetate succinimidyl ester (CFSE), or cell trace far-red (CTFR). Cells were then washed and mated with either unlabeled cells of a different mating type or with each other. In the case of wild type (WT) cells, no exchange of labeled protein was visible by fluorescence microscopy prior to pairing (Figure 3.1A,B). After pairing, however, content exchange was readily detected between labeled and unlabeled cells (Figure 3.1C,D), as was the reciprocal exchange of fluorescent protein between cells that were initially labeled either green or red (Figure 3.1E-H). When these paired cells separated after the completion of mating (12-16h post-mixing), the vast majority had both fluorescent markers, but were more intensely labeled one color or the other (Figure 3.1I-L). Finally, no exchange of labeled protein was seen when complementary mating types lacking the HAP2 gene (*ΔHAP2* strains) were crossed (Figure 3.1M-P). The latter result was entirely expected given the inability of these cells to form pores at the conjugation junction[8],

By observing cells at varying time points after mixing it was possible to determine the relative kinetics of membrane fusion events with respect to pairing in this system. As shown in Figure 3.1Q, pairs first became measurable 1h 15m after mixing labeled WT cells of complementary mating types (30˚C). By 1h 30min, 50% of all WT pairs had exchanged dye indicating that pores form rapidly following the adhesion of mating cells (Figure 3.1R). Notably,

the rate at which cells form pores (but not the rate of pairing) was significantly reduced when *ΔHAP2* strains were mated with a WT partner, and likely contributes to the reduced fertility of WT x *ΔHAP2* crosses described previously[8]

Quantitation of membrane fusion using flow cytometry

The ability to detect exchange of fluorescent proteins across the conjugation junction enabled the use of flow cytometry to quantitate the percentage of cells undergoing membrane fusion in large-scale mating cultures. As shown Figure 3.2A,B, when complementary WT mating types were cytoplasmically labeled, mixed, and acquired by flow cytometry, prior to pairing, two populations were seen, one $CFSE^{hi}$ and the other $CTFR^{hi}$ (Figure 3.2B). However, when the same cultures were fixed after the cells had completed mating and come apart, an entirely different pattern was observed (Figure 3.2C). In this case, the vast majority of cells (typically ~80%) contained both fluorescent tracers and fell into two equally sized populations, one brighter for CFSE and the other brighter for CTFR ("Mid" fluorescence gate in Figure 3.2C). These cells had clearly undergone membrane fusion and exchanged labeled protein during the mating process. In addition to these "Mid"-fluorescence events, three other populations were visible. The smallest (typically <10% of cells) was intensely labeled with both tracers ($CFSE^{hi}/CTFR^{hi}$, Figure 3.2C-E, upper right-hand "Pairs" gate) and likely represented either persistent pairs (i.e. cells that had paired and failed to come apart), or pairs that formed very late in mating cultures. Indeed, a few pairs were visible by microscopy even 16 h post-mixing, and, based on forward light scattering (FSC), the size distribution of individual events in the "Pairs" gate was consistently larger than that in the other gates (Figure 3.2F). The remaining two populations (comprising 10-15% of total cells) were single-labeled and expressed either $CFSE^{hi}$ or $CTFR^{hi}$ (Figure 3.2C). Cells in these populations had not acquired label from a mating partner, and were the expected number of cells that fail to generate true progeny in standard WT crosses[8].

Our interpretation of the flow cytometry results generated with WT matings was strongly corroborated in crosses with *ΔHAP2* strains. In the case of *ΔHAP2* x WT crosses (Figure 3.2D), double-labeled cells were present in the "Mid"-fluorescence gate, but were dramatically reduced compared to the same populations in WT x WT crosses (Figure 3.2C). Furthermore, in *ΔHAP2* x *ΔHAP2* matings (Figure 3.2E), virtually no double-labeled cells were present in the "Mid"-fluorescence gate, and the populations of single-labeled cells (which had not undergone fusion) were overrepresented relative to $CFSE^{hi}$ and $CTFR^{hi}$ populations in WT crosses (Figure 3.2C).

Based on changes of fluorescence intensity in the various populations pre- and post-mating, we determined that ~20% of labeled protein is reciprocally exchanged between WT cells regardless of the tracer or mating type background (Figure 3.2H,I). Interestingly, nearly the identical level of protein transfer was seen in crosses between *ΔHAP2* x WT strains (Figure 3.2I). This, and the reduced rate at which these cells fuse (Figure 3.1R) suggests that the initiation of pore formation at the conjugation junction rather than the total number of pores that form is the primary defect in *ΔHAP2* x WT crosses. It is also worth noting that in all instances in which complementary mating types were mixed, the populations of single-labeled cells that failed to fuse showed an average decline in fluorescence intensity of ~40% relative to the same populations of cells in cultures that were starved but not mixed (Figure 3.2J). This decline in fluorescence is likely attributable to "co-stimulation," a signaling event requiring physical contact of starved complementary mating types that alters gene expression and protein turnover to promote the cellular transition to mating competency[17, 18].

To establish a baseline for functional studies with HAP2 mutant constructs (see below), we performed multiple mass mating experiments with WT, *ΔHAP2*, and *ΔHAP2* rescue strains and determined the mean percentage of cells that fuse in each case (Figure 3.2G). Consistent with previous fertility data[8], the vast majority of cells in WT x WT crosses fuse, whereas the level is reduced approximately 75% when HAP2 is expressed in only one cell of a mating pair.

Figure 3.1 Conjugation leads to rapid exchange of labeled cytosolic proteins in mating *T. thermophila*. Live cells of complementary mating types were labeled with either carboxyfluorescein diacetate succinimidyl ester (CFSE) or Cell Trace™ Far Red (CTFR) and examined microscopically after fixation. Scale bars in all micrographs = 10 µm. **(A,B)** Overlay of phase and fluorescence images of labeled and unlabeled cells combined and fixed 15 min post-mixing showing either green or red labeling (but no exchange of fluorescent proteins). **(C,D)** Fluorescence images of labeled and unlabeled partners combined and fixed at 3.5 and 2.5 h post-mixing, respectively. Partial exchange of fluorescent proteins from labeled (bright) to unlabeled (faint) mating partners is seen**. (E-H)** Fluorescence Image of a wild type pair of cells in which both mating partners were separately labeled, combined and fixed 3.5 h post-mixing. Reciprocal exchange of labeled proteins is visible in the same mating pair viewed with either red (E) or green filter (F) filter sets. **(G)** Merged image of (E) and (F). **(H)** Phase image of the mating pair in (E-G). **(I-J)** Fluorescence images of labeled wild type cells 20 h after mixing. At this time point, pairs have come apart, but exconjugant progeny cells maintain a combination of the parental fluorescent markers. **(K)** Merged image of (I) and (J). **(L)** Phase image of the cells in (I-K). **(M-P)** A cross between Δ*HAP2* partners of complementary mating types with the same sequence of images as in (E-H). Note the absence of fluorescent protein transfer. **(Q)** Representative data from one experiment showing the kinetics of pairing for WT x WT (♦) and WT x Δ*HAP2* (□) crosses. **(R)** The kinetics of fusion in WT x WT (♦), WT x Δ*HAP2* (□) and Δ*HAP2* x Δ*HAP2* (**O**) crosses determined as the percentage of pairs showing visible transfer of fluorescent material at the indicated time points. Data are expressed as the mean +/- SEM for 3 and 4 independent experiments (♦ and □, respectively), and for 1 experiment (**O**).

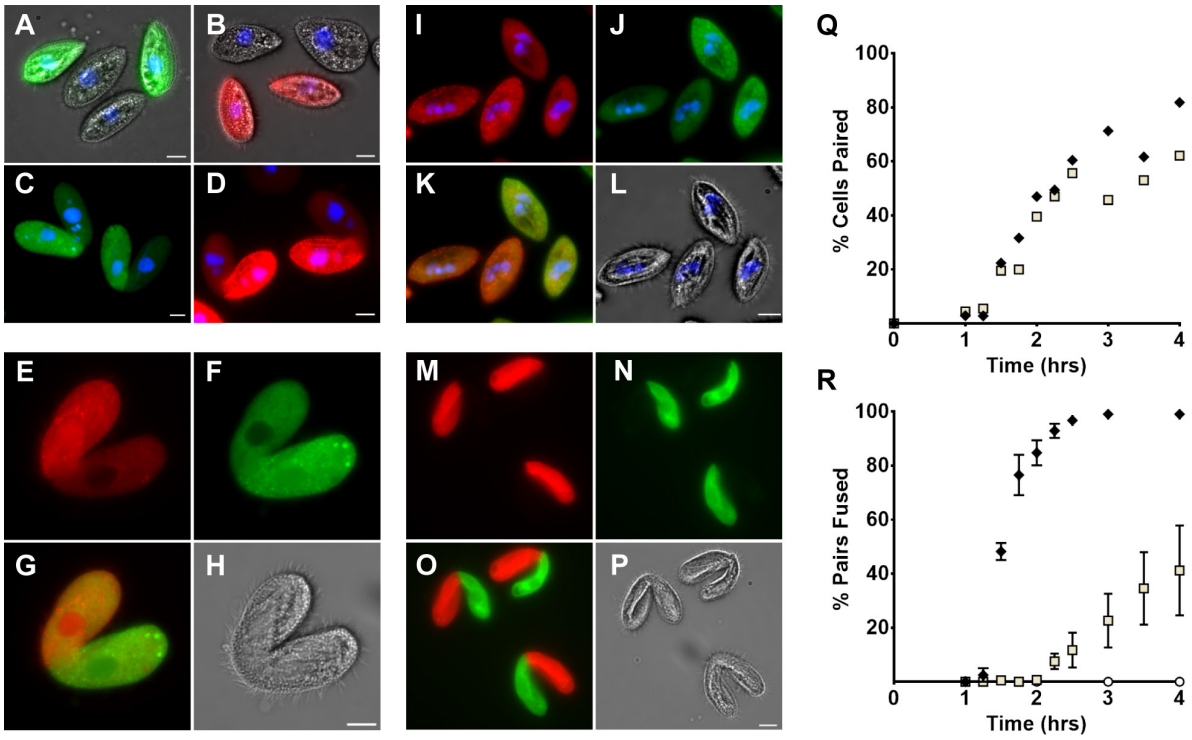
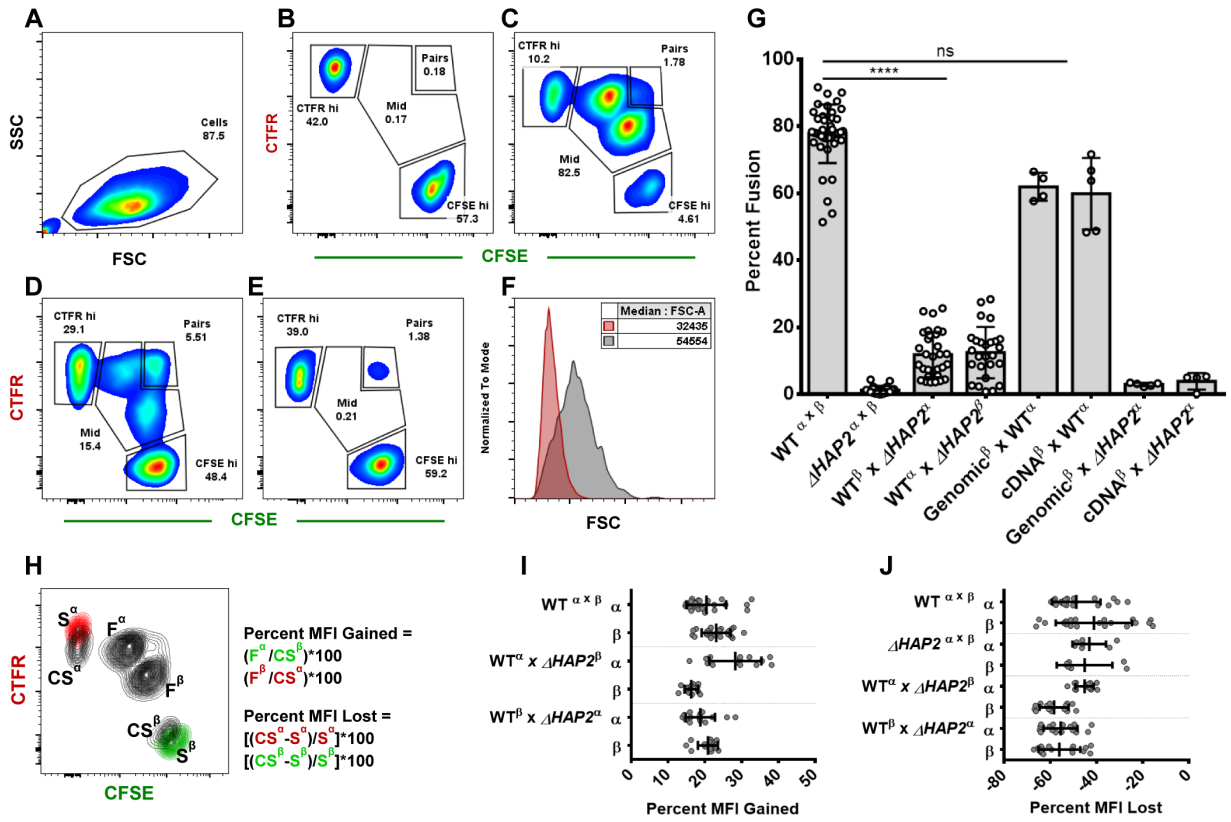

A
B
C
D
I
J
K
L
E
F
G
H
M
N
O
P
Q
% Cells Paired
100
80
60
40
20
0
1
2
3
4
Time (hrs)
R
% Pairs Fused
100
80
60
40
20
0
1
2
3
4
Time (hrs)

Figure 3.2 Quantitation of *T. thermophila* sexual cell fusion events by flow cytometry. Populations of WT and/or Δ*HAP2* cells of complementary mating types were labeled with either CFSE or CTFR, mixed together at a 1:1 ratio, and acquired at different time points after mating and fixation. Superscripts (α and β) denote mating types VII and VI, respectively. **(A)** Representative forward scatter (FSC) / side scatter (SSC) plot showing *Tetrahymena* cellular population being analyzed (circled). **(B)** A flow cytometry plot of labeled WT^{α} x WT^{β} ($WT^{\alpha \times \beta}$) cells fixed 15 min after mixing the complementary mating types. **(C)** The same culture as in (B), but instead fixed 16 h after mixing. **(D)** Representative plot of a WT^{α} x Δ*HAP2*$^{\beta}$ cross 16 h after mixing. **(E)** The flow cytometry plot of a Δ*HAP2*$^{\alpha}$ x Δ*HAP2*$^{\beta}$ (Δ*HAP2*$^{\alpha \times \beta}$) cross 16 h after mixing showing an absence of events in the mid-fluorescence gate. Numbers adjacent to the outlined areas in (a-e) indicated the percent of cells in these gates. **(F)** A representative histogram comparing the relative size of individual events (based on forward scatter) in the double-labeled $CFSE^{hi}$/$CTFR^{hi}$ gate (grey) versus the single-labeled $CTFR^{hi}$ gate (red). Median FSC intensity values for these two populations are shown in the inset. **(G)** The cumulative results of independent mass mating experiments including all biological replicates for different WT, Δ*HAP2*, and complementation strain crosses (circles represent the percentage of exconjugant cells in the mid-fluorescence gate from individual matings 16 h after mixing; bar represents mean and error bars +/- s.d.). “Genomic” or “cDNA” strains are designated according to which *HAP2* gene product was used to complement the Δ*HAP2* cell line during their construction. A one-sided Kruskal-Wallis test with a Dunn’s post-test found a significant difference (**** = $p<0.0001$) between the WT $^{\alpha \times \beta}$ cross and WT x *ΔHAP2* crosses, but no difference (ns= not significant) between the WT $^{\alpha \times \beta}$ cross and the genomic or cDNA complementation crosses. **(H)** Flow cytometry plots of cell populations from mated (dark grey) and unmated (red or green) cultures at the 16 h time point shown superimposed. Populations of double-labeled cells from the mated cultures are denoted (F), and single-labeled cells that had undergone “co-stimulation” but had not exchanged fluorescent protein are denoted (CS). Note that the populations with the highest fluorescence intensities (MFI) are the single-labeled starved cells from the cultures that had not been mated. The formulas used to measure the “Percent MFI Gained” (due to transfer of labeled protein from the opposite mating partner) and the “Percent MFI Lost” (due to co-stimulation) are indicated on the right. The calculations are color-coded to show the red or green MFI measurement that was used for each population (S, CS, F). **(I)** Chart showing the mean +/- s. d. of the percent MFI gained in each mating partner of a given cross, based on the upper formula in (H). The percent MFI gained in F populations was measured with respect to the MFI of the corresponding CS populations, as the co-stimulated partners would theoretically represent the starting fluorescence intensity prior to cellular fusion. Note that no substantial differences were seen in the amount of fluorescent protein exchanged between mating partners in WT x WT and *ΔHAP2* x WT crosses. **(J)** Chart showing the mean +/- s. d. of the percent MFI lost in each mating type of a given cross based on the lower formula in (a). Regardless of the parental cell lines used, a consistent reduction in the MFI was seen in mated cells that had not undergone fusion (CS) when compared to unmated starved cells (S).


A
SSC
FSC
Cells
87.5
B
CTFR
CFSE
CTFR hi
42.0
Mid
0.17
Pairs
0.18
CFSE hi
57.3
C
CTFR hi
10.2
Pairs
1.78
Mid
82.5
CFSE hi
4.61
D
CTFR
CFSE
CTFR hi
29.1
Pairs
5.51
Mid
15.4
CFSE hi
48.4
E
CTFR hi
39.0
Pairs
1.38
Mid
0.21
CFSE hi
59.2
F
Normalized To Mode
FSC
Median : FSC-A
32435
54554
G
Percent Fusion
ns
****
WT α x β
ΔHAP2 α x β
WTβ x ΔHAP2α
WTα x ΔHAP2β
Genomicβ x WTα
cDNAβ x WTα
Genomicβ x ΔHAP2α
cDNAβ x ΔHAP2α
H
CTFR
CFSE
Sα
Fα
CSα
Fβ
CSβ
Sβ
Percent MFI Gained =
(Fα/CSβ)*100
(Fβ/CSα)*100
Percent MFI Lost =
[(CSα-Sα)/Sα]*100
[(CSβ-Sβ)/Sβ]*100
I
WT α x β
WTα x ΔHAP2β
WTβ x ΔHAP2α
Percent MFI Gained
J
WT α x β
ΔHAP2 α x β
WTα x ΔHAP2β
WTβ x ΔHAP2α
Percent MFI Lost

Furthermore, complementation of *ΔHAP2* strains with either the genomic or cDNA versions of the *HAP2* gene rescued both the fertility defect[8] and the cell-cell fusion defect in *ΔHAP2* x WT crosses as expected (Figure 3.2G). Over-expression of HAP2 in WT cells, however, did not rescue fusion in crosses with *ΔHAP2* strains, reinforcing the idea that the presence of HAP2 on both sides of the conjugation junction (rather than the total amount of HAP2 expressed in any given mating pair) is crucial for efficient pore formation in the *Tetrahymena* system (Figure S-3.3A-D).

Predicted structural homologies between HAP2 and class II viral fusogens

As a starting point for the design of mutant constructs, we sought to gain insight into the overall architecture of *T. thermophila* HAP2 using template-based structural modeling programs including Phyre2[19, 20], RaptorX[21], and CPHmodels-3.0[22]. This approach uncovered high (>95%) confidence hits to class II viral fusogens from which predicted structures could be built. The known structure of a prototypical class II viral fusogen, the Dengue virus E glycoprotein (DENV)[23] (PDB ID: 1UZG) is shown in Figure 3.3A along with the predicted ectodomain structures for *T. thermophila* HAP2 generated by Phyre2 (Figure 3.3B) and RaptorX (Figure 3.3C).

In the case of Phyre2, the region of homology with DENV covered a 196 amino acid stretch immediately upstream of the consensus HAP2/GCS1 domain. This region had only 16% sequence identity but aligned closely to a 166-residue stretch of DENV at the level of predicted secondary structure (Figure 3.3D). CPHmodels-3.0 identified a similar partial structural homology to DENV (Figure S-3.5A). The RaptorX homology model predicted a structure for the entire HAP2 ectodomain (Figure 3.3C) based on a different class II fusion protein template, namely, the Rift Valley Fever virus glycoprotein C (PDB ID: 4HJ1)[24]. Together, these structural predictions of HAP2 showed three largely β-sheet-containing domains, analogous to domains I-III of the viral class II fusogens, and included a possible “fusion loop” located at the tip of domain II (circled in Figure 3.3A-C; boxed sequence in Figure 3.3D). The fusion loop in class

II viral fusogens inserts into endosomal membranes and is critical for the entry of viral genome into host cells[25, 26].

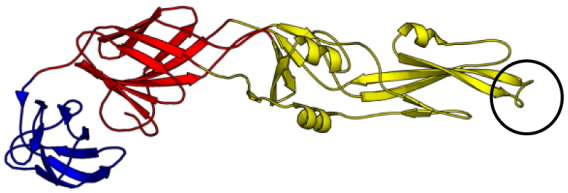
A total of 40 HAP2 orthologs[15] were submitted to Phyre2 batch processing[20] to determine the extent to which the predicted structural homology to class II viral fusion proteins is maintained across taxonomic groups. We found that 28 (~70%) had hits to class II viral fusogens (Table S-3.5). A subset of the 17 highest confidence hits to these viral proteins is shown in Figure 3E.

Mutational analysis of HAP2 function

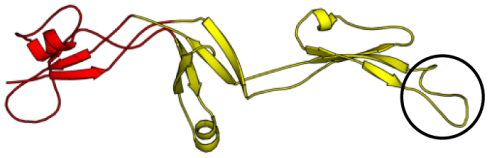
To address the predicted structural similarities between HAP2 and class II viral fusogens, we created cell lines carrying targeted mutations/deletions to various regions of *T. thermophila* protein and tested them for functional activity in crosses with WT cells using flow-cytometry. Large deletions to the extracellular region covering either the entire HAP2 domain (Δ281-329) or the region of alignment with Dengue Virus E glycoprotein identified by Phyre2 (Δ93-280), resulted in minimal fusogenic activity (Figure 3.4A,C). The functional relevance of these deletions was nevertheless unclear, given that both constructs were poorly expressed (Figure S-3.6). More interestingly, a much smaller, 28 amino acid deletion of the fusion loop itself (Δ152-179) had a profound effect on fusogenic activity, reducing it to the levels observed in ΔHAP2 x WT crosses (Figure 3.4A,C). When this deletion was repaired with the native sequence (“HAP2 FL Rescue”, Figure 3.4A,C), the activity was restored to WT levels, but it remained low when the deletion was complemented with a 17 amino acid sequence comprising the DENV fusion loop (“DENV FL Rescue”; Figure 3.4C)[27]. Along with these deletions, we explored additional targeted mutations in and around the predicted *T. thermophila* fusion loop. The majority of these mutations showed no effect on fusogenic activity (Figure 3.4C). These included alanine replacements of either LNL171-3 or R164 within the predicted loop itself, or replacement of FQY131-3 in a neighboring loop (Figure S-3.4B).

Figure 3.3 Homology modeling predicts a structural similarity between HAP2 and class II viral fusogens. The *T. thermophila* HAP2 primary sequence was submitted to template-based structural modeling platforms, Phyre2 and RaptorX. Known and predicted structures shown in panels **(A-C)** are colored by domain according to the convention used for class II fusion proteins: red = domain I; yellow = domain II; blue = domain III with black circles highlighting the known and predicted fusion loops. **(A)** The known structure of the Dengue Virus envelope glycoprotein ectodomain (DENV, PDB ID: 1UZG)[23]. **(B)** The Phyre2-predicted partial structure of the *T. thermophila* HAP2 ectodomain based on the template shown in (a). **(C)** The RaptorX-predicted *T. thermophila* HAP2 ectodomain structure based on the Rift Valley Fever Virus Glycoprotein C template (PDB ID: 4HJ1)[24]. **(D)** The Phyre2 generated primary and secondary structural alignment of *T. thermophila* HAP2 and the Dengue Virus Envelope glycoprotein. Sequence identities are shaded gray, alpha helices are indicated by green spirals and beta strands by blue arrows. The boxed region is the *T. thermophila* HAP2 sequence aligning to the viral envelope protein's fusion loop. **(E)** A table of 17 HAP2 orthologs from other species with the highest confidence hits to class II viral fusogens based on Phyre2 batch processing results (Top class II viral hits are listed as: % confidence, PDB ID of template envelope protein structure, and viral origin: DENV= Dengue Virus, TBEV= Tick Borne Encephalitis Virus, WNV= West Nile Virus).

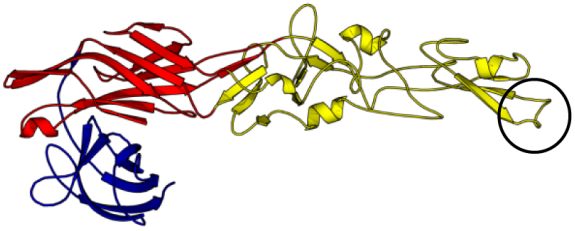
A



B



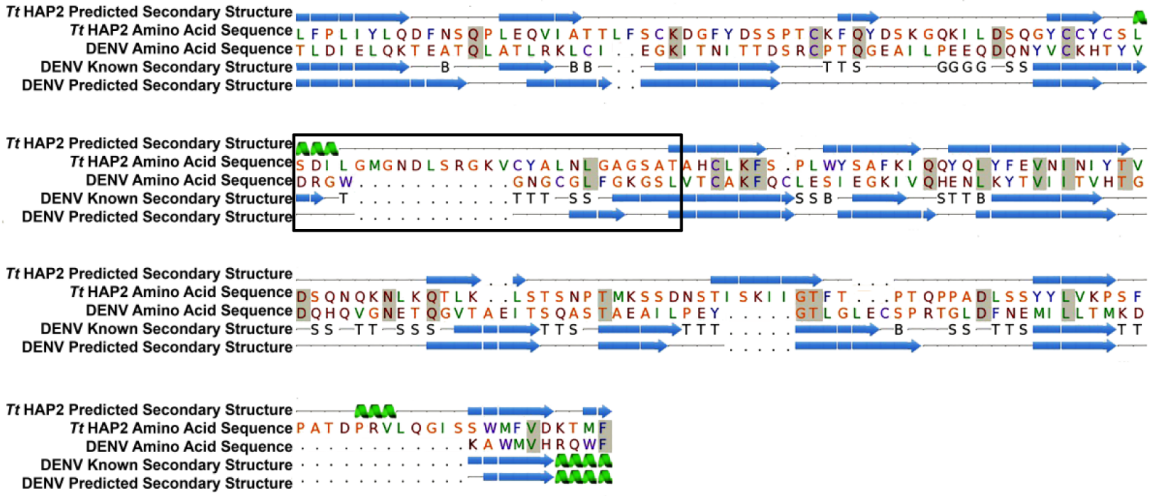
C



E

	Organism	Top CII Viral hit
Viridiplantae	Chlorella variabilis (unicellular green algae)	95.0%, 4B03, DENV
	Coccomyxa subellipsoidea (green algae)	90.9%, 3UAJ, DENV
	Lilium longiflorum (easter lily)	96.3%, 3UAJ, DENV
	Arabidopsis thaliana (mouse ear cress)	92.9%, 1URZ, TBEV
Excavata	Trypanosoma brucei (sleeping sickness)	96.7%, 1UZG, DENV
	Trypanosoma cruzi (Chagas disease)	95.1%, 3UAJ, DENV
	Strigomonas culicis (mosquito parasite)	91.1%, 1UZG, DENV
	Angomonas deanei (insect parasite)	83.9%, 2OF6, WNV
	Phytomonas sp. (isolate EM1) (plant parasite)	94.9%, 1SVB, TBEV
	Leishmania major (cutaneous leishmaniasis)	93.3%, 1UZG, DENV
	Naegleria gruberi (ameboflagellate)	93.3%, 3UAJ, DENV
Amoebozoa	Acanthamoeba castellanii (soil amoeba)	96.1%, 3UAJ, DENV
Alveolata	Tetrahymena thermophila (ciliated protozoan)	94.9%, 1UZG, DENV
	Ichthyophthirius multifiliis (ciliated fish parasite)	95.2%, 3UAJ, DENV
	Paramecium tetraurelia (ciliated protozoan)	87.3%, 3UAJ, DENV
Opisthokonta	Monosiga brevicollis (Choanoflagellate)	96.2%, 1UZG, DENV
	Nematostella vectensis (starlet sea anemone)	96.3%, 1UZG, DENV

D



However, substitutions of serine for the first two cysteine residues (CC147-148SS) in a highly conserved cysteine motif that precedes the loop resulted in a dramatic decline in the percentage of cells capable of fusion (Figure 3.4C). Notably, cysteine residues in the cognate region of class II viral fusogens participate in the formation of disulfide bonds that are thought to stabilize the fusion loop [28, 29]. In contrast with the large deletions within the ectodomain (above), mutant constructs in the region of the fusion loop were expressed and, in almost all cases, correctly localized to the conjugation junction of mating *T. thermophila* (Figure S-3.6).

In addition to mutations to extracellular region, alterations to the cytosolic domain of *T. thermophila* HAP2 were also constructed (Figure 3.4B). These included serine substitutions for cysteine residues in some or all predicted palmitoylation sites (C_5 and C_8); deletion of a stretch of highly basic amino acids (Δ580-596); and deletion of almost the entire cytosolic region (beginning at residue 580). The corresponding mutant proteins were all expressed and correctly localized (Figure S-3.6), and while matings with the C_8 substitution showed a slight, but statistically significant decrease in fusogenic activity when compared with WT x WT crosses, the other cytosolic alterations showed no measurable effects (Figure 3.4D).

Biophysical evidence for interactions between the HAP2 fusion loop and membranes

The effects of mutations to the HAP2 fusion loop described above, coupled with the known importance of this domain in the activity of class II viral fusogens, clearly suggested that HAP2 and the viral proteins catalyze membrane fusion events through a similar mechanism. To begin to address this question experimentally we attempted to induce membrane fusion by expressing *Tetrahymena* HAP2 in heterologous systems, namely, pseudotyped viruses and mammalian tissue culture cells. By and large, these efforts were unsuccessful (Figure S-3.7). Nevertheless, biophysical studies clearly demonstrated the ability of the predicted HAP2 fusion loop to interact directly with model membranes, an inherent property of viral fusion peptides.

As revealed by circular dichroism spectroscopy, a synthetic peptide corresponding to the fusion loop of *T. thermophila* HAP2 adopts a partially (~30%) β-strand containing structure in

the presence of small unilamellar vesicles, but is essentially a random coil in solution (Figure 3.5B). This alteration in secondary structure mimics that observed for a WT DENV fusion peptide under the same conditions, and suggests that, as in the case of the viral peptide, the fusion loop of *T. thermophila* HAP2 can bind to membranes.

We then applied electron spin resonance (ESR) spectroscopy to determine whether the predicted HAP2 fusion loop can insert directly into lipid bilayers by measuring changes to the membrane-order parameter, S_o, of spin-labeled lipids in model membranes in the presence or absence of synthetic peptides. Increased membrane ordering in the presence of viral fusion peptides has been attributed to a membrane dehydration effect in which loosely bound water in the inter-bilayer space moves to the bulk water phase. Such peptide-induced changes are thought to be functionally significant as they can lower the energy barrier between closely apposed membranes allowing fusion to occur[30–33]. Indeed, previous ESR studies with fusion peptides from different class I fusogens (HIV gp41 and influenza hemagglutinin) suggest that this membrane ordering effect is a general phenomenon, as well as a critical step for viral membrane fusion[31, 32, 34, 35]. Here we incorporated two different spin-labeled lipids, DPPTC and 5PC, into model membranes to detect peptide-induced changes in membrane structure at both headgroup (membrane surface) and acyl chain (hydrophobic bilayer interior) regions respectively[31]. As shown in Figure 3.5C,D, increasing the peptide:lipid ratio from 0 to 2% resulted in substantial increases in membrane ordering at both the headgroup (Figure 3.5C) and acyl chain regions (Figure 3.5D) when synthetic peptides corresponding to the predicted HAP2 and DENV WT fusion loops were used, while little increase was seen with a non-interacting mutant peptide, DENV W101A[36]. Furthermore, the roughly S-shaped curves of S_o as a function of increasing WT peptide concentrations suggests cooperativity in the membrane ordering effect, consistent with the requirement for class II proteins to oligomerize for efficient fusion to occur[37].

Figure 3.4 Sequence elements important for *T. thermophila* HAP2 function. Altered versions of the *HAP2* gene were placed under control of the endogenous promoter at the *HAP2* locus in a *ΔHAP2* cell line. Mutant cell lines were then mated with a wild type (WT) partner, and the percentage of cells undergoing fusion determined by flow cytometry. Mutant cell lines that showed levels of fusion equivalent to WT $^{\alpha \times \beta}$ crosses (~80%) were considered functional HAP2 constructs. **(A)** Diagrams of truncations / mutations to the ectodomain. The region in and around the predicted fusion loop is expanded, and amino acids targeted for mutations are highlighted in black, while those deleted in the fusion loop truncation are shown in bold, black, italicized lettering. **(B)** Diagrams of truncations / mutations to the cytosolic domain. The region in and around the poly-basic stretch (underlined) is expanded, and the potentially palmitoylated cysteine residues targeted for mutations are highlighted in black. The numbers in (A) and (B) refer to the numerical positions and/or range of truncated amino acids relative to the full-length HAP2 protein sequence. Abbreviations are: SP = signal peptide; DENV = Phyre2-predicted Dengue Virus envelope protein region of homology; FL = fusion loop; H/G = HAP2/GCS1 domain; TM = transmembrane domain; B = poly-basic domain; HA = influenza hemagglutinin epitope tag; FLAG-10xHis = epitope tag. **(C, D)** Bar charts showing the mean percentage +/- s.d. of exconjugant cells in the mid-fluorescence gate (cells that had undergone fusion) after mating as determined by flow cytometry. Circles represent fusion data from individual matings 16 h after mixing for the various constructs. A one-sided Kruskal-Wallis with Dunn's post test found no significant differences between the WT $^{\alpha \times \beta}$ cross (Figure 3.2G) and the *HAP2* mutant crosses Δ510-513, HAP2 FL Rescue, FQY131-3AAA, R164A, LNL171-3AAA, C5→S, ΔBasic Domain, and ΔC' term. A modest, yet statistically significant reduction (p=0.0011) in the percentage of fusion was observed for C8→S mutants when compared with the WT $^{\alpha \times \beta}$ cross. Likewise, no significant differences were found between the WT$^{\alpha}$ x *ΔHAP2*$^{\beta}$ cross (shown in Figure 3.2G) and the *HAP2* mutant crosses ΔHAP2 domain, ΔDENV region, ΔFusion Loop, DENV FL Rescue, and CC147-8SS. Sample sizes for each cross are listed in the methods.

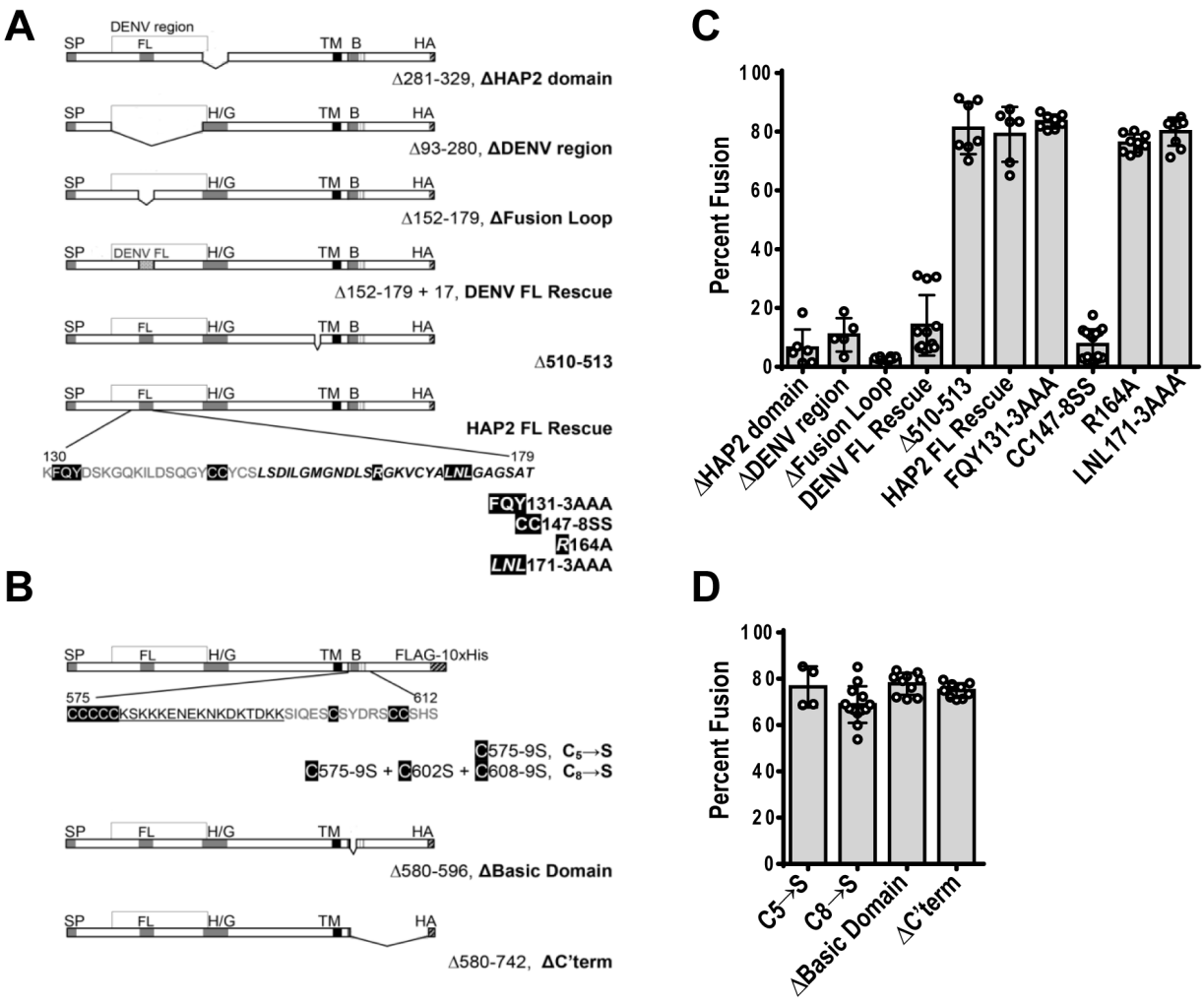
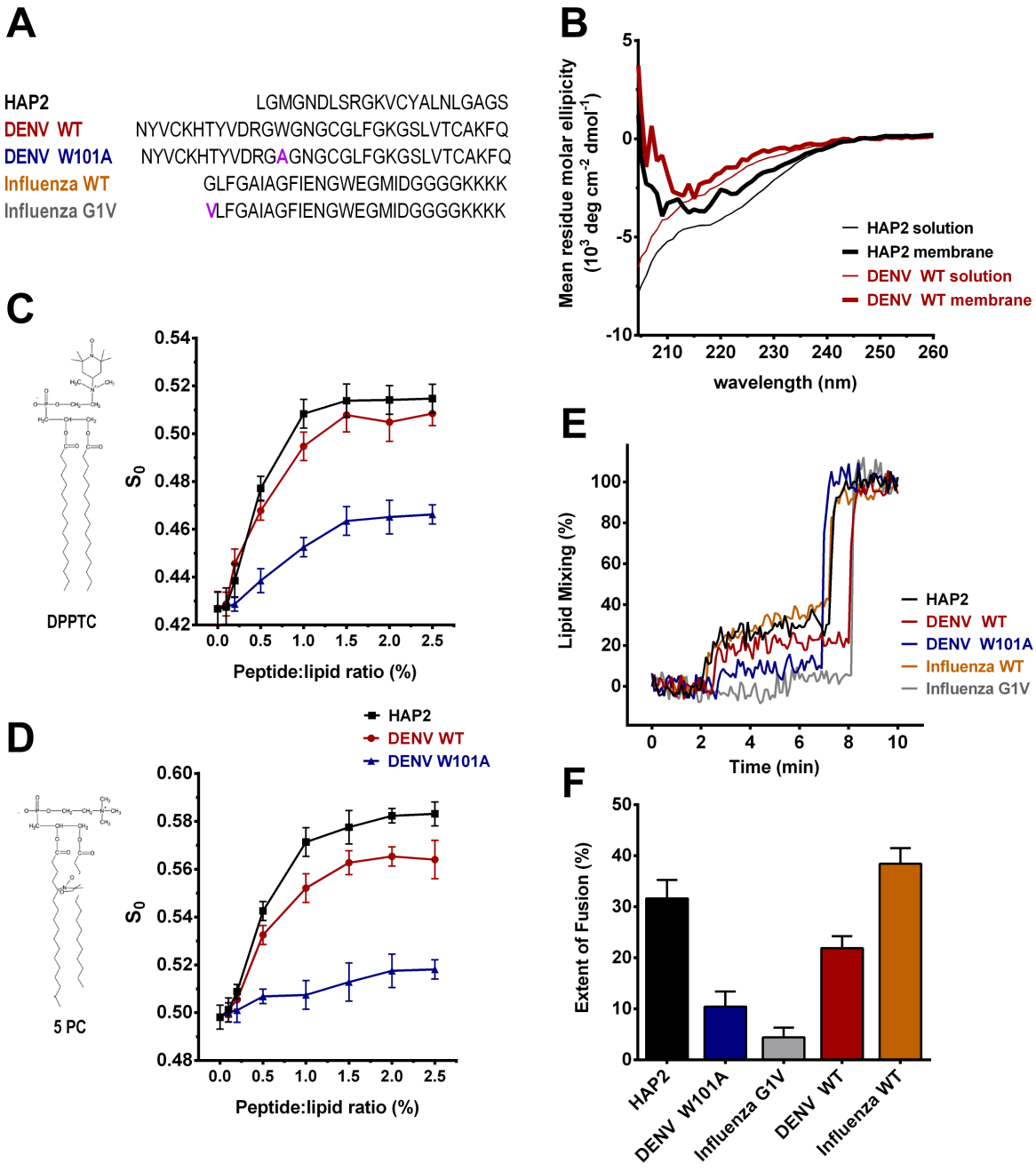

A
DENV region
SP
FL
TM
B
HA
Δ281-329, ΔHAP2 domain
H/G
Δ93-280, ΔDENV region
Δ152-179, ΔFusion Loop
DENV FL
Δ152-179 + 17, DENV FL Rescue
Δ510-513
HAP2 FL Rescue
130
179
KFQYDSKGQKILDSQGYCCYCSLSDILGMGNDLSRGKVCYALNLGAGSAT
FQY131-3AAA
CC147-8SS
R164A
LNL171-3AAA
B
FLAG-10xHis
575
612
CCCCCKSKKKENEKNKDKTDKKSIQESCSYDRSCCSHS
C575-9S, C5→S
C575-9S + C602S + C608-9S, C8→S
Δ580-596, ΔBasic Domain
Δ580-742, ΔC'term
C
Percent Fusion
100
80
60
40
20
0
ΔHAP2 domain
ΔDENV region
ΔFusion Loop
DENV FL Rescue
Δ510-513
HAP2 FL Rescue
FQY131-3AAA
CC147-8SS
R164A
LNL171-3AAA
D
Percent Fusion
100
80
60
40
20
0
C5→S
C8→S
ΔBasic Domain
ΔC'term

Figure 3.5 Interaction of the *T. thermophila* HAP2 fusion peptide with model membranes. **(A)** Amino acid sequences of synthetic peptides for the predicted *T. thermophila* HAP2 fusion loop (HAP2), the wild type Dengue Virus fusion loop (DENV WT), a mutant version of the DENV peptide (DENV W101A), the wild type Influenza virus fusion peptide (Influenza WT), and a mutant version of the Influenza fusion peptide (Influenza G1V) are shown. Amino acid substitutions in the mutant peptides that reduce fusogenic activity are indicated in purple letters. **(B)** Circular dichroism spectra of the DENV and predicted *Tetrahymena* HAP2 fusion loop peptides (2 μg/mL; pH 5) in the presence (thick line) or absence (thin line) of small unilamellar vesicles. **(C)** The head group (DPPTC) and **(D)** acyl chain (5PC) spin-labeled lipids (left), are shown next to their corresponding electron spin resonance (ESR) plots (right). ESR plots depict the order parameter (S_0) of spin-labeled lipids within multilamellar liposome vesicles (y-axis) plotted as a function of increasing peptide to lipid ratio (x-axis). Data points and error bars represent the mean +/- s.d. for 2 (DENV W101A) or 3 (DENV WT and HAP2) independent experiments. **(E)** Raw data from a representative lipid mixing experiment showing R18 fluorescence dequenching (y-axis) over time (x-axis). Synthetic fusion peptides were added to a mixed population of R18-quenched and unlabeled liposomes at ~2 min, followed by Triton X-100 at 7-8 min to establish maximum dequenching values for normalization purposes. **(F)** Bar chart showing the mean and s.d. (error bars) for normalized percent lipid mixing data from 3 independent experiments. All measurements were made at 25˚C and membrane compositions consisted of POPC:POPG:Chol=5:2:3.


A
HAP2 LGMGNDLSRGKVCYALNLGAGS
DENV WT NYVCKHTYVDRGWGNGCGLFGKGSLVTCAKFQ
DENV W101A NYVCKHTYVDRGAGNGCGLFGKGSLVTCAKFQ
Influenza WT GLFGAIAGFIENGWEGMIDGGGGKKKK
Influenza G1V VLFGAIAGFIENGWEGMIDGGGGKKKK
B
Mean residue molar ellipicity (10^3 deg cm^{-2} dmol^{-1})
wavelength (nm)
HAP2 solution
HAP2 membrane
DENV WT solution
DENV WT membrane
C
DPPTC
S_0
Peptide:lipid ratio (%)
D
5 PC
HAP2
DENV WT
DENV W101A
S_0
Peptide:lipid ratio (%)
E
Lipid Mixing (%)
Time (min)
HAP2
DENV WT
DENV W101A
Influenza WT
Influenza G1V
F
Extent of Fusion (%)
HAP2
DENV W101A
Influenza G1V
DENV WT
Influenza WT

These data indicate that the respective WT fusion peptides can insert into membranes, and supports the idea that the *T. thermophila* HAP2 fusion loop participates in membrane fusion during mating.

To test the fusogenic capacity of the predicted HAP2 fusion loop directly, we conducted lipid-mixing assays with the synthetic peptide from *Tetrahymena* HAP2 and known fusion peptides from both class I and class II viral envelope proteins. Lipid-mixing assays are based on the fusion-sensitive fluorescence emission of the lipophilic dye, R18, which becomes dequenched upon the merger of labeled and unlabeled large unilamellar vesicles (LUVs). Representative curves of raw data from one of three independent experiments (Figure 3.5E) show the increase in R18 fluorescence of LUVs between 2 and 7 min following addition of fusion peptide. Figure 3.5F shows the normalized aggregate results of these experiments. As expected, the non-fusogenic mutant peptides, DENV W101A, and G1V from the Influenza hemagglutinin (Figure 3.5A), promoted only low levels of lipid mixing (10% and 4%, respectively). However, like the WT viral fusion peptides, the synthetic peptide from *T. thermophila* HAP2 promoted high levels of lipid mixing (32% compared to 22% for DENV WT and 38% for Influenza WT peptides). Taken together these data indicate that a synthetic peptide corresponding to the predicted fusion loop of *T. thermophila* HAP2 is capable of interacting directly with membranes, inducing membrane ordering and promoting vesicle fusion.

Discussion

Measurement of protein exchange across the conjugation junction by flow cytometry offered a useful method to assay pore formation in mating cultures of *T. thermophila,* and to rapidly probe the functional role of individual amino acids and peptide motifs within the *Tetrahymena* HAP2 protein. Using this assay, we found that large alterations to the cytosolic domain had little to no effect on membrane pore formation in the *Tetrahymena* system. While this conflicts somewhat with data from other systems[14–16] it may suggest that the sequence

requirements in this region of the protein differ in different organisms. By contrast, we found that alterations to the ectodomain had significant effects on the ability of mating *T. thermophila* to form fusion pores. Template-based structural prediction modeling indicated that this region of the protein has an extended shape with three, largely β-sheet-containing folds along with a predicted fusion loop that are highly reminiscent of class II viral fusion proteins (Figure 3.3B-D). A similar topology extends to HAP2 orthologs from a wide array of other species (Figure 3.3E, Table S-3.5) and has now been validated by X-ray crystallography of the *C. reinhardtii* HAP2 ectodomain (Fedry et al., submitted). Together with the functional data described here, these studies provide overwhelming evidence that HAP2 is a bona fide membrane fusogen.

Current findings suggest two primary mechanisms by which HAP2 could drive gamete fusion, one used by class II viral fusogens for the invasion of host cells, and the other utilized by the structurally related developmental fusogen, EFF-1, to mediate syncytia formation in embryos and larvae of the nematode worm *Caenorhabditis elegans*[38, 39]. In the first case, the low pH of the endosomal compartment triggers exposure and insertion of a hydrophobic fusion loop into the outer leaflet of endosomal membranes[26]. Trimerization and structural rearrangements of the protein then draw apposed cellular and viral membranes into close proximity allowing fusion to occur[25, 40]. The developmental fusogen, EFF-1, on the other hand, lacks an obvious fusion peptide, and while adopting the same overall 3D fold as class II viral fusogens, is thought to rely primarily on conformational changes following trans-trimerization of monomers on apposed membranes to drive cell-cell fusion[38]. Consequently, while class II viral fusogens are initially present on only one membrane, EFF-1 is required on the surfaces of both apposed membranes in order to mediate lipid mixing[41]. With respect to these mechanisms, HAP2 appears sufficient to catalyze membrane fusion when present on only one membrane (i.e. on male gametes) in sexually dichotomous species. Furthermore, in the case of *T. thermophila* (which expresses HAP2 in all mating types), fusion can still occur, albeit at reduced levels, when the protein is expressed in only one cell of a mating pair (Figure 3.2D,G). Perhaps more importantly, it now

appears that HAP2 orthologs contain functional fusion loops. As shown here, membrane pore formation in mating *T. thermophila* is blocked when the predicted fusion loop is deleted from the native protein (Figure 3.4C) and a synthetic peptide corresponding to this region of *Tetrahymena* HAP2 can interact directly with model membranes and mediate lipid mixing to the same extent as viral fusion peptides (Figure 3.5C-F).

While the current data argue that HAP2 and the class II viral proteins act by similar mechanisms, the reduced levels of fusion observed when *T. thermophila* HAP2 is expressed in only one cell of a mating pair opens the possibility that these cells utilize a hybrid mechanism involving both insertion of a fusion loop (as in the case of the viral proteins) and trans-interactions between HAP2 on apposed membranes (as has been proposed for EFF-1) to bring about efficient membrane mixing. A more complete understanding of the mechanisms underlying HAP2-mediated membrane fusion would clearly benefit from alternative approaches used previously to dissect class II protein function, in particular, fusion assays with pseudotyped viruses and/or cultured cells expressing recombinant forms of the protein[25, 42]. *T. thermophila* HAP2 can be efficiently expressed in mammalian tissue culture cells and incorporated into retroviral pseudotyped particles (Figure S-3.7). Nevertheless, the *Tetrahymena* protein failed to mediate entry of pseudotyped virus particles, and while the results with mammalian tissue culture cells was intriguing (Figure S-3.7), additional studies will be required to demonstrate this more convincingly and identify the specific environmental factors necessary for triggering fusion events.

Regardless of the precise mechanism underlying HAP2-mediated membrane fusion, the evolutionary relationship between HAP2 and class II viral proteins is clearly interesting. While it is possible these proteins arose through convergent evolution, the overall topologies and near identical folding patterns of HAP2 and class II viral fusogens (Fedry et al., submitted) makes this improbable[38]. The alternative hypothesis, namely, that they evolved from a common ancestor is certainly more plausible but leaves open the question of which came first. Recent evidence

that similar class II structures are present in phylogenetically distinct virus families has suggested that the coding elements for these proteins arose independently in different viral lineages through the capture of a cellular gene encoding either a developmental fusogen with class II structure, or a bygone viral fusogen piggybacking in the host genome[24, 43–47]. HAP2, on the other hand, is present within the basal lineages of all major branches of the eukaryotic tree of life and likely dates to the last common ancestor of all eukaryotes. This would make it the oldest class II fusogen that we know of and thus a strong candidate as the ancestral fusogen from which other class II proteins evolved. At the same time, the existence of viruses pre-dates the evolution of eukaryotic sex[48–50], and it is equally plausible that HAP2 originated with a virus, was exapted for use in gamete cell fusion early in the course of eukaryotic evolution, and perhaps then reacquired by modern viruses. Invasion of eukaryotic genomes by viruses is widespread[43, 45, 47], and there is clear evidence that genes for viral fusogens have taken on new functions in the case of mammalian syncytins, which are of retroviral origin, and promote cell-cell fusion during placentation in diverse species[51, 52]. While their evolutionary relationship is less clear, eukaryotic SNAREs and class I viral fusogens also appear to be structurally related, and may offer yet another example of the repurposing of viral fusion proteins[46, 53, 54].

The strict requirement for cell-cell fusion in sexual reproduction combined with the ancient lineage of HAP2 and its role in fertility in a broad range of taxa, argues persuasively for the involvement of this protein in the origin of eukaryotic sex[55, 56]. This argument becomes all the more interesting if HAP2 arose from a virus or related parasitic DNA element such as a transposon. First, it would suggest that a key step (if not the key step) in fertilization was made possible by a virus. Absent that step, sex, and the diversity of life that it spawned (including man), may never have evolved. Second, a role for parasitic DNA in the origin of eukaryotic sex has long been argued. As proposed originally by Donal Hickey, sex may have arisen as a byproduct of selective pressure on some hypothetical fragment of selfish DNA to spread

horizontally from cell to cell, thus favoring its survival[57, 58]. An endogenous viral element or related fragment of parasitic DNA could do this by attaining the capacity to promote cell-cell fusion. Theoretically this could have occurred through natural selection on a given DNA element (for example, by evolving a coding sequence for a membrane fusogen), or more simply, through the capture of a gene for a viral fusogen following the infection of some early eukaryotic cell. In either case, the acquisition of such a coding sequence would have served as the starting point for additional evolutionary tinkering to deal with the consequences of fusion, leading eventually to the emergence of karyogamy, meiosis, recombination, and all the modern manifestations of eukaryotic sex[56].

While HAP2 is deeply rooted in the eukaryotic tree of life, a number of taxonomic lineages, including vertebrates and fungi, lack HAP2 orthologs. Such sexual species have clearly found an alternative means to accomplish gamete membrane fusion, but whether they use novel proteins, or variants of currently known fusion proteins remains to be determined. To begin to address this question, we have probed mammalian genomes with appropriate search engines, such as BackPhyre[20], and identified open reading frames predicted to encode elements of the class II fusogen fold. Moving forward, it will be interesting to determine whether such sequences are the remnants of viral genes for class II proteins[43], specify novel cellular proteins that catalyze membrane fusion events (perhaps even replacing HAP2), or represent instances in which the class II fold has taken on entirely new functions.

Materials and methods

***Tetrahymena* strains and culture conditions**. *Tetrahymena thermophila* strains were obtained from the Tetrahymena Stock Center, Cornell University (https://tetrahymena.vet.cornell.edu/). All established and newly created cell lines are described in Table S-3.1. For routine growth, cells were incubated at 30˚C in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 33.3 μM $FeCl_3$) on a platform shaker at ~100 rpm. For mating studies, log phase cells of complementary mating types were starved in 10 mM Tris buffer (pH 7.5) for up to 48 h, then mixed in equal numbers to a final concentration of 2 × 10^5 cells/mL at 30˚C. For somatic (macronuclear) transformation, target cells were grown to late log phase (~1 × 10^6 cells/mL) in NEFF medium and starved overnight in 10 mM Tris buffer (pH 7.5) for 24 h at ~2 × 10^5 cells/mL prior to biolistic transformation and drug selection (see below, *T. thermophila* strain construction).

***T. thermophila* strain construction.** Cell lines used for the creation of HAP2 mutations / truncations at the endogenous *T. thermophila* HAP2 gene locus are designated *ΔHAP2-428* clone 5 and *ΔHAP2-427* clone 6 and are derivatives of the heterokaryon strains CU428.2 and CU427.4[59]. These strains express complementary mating types (VII and VI respectively) and lack the entire HAP2 coding sequence in the macronucleus[8]. Relevant gene constructs were introduced into the knockout cell lines using a gene-gun (see below) and stable drug-resistant transformants were selected by growth in cycloheximide following homologous recombination at the HAP2 locus (Table S-3.1).

Mutant HAP2 gene constructs were prepared using either overlap PCR, or in some cases, site-directed mutagenesis with a Q5® Site-Directed Mutagenesis kit (New England BioLabs). A full list of PCR primers and their use in gene construction is provided in Table S-3.2. In all cases, epitope tags (either HA or Flag-His) were added to the 3’ ends of the HAP2 cDNA constructs by PCR to permit localization of the recombinant gene products. All PCR reactions

were carried out with Phusion High-Fidelity *Taq* DNA Polymerase (ThermoFisher). HAP2 PCR products were gel purified and cloned into a previously constructed pCR™4Blunt-*TOPO*® plasmid vector backbone (ThermoFisher), which had been modified to contain ~1000 bp of 5' and 3' flanking sequences from the *T. thermophila* HAP2 gene[8]. *Bam*HI and *Kpn*I restriction sites situated between the *HAP2* 5' and 3' flanks were used for directional cloning of coding sequences for the mutated/truncated *HAP2* gene products[8]. The modified vector also contained a pHrpl29-B cycloheximide resistance cassette within the 3' flanking sequence 387 bp downstream of the *Kpn*I restriction site. Purified HAP2 PCR fragments and vector DNA were digested with *Bam*HI and *Kpn*I and ligated with T4 DNA ligase (New England BioLabs,) prior to transformation and amplification in E. cloni 10G competent cells (Lucigen). After transformation, all plasmid DNA HAP2 sequences were verified by sequencing (Cornell Biotechnology Resource Center) to confirm correct gene construction.

For transformation into *T. thermophila*, plasmid DNA harboring relevant inserts was purified, linearized by digestion with *Eco*RI, and introduced into the macronucleus of *ΔHAP2* parental strains via biolistic bombardment[60] using a PDS-1000/He Biolistic Particle Delivery System (Bio-Rad). Positive transformants were selected in NEFF medium containing 25 μg/mL cycloheximide, and supplemented with 1.25 μg/mL Fungizone, 250 μg/mL Streptomycin, and 250 μg/mL Penicillin G. Transformed clones were then pushed to complete macronuclear replacement for the target construct via growth in NEFF medium containing increasing concentrations of cycloheximide (up to 50 μg/mL). Genomic DNA was extracted with phenol:chloroform:iso-Amyl alcohol (25:24:1, VWR), PCR amplified and the resulting products analyzed by agarose gel electrophoresis and, in some cases, sequencing.

In addition to creating strains in which mutant gene constructs were targeted to the endogenous HAP2 locus, a cell line that over-expresses HAP2 transcripts in an otherwise wild type HAP2 background was generated using the stable, high-copy ribosomal DNA vector pTRAS (Tetragenetics). For this strain, a full-length HAP2 cDNA was amplified with the primer

pair Hap2BamH1for/GSP2Kpn1rev (Table S-3.2). The resulting PCR product was then digested with *Bam*HI and *Kpn*I, gel purified, and ligated to a *Kpn*I-digested fragment of synthetic DNA containing the following multi-sequence tag: *Kpn*I restriction site; PreScission S protease cleavage sequence; Streptavidin Binding Peptide sequence; and, 6× His tag. The ligation product was then purified and subjected to PCR for the further addition of a 3' HA tag and *Sac*I restriction site using the primer pairs BamHIHAP2for/HA-tag Rev (Table S-3.2). This PCR product was gel purified and restricted with *Bam*HI and *Sac*I for ligation into the shuttle vector pTIEV4 (Tetragenetics). The *HAP2-tags*-pTIEV4 plasmid was amplified in E. cloni 10G (Lucigen), purified, and then cut at a unique *Not* I site. The insert was then sub-cloned into pTRAS, downstream of a cadmium-inducible promoter from the *MTT1* gene of *T. thermophila* [61]. Finally, the recombinant vector was biolistically transformed into 8-10 h mating cultures of *Tetrahymena* strains CU428 x B2086 as described above. Macronuclear transformants were selected by growth in NEFF medium containing increasing concentrations of paromomycin (up to 800 µg/mL) and then frozen in liquid nitrogen[62].

Flow cytometry assays for cell-cell fusion. Complementary mating types of *T. thermophila* were grown and placed in starvation medium (10 mM Tris buffer, pH 7.5) for 24 h to ensure all cells had completed asexual division and were arrested at the G1 stage[63]. Cells were then washed once by centrifugation at 400 × g and resuspended in 0.1 × PBS prior to labeling. All labeling reactions and centrifugation steps were carried out in 15 mL glass conical tubes. One mL of 0.1 × PBS containing 7×10^6 cells was combined with 1 mL of the same buffer containing either 20 µM CFSE (Affymetrix eBioscience) or 10 µM CTFR (Life Technologies), and then incubated in the dark for 5 min at room temperature (RT), or 15 min at 30˚C for CFSE and CTFR, respectively. Ten mL of NEFF media was then immediately added to quench excess unbound label, and cells were washed and resuspended in 10 mL of 10 mM Tris (pH 7.5). Cells were maintained overnight in the dark at 30˚C, and the following day, washed again (10 mL of 10 mM Tris) and counted. Equal numbers of cells of each mating type were then combined in

100 mm × 10 mm petri dishes (0.5–2 × 10^6 total cells/dish), and allowed to mate for 16-20 h at 30˚C in a darkened incubator. Following mating, exconjugant cells were centrifuged (350-400 × g) and fixed with IC Fixation buffer (Affymetrix eBioscience) in the dark for 20 min at RT, then resuspended in 1 × PBS containing 0.3% BSA prior to acquisition on a BD FACSCanto™ II Flow Cytometer.

For each mating reaction a minimum of 30,000 events were acquired. Data were analyzed using FlowJo software (FlowJo LLC). Unmated, single-labeled control cell cultures were fixed at the same time as mating cultures and subjected to flow cytometry as above. Labeled populations from the unmated cultures served as guides for drawing gates around double-labeled populations from experimental mating cultures. The starved unmated populations were also used to estimate fluorescence loss due to “co-stimulation” (Figure 3.2) after it became clear these populations had higher fluorescence intensity than the parallel, single-labeled populations in mating cultures. Bar charts displaying the frequency of “Mid”-fluorescence events (percent of cells that had undergone fusion) in each cross were generated using Prism Software (GraphPad Inc.).

Measurement of cell-cell pairing and membrane fusion kinetics. Complementary mating types of *T. thermophila* were labeled and mixed together to initiate mating as described above. Samples from the mating cultures were then collected and fixed at the indicated time points within the first 4 h after mixing. Cells were observed under phase and fluorescence optics to determine the percent of cells in pairs (% pairing), as well as the percentage of pairs that had visibly exchanged fluorescent content (% fusion). For % pairing, at least 100 “subjects” (pairs or single cells) were counted for each time point and the percentage calculated as the number of cells in pairs over the total number of cells counted multiplied by 100. For % fusion, at least 50 pairs were counted for each time point with their fusion status determined based on visual detection of fluorescent tracer in only one (not fused) versus both (fused) cells of a mating pair. The % fusion was calculated as the number of pairs fused over the total number of pairs

counted multiplied by 100. It should be noted that at the earliest time points (≤1 h) there were fewer than 50 pairs available to count and so measurements of fusion at these times were less robust. Cell counts, as well as the fluorescence micrographs shown in Figure 3.1, were obtained using a Zeiss Axio Imager M1 microscope equipped with an AxioCamMR3 camera. Prism software was used to generate scatter plots showing the kinetics of cellular pairing/fusion over time.

Template-based structural homology modeling. Initial protein homology modeling studies were conducted in 2013 using the Protein Homology/analogY Recognition Engine V2.0 (Phyre2) and the full-length amino acid sequence of *T. thermophila* HAP2 (accession number: KJ629172) as query. Phyre2 uses advanced remote homology detection algorithms involving alignments of amino acid sequences and predicted secondary structures to identify template hits within a hidden Markov model database of known structures. The identified top-scoring hits help form an alignment for the construction of a crude backbone structure while loop modeling and side chain placement are subsequently applied to finalize the predicted structure[20]. Follow-up Phyre2 searches were carried out between 2014-2016 after other relevant class II structures were published (e.g. EFF-1), but yielded equivalent results. A Phyre batch processing query using 40 published[15] HAP2 sequences was submitted in June 2015 (accession numbers of sequences listed in Table S-3.5). Results yielded the top 20 ranked hits to known structures for each species' version of HAP2, as well as a confidence level for the prediction, a structural ID and template name, query start/stop sites of the aligned sequence, and a predicted structure of each ortholog based on the top ranking template hit. A partial list summarizing the top three ranking hits from both viral and non-viral template structures is shown in Table S-3.5, as well as a list of the top viral hits in Figure 3.3E.

Submission of the *T. thermophila* HAP2 sequence to the RaptorX[21,64], CPHmodels3.0[22] and HHpred[65] (Figure S-3.4C) template-based homology modelers occurred between 2015-2016. Images of predicted structural models were generated using

PyMol (Schrodinger, LLC). In all cases, domains I, II, and III were shaded red, yellow, and blue according to the boundary locations for these domains inferred from the generated alignments of *T. thermophila* HAP2 and known class II fusion protein structures.

Lipids and peptides used for biophysical studies. The lipids POPC, POPG, the chain spin label, 5PC, and the head group spin label, dipalmitoylphosphatidyl-tempo-choline (DPPTC), were purchased from Avanti® Polar Lipids. Cholesterol was purchased from Sigma. HAP2 and Dengue virus fusion peptides were synthesized by ChinaPeptides Co., Ltd., and the Influenza fusion peptides by SynBioSci Co. The sequences of wild type and mutant peptides, along with the structures of the spin labeled lipids are shown in Figure 3.5A,C,D.

Preparation of membrane vesicles. The volumes of POPC, POPG, cholesterol and 0.5% (mol:mol) spin-labeled lipids in chloroform were mixed according to a 5:2:3 vol:vol ratio of POPC:POPG:Chol and dried by N_2 flow. The mixture was evacuated in a vacuum drier overnight to remove any trace of chloroform. To prepare multilamellar vesicles (MLVs), the lipids were resuspended and fully hydrated using 1 mL of pH 5 buffer (5 mM HEPES, 10 mM MES, 150 mM NaCl, pH 5) at room temperature for 2 h. To prepare small unilamellar vesicles (SUVs), the lipids were resuspended in pH 5 buffer and sonicated in ice bath for 20 min. To prepare large unilamellar vesicles (LUVs), the lipids were frozen and thawed 5 times before they were extruded in an Avanti extruder through a membrane with 100 nm pore size.

Electron spin resonance (ESR) spectroscopy and nonlinear least-squares fit of ESR spectra. To prepare the samples for lipid ESR study, a stock solution of the Fusion Peptide (FP) (1 mg/mL) was added to the lipid POPC:POPG:Chol=5:2:3 MLV dispersion (above) at the experimentally indicated ratios. After 20 min of incubation, the dispersion was centrifuged at 13,000 rpm for 10 min. The pellet was transferred to a quartz capillary tube for ESR measurement. ESR spectra were collected on an ELEXSYS ESR spectrometer (Bruker Instruments) at X-band (9.5 GHz) at 25˚C using a N2 Temperature Controller (Bruker Instruments). The ESR spectra from the labeled lipids were analyzed using the NLLS fitting

program based on the stochastic Liouville equation[66] using the MOMD (Microscopic Order Macroscopic Disorder) model as in previous studies[31,67–69]. The fitting strategy is the same as previously reported[33]. S_0 is defined as follows: $S_0=<D_{2,00}>=<1/2(3\cos^2\theta-1)>$, where $D_{2,00}$ is the Wigner rotation matrix elements and θ is the polar angle for the orientation of the rotating axes of the nitroxide bonded to the lipid relative to the director of the bilayer, i.e. the preferential orientation of lipid molecules[67,70], and the angular brackets imply ensemble averaging. S_0 indicates how well the chain segment to which the nitroxide is attached, is aligned along the normal to the lipid bilayer[33].

Circular dichroism spectroscopy (CD). Fusion peptides (0.2 mg/mL in pH 5 solution) were mixed with SUVs composed of POPC:POPG:Chol=5:2:3 at a ratio of 1:100 peptide:lipid at room temperature for 10 min before measurement. The CD spectra were collected at 25˚C on an AVIV Model 202-01 Circular Dichroism Spectrometer (AVIV biomedical Inc.). The signals from pure SUVs or pure solution were subtracted from the sample spectra as blanks. The CD spectra were analyzed using K2D3[71].

Fluorescence dequenching assays. The protocol for fluorescence dequenching assays to monitor vesicle fusion was adopted from a previous study[72]. Fluorescently labeled LUVs (2.5 μM, final concentration) containing 2% Octadecyl Rhodamine B chloride (R18, Molecular Probe, ThermoFisher Scientific) and unlabeled LUV (22.5 μM, final concentration) were mixed in 1 mL of pH 5 buffer. Fusion peptides were then added from concentrated stock solutions to give a 1 μM final concentration of each peptide. 10% Triton X-100 was added to achieve a 1% final concentration after fusion reactions were complete. The fluorescence spectra were collected on a Varian Cary Eclipse Fluorescence Spectrometer (Agilent Technologies). Fluorescence intensities of the samples before addition of fusion peptides and after the addition of Triton X-100 were used to set the baseline (0%) and 100% fusion levels, respectively. The fluorescence yields of the experimental samples were normalized to these levels to determine % lipid

mixing[73]. Fluorescence intensity variations due to volume changes were corrected in each case. All experiments were performed at least 3 times and representative curves are shown.

Statistical information. All statistical tests were performed using Prism 7 software (GraphPad Inc.). Sample sizes for the functional analyses of HAP2 mutant strains by flow cytometry (Figures 3.2G and 3.4B,D) are listed in parentheses below with the total number of biological replicates over all experiments (that is, total number of individual matings performed) listed first, followed by the total number of independent experiments performed for each cross: $WT^{\alpha \times \beta}$(42, 17); Δ*HAP2*$^{\alpha \times \beta}$(20, 8); WT^{β} x Δ*HAP2*$^{\alpha}$ (31,10); WT^{α} x Δ*HAP2*$^{\beta}$ (26,10); Genomic $^{\beta}$ x WT^{α} (4, 2); Genomic $^{\beta}$ x Δ*HAP2*$^{\alpha}$ (5, 3); cDNA $^{\beta}$ x WT^{α}(5, 4); Δ*HAP2* domain (6, 5); ΔDENV region (5, 4); ΔFusion Loop(10, 5); DENV FL Rescue (12, 3); Δ510-513 (7, 3); HAP2 FL Rescue (6, 3); FQY131-3AAA (8, 3); CC147-8SS (16, 4); R164A (9, 3); LNL171-3AAA (9, 3); $C_5 \rightarrow S$ (4, 3); $C_8 \rightarrow S$ (13, 5); ΔBasic Domain (10, 3); ΔC'term (9, 3). No pre-determined power analyses were performed for calculating sample sizes. The criterion for including any given cross in cell fusion analysis was a pairing frequency of >60% in mating cultures 3 h post mixing. No randomization was applied to samples and the investigators were not blinded to sample identity during the experiments.

A one-sided non-parametric Kruskal Wallis test with Dunn's multiple comparisons post-test was applied to the cellular fusion data because not all data sets passed the D'Agostino-Pearson Omnibus normality test. Associated P values and standard deviations (as measurements of variance) are shown in the text, figures, and figure legends. From these tests, there was a significant difference found in the percent fusion for data from the 8 different crosses in Figure 3.2G (H_7= 116.3, P<0.0001). For Figure 3.4C, a significant difference was also found in comparisons to either data from $WT^{\alpha \times \beta}$ or WT^{α} x KO^{β} crosses respectively (H_{11}=123.2 and H_{10}=85.78, and P<0.0001). Similarly, when the groups shown in Figure 3.4D were compared to data from $WT^{\alpha \times \beta}$ and WT^{α} x KO^{β} crosses respectively, significant differences were also found (H_4=14.92, P=0.0049 and H_4=48.17, and P<0.0001).

Data availability. All relevant data are available from the authors. Specifically, HAP2 mutant cell lines have been deposited in the *Tetrahymena* Stock Center at Cornell University and are available for use by the community at-large. GenBank Accession numbers to relevant sequences used in this study are shown in Table S-3.5. The entire Phyre batch processing data set in its raw form along with the corresponding Phyre2-predicted structures and homology models from other template-based prediction tools will be made available upon request to the corresponding author.

REFERENCES

1. Johnson, M.A., Besser, K. von, Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z., and Preuss, D. (2004). Arabidopsis hapless mutations define essential gametophytic functions. Genetics *168*, 971–982.

2. Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). Generative Cell Specific 1 is essential for angiosperm fertilization. Nat. Cell Biol. *8*, 64–71.

3. Wong, J.L., and Johnson, M.A. (2010). Is HAP2-GCS1 an ancestral gamete fusogen? Trends Cell Biol. *20*, 134–141.

4. Blagborough, A.M., and Sinden, R.E. (2009). Plasmodium berghei HAP2 induces strong malaria transmission-blocking immunity in vivo and in vitro. Vaccine *27*, 5187–5194.

5. Ebchuqin, E., Yokota, N., Yamada, L., Yasuoka, Y., Akasaka, M., Arakawa, M., Deguchi, R., Mori, T., and Sawada, H. (2014). Evidence for participation of GCS1 in fertilization of the starlet sea anemone Nematostella vectensis: Implication of a common mechanism of sperm–egg fusion in plants and animals. Biochem. Biophys. Res. Commun. *451*, 522–528.

6. Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization. Nat. Cell Biol. *8*, 64–71.

7. Besser, K. von, Frank, A.C., Johnson, M.A., and Preuss, D. (2006). Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development *133*, 4761–4769.

8. Cole, E.S., Cassidy-Hanley, D., Pinello, J.F., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Jr, T.G., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. Curr. Biol. CB *24*, 2168–2173.

9. Liu, Y., Tewari, R., Ning, J., Blagborough, A.M., Garbom, S., Pei, J., Grishin, N.V., Steele, R.E., Sinden, R.E., Snell, W.J., *et al.* (2008). The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in Chlamydomonas and Plasmodium gametes. Genes Dev. *22*, 1051–1068.

10. Liu, Y., Misamore, M.J., and Snell, W.J. (2010). Membrane fusion triggers rapid degradation of two gamete-specific, fusion-essential proteins in a membrane block to polygamy in Chlamydomonas. Dev. Camb. Engl. *137*, 1473–1481.

11. Hirai, M., Arai, M., Mori, T., Miyagishima, S.Y., Kawai, S., Kita, K., Kuroiwa, T., Terenius, O., and Matsuoka, H. (2008). Male fertility of malaria parasites is determined by GCS1, a plant-type reproduction factor. Curr. Biol. CB *18*, 607–613.

12. Kawai-Toyooka, H., Mori, T., Hamaji, T., Suzuki, M., Olson, B.J.S.C., Uemura, T., Ueda, T., Nakano, A., Toyoda, A., Fujiyama, A., *et al.* (2014). Sex-Specific Posttranslational Regulation of the Gamete Fusogen GCS1 in the Isogamous Volvocine Alga Gonium pectorale. Eukaryot. Cell *13*, 648–656.

13. Okamoto, M., Yamada, L., Fujisaki, Y., Bloomfield, G., Yoshida, K., Kuwayama, H., Sawada, H., Mori, T., and Urushihara, H. (2016). Two HAP2-GCS1 homologs responsible for gamete interactions in the cellular slime mold with multiple mating types: Implication for common mechanisms of sexual reproduction shared by plants and protozoa and for male-female differentiation. Dev. Biol. *415*, 6–13.

14. Wong, J.L., Leydon, A.R., and Johnson, M.A. (2010). HAP2(GCS1)-dependent gamete fusion requires a positively charged carboxy-terminal domain. PLoS Genet. *6*, e1000882.

15. Liu, Y., Pei, J., Grishin, N., and Snell, W.J. (2015). The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in Chlamydomonas and regulates the fusion reaction. Development *142*, 962–971.

16. Mori, T., Hirai, M., Kuroiwa, T., and Miyagishima, S.Y. (2010). The functional domain of GCS1-based gamete fusion resides in the amino terminus in plant and parasite species. PloS One *5*, e15957.

17. Cole, E.S. (2000). The Tetrahymena Conjugation Junction. Madame Curie Biosci. Database. Available at: http://www.ncbi.nlm.nih.gov/books/NBK6002/.

18. Nasir, A.M., Yang, Q., Chalker, D.L., and Forney, J.D. (2015). SUMOylation Is Developmentally Regulated and Required for Cell Pairing during Conjugation in Tetrahymena thermophila. Eukaryot. Cell *14*, 170–181.

19. Kelley, L.A., and Sternberg, M.J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. *4*, 363–371.

20. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

21. Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., and Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. *7*, 1511–1522.

22. Nielsen, M., Lundegaard, C., Lund, O., and Petersen, T.N. (2010). CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. Nucleic Acids Res. *38*, W576–W581.

23. Modis, Y., Ogata, S., Clements, D., and Harrison, S.C. (2005). Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. J. Virol. *79*, 1223–1231.

24. Dessau, M., and Modis, Y. (2013). Crystal structure of glycoprotein C from Rift Valley fever virus. Proc. Natl. Acad. Sci. U. S. A. *110*, 1696–1701.

25. Kielian, M., and Rey, F.A. (2006). Virus membrane-fusion proteins: more than one way to make a hairpin. Nat. Rev. *4*, 67–76.

26. Zaitseva, E., Yang, S.T., Melikov, K., Pourmal, S., and Chernomordik, L.V. (2010). Dengue virus ensures its fusion in late endosomes using compartment-specific lipids. PLoS Pathog. *6*, e1001131.

27. Seligman, S.J. (2008). Constancy and diversity in the flavivirus fusion peptide. Virol. J. *5*, 27.

28. Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C., and Harrison, S.C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. Nature *375*, 291–298.

29. DuBois, R.M., Vaney, M.-C., Tortorici, M.A., Kurdi, R.A., Barba-Spaeth, G., Krey, T., and Rey, F.A. (2013). Functional and evolutionary insight from the crystal structure of rubella virus protein E1. Nature *493*, 552–556.

30. Chernomordik, L.V., and Kozlov, M.M. (2008). Mechanics of membrane fusion. Nat. Struct. Mol. Biol. *15*, 675–683.

31. Ge, M., and Freed, J.H. (2003). Hydration, Structure, and Molecular Interactions in the Headgroup Region of Dioleoylphosphatidylcholine Bilayers: An Electron Spin Resonance Study. Biophys. J. *85*, 4023–4040.

32. Lai, A.L., and Freed, J.H. (2014). HIV gp41 fusion peptide increases membrane ordering in a cholesterol-dependent fashion. Biophys. J. *106*, 172–181.

33. Lai, A.L., and Freed, J.H. (2015). The Interaction between Influenza HA Fusion Peptide and Transmembrane Domain Affects Membrane Structure. Biophys. J. *109*, 2523–2536.

34. Dimitrov, A.S., Rawat, S.S., Jiang, S., and Blumenthal, R. (2003). Role of the fusion peptide and membrane-proximal domain in HIV-1 envelope glycoprotein-mediated membrane fusion. Biochemistry (Mosc.) *42*, 14150–14158.

35. Qiao, H., Armstrong, R.T., Melikyan, G.B., Cohen, F.S., and White, J.M. (1999). A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype. Mol. Biol. Cell *10*, 2759–2769.

36. Melo, M.N., Sousa, F.J.R., Carneiro, F.A., Castanho, M.A.R.B., Valente, A.P., Almeida, F.C.L., Da Poian, A.T., and Mohana-Borges, R. (2009). Interaction of the Dengue Virus Fusion Peptide with Membranes Assessed by NMR: The Essential Role of the Envelope Protein Trp101 for Membrane Fusion. J. Mol. Biol. *392*, 736–746.

37. Stauffer, F., Melo, M.N., Carneiro, F.A., Sousa, F.J.R., Juliano, M.A., Juliano, L., Mohana-Borges, R., Poian, A.T.D., and Castanho, M.A.R.B. (2008). Interaction between dengue virus fusion peptide and lipid bilayers depends on peptide clustering. Mol. Membr. Biol. *25*, 128–138.

38. Perez-Vargas, J., Krey, T., Valansi, C., Avinoam, O., Haouz, A., Jamin, M., Raveh-Barak, H., Podbilewicz, B., and Rey, F.A. (2014). Structural basis of eukaryotic cell-cell fusion. Cell *157*, 407–419.

39. Mohler, W.A., Shemer, G., Campo, J.J. del, Valansi, C., Opoku-Serebuoh, E., Scranton, V., Assaf, N., White, J.G., and Podbilewicz, B. (2002). The type I membrane protein EFF-1 is essential for developmental cell fusion. Dev. Cell *2*, 355–362.

40. Sapir, A., Avinoam, O., Podbilewicz, B., and Chernomordik, L.V. (2008). Viral and developmental cell fusion mechanisms: conservation and divergence. Dev. Cell *14*, 11–21.

41. Avinoam, O., Fridman, K., Valansi, C., Abutbul, I., Zeev-Ben-Mordehai, T., Maurer, U.E., Sapir, A., Danino, D., Grunewald, K., White, J.M., *et al.* (2011). Conserved eukaryotic fusogens can fuse viral envelopes to cells. Science *332*, 589–592.

42. Dubé, M., Rey, F.A., and Kielian, M. (2014). Rubella virus: first calcium-requiring viral fusion protein. PLoS Pathog. *10*, e1004530.

43. Katzourakis, A., and Gifford, R.J. (2010). Endogenous Viral Elements in Animal Genomes. PLoS Genet. *6*, e1001191.

44. Lavialle, C., Cornelis, G., Dupressoir, A., Esnault, C., Heidmann, O., Vernochet, C., and Heidmann, T. (2013). Paleovirology of “syncytins”, retroviral env genes exapted for a role in placentation. Philos. Trans. R. Soc. B Biol. Sci. *368*, 20120507–20120507.

45. Robinson, L.R., and Whelan, S.P.J. (2016). Infectious Entry Pathway Mediated by the Human Endogenous Retrovirus K Envelope Protein. J. Virol. *90*, 3640–3649.

46. Modis, Y. (2014). Relating structure to evolution in class II viral membrane fusion proteins. Curr. Opin. Virol. *5*, 34–41.

47. Frame, I.G., Cutfield, J.F., and Poulter, R.T. (2001). New BEL-like LTR-retrotransposons in Fugu rubripes, Caenorhabditis elegans, and Drosophila melanogaster. Gene *263*, 219–230.

48. Koonin, E.V., Senkevich, T.G., and Dolja, V.V. (2006). The ancient Virus World and evolution of cells. Biol. Direct *1*, 29.

49. Koonin, E.V. (2009). On the origin of cells and viruses: primordial virus world scenario. Ann. N. Y. Acad. Sci. *1178*, 47–64.

50. Cavalier-smith, T. (2002). Origins of the machinery of recombination and sex. Heredity *88*, 125–141.

51. Blond, J.-L., Lavillette, D., Cheynet, V., Bouton, O., Oriol, G., Chapel-Fernandes, S., Mandrand, B., Mallet, F., and Cosset, F.-L. (2000). An Envelope Glycoprotein of the Human Endogenous Retrovirus HERV-W Is Expressed in the Human Placenta and Fuses Cells Expressing the Type D Mammalian Retrovirus Receptor. J. Virol. *74*, 3321–3329.

52. Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S., *et al.* (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature *403*, 785–789.

53. Donald, J.E., Zhang, Y., Fiorin, G., Carnevale, V., Slochower, D.R., Gai, F., Klein, M.L., and DeGrado, W.F. (2011). Transmembrane orientation and possible role of the fusogenic peptide from parainfluenza virus 5 (PIV5) in promoting fusion. Proc. Natl. Acad. Sci. U. S. A. *108*, 3958–3963.

54. Skehel, J.J., and Wiley, D.C. (1998). Coiled coils in both intracellular vesicle and viral membrane fusion. Cell *95*, 871–874.

55. Goodman, C.D., and McFadden, G.I. (2008). Gamete fusion: key protein identified. Curr. Biol. CB *18*, R571-3.

56. Speijer, D., Lukeš, J., and Eliáš, M. (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc. Natl. Acad. Sci. *112*, 8827–8834.

57. Hickey, D.A. (1982). Selfish DNA: A Sexually-Transmitted Nuclear Parasite. Genetics *101*, 519–531.

58. Hickey, Donal A., and Rose, Michael R. (1988). The role of gene transfer in the evolution of eukaryotic sex. In The Evolution of sex: an examination of current ideas, R. E. Michod and B. R. Levin, eds. (Sunderland, Mass: Sinauer Associates), pp. 161–175.

59. Bruns, Peter J, and Brussard, Trudy B. (1974). Positive selection for mating with functional heterokaryons in Tetrahymena pyriformis. Genetics *78*, 831–841.

60. Cassidy-Hanley, D., Bowen, J., Lee, J.H., Cole, E., VerPlank, L.A., Gaertig, J., Gorovsky, M.A., and Bruns, P.J. (1997). Germline and somatic transformation of mating Tetrahymena thermophila by particle bombardment. Genetics *146*, 135–147.

61. Shang, Y., Song, X., Bowen, J., Corstanje, R., Gao, Y., Gaertig, J., and Gorovsky, M.A. (2002). A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. *99*, 3734–3739.

62. Cassidy-Hanley, D., Smith, H.R., and Bruns, P.J. (1995). A simple, efficient technique for freezing Tetrahymena thermophila. J. Eukaryot. Microbiol. *42*, 510–515.

63. Wolfe, J. (1976). G1 arrest and the division/conjugation decision in Tetrahymena. Dev. Biol. *54*, 116–126.

64. Peng, J., and Xu, J. (2011). Raptorx: Exploiting structure information for protein alignment by statistical inference. Proteins Struct. Funct. Bioinforma. *79*, 161–171.

65. Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. *33*, W244-248.

66. Budil, David E., Lee, Sanghyuk, Saxena, Sunil, and Freed, Jack H. (1996). Nonlinear-Least-Squares Analysis of Slow Motion EPR Spectra in One and Two Dimensions Using a Modified Levenberg-Marquardt Algorithm. J. Magn. Reson. *Series A*, 155–189.

67. Ge, M., and Freed, J.H. (2009). Fusion Peptide from Influenza Hemagglutinin Increases Membrane Surface Order: An Electron-Spin Resonance Study. Biophys. J. *96*, 4925–4934.

68. Ge, M.T., Costa-Filho, A., Gidwani, A., Holowka, D., Baird, B., and Freed, J. (2001). The structure of bleb membranes of RBL-2H3 cell is heterogenous: An ESR study. Biophys. J. *85*, 4023–4040.

69. Smith, A.K., and Freed, J.H. (2009). Determination of tie-line fields for coexisting lipid phases: an ESR study. J. Phys. Chem. B *113*, 3957–3971.

70. Liang, Z., and Freed, J.H. (1999). An Assessment of the Applicability of Multifrequency ESR to Study the Complex Dynamics of Biomolecules. J. Phys. Chem. B *103*, 6384–6396.

71. Louis-Jeune, C., Andrade-Navarro, M.A., and Perez-Iratxeta, C. (2012). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins Struct. Funct. Bioinforma. *80*, 374–381.

72. Ravoo, B.J., Weringa, W.D., and Engberts, J.B. (1999). Membrane fusion in vesicles of oligomerizable lipids. Biophys. J. *76*, 374–386.

73. Lai, A.L., Moorthy, A.E., Li, Y., and Tamm, L.K. (2012). Fusion Activity of HIV gp41 Fusion Domain Is Related to Its Secondary Structure and Depth of Membrane Insertion in a Cholesterol-Dependent Fashion. J. Mol. Biol. *418*, 3–15.

Supplemental figures

Table S-3.1 *Tetrahymena thermophila* strains.

Strain[a]	Genotype (micronucleus)	Genotype; Phenotype (macronucleus)
CU428.2	*CHX1/CHX1*; *mpr1-1/mpr1-1*	*MPR1, CHX1;* mp-s, cy-s, VII
CU427.4	*chx1-1/chx1-1;MPR1/MPR1*	*MPR1, CHX1;* mp-s, cy-s, VI
ΔHAP2-428	*CHX1/CHX1; mpr1-1/mpr1-1*	*MPR1,CHX1,hap2-1[Δ::neo4];* mp-s, cy-s, pm-r, VII
ΔHAP2-427	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4];* mp-s, cy-s, pm-r, VI
HAP2cDNAResc427	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2cDNA (3'cy2)];* mp-s, pm-r, cy-r, VI
HAP2genomicResc427	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-3 (3'cy2)];* mp-s, pm-r, cy-r, VI
^ΔHAP2 Domain	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N2(3'cy2)]; mp-s, pm-r, cy-r, VI*
^ΔDENV Region	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N3(3'cy2)]; mp-s, pm-r, cy-r, VI*
^ΔFusion Loop	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N5(3'cy2)]; mp-s, pm-r, cy-r, VI*
^DENV FL Rescue	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N4(3'cy2)]; mp-s, pm-r, cy-r, VI*
^Δ510-513	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N6(3'cy2)]; mp-s, pm-r, cy-r, VI*
^HAP2 FL Rescue	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N5R(3'cy2)]; mp-s, pm-r, cy-r, VI*
^FQY131-3AAA	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N71(3'cy2)]; mp-s, pm-r, cy-r, VI*
^CC-147-8SS	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N72(3'cy2)]; mp-s, pm-r, cy-r, VI*
^R164A	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N73(3'cy2)]; mp-s, pm-r, cy-r, VI*
^LNL171-173AAA	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-N74(3'cy2)]; mp-s, pm-r, cy-r, VI*
^C_5→S	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-C3(3'cy2)]; mp-s, pm-r, cy-r, VI*
^C_8→S	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-C5(3'cy2)]; mp-s, pm-r, cy-r, VI*
^ΔBasic Domain	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-C2(3'cy2)]; mp-s, pm-r, cy-r, VI*
^ΔC' term	*chx1-1/chx1-1; MPR1/MPR1*	*MPR1,CHX1, hap2-1[Δ::neo4/hap2-C1(3'cy2)]; mp-s, pm-r, cy-r, VI*

[a]CU428.2 and CU427.4 are functional heterokaryons that are phenotypically sensitive (s) to 6-methylpurine (mp) and cycloheximide (cy), respectively, due to markers in the macronucleus, and homozygous for resistance (r) to 6-methylpurine (CU428.2) or cycloheximide (CU427.4) in the micronucleus. The listed cell lines were generated from inbred B strains of *T. thermophila*. Strains marked with a ^ were created during the course of this project and were made by introducing mutant constructs into the ΔHAP2-427 cell line.

Table S-3.2 PCR Primers.

Primer Name[a]	Sequence 5' → 3'	Strain Use[b]
BamHIHap2For	GATTACggatccATGAAATTTTTGGCTTTTGG	All Truncation Mutants
TtFLHap2HAKpnIR	TAGTACggtaccTCAAGCATAATCAGGAACAT CATAAGGATATTCAATTAGTAGATAGAGAG GAGATGTTG	For adding HAP2 C'term HA tag
LongFLAGHISHap2r ev	AATCGGTACCTCAGTGGTGATGGTGGTGG TGGTGGTGATGGTGTTTATCGTCGTCGTCT TTATAATCACGCGTTTCAATTAGTAGATAGA GAGGAGATG	C_5→S, C_8→S
C1Hap2HAKpnIR	GTACggtaccTCAAGCATAATCAGGAACATCA TAAGGATAGCAACAGCAACAGCAAAG	ΔC'term
C2HyHAP2for	CTTTGCTGTTGCTGTTGCTCAATTTAAGAAA GTTGCAGTTATGA	ΔBasic Domain
C2HyHAP2rev	TCATAACTGCAACTTTCTTAAATTGAGCAAC AGCAACAGCAAAG	ΔBasic Domain
C3HyHAP2for	GGTTTCTTTCTTCTTCTTCTTCTAAATCTAA GAAAAAGGAAAATGAAAAAAATAAAG	C_5→S
C3HyHAP2rev	TCTTAGATTTAGAAGAAGAAGAAGAAAGAA ACCGAATAATAGGGACTAAATAC	C_5→S
C5HyHAP2for	GAAAGTTCTAGTTATGATAGATCGTCTTCTT CACATTCAATATCATAGTCATATTAGG	C_8→S
C5HyHAP2rev	GAATGTGAAGAAGACGATCTATCATAACTA GAACTTTCTTAAATTGATTTTTTATCTG	C_8→S
N2hyHAP2for	TCTTTAAGGGATAAGTTCTTGGAGTGATCT CATTCTTTTATCATAAAAT	ΔHAP2 Domain
N2hyHAP2rev	ATTTTATGATAAAAGAATGAGATCACTCCAA GAACTTATCCCTTAAAGA	ΔHAP2 Domain
N3HYHAP2for	GGTTACTAAATCTCCAGTGACTGCTATGTT TGTTGACAAAACCATG	ΔDENV Region
N3hyHAP2rev	CATGGTTTTGTCAACAAACATAGCAGTCAC TGGAGATTTAGTAACC	ΔDENV Region
N4HYHAP2for	TAATGGTTGTGGTTTATTTGGTAAAGGTTCT TTAGCTCACTGCCTAAAATTCAGTC	DENV FL Rescue
N4hyHAP2rev	CCAAATAAACCACAACCATTACCCCAACCT CTATCAACAGAGCAATAGCAGCAATAACC	DENV FL Rescue
N5HYHAP2for	GGTTATTGCTGCTATTGCTCTGCTCACTGC CTAAAATTCAGTC	ΔFusion Loop
N5hyHAP2rev	GACTGAATTTTAGGCAGTGAGCAGAGCAAT AGCAGCAATAACC	ΔFusion Loop
N5RHyHAP2for1	GTGCTATGCTCTAAATCTTGGTGCAGGATC AGCAACAGCTCACTGCCTAAAATTCAGTC	HAP2 FL Rescue
N5RHyHAP2for2	ATTAGGCATGGGTAATGATCTATCAAGAGG TAAAGTGTGCTATGCTCTAAATCTTGGTG	HAP2 FL Rescue
N5RhyHAP2rev	CTTGATAGATCATTACCCATGCCTAATATAT CTGATAGAGAGCAATAGCAGCAATAACC	HAP2 FL Rescue
N6HyHAP2for	GGAAATAACACAGGCTAGTAATAATCAAAG TCATCCTAATCCTGCAGTG	Δ510-513
N6HyHAP2rev	CACTGCAGGATTAGGATGACTTTGATTATT ACTAGCCTGTGTTATTTCC	Δ510-513
N7_R164A_F	TGATCTATCAGCTGGTAAAGTGTGCTATGC	R164A

N7_R164A_R	TTACCCATGCCTAATATATC	R164A
N7_CC147SS_F	CTAAGGTTATAGTAGTTATTGCTCTCTATCAGATATATTAG	CC147-8SS
N7_CC147SS_R	CTGTCCAAAATCTTTTAGCC	CC147-8SS
N7_FQY131AAA_F	AGCTGACTCTAAAGGCTAAAAGATTTTG	FQY131-3AAA
N7_FQY131AAA_R	GCAGCTTTACAAGTTGGACTACTATC	FQY131-3AAA
N7_LNL171AAA_F	AGCTGGTGCAGGATCAGCAACA	LNL171-3AAA
N7_LNL171AAA_R	GCAGCAGCATAGCACACTTTACCTC	LNL171-3AAA
HAP2 5'FlankFor	GTTATTTTCAGCATCTTCTTTCATTTG	Genotyping
HAP2 3'FLANK REV	ATCTCTTCTGATCATAGAGCACC	Genotyping
RV3rev	GACATTAAAGCAAGTTAAGCATAAATAAAG	All Truncations
Mtt1RevSeq	AATACGAAACTGATTTTATGCAA	Genotyping
PreStagsForKpnI	GTAATCggtaccGAAGTTTTGTTCCAAGGTCCC	Over-expression HAP2
HistagRevSacI	GATTACgagctcTCAGTGGTGGTGGTGGTGGT	Over-expression HAP2

[a]All primers used in this project were supplied by Sigma-Aldrich. Primers were stored as 100 μM stock solutions, and diluted to 10 μM working solutions prior to use in PCR reactions. Melting temperatures were determined using Modified Breslauer's thermodynamics, dH and dS parameters as recommended by the manufacturer. Added restriction sites are shown in lower case letters.

[b]All constructs made with these primers were subjected to Sanger sequencing to verify accuracy.

Figure S-3.3 Over-expression of HAP2 in *T. thermophila.* In WT x *ΔHAP2* crosses the kinetics of fusion, as well as the final percentage of cells capable of fusion was significantly reduced. To determine whether these effects were due to a reduction in the total amount of HAP2 expressed in a given mating pair, or its pattern of expression in only one mating partner, we over-expressed a full-length HA-tagged version of HAP2 cDNA in a WT partner and mated these cells with a *ΔHAP2* knockout strain. We generated the over-expression strain by cloning the tagged HAP2 cDNA into a stable high-copy ribosomal DNA vector and used a robust cadmium-inducible promoter to drive the expression of HAP2 transcripts. Over-expressing cells (designated, O.E. HAP2) were induced with 0.1 μg/mL $CdCl_2$ 30 min prior to mating. **(A)** Bar chart showing the percent fusion in crosses between the O.E. HAP2 strain and either a WT or *ΔHAP2* partner. The data for WT^{β} x $HAP2^{\alpha}$ crosses from Fig. 2g is shown for comparison. HAP2 over-expression had no effect on percent fusion in crosses with the WT partner, and slightly diminished fusion in crosses with the *ΔHAP2* strain. **(B)** A single experiment showing that the kinetics of fusion in O.E. HAP2 x *ΔHAP2* cross were similar to WT x *ΔHAP2* matings (see Figure 3.1R for comparison). To verify that HAP2 was in fact overexpressed and correctly localized in the O.E. strain we examined these cells by fluorescence microscopy and Western blotting following induction with $CdCl_2$. **(C)** Fluorescence (below) and merged bright field-fluorescence images (above) showing representative single and paired cells from an O.E. HAP2 x WT mating culture fixed 3 h after mixing complementary mating types and immunolabled with anti-HA antibodies. The co-stimulated cell on the left, and the mating pair on the right showed expression and correct localization of the over-expressed recombinant HAP2 protein at the anterior tip of the co-stimulated cell, and the conjugation junction of mating cells, respectively. **(D)** A time course of recombinant HAP2 expression in unmated O.E. HAP2 cells in 10 mM Tris buffer induced with 0.1 μg/mL $CdCl_2$ for 10 h as detected by Western blotting. A strong signal just above the expected size of the O.E. HAP2 protein (arrow) was readily detected in cell lysates of the O.E. HAP2 strain beginning at ~ 1 h post induction. It is worth noting that when the same epitope-tagged version of the HAP2 cDNA was placed under the control of the endogenous promoter at the HAP2 locus, mating cultures showed signals by immunofluorescence microscopy (Figure S-3.6) but not by Western blotting suggesting the protein may be highly sensitive to degradation following cellular lysis.

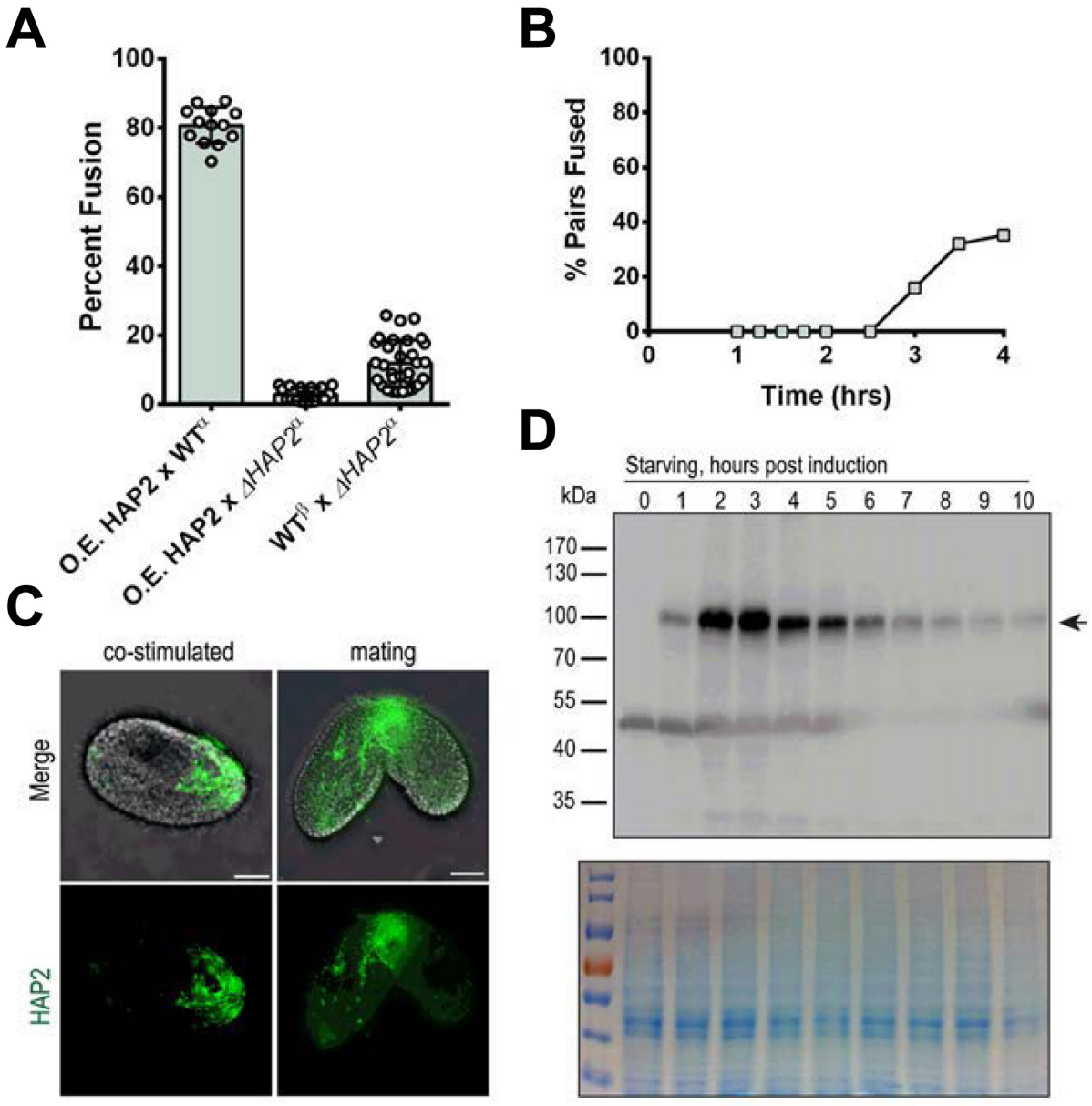
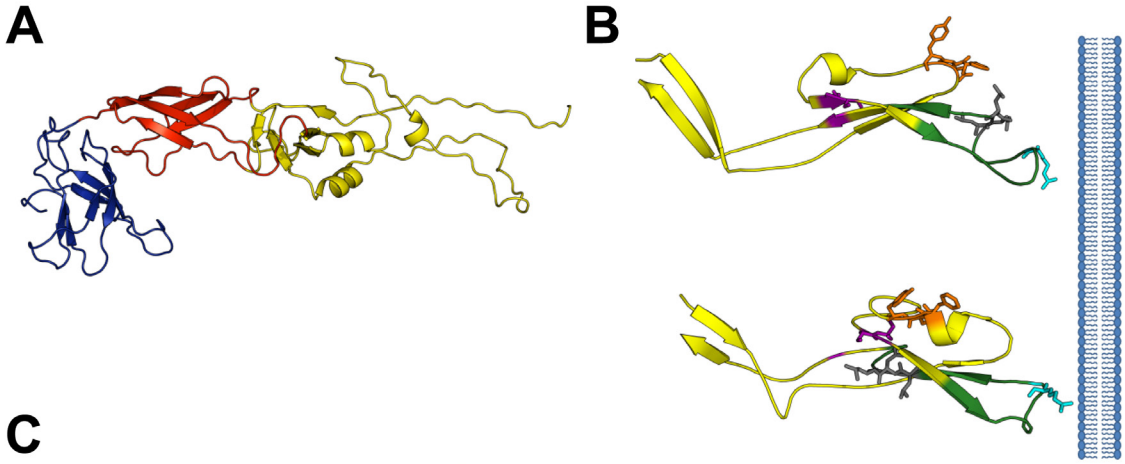
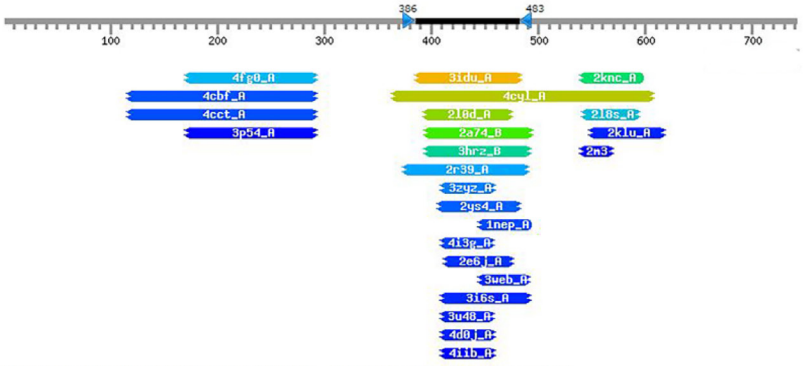

A
Percent Fusion
100
80
60
40
20
0
O.E. HAP2 x WTα
O.E. HAP2 x ΔHAP2α
WTβ x ΔHAP2α
B
% Pairs Fused
100
80
60
40
20
0
0
1
2
3
4
Time (hrs)
C
co-stimulated
mating
Merge
HAP2
D
Starving, hours post induction
kDa
0 1 2 3 4 5 6 7 8 9 10
170
130
100
70
55
40
35

Figure S-3.4 Template-based homology modeling of HAP2. (A) A CPHmodels3.0 generated partial structure of the *T. thermophila* HAP2 ectodomain based on the DENV template PDB ID: 1UZG (aligned residues are 172-478). Domains follow the convention for class II viral fusogens and are color-coded as follows: domain I is red, domain II is yellow, and domain III is blue. **(B)** Magnified views of the locations of the tested site-directed mutations within and near the HAP2 fusion loop shown on the Phyre2 (top, residues 107-193) and RaptorX (bottom, residues 96-193) predicted structures and juxtaposed to a cartoon membrane. The region of the fusion loop that was truncated is shown in green. Site-directed mutations are shown as sticks with CC147-8 in magenta, FQY131-3 in orange, R164 in cyan, and LNL171-3 in grey. The cysteines predicted to form disulfide bonds with residues 147-8 are shaded magenta (but not shown as sticks). Depending on loop orientation, which is difficult for template-based modelers to predict, the Phyre2 and RaptorX models show a second loop (containing the FQY131-3 mutation) that might approach the lipid bilayer. Nevertheless, alteration of the FQY motif in this loop had no affect on the fusogenic activity. **(C)** A screenshot of HHpred top hit results to the *T. thermophila* HAP2 protein sequence. This was the only template-based modeling tool we tested that identified a homology between HAP2 and EFF-1, the developmental cell-cell fusogen from *C. elegans.*



C



Query Sat_May_21_02:30:53_+0200_2016 (seq=MKFLAFGLIY...STSPLYLLIE Len=742 Neff=5.6 Nseqs=150)
Parameters score SS:yes search:local realign with MAP:no

No	Hit	Prob	E-value	P-value	Score	SS	Cols	Query HMM	Template HMM
1	3idu_A Uncharacterized protein	84.4	10	0.00029	34.9	11.1	92	386-483	12-104 (127)
2	4cyl_A EFF-1A; cell adhesion,	80.7	3.4	9.3E-05	47.5	7.3	224	364-606	376-627 (658)
3	2l0d_A Cell surface protein; s	78.4	12	0.00032	33.7	9.0	77	393-474	6-84 (114)
4	2a74_B Complement component C3	73.7	74	0.002	31.1	14.0	97	394-493	82-185 (188)
5	2knc_A Integrin alpha-IIB; tra	63.8	11	0.0003	30.6	4.5	43	540-598	12-54 (54)
6	3hrz_B Cobra venom factor; ser	61.2	1.4E+02	0.0037	30.7	13.5	95	394-491	83-183 (252)
7	2l8s_A Integrin alpha-1; trans	57.2	11	0.00031	30.5	3.5	38	542-593	11-48 (54)
8	4fg0_A Polyprotein; viral enve	55.7	57	0.0016	36.6	9.9	100	171-292	105-223 (407)
9	2r39_A FIXG-related protein; s	54.0	1.5E+02	0.004	26.9	13.7	103	374-489	11-114 (118)
10	3zyz_A Beta-D-glucoside glucoh	50.8	36	0.00098	40.8	8.0	49	410-458	616-672 (713)
11	2ys4_A Hydrocephalus-inducing	48.1	69	0.0019	28.4	7.8	70	407-482	41-110 (122)

Table S-3.5 Phyre2 batch processing results.

Organism name, accession #[a]	CII viral envelope hits[b]	Non viral envelope hits
Gonium pectorale, BAO57178.1	11 (61.3%, 2OF6, WNV) 14 (59.4%, 3UAJ, DENV)	1 (92.0%, 4FXK, complement c4-a) 2 (82.4%, 3BGA, beta-galactosidase) 3 (82.3%, 1CZD, DNA POL protein g45)
Chlorella variabilis, XP_005851393.1	1 (95.5%, 4B03, DENV) 2 (94.9%, 1UZG, DENV) 4 (94.7%, 1SVB, TBEV)	9 (89.0%, 2X41, beta-glucosidase) 10 (85.7%, 2E6J, papd-like domain, hydin) 11 (84.5%, 3QBT, INPP5 ocrl-1)
Volvox carteri, XP_002952884.1	5 (42.9%, 2OF6, WNV) 11 (23.6%, 4CBF, DENV)	1 (58.0%, 2R39, protein from V. parhaemolyticus) 2 (51.4%, 2E6J, papd-like domain, hydin) 3 (46.8%, 2RNR, TFiih complex p62)
Chlamydomonas reinhardtii, ABO29824.2		1 (88.5%, 4FXK, complement c4-a) 2 (83.9%, 2E6J, papd-like domain, human hydin) 3 (81.4%, 1JZ8, beta-galactosidase)
Helicosporidium sp, KDD77085.1	3 (36.9%, 2OF6, WNV) 4 (36.9%, 4CBF, DENV)	1 (39.7%, 2M0G, splicing factor 1) 5 (23.9%, 1M2s, K-channel blocking toxin bmtx3) 6 (23.0%, 4QE0, duf5043 from B. uniformis)
Coccomyxa subellipsoidea, XP_005651045.1	1 (90.9%, 3UAJ, DENV) 3 (80.9%, 2OF6, WNV) 6 (75.1%, 1UZG, DENV)	4 (79.0%, 2E6J, papd-like domain, hydin) 5 (76.8%, 2X41, beta-glucosidase) 13 (66.0%, 4FXK, complement c4-a)
Lilium longiflorum, BAE71142.1	1 (96.3%, 3UAJ, DENV) 2 (94.3%, 1OK8, DENV) 3 (93.3%, 1SVB, TBEV)	9 (84.7%, 4IID, beta-glucosidase 1) 11 (77.9%, 2E6J, papd-like domain, hydin) 20 (64.9%, 2L0D, m. acetivorans protein)
Zea mays, NP_001307741	8 (59.0%, 3UAJ, DENV)	1 (91.3%, 2XC8, B. subtilis spp1 phage) 2 (80.7%, 2L0D, m. acetivorans protein) 3 (79.6%, 4FXK, complement c4-a)
Arabidopsis thaliana, NP_192909.2	1 (92.9%, 1URZ, TBEV) 2 (91.5%, 1OK8, DENV) 6 (89.6%, 1SVBA, TBEV)	3 (90.0%, 2KL6, cardb domain of p. furiosus) 4 (89.8%, 1JZ8, beta-galactosidase) 8 (84.6%, 4FXK, complement c4-a)
Cyanidioschyzon merolae, XP_005536505.1		1 (85.6%, 2X41, beta-glucosidase) 6 (69.0%, 3QIS, INPP5 ocrl-1) 9 (65.8%, 2E6J, pap-d like domain, hydin)
Galdieria sulphuraria, XP_005708101.1		1 (92.7%, 1W8O, bacterial sialidase) 2 (89.6%, 4FXK, complement c4-a) 3 (87.4%, 3ZZ1, beta-d-glucoside glucohydrolase)
Trypanosoma brucei, XP823296.1	1 (96.7%, 1UZG, DENV) 2 (96.3%, 1SVB, TBEV) 5 (95.1%, 2OF6, WNV)	11 (86.6%, 2X41, beta-glucosidase) 14 (60.6%, 2MI2, protein translocase protein tatb) 17 (56.7%, 2E6J, papd-like domain, hydin)
Trypanosoma cruzi, XP_814894.1	1 (95.1%, 3UAJ, DENV) 2 (92.7%, 2OF6, WNV) 3 (91.6%, 1SVB, TBEV)	9 (81.7%, 1JZ8, beta-galactosidase) 15 (57.9%, 3QBT, INPP5 ocrl-1) 19 (45.7%, 2XC8, B. subtilis spp1 phage)
Strigomonas culicis, EPY22600.1	1 (91.1%, 1UZG, DENV) 2 (90.4%, 1SVB, TBEV) 3 (90.4%, 2OF6, WNV)	5 (86.4%, 1JZ8, beta-galactosidase) 6 (80.7%, 2E6J, papd-like domain, hydin) 12 (60.9%, 2XC8, B. subtilis spp1 phage)
Angomonas deanei, EPY38446.1	3 (83.9%, 2OF6, WNV)	1 (88.1%, 1JZ8, beta-galactosidase) 2 (85.9%, 2XC8, B. subtilis spp1 phage)
Phytomonas sp. Isolate EM1, CCW64758.1	1 (94.9%, 1SVB, TBEV) 2 (93.8%, 2OF6, WNV) 3 (92.0%, 1UZG, DENV)	8 (86.2%, 1JZ8, beta-galactosidase) 12 (73.1%, 4FCK, complement c4-a) 16 (62.3%, 2E6J, papd-like domain, human hydin)

Leishmania major, XP_003722443.1	1 (93.3%, 1UZG, DENV) 4 (88.7%, 1SVB, TBEV) 5 (86.3%, 2OF6, WNV)	7 (81.6%, 2E6J, papd-like domain, human hydin) 8 (76.5%, 2X41, beta-glucosidase) 11 (71.5%, 3QBT, INPP5 ocrl-1)
Naegleria gruberi, XP_002674350.1	1 (93.3%, 3UAJ, DENV) 4 (90.3%, 1SVB, TBEV) 6 (88.6%, 2OF6, WNV)	9 (73.1%, 2KNC, integrin alpha-iib) 12 (66.8%, 2KL6, cardb domain of p. furiosus) 13 (61.7%, 2V5Y, tyrosine-protein phosphatase mu)
Physarum polycephalum, BAE71144.1		1 (86.3%, 1JZ8, beta-galactosidase) 2 (85.4%, 4FXK, Complement c4-a alpha chain) 3 (73.3%, 2V5Y, tyrosine-protein phosphatase mu)
Dictyostelium fasciculatum, XP_004359139.1	19 (39.6%, 4B03, DENV)	1 (88.2%, 1JZ8, beta-galactosidase) 8 (59.1%, 2JE8, beta-mannosidase) 9 (58.4%, 2XC8, B. subtilis spp1 phage)
Acanthamoeba castellanii, XP_004341525.1	1 (96.1%, 3UAJ, DENV) 2 (95.3%, 2OF6, WNV) 3 (94.8%, 1SVB, TBEV)	11 (72.8%, 1JZ8, beta-galactosidase) 13 (67.6%, 2E6J, papd-like domain, human hydin) 16 (64.8%, 2RNR, TFiih complex p62)
Theileria parva, XP_764209.1		1 (60.7%, 1JZ8, beta-galactosidase) 6 (35.1%, 3QBT, INPP5 ocrl-1) 8 (29.1%, 4FXK, complement c4-a)
Toxoplasma gondii, EPT31063.1		1 (81.2%, 1JZ8, beta-galactosidase) 2 (73.2%, 2MI2, translocase protein tatb) 3 (55.1%, 2VRS, capsid protein avian reovirus)
Plasmodium berghei, XP_676900.1		1 (57.9%, 1JZ8, beta-galactosidase) 3 (32.5%, 1A87, colicin n) 8 (16.4%, 2XC8, B. subtilis spp1 phage)
Plasmodium falciparum, XP_001347424.1	2 (37.7%, 3J2W, CHIKV)	1 (56.2%, 1JZ8, beta-galactosidase) 4 (37.2%, 4FXK, complement c4-a) 5 (34.7%, 1A87, colicin n)
Oxytricha trifallax, EJY77656.1		1(98.7%, 4CGK, protein PCSB from S.pneumoniae) 2(98.4%, 3VKG, dynein heavy chain) 5(97.8%, 4L1B, PI3K regulatory subunit alpha)
Tetrahymena thermophila, KJ629172	1 (94.9%, 1UZG, DENV) 3 (89.7%, 2OF6, WNV) 7 (78.4%, 1SVB, TBEV)	8 (76.5%, 1JZ8, beta-galactosidase) 10 (55.6% 4FXK, complement c4-a) 11 (52.5%, 2E6J, papd-like domain, human hydin)
Ichthyophthirius multifiliis [c]	1 (95.2%, 3UAJ, DENV) 3 (91.1%, 1SVB, TBEV) 9 (78.7%, 2OF6, WNV)	4 (84.7%, 1JZ8, beta-galactosidase) 14 (48.2%, 2KL6, cardb domain of p. furiosus) 15 (47.4%, 2E6J, papd-like domain, human hydin)
Paramecium tetraurelia, XP_001431224.1	1 (87.3%, 3UAJ, DENV) 3 (76.9%, 2OF6, WNV) 4 (72.9%, 1UZG, DENV)	2 (77.3%, 2MKV, Na/K-transporting atpase) 5 (72.2%, 2E6J, papd-like domain, human hydin) 7 (65.4%, 2JO1, phospholemman)
Capsaspora owczarzaki, XP_004343268.1		1 (84.7%, 2XC8, B. subtilis spp1 phage) 2 (82.3%, 1JZ8, beta-galactosidase) 3 (82.2%, 4FXK, complement c4-a)
Salpingoeca rosetta, XP_004989263.1	1 (79.5%, 3UAJ, DENV) 5 (64.6%, 2OF6, WNV) 11 (51.9%, 1URZ, TBEV)	3 (68.1%, 2PBD, phosphoprotein) 4 (66.9%, 2LFT, human prion protein with e219k) 7 (60.1%,1LNZ, spo0b-associated gtp-binding protein)
Monosiga brevicollis, XP_001746497.1	1 (96.2%, 1UZG, DENV) 2 (95.0%, 4B03, DENV) 3 (94.7%, 3UAJ, DENV)	11 (63.2%, 1QK6, huwentoxin-i) 15 (33.8%, 2IEC, protein in m. kandleri) 16 (33.0%, 2V5Y, tyrosine-protein phosphatase mu)
Nematostella vectensis, XP_001628495.1	1 (96.3%, 1UZG, DENV) 2 (95.5%, 1SVB, TBEV) 4 (95.0%, 2OF6, WNV)	11 (78.6%, 2X41, beta-glucosidase) 15 (52.6%, 2E6J, papd-like domain, human hydin) 16 (42.0%, 2R39, protein from V. parahaemolyticus)

Hydra vulgaris, ABN45755.1	3 (73.0%, 2OF6, WNV) 6 (69.2%, 1UZG, DENV)	1 (90.3%, 1JZ8, beta-galactosidase) 5 (70.7%, 2E6J, papd-like domain, human hydin) 13 (49.9%, 4AK2, heparin-binding protein)
Tribolium castaneum, EFA06462.1	1 (62.7%, 1SVB, TBEV) 10 (30.2%, 1UZG, DENV)	2 (48.7%, 3MU3, chicken md-1) 3 (43.4%, 4CCV, histidine-rich glycoprotein) 4 (41.7%, 2CG7, fibronectin)
Acyrthosiphon pisum, XP_003245993.2	3 (36.6%, 1UZG, DENV) 6 (32.7%, 1SVB, TBEV)	1 (52.4%, 2YS4, papd-like domain, human hydin) 2 (47.1%, 1XSZ, GEF ralf) 9 (27.6%, 3U6X, phage tp901-1 baseplate tripod)
Capitella teleta, ELU07639.1	1 (77.6%, 1SVB, TBEV) 2 (75.0%, 2OF6, WNV) 3 (74.4%, 1UZG, DENV)	7 (63.0%, 4GWM, hydrolase, promeprin beta) 8 (53.0%, 3MU3, chicken md-1) 9 (47.0%, 3DUE, periplasmic protein from B. vulgatus)
Pediculus humanus corporis, XP_002429972.1		1 (87.1%, 1GM6, boar salivary lipocalin) 2 (84.0%, 1EW3, allergen equ c1 3 (82.5%, 3MU3, chicken md-1)
Drosophila melanogaster, NP_001034068.2		1 (55.8%, 1JRJ, exendin-4) 2 (45.1%, 1D0R, glucagon-like peptide) 3(33.8%, 2CCT, zinc finger domain of DnaJ)
Saccoglossus kowalevskii, XP_006821859.1		1 (50.4%, 1GYG, alpha-toxin from C. perfringens) 2 (48.0%, 3DKB, tumor necrosis factor a20) 4 (24.7%, 1TM9, protein from Mycoplasma genitalium)
Apis mellifera, XP_006565646.1		1(81.1%, 3QBT, INPP5 ocrl-1) 4 (71.6%, 3MU3, chicken md-1) 11(50.3%, 2F61, beta-glucosidase)

[a]The name of HAP2 containing organisms (top) and GenBank accession numbers of their respective HAP2 orthologs (bottom). The identities of HAP2 orthologs were based on the findings of previous studies [S1].

[b]The Phyre2 hits to each HAP2 ortholog are listed as Hit Rank (Confidence%; PDB ID; name of protein). The top three ranking hits (out of 20 total) to different templates in each category (viral/non-viral) are shown with a hit rank of 1 being the best-scoring template from which the predicted structure was built. Viral envelope protein hits are abbreviated based on the name of the virus from which they were derived (DENV = Dengue Virus E glycoprotein; TBEV = Tick Borne Encephalitis Virus envelope glycoprotein; WNV = West Nile Virus Envelope glycoprotein; and, CHIKV = Chikungunya virus envelope protein). All species that had a hit to a viral envelope structure are shaded grey.

[c]The hits to the *Ichthyophthirius multifiliis* HAP2 ortholog were determined through the Phyre2 processing portal.

Figure S-3.6 Immunofluorescence localization of mutated versions of HAP2. HAP2 localizes to the conjugation junction of *T. thermophila*, a region where cells of complementary mating types adhere and form membrane pores. In each panel, localization of a C'-terminal HA- or FLAG-tagged version of HAP2 is shown in crosses between WT cells and strains harboring mutations/truncations to the *HAP2* coding sequence. In all cases, mutated/truncated gene constructs were targeted to the *HAP2* locus and expressed under the control of the endogenous promoter. Cells were fixed 2.5-5 h after mixing of complementary mating types, then permeabilized and immunolabeled with anti-HA or anti-FLAG antibodies. In some cases, nuclei were labeled 30 min prior to fixation with Hoechst 33258. Each row of paired images shows a merged fluorescence-bright field image (above), and the fluorescence image alone (below), for a representative mating pair from crosses containing the indicated mutated/truncated construct (as labeled above the paired images). No signal was seen in matings between cells that lacked epitope-tagged HAP2. Scale bars are 10μm.

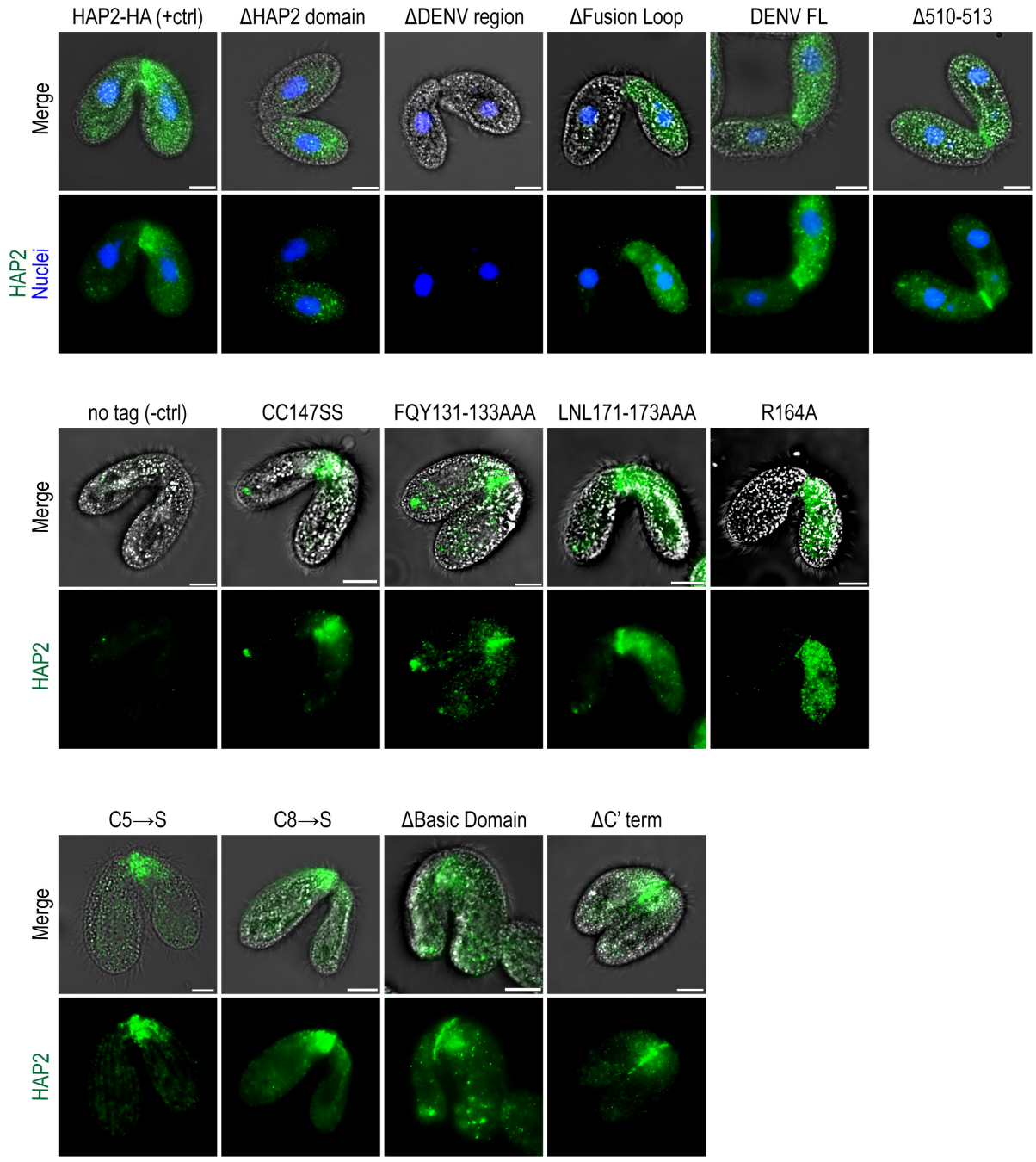
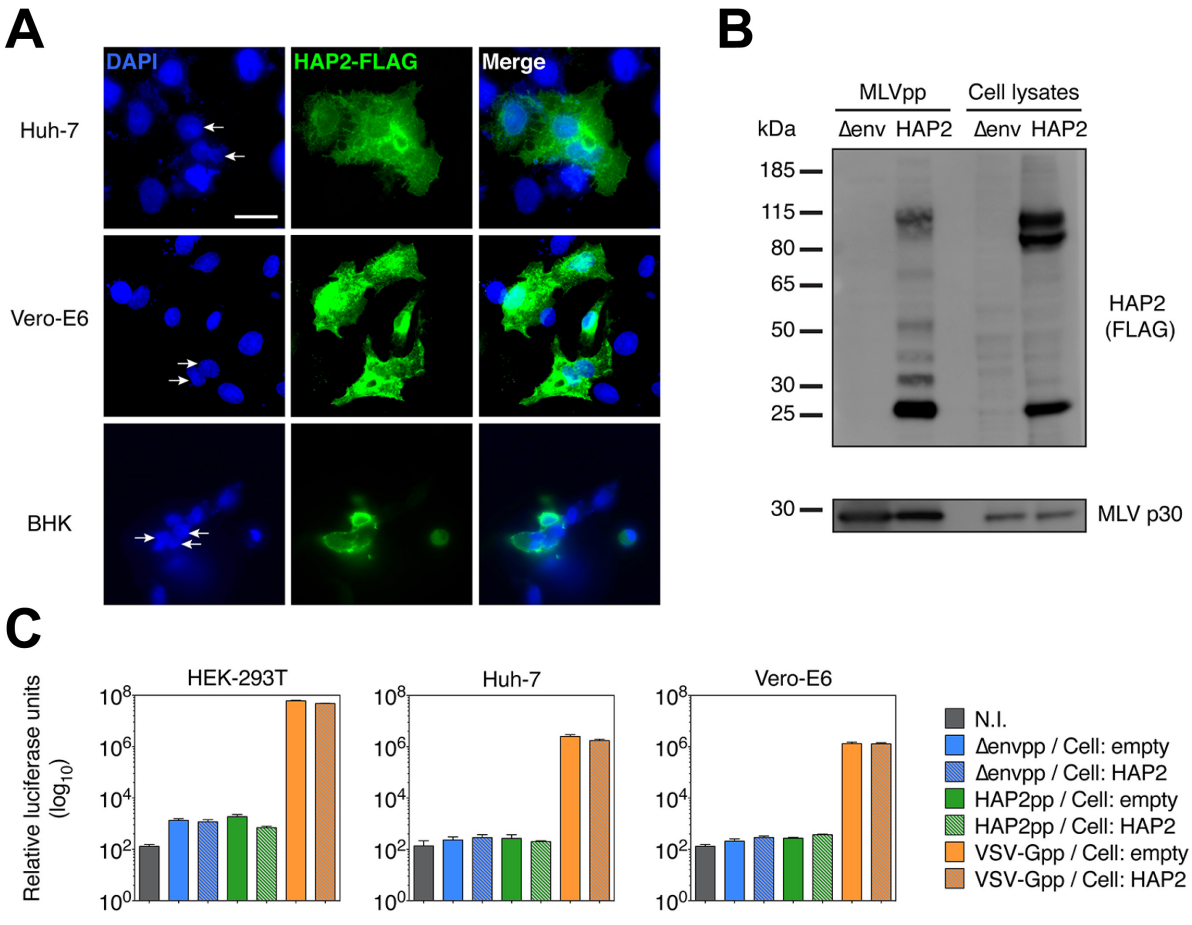

HAP2-HA (+ctrl)
ΔHAP2 domain
ΔDENV region
ΔFusion Loop
DENV FL
Δ510-513
Merge
HAP2
Nuclei
no tag (-ctrl)
CC147SS
FQY131-133AAA
LNL171-173AAA
R164A
Merge
HAP2
C5→S
C8→S
ΔBasic Domain
ΔC' term
Merge
HAP2

Figure S-3.7 HAP2 fusion assays in heterologous systems. (A) Immunofluorescence microscopy analysis of HAP2-expressing mammalian cell lines. Transfection of human liver (Huh-7), African green monkey (Vero-E6), and baby hamster kidney (BHK) cells with a codon-optimized *T. thermophila HAP2* gene fused to a 3x FLAG epitope tag resulted in detectable protein expression in all lines tested via immunofluorescence. Transfected cells were fixed, permeabilized and immunolabeled with anti-FLAG antibodies, followed by DAPI staining to label cell nuclei. The HAP2 protein appeared to be expressed on the plasma membrane, as the outline of cells, including the filopodia and membrane ruffles of the Huh-7 cells, could be well delineated after labeling. Arrows show nuclei present in small, HAP2-expressing multi-nucleated cells in all threecell lines. Scale bar is 25 μm. **(B)** Western blot analysis of HAP2 protein expression in transfected cells and HAP2 incorporation into pseudotyped particles. To determine whether HAP2 function was sufficient to fuse viral envelopes to cells we generated murine leukemia virus (MLV) retroviral pseudotyped particles that incorporated heterologous HAP2 (HAP2pp). HEK-293T cells were co-transfected to produce either *T. thermophila* HAP2-3xFLAG pseudotyped particles (HAP2pp) or no envelope protein control particles (Δenvpp). After supernatant harvesting, the pseudotyped particle producer cells were lysed and analyzed by Western blot along with the concentrated pseudotyped particles. HAP2-3xFLAG was detected using an anti-FLAG tag antibody, and MLV p30 capsid protein (loading control) was detected using the anti-MLV p30 capsid antibody (4B2). In both HAP2-transfected cell lysates and in HAP2pp, the HAP2 protein could be detected. In cell lysates, bands were detected at ~115, ~100, and ~25 kDa, whereas in HAP2pp only the ~115 and 25 kDa bands, were seen, indicating preferential incorporation of certain forms of HAP2 into the particles. Several bands of intermediate sizes were also observed suggesting HAP2 sensitivity to proteolysis. **(C)** Infectivity assays with HAP2pp. Pseudotyped particles contain a luciferase gene that integrates into the target cell's genome if successful viral membrane fusion and entry has occurred. Assaying for luciferase activity after application of particles allowed for determination of the infectivity, and thus fusogenicity, of each type of pseudotyped particle analyzed. Cell lines used to test the infectivity of these particles were HEK-293T, Huh-7, and Vero-E6 cells that were either transfected with an empty vector, or a vector designed to express HAP2-3xFLAG (as in [A]). Twenty-four hours post transfection, HAP2pp, Δenvpp, and VSV-G pseudotyped particles (positive control particles pseudotyped with the VSV-G envelope glycoprotein) were used to inoculate cells. Seventy-two hours post infection, cells were lysed and luciferase activity was determined using a luminometer. Note the relatively low levels of luciferase activity in the control Δenvpp and HAP2pp infected samples (~10^2-10^3 relative luciferase units, RLU).


A
DAPI
HAP2-FLAG
Merge
Huh-7
Vero-E6
BHK
B
MLVpp
Cell lysates
kDa
Δenv HAP2
Δenv HAP2
185
115
80
65
50
30
25
HAP2
(FLAG)
30
MLV p30
C
Relative luciferase units
(log10)
HEK-293T
Huh-7
Vero-E6
N.I.
Δenvpp / Cell: empty
Δenvpp / Cell: HAP2
HAP2pp / Cell: empty
HAP2pp / Cell: HAP2
VSV-Gpp / Cell: empty
VSV-Gpp / Cell: HAP2

Supplemental Materials and Methods

Immunofluorescence microscopy. Wild type cells (*T. thermophila* strain CU428.2 [mating type VII]) and cell lines harboring tagged HAP2 constructs (strain CU427.4, mating type VI) were cultured separately, starved and then combined to initiate mating as described above. At time points indicated in the figure legend, mating pairs were washed in 20 mM HEPES buffer, pH 7.4, then fixed for 20 min at room temperature by gentle addition of IC Fixation Buffer (Affymetrix eBioscience) at a ratio of 1:1 with resuspended cells. In some cases, Hoechst 33258 (Invitrogen) was added to the media 30 min prior to fixation for localization of macro- and micronuclei. All centrifugation steps were carried out at 350 × g. After fixation, cells were again centrifuged and resuspended in 1 × Permeabilization Buffer (Affymetrix eBioscience), blocked in PBS containing 3% BSA, and incubated overnight at 4˚C with the addition of either mouse anti-HA (anti-HA.11, BioLegend, formerly Covance) or rabbit anti-FLAG (Rockland, Inc.) antibodies at a dilution of 1/1000, followed by a 1 h incubation in secondary Alexa 488-conjugated goat anti-mouse and goat anti-rabbit antibodies respectively, at a dilution of 1/1000 (Life Technologies).

For immunofluorescence analyses of HAP2-3 × FLAG protein expression in mammalian cell lines, cells were washed in PBS buffer, and fixed for 1 h in PBS containing 4% paraformaldehyde. Cells were then permeabilized with Triton X-100, blocked with normal goat serum, and immunolabeled using first an anti-FLAG antibody (FLAG M2 antibody, Sigma), and next, a goat anti-mouse Alexa488 antibody (Life Technologies) at dilutions of 1/1000 and 1/500 respectively. Upon mounting, nuclei were stained by addition of Fluoromount G with DAPI (Electron Microscopy Sciences).

In both cases (i.e. mating *T. thermophila,* and HAP2-transfected mammalian cell lines), slides were analyzed for HAP2 expression and localization using either the 63× or 100× objectives on a Zeiss Axio Imager M1 microscope equipped with an AxioCamMR3 camera.

Western blot analysis. *T. thermophila* cells modified to over express HAP2 were induced to

express a C-terminally HA-tagged HAP2 through the addition of 0.1 µg/mL $CdCl_2$ to the 10 mM Tris starvation medium. Cell pellets (2×10^6 cells) were prepared every hour after induction for 10 h and 100 µL of 10 × Roche cOmplete EDTA-free Protease Inhibitor Cocktail (Roche) was added prior to freezing. Frozen cell pellets were resuspended in 100 µL of 2x sodium dodecyl sulfate (SDS) loading buffer and boiled for 2 min before addition of DTT. Protein samples (equivalent to ~1.5×10^5 cells) were separated on a 10% poly-acrylamide SDS gel in Tris-Glycine running buffer, and transferred to a 0.45 µm nitrocellulose membrane (Bio-Rad) for 1 h at 90 V. The blot in Figure S-3.3D was blocked in a PBS-Tween containing 5% milk solution and probed with a 1/1000 dilution of anti-HA antibody (anti-HA.11, BioLegend, formerly Covance), followed by HRP-conjugated secondary anti-mouse IgG antibody at 1/1000 (Southern Biotech). Signals were developed using a SuperSignal West Pico ECL kit and images acquired using a Syngene gel imager (Synoptics Ltd.).

For HAP2pp and ∆envpp conditions (Figure S-3.7A), co-transfected cells were lysed using with 1 × radioimmunoprecipitation assay (RIPA) buffer (EMD Millipore) containing protease inhibitor cocktail (Roche). HAP2pp and ∆envpp pseudotyped particles were ultracentrifuged at 42,000 rpm for 2 h at 4°C, using a TLA-55 rotor with an Optima-MAX-E centrifuge (Beckman-Coulter). Viral pellets were resuspended in PBS. Lithium dodecyl sulfate (LDS) loading buffer and DTT were added to cell lysates and concentrated viral solutions, which were then heated at 95°C for 5 min. Protein samples were separated on a NOVEX Bis-Tris gel (Life Technologies) and transferred on a polyvinylidene fluoride membrane (GE Healthcare). Detection of HAP2-FLAG was performed using the mouse anti-FLAG tag antibody (FLAG M2 antibody, Sigma) and MLV capsid detection was performed using the mouse monoclonal anti-MLV capsid p30 (4B2, Abcam). Detection of Western blot signal was performed using an ECL kit (Pierce) and image acquisition was performed using an LAS-3000 imager (FujiFilm).

Expression of *T. thermophila* HAP2 in mammalian cell lines. A synthetic, codon-optimized [S2] version of the full length *T. thermophila* HAP2 gene was made (Genscript) with a 5'- *Hind*III

restriction site and Kozak consensus sequence prior to the start codon, and a 3 × FLAG epitope tag followed by a stop codon and an *Eco*RI restriction site at the 3'-end. The codon-optimized HAP2 gene was digested at the aforementioned restriction sites, purified, and ligated to a similarly digested and purified pcDNA™3.1(+) mammalian expression vector (kindly provided by G. Whittaker). Resulting plasmids were verified by restriction analysis and sequencing (Cornell Biotechnology Resource Center). Purified plasmid DNA was concentrated by ethanol precipitation to ~1 μg/μL and used for transient transfections of mammalian cells.

Mammalian cell cultures of HEK-293T cells (ATCC), Huh-7 cells (Japan Health Science Research Resources Bank, Japan), Vero-E6 cells (ATCC), and BHK-21 cells (kindly provided by Mark Whitt) were maintained at 37˚C 5% CO_2 in DMEM (Corning) supplemented with 10% fetal bovine serum (ThermoFisher), 10 mM HEPES (Corning), 100 IU/mL penicillin and 100 μg/mL streptomycin (Corning). For immunofluorescence analysis of transfected cells, 3-12.5 × 10^4 Huh-7, Vero-E6, or BHK-21 cells were seeded in microscopy chamber slides (EMD Millipore) and incubated for 18 h at 37°C in a 5% CO_2 incubator. Cell supernatants were gently aspirated and replaced with 100 μL of warm Opti-MEM (ThermoFisher) media, then transfected by the further addition of 25 μL of an Opti-MEM-Lipofectamine 2000 (ThermoFisher) mixture containing the plasmid DNA encoding HAP2 or a pCAGGS empty vector control at a final concentration of 4 ng/μL, and incubated at 37°C in a 5% CO_2 incubator for 6 h. Transfection medium supernatant was then gently removed and replaced with 100 μL of warm DMEM supplemented with 10% fetal bovine serum and 10 mM HEPES but without penicillin/streptomycin and incubated 24 h at at 37°C in a 5% CO_2 incubator.

Pseudotyped particle production. Murine Leukemia Virus (MLV)-based HAP2-pseudotyped particles (HAP2pp) were generated as previously described [S3]. 1 × 10^6 HEK-293T cells were seeded in six-well plates and incubated at 37°C in a 5% CO_2 incubator for 18 h. Cells were co-transfected with pcDNA-HAP2-FLAG plasmid (HAP2pp), or pCAGGS empty vector control (Δenvpp), or VSV-G encoding plasmid (VSV-Gpp), along with MLV Gag-Pol packaging

construct and the MLV transfer vector (encoding a luciferase reporter gene), using Lipofectamine 2000 transfection reagent (Life Technologies) according to manufacturer's instructions. The cells were incubated at 37°C in a 5% CO_2 incubator for 48 h. Supernatants containing released pseudotyped particles were harvested and filtered through 0.45 µm membranes and stored at -80˚C.

Pseudotyped particle infection. For infection assays, 2.5×10^5 HEK-293T, or Huh-7, or Vero-E6 cells were seeded in 24-well plates and incubated at 37°C in a 5% CO_2 incubator for 18 h. Cells were transfected with either pCAGGS empty-vector control or pcDNA-HAP2-FLAG plasmid and incubated at 37°C for 24 h at 37°C in a 5% CO_2 incubator. The cells were washed with PBS, and 200 µL of pseudotyped particles were added to cells and incubated at 37°C in a 5% CO_2 incubator for 2 h. Complete medium was then added and cells were incubated at 37°C in a 5% CO_2 incubator for 72 h, after which luciferase activity was measured using Luciferase Assay Kit (Promega), and luminometer readings performed with a Glomax 20/20 system (Promega). Experimental values were plotted using Prism 7 (GraphPad) and are average relative luciferase units of three replicates (n = 3) with error bars representing standard deviation (s.d.).

Supplemental references and notes

S1. Liu, Y., Pei, J., Grishin, N., and Snell, W.J. (2015). The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in Chlamydomonas and regulates the fusion reaction. Development *142*, 962–971.

S2. For codon optimization of *Tetrahymena thermophila* HAP2, the amino acid sequence was first codon optimized for expression in *Homo sapiens* (GenScript) and was then submitted to the graphical codon usage analyzer (http://gcua.schoedl.de/sequential_v2.html) and manually altered (relative adaptiveness values of no less <40%) to allow for optimal expression in both insect (*Drosophila melanogaster)* and human (*Homo sapiens*) cells.

S3. Bartosch B., Dubuisson J., and Cosset F.L. (2003). Infectious hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes. J Exp Med *197*, 633-642.

Chapter four

The *Tetrahymena* conjugation-junction specific protein ZFR1 does not function in sexual cell fusion

Abstract

The fusion of individual gamete cells during sexual fertilization is a tightly regulated act fundamental to the perpetuation of eukaryotic life that unfortunately, we still know very little about. Here, to better understand the interplay of molecular components controlling cellular fusion during fertilization, we tested the possibility that a recently identified fertilization-specific zinc finger protein, ZFR1, may act in concert with the gamete fusion protein, HAP2, to drive membrane fusion events during sexual conjugation of the model ciliated protist, *Tetrahymena thermophila*. We found that when both *Tetrahymena* mating partners lacked *ZFR1,* pair fertility declined to ~20% (similar to previous findings), but that this loss of fertility was not due to the improper localization of HAP2 or defects in membrane fusion between mating cells. Interestingly, despite the known instability of the conjugation junction structure in *ZFR1* knock-out mating partners, we found that these cells continue to form pairs very late into conjugation compared to their wild-type counterparts. Together, these data solidify support for a post-fusion block to *Tetrahymena* fertility in Δ*ZFR1* cells, and are suggestive of a role for *ZFR1* in the functional regulation of developmental signaling cascades and/or junctional membrane recognition and adhesion.

Introduction

In order to construct a mechanistic model for how cell-cell fusion is accomplished by the gamete membrane fusion protein HAP2 during fertilization it is necessary to also know the cellular regulators of HAP2. It is already understood that HAP2 transcription, proteolysis, and intracellular trafficking are tightly controlled in plants and algae, limiting the amount of time HAP2 is expressed and membrane localized for its function in fusion[1–3]. Considering the drastic consequences to fitness that would ensue if a gamete fusogen were dysregulated [4], we reasoned that certain conjugation-junction localized proteins in *Tetrahymena thermophila* may exert their functions through control of HAP2-mediated fusion pore formation.

Many proteins are known to localize to the nuclear exchange junction during *Tetrahymena thermophila* sexual conjugation. Based on expression and junctional localization alone, proteomic studies on the secretome of starved mating-competent *Tetrahymena*[5] and methods for specifically isolating the proteins embedded within the entire conjugation junctional membrane itself [6], have added to a growing list of potential HAP2-interactors. This list includes, among others, the protein fenestrin, a possible signaling molecule or transcriptional regulator that has an implied structural role in the conjugation junction[6]; EF1-α, a ribosomal elongation factor with a possible 'moonlighting' function in actin bundling and extracellular matrix remodeling [7,8]; TCBP-25, an EF-hand Ca^{2+} binding protein[9]; FTT18, a 14-3-3 signaling protein that could bind and regulate any number of enzymes or transmembrane proteins[10,11]; and, ZFR1, a zinc finger protein recently demonstrated to affect fertility and the mechanical stability of mating pairs [12].

To begin to explore the possibility that one of these proteins may be involved with HAP2-fusogenic function we focused our efforts on ZFR1, testing whether it functions in tandem with HAP2 in the establishment of fusion pores, or has a regulatory role in HAP2 localization to the junctional membrane. The ZFR1 protein itself is 514 amino-acid cytosolic protein (no signal peptide or predicted transmembrane domains) with an amino terminal B-box and a C'-terminal

hydrophobic domain. A previous study of ZFR1 predicted that it functions late in conjugation (8-10 hrs), possibly in membrane remodeling after nuclear exchange [12]. To us, this protein stood out among potential HAP2-interactors for its specific localization to the conjugation junction, and requirement for cell-cell adhesion around the time of prezygotic nuclear division. In previous studies of *T. thermophila* HAP2, knockout cell lines were infertile and compromised in their ability to form stable pairs at about the same time point during mating[13]. In the HAP2 study, the observed defects in pair stability were coincident with a lack of membrane fusion pore formation at the conjugation junction.

Zinc finger domain-containing proteins have diverse functions within the cell ranging from transcriptional activation to protein folding, lipid binding and protein turnover[14,15]. Not surprisingly, zinc finger proteins have also been reported to impact fertility in systems other than *Tetrahymena.* For example, spermatogenesis in the nematode worm, *Caenorhabditis elegans,* relies on Spe10, a tetraspanin containing a DHHC-CRD zinc finger motif (palmitoyl transferase), for the proper maturation and delivery of lysosome-related fibrous body membranous organelles to the developing spermatids[16]. Since *Tetrahymena* does not make true gametes, it is difficult to envision how a protein with this type of function might be involved in fertilization, although "gametic" pronuclear maturation represents at least one possibility. Aside from Spe10, several other zinc-finger domain containing proteins have been discovered to have key roles in fertility and/or development in other organisms[17–19].

Results

ZFR1 deletion does not effect HAP2 localization or cell-cell fusion during mating.

As indicated above, previous studies in *T. thermophila* had shown that deletion of *ZFR1* leads to mechanical instability of mating pairs as well as low fertility. Because HAP2 deletion strains have a remarkably similar phenotype, we hypothesized that the ZFR1 might play a role in membrane pore formation at the conjugation junction and that the HAP2 and ZFR1 gene

products might interact. To examine these possibilities we created *ΔZFR1* knockout cell lines in three different mating types of *T. thermophila* and engineered one of these to express a cadmium-inducible HAP2::GFP reporter construct (Figure 4.3A,B and methods). Mating these cells either to each other or to WT cells allowed us to examine the role of ZFR1 in membrane pore formation during conjugation, as well as test for possible interactions between ZFR1 and HAP2 based on localization of the GFP-tagged HAP2 gene product.

As shown in Figure 4.1A, deletion of *ZFR1* had no effect on HAP2 localization at the conjugation junction. Indeed in crosses between *ΔZFR1* cell lines of different mating types, the localization and signal intensity of the HAP2::GFP fusion protein appeared at least as great as in previous studies with GFP-tagged HAP2 in cell lines harboring the wild-type *ZFR1* gene[13]. These results indicated that ZFR1 is not involved in HAP2 trafficking to the conjugation junction in mating *Tetrahymena,* and that it is likely these two gene products do not interact.

To examine the role of ZFR1 in membrane fusion pore formation, we used flow cytometry to measure the exchange of fluorescently labeled proteins between mating partners in crosses between Δ*ZFR1* knockout cell lines as compared to WT crosses (as described in Chapter 3). Here again we found only a modest reduction in the percent of cell-cell fusion in crosses between Δ*ZFR1* strains relative to WT matings (Figure 4.1B), with the difference likely attributable to the reduced pair stability observed in matings of Δ*ZFR1* cell lines [12]. These results showed that ZFR1 is not involved in HAP2-mediated cell-cell fusion during *Tetrahymena* sexual conjugation.

Δ*ZFR1* cells have abnormally high pairing frequencies late into conjugation.

Much as with *ΔHAP2* knockout strains (Chapter 2), mating pairs of *ΔZFR1* knockout strains are easily disrupted by mechanical agitation[12]. Based on these results, we decided to further test *ΔZFR1* cell pair stability in undisturbed mating cultures to establish the time course of normal cell-cell interactions in these crosses. As shown in Figure 4.2, although *ΔZFR1* x

ΔZFR1 crosses had lower overall pairing frequencies throughout the duration of conjugation when compared with WT x WT or Δ*ZFR1* x WT matings, interestingly, *ΔZFR1* crosses maintained this mediocre pairing frequency into very late time points during conjugation. Rather than synchronously disengaging from one another at the 11-12 hour time point as with WT cells[20], we instead saw that Δ*ZFR1* x Δ*ZFR1* crosses continued to pair well past 12 hrs, with only a gradual drop in pairing frequency at later time points. These results indicate one of two possibilities: either (1) Δ*ZFR1* cells stay paired well past the time when WT cell pairs normally come apart, or (2) Δ*ZFR1* pairs are in a state of constant flux between coming apart and pairing up again due to the either to a weakness in adhesion events involved in pairing or an inappropriate signal allowing complementary mating types to continue to recognize one another and pair at these late time points. Regardless, the persistence of *ΔZFR1* cell pairing so late in mating was an unexpected and potentially interesting phenotype.

Genetic evidence that Δ*ZFR1* strains undergo cross-fertilization.

In addition to normal exchange haploid pronuclei resulting in the formation of true progeny, *Tetrahymena* is capable of cyotogamy, a low frequency self-fertilization event in which migratory and stationary pronuclei of starved cells undergo fusion to form a zygotic micronucleus. Such self-fertilization events have been observed to normally occur in 2-12% of starved mixed cells of different mating types in WT mating cultures [13]. The previous study of Δ*ZFR1* progeny formation[12] did not allow for rigorous testing of normal cross- versus self-fertilization events because genetic markers for both parental cell lines were lacking. Thus, it was unclear whether the reported frequency of “progeny” development [12] (~19%) was the result of an unusually high rate of self-fertilization, or just a low rate of normal, cross-fertilization.

We were able to resolve this issue in the current study by constructing our macronuclear (Mac) ∆*ZFR1* knockout cell lines (above) from available “heterokaryon” strains (Figure 4.3A,B and methods) that carry different drug resistance markers (either 6-methylpurine [mpr] or cycloheximide [cyr]) in their transcriptionally silent micronuclei (Mic). Because the ∆*ZFR1* Mac in

both of these cell lines are sensitive to both mpr and cyr, productive matings of these lines allow their drug-resistant Mics to be brought into expression in the resultant progeny, allowing us to distinguish between cross-fertilization (true progeny) and self-fertilization (cytogamy). This was done by growing the post-mating synclone cultures sequentially in media containing 6-methylpurine and cycloheximide. Only true, cross-fertilized progeny survive both drugs (*R/R*, resistant mpr, resistant cyr).

To test Δ*ZFR1* deletion strains for their capacity to generate cross-fertilized progeny, cells of complementary mating types were starved and mated. Individual mating pairs were then hand-isolated and their exconjugant synclones were grown in the presence or absence of drug (see Chapter 2, Figure S-2.3). As shown in Figure 4.3C, the overall survival of Δ*ZFR1* x Δ*ZFR1* mating pairs in the absence of drug was the same as in WT matings. More importantly, the frequency of self-fertilization events appeared normal (~9.2%), and only 22.5 % of pairs isolated from a Δ*ZFR1* x Δ*ZFR1* matings resulted in true, “cross-fertilized” progeny. In light of the fact that ~60% of the Δ*ZFR1* mating culture were able to form fusion pores (Figure 4.1B), the substantial decrease in cross-fertilized progeny development in Δ*ZFR1* relative to WT crosses clearly indicates that the defect in Δ*ZFR1* strain fertility occurs downstream of membrane fusion.

Figure 4.1 Localization of HAP2 and cell-cell fusion in *ZFR1* deletion strains.
(A) A cell line in which the endogenous *ZFR1* gene had been deleted (Δ*ZFR1)* was induced with 0.1μg/mL $CdCl_2$ to express an *Mtt1*-promoter-driven *HAP2::GFP* construct, and then crossed with another Δ*ZFR1* cell line of a different mating type. After 5.5 hours of mating, nuclei were labeled for 30 min with Hoechst 33258 and cells were fixed and examined for the location of HAP2::GFP. Two representative pairs are shown with one fluorescence-alone image (left) and one merged fluorescence-bright-field image (right) for each pair. Scale bars are 10μm. **(B)** Populations of wild-type (WT) and Δ*ZFR1* cell lines of different mating types (α=Mating Type VII, β=Mating Type VI) were labeled with either CTFR or CFSE, mixed together at 1:1 ratio, and acquired with a BD FACS Canto flow cytometer 16 h after mating and fixation. The percent of total events (cells) that had exchanged fluorescent label (as described in Chapter 3) was quantified for each individual mating performed (open circles) and is plotted with error bars representing ± s.d. and the bar representing the mean of samples collected for a given cross.

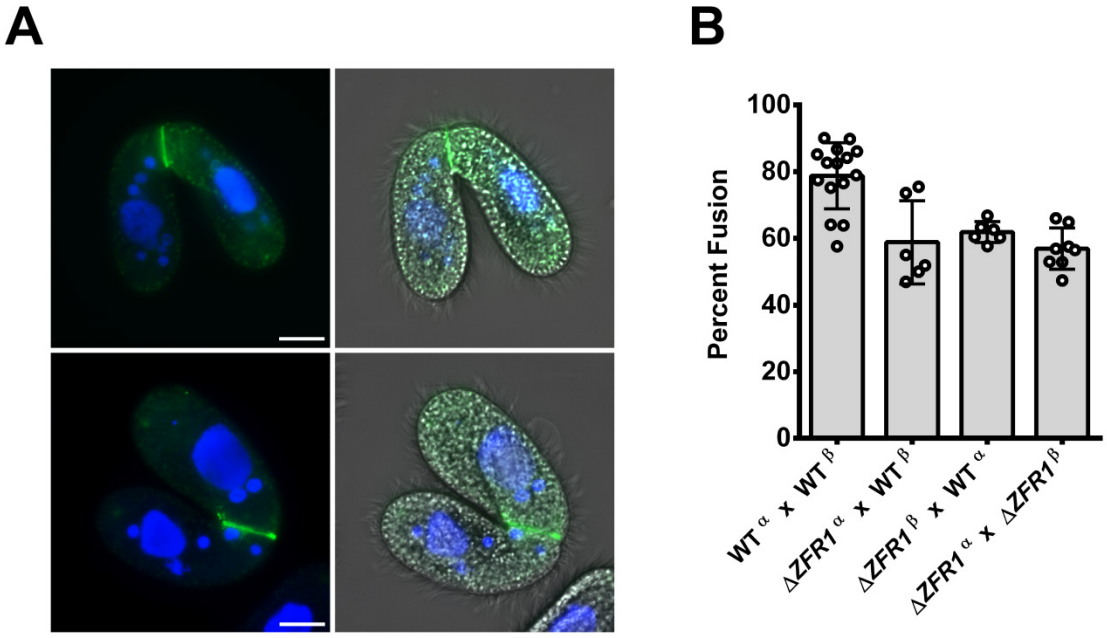

A
B
Percent Fusion
100
80
60
40
20
0
WT α x WT β
ΔZFR1 α x WT β
ΔZFR1 β x WT α
ΔZFR1 α x ΔZFR1 β

Figure 4.2 Pairing frequencies during conjugation of *ZFR1* deletion strains.
WT and Δ*ZFR1* cell lines of different mating types (α=Mating Type VII, β=Mating Type VI), were starved and mixed together at a 1:1 ratio at time zero. All cultures were placed in a stationary incubator at 30ºC for the duration of mating. The frequency of cells in pairs was measured every 30 minutes (closed squares) by carefully withdrawing cells from the culture vessel, adding fixative, and placing them on microscope slides for counting (>100 "objects" / time point) under bright field visualization. Data points from the WT^{α} x WT^{β} (black line) and $\Delta ZFR1^{\alpha}$ x WT^{β} (green line) crosses are each representative of one independent mating. Data points for the $\Delta ZFR1^{\alpha}$ x $\Delta ZFR1^{\beta}$ mating (blue line) represent mean data from two independent experimental matings.

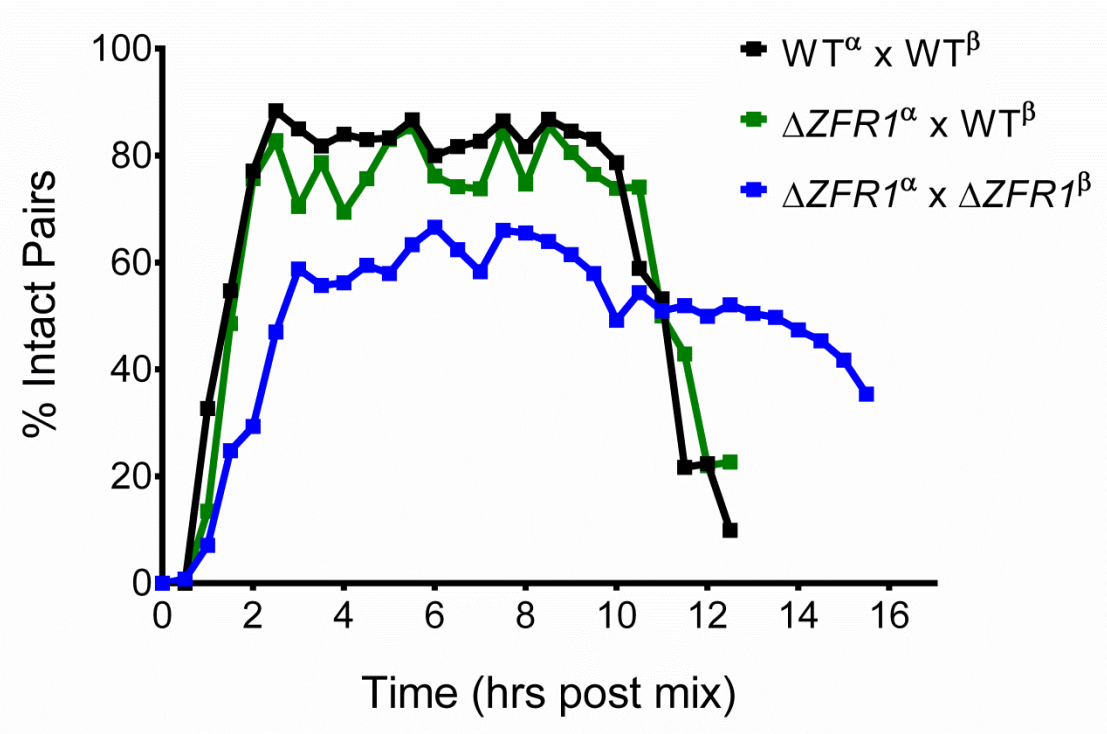
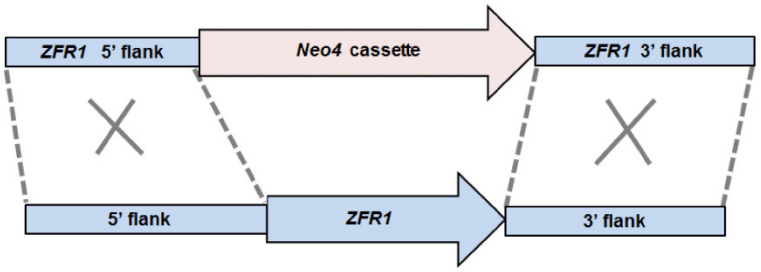
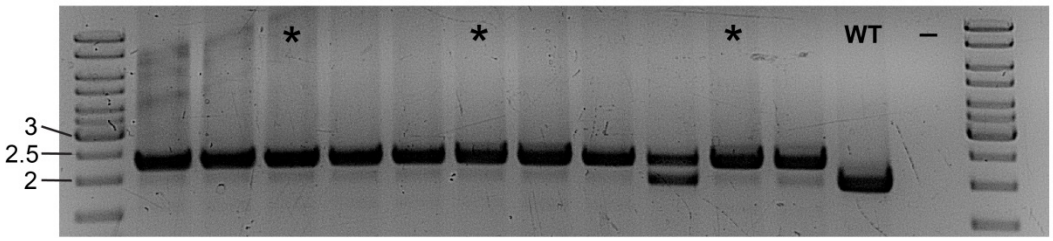

100
80
60
40
20
0
% Intact Pairs
0
2
4
6
8
10
12
14
16
Time (hrs post mix)
WT$^{\alpha}$ x WT$^{\beta}$
Δ*ZFR1*$^{\alpha}$ x WT$^{\beta}$
Δ*ZFR1*$^{\alpha}$ x Δ*ZFR1*$^{\beta}$

Figure 4.3 Construction and fertility testing of *ΔZFR1* strains. (A) A schematic of showing the construct made for replacement of the *ZFR1* locus with a neo4 cassette conferring resistance to paromomycin, X's denote sites of homologous recombination between the plasmid DNA sequence biolistically-introduced and the 5' and 3' flanks surrounding the *ZFR1* open reading frame. **(B)** PCR amplification across the predicted recombination site of wild-type and transformed *Tetrahymena* strains showing replacement of the endogenous sequence (2Kb amplicon), with the much longer resistance cassette (~2.4Kb amplicon). **(C)** Effects of *ZFR1* deletion on mating success. Reciprocal crosses (x) between *ΔZFR1* strains were carried out to corroborate previously obtained results[12]. For each cross, pairs were individually isolated into drops to establish synclones and phenotypically characterized by individual pair survival and progeny drug resistance phenotypes to different micronuclear markers present in both parental cells according to the same rubric used in Chapter one, figure S3. From these tests, individual progeny synclones fell into three categories based on their resistance (R) and sensitivity (S) to the drugs cycloheximide and 6-methylpurine: "Cross-fertilizers"; R/R, were clones that successfully completed sexual conjugation and made a new macronucleus. "Self-fertilizers"; R/S or S/R, were clones where nuclear exchange was blocked and only one parent developed a new macronucleus. "Back-outs"; S/S, were completely unsuccessful in sex and development and retained their parental macronucleus. In each cell, the percentage of pairs displaying a given phenotype is shown on the left next to parentheses containing the raw data used in that calculation (# of pairs with that given outcome / total number of pairs).

A



B



C

Mating	% Survival	% "Cross-fertilizers" *R/R*	% "Self-fertilizers" *R/S* or *S/R*	% "Back-outs" *S/S*
Δ*ZFR1*$^{\alpha}$ x Δ*ZFR1*$^{\beta}$	87 (552/634)	22.5 (124/552)	9.2 (51/552)	68.3 (377/552)

Discussion

We originally initiated this study to test whether ZFR1, a protein that localizes to the conjugation junction of mating *T. thermophila*, plays a role in directing HAP2-mediated membrane fusion during fertilization. We found, however, that in the absence of *ZFR1* expression, a GFP-tagged version of HAP2 still correctly localized to the conjugation junction (Figure 4.1A) and that mating cells were able to form fusion pores (Figure 4.1B). Additionally, we found that Δ*ZFR1* matings consistently paired at lower frequencies and quite surprisingly, continued to pair for longer than their wild-type counterparts (Figure 4.2).

In light of recent studies on HAP2 [13] one might be led to believe that membrane fusion is absolutely required for the formation of mechanically stable pairs, but these results highlight the important point that membrane pore formation is not sufficient to ensure pair stability during *Tetrahymena* mating. This is evidenced by the fact that Δ*ZFR1* pairs can fuse, but still have substantial issues with pair stability throughout conjugation[12], indicating that even when pairs are ratcheted together by fusion pores, an additional factor(s) is still necessary to prevent them from coming apart.

The experiments described here demonstrating the absence of a functional role for ZFR1 in membrane pore formation support previous studies [12] which showed that the absence of *ZFR1* expression appears to block *Tetrahymena* fertilization around ~8 -10 hours into mating. The curious persistent pairing phenotype which extends late into mating further supports this. Although not described here, bioinformatic analysis of the ZFR1's putative protein domain arrangement through sequence and structure-based homologies (psi-BLAST & Phyre2 hits, data not shown) raise the intriguing possibility that ZFR1 could be acting as an ubiquitin ligase during *Tetrahymena* mating. In this regard, it is worth noting that defects in similar protein-modification pathways within the cell, namely, the sumoylation pathway, have also recently been noted to have lower pair stability during *Tetrahymena* conjugation, a result reminiscent of those observed for Δ*ZFR1* matings[21].

Together, these findings open new lines of questioning as to the nature and control of cellular adhesion and recognition proteins during *Tetrahymena* mating: What factors in addition to membrane fusion are necessary to hold mating cells together during conjugation? What is the cellular "off-switch" for mating pair recognition and adhesion at the 12-hour time point of *Tetrahymena* conjugation? A role for protein-degradation pathways in either the maturation or regression of the conjugation junction seems intuitively logical in *Tetrahymena* mating, as the dynamic remodeling of the protein landscape at the conjugation-junction is dependent on proper intra- and intercellular signaling throughout mating, and is ultimately necessary for the timely recognition, adhesion, fusion, and finally, separation of cellular mating partners.

Materials and methods

***Tetrahymena* strains and culture conditions.** *Tetrahymena* strains were obtained from the Tetrahymena Stock Center, Cornell University. Cells were grown at 30°C in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 33.3μM $FeCl_3$) on a platform shaker at ~ 100 rpm. For mating, log phase cells were starved at ~2 $\times 10^5$ cells/ml for 12-18 hrs in 10 mM Tris buffer (pH 7.5), and mixed in equal concentration at 30°C. For somatic transformation, target cells were grown to mid-to-late log stage (~6 $\times$ 10^5 – 1 $\times$ 10^6 cells/ml) in NEFF medium, and starved overnight in 10 mM Tris buffer (pH 7.5) at ~2 $\times$ 10^5 cells/ml prior to biolistic transformation[22,23]. Following transformation, cells were transferred to NEFF with 250μg/ml Penicillin G, 250 μg/ml Streptomycin, and 1.25μg/ml Fungizone, and grown at 30°C.

***Tetrahymena ZFR1* deletion strain construction:** The endogenous *ZFR1* gene was knocked out in the macronucleus of *T. thermophila* strains CU428.2, CU427.4, and HAP2::GFP[13] by homologous recombination using a "Neo4" paromomycin resistance selection cassette (Figure 4.3A). The knockout construct was made by separately amplifying 5'- and 3'-flanking regions of the ZFR1 gene via PCR using *T. thermophila* genomic DNA as template and primer pairs ZFR1 5'Flankfor/ZFR1 5'flankrevXhoI and ZFR1 3'FLANKFORSacI

/3'FLANK REV, respectively (see below). The Neo resistance cassette was amplified from the Neo4 vector using the primer pair Neo4for2XhoI/Neo4cyrevSacI. The three PCR products were cut with the restriction enzymes X*ho*I, S*ac*I or both as appropriate, then gel purified and sequentially ligated in reactions using T4 DNA ligase. The final ligation product was gel purified, blunt-end cloned into PCR4BLUNT-TOPO and sequenced. The resulting plasmid was linearized and transformed into the somatic macronucleus of *Tetrahymena thermophila*. Clones that were completely replaced in the endogenous *ZFR1* locus with the Neo4 selection cassette were selected by serial passage of clones in NEFF growth medium containing incrementally increasing concentrations of paromomycin (up to 800µg/mL). Complete replacement of the endogenous ZFR1 gene was verified by PCR amplification of genomic DNA using primers spanning the Neo4 insertion site (ZFRrescueF1 and ZFRrescuerev), with complete knockout clones showing only one PCR fragment of the size expected for the knockout construct (Figure 4.3B). Knockout strains generated and used in experiments (PCR results with * in Figure 4.3B) were frozen under the names: *ZFR1KOinCU428cl.51, ZFR1KOinCU427cl.3,* and *GFPHAP2ZFR1KOcl.2.* The GenBank accession number for the ZFR1 protein sequence is XP_001029830 (TTHERM_01285910).

Primers used: ZFR1 5'Flankfor 5' AGAAGAAAAGTATGACAAAATTAACTG3'

ZFR1 5'flankrevXhoI 5' TAAGTACTCGAGTGACCTGAAAATATCTTAGCTAGC 3'

ZFR13'FLANKFORSacI 5' CTTATGGAGCTCGATAAAATTAAGTTCTATCTATGTTCATGAAC3'

ZFR1 3' FLANK REV 5' AGAATGCATCGAAAGATTTAAAACC3'

Neo4for2XhoI 5' GACTTACTCGAGAATAAGGGTTTTGAATAACTCCT 3'

Neo4cyrevSacI 5' ATTCTAGAGCTCTGCATTTTTCCAGTAAAAATTTGA 3'

ZFRrescueF1 5' GAAGAAACAAATAACTAACTATTGAATTCT 3'

ZFRrescuerev 5' AACAATTTTAAACATTAGCGAACT 3'

***ΔZFR1* Pair Stability Assay.** Complementary mating types (called "α" and "β" for CU428 and CU427 respectively) of Δ*ZFR1* cells were starved in 10mM Tris pH 7.5 and

mixed together at a 1:1 ratio in 50 mm × 10 mm petri dishes prior to incubation at 30ºC for the duration of mating. The final concentration of the 4 mL mating mixture was 2 x 10^5 cells/mL. At 30 minutes intervals after mixing, ~10-15 µl of the *ΔZFR1*$^{\alpha}$ *x ΔZFR1*$^{\beta}$ mating mixture was gently removed and fixed with an equal volume of IC fix buffer (eBioscience). Slides were immediately prepared from this fixed sample and visualized using the bright field optics of a model # 131-SP National Microscope. To lessen eye-strain during counting, enumeration of pairs and single cells was assisted by viewing the slide via a computer running the Motic Images Plus 2.0 software and projecting a live feed from the microscope's eyepiece through a MoticCam3.0 camera.

Genetic analysis of fertilization events in *ΔZFR1* matings via drug selection. Mating success as determined by "cross-fertilization" in *ΔZFR1* crosses was measured as described in Chapter 2, Figure S-2.3. Briefly, *ΔZFR1* cell lines were constructed carrying different drug-resistance markers in their transcriptionally silent germline micronuclei. *ΔZFR1*$^{\alpha}$ carried a dominant 6-methylpurine resistance allele (Mpr) in its micronucleus and *ΔZFR1*$^{\beta}$ carried a dominant cycloheximide resistance allele (Cyr) in its micronucleus. Upon mating, these germline resistance alleles were brought into expression in the newly developing macronucleus of progeny cells. If progeny were not formed, than the parental cells remain sensitive to both cycloheximide and 6-methyl-purine. Single pairs were isolated from a *ΔZFR1*$^{\alpha}$ *x ΔZFR1*$^{\beta}$ cross at 6 h and 9.5 h post mix. Results of the 6 hr and 9.5 hr pair isolations were similar, and data from these two independent mating experiments were combined for the results shown in Figure 4.3C. After pairs were isolated, they were allowed to complete mating, and exconjugants cells were grown in hanging drops to establish "synclones," or mixed populations of the karyonoidal descendants from each exconjugant. These synclones were then subjected to sequential drug testing in first cycloheximide, then 6-methyl-purine. True "cross-fertilized" progeny were defined as those synclones that were resistant to both drugs. Self-fertilizers or "cytogamonts" show resistance to only one drug. Parental cells which were unable to complete sexual conjugation and maintained their drug-sensitive Mac were not resistant to either drug.

Flow cytometry assays for cell-cell fusion: Flow cytometry fusion assays were carried out as described in Chapter 3. Briefly, complementary mating types of the indicated *T. thermophila* cell lines were grown and placed in starvation medium (10 mM Tris buffer, pH 7.5) labeled with amine reactive dyes, and subjected to mating and subsequent flow cytometry acquisition. 7×10^6 cells were washed once in 0.1 × PBS and resuspended 1 mL of the same buffer followed by 1mL of 0.1 x PBS containing either 20 µM CFSE (eBioscience) or 10 µM CTFR (Life Technologies). Cells were then incubated in the dark for 5 min at room temperature (RT), or 15 min at 30˚C for CFSE and CTFR, respectively. NEFF media was immediately added to quench unincorporated label and cells were washed and resuspended in 10 mM Tris (pH 7.5). Samples were maintained in a 30˚C incubator overnight, then washed again the following day prior to mating the cells at 30˚C in petri dishes ($0.5–2 \times 10^6$ total cells/dish). After mating (16-20 hr post-mixing) exconjugant cells were washed in 10mM Tris, fixed with IC Fixation buffer (eBioscience), washed once more in 1 × PBS, and resuspended in 1 × PBS containing 0.3% BSA and acquired on a BD FACSCanto™ II Flow Cytometer. A minimum of 30,000 events were collected for each mating reaction and data analyzed with FlowJo software (FlowJo LLC). The criterion for inclusion in the analyses was a % pairing $\geq$ 60% at 3 hr post mixing.

Fluorescence microscopy. For cellular imaging of the localization of GFP-tagged HAP2 in Δ*ZFR1* matings, starved cells harboring the *HAP2::GFP* construct were induced to express HAP2-GFP by addition of 0.1ug/mL $CdCl_2$ to the 10mM Tris pH 7.5 starvation medium 30 min before mixing with starved Δ*ZFR1* of the complementary mating type. Additional $CdCl_2$ was added to the starvation medium containing the mixed, mating population to bring the final concentration up to 0.1ug/mL. Mating cells were fixed at the time points indicated in the text by the addition of IC fix buffer for 20min at RT. Fixed cells were washed once with 0.1x PBS - 0.3%BSA buffer and immediately imaged using a Zeiss Axio Imager M1 microscope equipped with an AxioCamMR3 camera.

REFERENCES

1. Sprunck, S., Rademacher, S., Vogler, F., Gheyselinck, J., Grossniklaus, U., and Dresselhaus, T. (2012). Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. Science *338*, 1093–1097.

2. Liu, Y., Misamore, M.J., and Snell, W.J. (2010). Membrane fusion triggers rapid degradation of two gamete-specific, fusion-essential proteins in a membrane block to polygamy in Chlamydomonas. Dev. Camb. Engl. *137*, 1473–1481.

3. Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M., and Twell, D. (2011). The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. Plant Cell *23*, 534–549.

4. Johnson, M.A. (2010). Fertilization: monogamy by mutually assured destruction. Curr. Biol. CB *20*, R571-3.

5. Madinger, C.L., Collins, K., Fields, L.G., Taron, C.H., and Benner, J.S. (2010). Constitutive secretion in Tetrahymena thermophila. Eukaryot. Cell *9*, 674–681.

6. Cole, E.S., Anderson, P.C., Fulton, R.B., Majerus, M.E., Rooney, M.G., Savage, J.M., Chalker, D., Honts, J., Welch, M.E., Wentland, A.L., *et al.* (2008). A proteomics approach to cloning fenestrin from the nuclear exchange junction of tetrahymena. J. Eukaryot. Microbiol. *55*, 245–256.

7. Dallo, S.F., Kannan, T.R., Blaylock, M.W., and Baseman, J.B. (2002). Elongation factor Tu and E1 beta subunit of pyruvate dehydrogenase complex act as fibronectin binding proteins in Mycoplasma pneumoniae. Mol. Microbiol. *46*, 1041–1051.

8. Bunai, F., Ando, K., Ueno, H., and Numata, O. (2006). Tetrahymena eukaryotic translation elongation factor 1A (eEF1A) bundles filamentous actin through dimer formation. J. Biochem. (Tokyo) *140*, 393–399.

9. Hanyu, K., Takemasa, T., Numata, O., Takahashi, M., and Watanabe, Y. (1995). Immunofluorescence localization of a 25-kDa Tetrahymena EF-hand Ca(2+)-binding protein, TCBP-25, in the cell cortex and possible involvement in conjugation. Exp. Cell Res. *219*, 487–493.

10. Kilburn, C.L., Pearson, C.G., Romijn, E.P., Meehl, J.B., Jr, T.H.G., Culver, B.P., 3rd, J.R.Y., and Winey, M. (2007). New Tetrahymena basal body protein components identify basal body domain structure. J. Cell Biol. *178*, 905–912.

11. Bunney, T.D., Boer, A.H.D., and Levin, M. (2003). Fusicoccin signaling reveals 14-3-3 protein function as a novel step in left-right patterning during amphibian embryogenesis. Development *130*, 4847–4858.

12. Xu, J., Tian, H., Wang, W., and Liang, A. (2012). The zinc finger protein Zfr1p is localized specifically to conjugation junction and required for sexual development in Tetrahymena thermophila. PloS One *7*, e52799.

13. Cole, E.S., Cassidy-Hanley, D., Pinello, J.F., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Jr, T.G., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. Curr. Biol. CB *24*, 2168–2173.

14. Laity, J.H., Lee, B.M., and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struct. Biol. *11*, 39–46.

15. Gamsjaeger, R., Liew, C.K., Loughlin, F.E., Crossley, M., and Mackay, J.P. (2007). Sticky fingers: zinc-fingers as protein-recognition motifs. Trends Biochem. Sci. *32*, 63–70.

16. Gleason, E.J., Lindsey, W.C., Kroft, T.L., Singson, A.W., and L'hernault, S.W. (2006). spe-10 encodes a DHHC-CRD zinc-finger membrane protein required for endoplasmic reticulum/Golgi membrane morphogenesis during Caenorhabditis elegans spermatogenesis. Genetics *172*, 145–158.

17. Davies, B., Hatton, E., Altemose, N., Hussin, J.G., Pratto, F., Zhang, G., Hinch, A.G., Moralli, D., Biggs, D., Diaz, R., *et al.* (2016). Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. Nature *530*, 171–176.

18. Zhang, S., Qiu, W., Wu, H., Zhang, G., Huang, M., Xiao, C., Yang, J., Kamp, C., Huang, X., Huellen, K., *et al.* (2001). The shorter zinc finger protein ZNF230 gene message is transcribed in fertile male testes and may be related to human spermatogenesis. Biochem. J. *359*, 721–727.

19. Ramos, S.B.V., Stumpo, D.J., Kennington, E.A., Phillips, R.S., Bock, C.B., Ribeiro-Neto, F., and Blackshear, P.J. (2004). The CCCH tandem zinc-finger protein Zfp36l2 is crucial for female fertility and early embryonic development. Development *131*, 4883–4893.

20. Martindale, D.W., Allis, C.D., and Bruns, P.J. (1982). Conjugation in Tetrahymena thermophila. A temporal analysis of cytological stages. Exp. Cell Res. *140*, 227–236.

21. Nasir, A.M., Yang, Q., Chalker, D.L., and Forney, J.D. (2015). SUMOylation Is Developmentally Regulated and Required for Cell Pairing during Conjugation in Tetrahymena thermophila. Eukaryot. Cell *14*, 170–181.

22. Bruns, P.J., and Cassidy-Hanley, D. (2000). Biolistic transformation of macro- and micronuclei. Methods Cell Biol. *62*, 501–512.

23. Seashell Technology - Carrier Particle Protocols for Plasmid DNA Available at: http://www.seashelltech.com/protocols.shtml [Accessed October 24, 2016].

Chapter five

Evidence of a cryptic sexual stage in the freshwater fish parasite, *Ichthyophthirius multifiliis*[1]

[1] Figure 5.1 in this chapter was adapted from a figure in Coyne *et al.* © 2011, Genome Biology, and is used here with the permission of the corresponding author.

Abstract

HAP2 is an evolutionarily conserved protein known to play a critical role in the gamete membrane fusion of a wide range of eukaryotes. Expression of *HAP2* occurs only during the sexual stage of the life cycle in male gametes, and elimination of the *HAP2* gene blocks fertilization. Because HAP2's function is restricted to sexually active cells, the expression profile of *HAP2* has a potential to shed light on the presence of a sexual stage in difficult to study eukaryotic microbes. This includes the parasitic ciliate *Ichthyopthirius multifiliis*, a common, obligate parasite of freshwater fish. Recent sequencing of the *I. multifiliis* genome has permitted the identification of a *HAP2* ortholog of in this species. After mapping the corresponding *HAP2* transcript using 5' RACE, we looked for clues to help pinpoint the potential sexual stage of this parasite by examining *HAP2* expression patterns throughout the life cycle. We found that the infective theront stage of the life cycle is where *HAP2* mRNA is most abundantly expressed, with the possibility of differential RNA processing of the *HAP2* message at other parasitic stages also limiting the presence of the mature transcript. We then replaced the HAP2 ectodomain of a related ciliate, *T. thermophila,* with the *I. multifiliis* HAP2 ectodomain and found that the *I. multifiliis* protein was unable to rescue sexual cell fusion events using assays developed in *T. thermophila*. Together these results reinforce the species-specific N'-terminal sequence conservation requirements for HAP2 previously reported in plants, and provide further support for the presence of a sexual theront stage in this commercially important fish pathogen.

Introduction

The pathogenic ciliate Ichthyophthirius multifiliis is the causative agent of “white-spot disease”, a highly contagious, often lethal disease of freshwater fish that is a major problem for commercial fish producers and aquarium hobbyists alike, resulting in significant economic loss[1,2]. Currently there is no preventative treatment and no effective cure for large-scale outbreaks. Although the observable I. multifiliis life cycle is well documented, the presence of a sexual stage has not been definitively shown[1], and a more resolute understanding of the life cycle could lead to better therapeutic interventions. The life cycle of this obligate parasite consists of morphologically distinct stages both on and off the fish, called the trophont and tomont, tomite and theront, respectively (Figure 5.1)[1,3]. Here, we provide data to support the idea that theronts may be a sexual stage in this parasite’s life cycle based on transcriptional upregulation of HAP2, a key gene encoding the sexual fusogen necessary for fertilization events in a broad array of species.

Based on the wide representation of meiotic genes in extant eukaryotes, it is now strongly suspected that many originally assumed asexual eukaryotic microbes, especially pathogenic species[4,5], may contain the presence of a sexual stage[6,7]. In this respect, it is possible that many ciliate species, including Ichthyophthirius, are also secretively sexual[7]. Even though it has yet to be definitively shown, previous studies on this important fish pathogen do contain evidence to support that sexual conjugation may be possible[8,9]. Like many ciliates, Ichthyophthirius is binucleate, and perhaps the most suggestive evidence for sex in the I. multifiliis is its maintenance of a micronucleus[1]. The macronuclei of ciliates are transcriptionally expressed, whereas the micronucleus houses the germline genetic information and is only utilized during sexual development. Amicronucleate ciliates species are common[10], but they only vegetatively reproduce and are incapable of sexual fertilization due to their lack of a micronucleus[11]. Also, following extended periods of clonal growth, ciliates

tend to eventually lose functionality and mass in their micronucleus, in what is known as germinal senescence[12], sometimes causing cell death. However, micronuclear loss has not occurred in Ichthyophthirius. Furthermore, other cytological data has captured some provocative images of what appears to be a theront fusing to a trophont on a live fish[8].

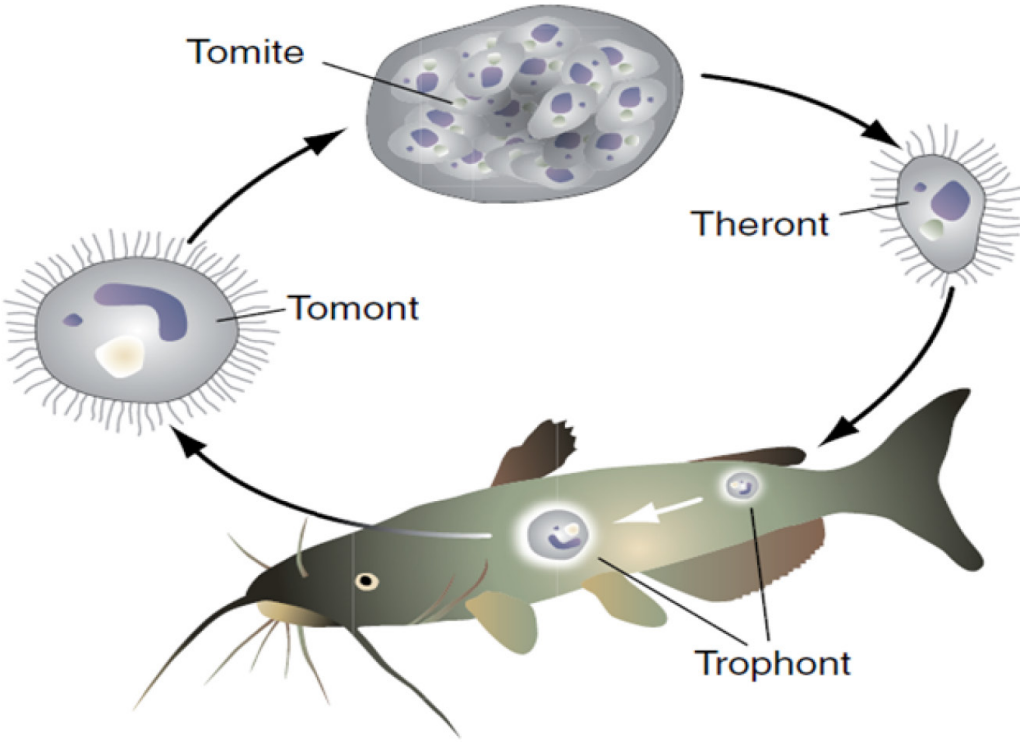
With the sequencing of the I. multifiliis genome[3] and generation of expressed sequence tag libraries[13–15], genetic evidence has also emerged for the presence of sex-specific machinery. Homologs of essential meiotic genes[16] and ciliate conjugation-specific genes[9] have now been cataloged in I. multifiliis. A study of stage-specific parasite gene expression also found that transcripts associated with the gene ontology categories of reproduction and development were four times more prevalent in theront than trophont samples[13]. Finally, molecular genetic diversity characterization of sampled parasite populations has found levels of variation high enough to imply that I. multifiliis has been reproducing sexually[9].

Despite such suggestive evidence, it has been difficult to observe sex in Ichthyophthirius. According Dunthorn and Katz[7], sex is easily missed in these ciliates for three reasons; first, inappropriate laboratory conditions, second, lack of obvious sexual morphological features, and third, the facultative nature of sex in ciliates means it is not always a regular occurance. These same problems apply to I. multifiliis, as its obligate parasitic lifestyle requires serial passage on live fish for maintenance in the laboratory and methods have yet to be established for axenic culture, making it very difficult to observe possible sexual events in vivo. In order to advance the characterization of a possible sexual life cycle, it will be necessary to first identify the parasitic stage at which sex is likely to occur.

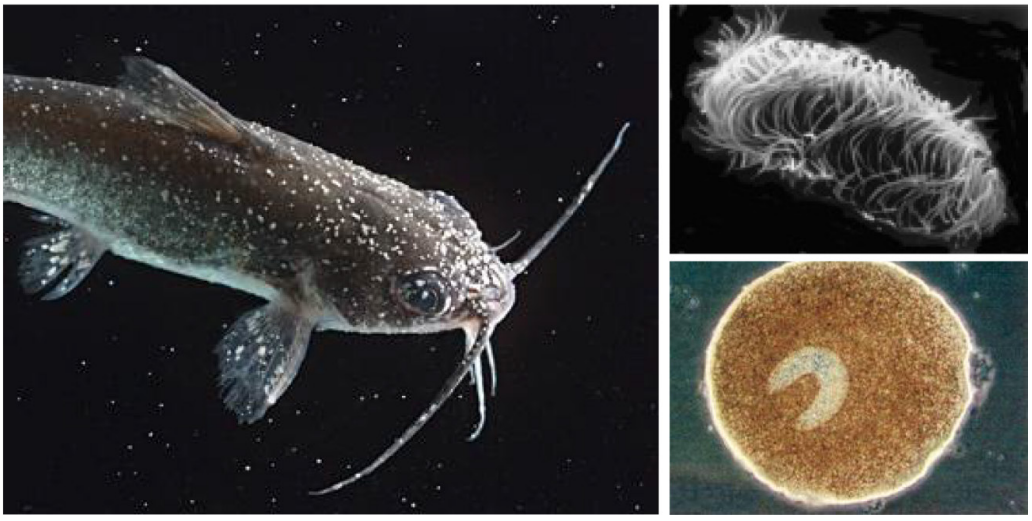
The gene, Hapless 2 (HAP2), is one potential marker for sexual reproduction in species with cryptic sexual lifestyles. So far, HAP2 has been studied in plants[17], green algae[18,19].

Figure 5.1 The known life cycle of *Ichthyophthirius multifiliis*. (A) Infective theronts burrow through surface mucus, infect the epidermis, and then begin to metamorphose into feeding trophonts. Excessive cellular hypertrophy during 4-7 days of growth eventually culminates in parasite egress from the host as tomonts, which swim briefly before adhering to a substrate and secreting an encompassing gelatinous capsule. Mitotic division within this protective enclosure leads to the generation of hundreds of tomites, which within 24hrs at room temperature escape the capsule as infective theronts. If theronts fail to infect fish they die within 1 to 2 days. This image was adapted from Coyne et al., Genome Biology 2011[3] and is used with the permission of the corresponding author. **(B)** Images of *I.multifiliis* parasitic stages. An infected channel catfish is shown (left). Each individual white spot on the fish is a single trophont. On the right, representative images of the theront (top) and trophont stage (bottom) parasites are shown. Note that the apparent difference in cilia length between the theront and trophont is an illusion, cilia are actually the same length in both cases, the trophont is just a much larger cell. Images courtesy of Dr. Theodore Clark.

A



B



cnidarians[20,21], insects[22], amoebae[23], and protozoa[19,24,25] where it's been found to be a critical protein for gamete membrane fusion during sex, with expression on male gametes correlating with the sexual stage of the organism's life cycle[6,26]. The gene itself encodes a single pass, type 1 transmembrane protein whose extracellular amino terminus is highly conserved at the primary sequence level. In a study by Wong et. al[27], the exchange of the Arabidopisis amino terminus with that of a closely related species (Sisymbrium irio, 89% amino acid sequence identity) produced a functional protein. However, when the HAP2 ectodomain from a more divergent species was substituted instead (Oryza sativa, 59% identical) the resulting protein did not function in Arabidopisis fertilization. These findings give a sense of the functional malleability of the HAP2 ectodomain between different species, but similar studies have yet to be attempted outside of plants.

HAP2 homologs are also present in the phylum Ciliophora[24], whose members are predominantly isogametic, with only a very few species having what could be thought of as male/female (anisogamous) gametes. A majority of ciliate species actually do not make true gametes at all, instead producing gametic –like pronuclei, in cells of multiple distinct mating types[28]. In this regard, HAP2 expression and function has been studied in a close relative of Ichthyophthirius, namely T. thermophila, where it has also been found to be restricted to the sexual conjugation phase of the life cycle and to be necessary for fertilization[24]. But Tetrahymena differs slightly from other anisogamous systems in that all seven mating types (or sexes) express HAP2. This expression starts right after opposite mating types are initially mixed together, and continues for the next three hours at high levels regardless of whether or not the expressing Tetrahymena cells end up finding a mating partner or even carry through with the conjugation process. Other genetic markers of sex in ciliates, such as those involved in ciliate-specific progeny development[9] or in meiosis detection toolkits[29] are expressed only after sexually-activated (or costimulated) cells have found a partner and begun the conjugation process, which in an organism that may only facultatively sexually reproduce could lead to

negative or difficult to interpret results. As a broadly conserved gene expressed early and universally in the mating process, this makes HAP2 expression a better, less confounding indicator of a sexual life stage in Ichthyophthirius.

Results

Identification and 5' mapping of the *I. multifiliis HAP2* coding sequence

A TblastN search of the *Ichthophthirius multifiliis* genome database[3,30] using the *T. thermophila* HAP2 protein sequence[24] as query uncovered a high similarity alignment (score=396, e-value= e^{-111}) with a hypothetical protein encoded by the gene identifier IMG5_026080. This hypothetical protein was predicted to contain the conserved protein domain *HAP2-GCS1* (aka. NCBI #cl11296, Pfam#PF10699)[31,32], but was shorter (525 amino acids) than expected given the lengths of other HAP2 orthologs.

Due to past issues with computational gene calls in ciliate genomes, we decided to perform 5' Rapid Amplification of cDNA Ends (5'RACE) to identify the transcriptional start site of the IMG5_026080 coding sequence, henceforth referred to as "*Im HAP2*." The AT-rich nature of the ciliate genomes slightly increases the difficulty of RACE procedures in these organisms, which has led to the development of elegant protocols utilizing oligonucleotide ligation to overcome these hurdles[33]. We adopted a slight modification of this approach outlined originally by Eyal et. al.[34] in which a gene-specific primed theront *Im HAP2* cDNA was circularized via intramolecular ligation with T4 RNA ligase. Inverse PCR on circularized products identified the transcriptional start site of the *Im HAP2* gene as being 12 nucleotides upstream of a newly predicted start codon. Our new annotation of the *Im HAP2* coding sequence (Figure 5.2A) adds one intron and 231 nucleotides (77 amino acids) to the originally-predicted *Im HAP2* coding sequence, increasing the total expected length of the gene product to 602 amino acids.

ClustalW alignment of HAP2 sequence

CLUSTAL[35,36] alignment of the predicted amino acid sequences of *Ichthyophthirius* and *Tetrahymena* HAP2 are shown in Figure 5.2B. Overall, the degree of amino acid sequence conservation was high, with 46% identity between the ectodomains of both proteins. Surprisingly, *Im* HAP2 is predicted to encode a much shorter cytosolic domain than *T. thermophila* HAP2 (14 vs. 180 amino acids respectively). In addition to this, while *Im* HAP2 conserves a poly-basic region in its cytosolic domain, it lacks any of the characteristic cysteine residues present in other HAP2 orthologs. Both the poly-basic regions and conserved cysteine motifs in the C' terminus have been found to be important for HAP2 function in some species[27,37], but in a studies of *Plasmodium*[38] and *Tetrahymena* HAP2 (submitted), these sequence elements were found to have little-to-no functional impact on fusion.

HAP2 Stage-Specific Gene Expression

To quantify *Im HAP2* expression throughout the *I. multifiliis* life cycle trophonts were collected from juvenile catfish at 7-9 day post-infection, and either immediately processed or allowed to mature in carbon-filtered water at room temperature to either the tomite or theront stage before extraction of total RNA. cDNAs made from these samples were used as template for quantitative PCR. As shown in Figure 5.3, theronts appeared to express the highest levels of *Im HAP2* mRNA,. In this Figure, the results from two independent samples of theront RNA were compared with the average transcript levels from multiple biological replicates of trophont and tomite RNAs. A one-way Kruskal Wallis test found that one of the theront samples had significantly higher expression than all other samples at $p<0.0001$. Although the mean of the second theront sample was about 2 fold higher than that of tomite and trophont samples collected, this difference was not significant. Currently, the reason for the large discrepancy observed between the two theront samples is unknown.

Figure 5.2 Identification of the *I. multifiliis HAP2* ortholog. The *HAP2* ortholog from *I. multifiliis*, whose cDNA sequence was determined using 5'RACE. **(A)** The complete cDNA sequence of *Im HAP2*. 5' RACE identified an earlier start codon and sequence (grey shaded nucleotides) that were missed by gene-call algorithms. **(B)** A CLUSTAL 2.1 alignment of the entire *I. multifiliis* (Im) and *T. thermophila (*Tt) HAP2 amino acid sequence. Symbols below aligned amino acids indicate degree of similarity: ***** indicates identity, **:** is strong conservation, **.** is weak conservation, and " " is no conservation. In both (A) and (B), highlighted sections indicate the positions of important HAP2 sequence elements, including the signal sequence (green), fusion loop (red), HAP2 domain (yellow), transmembrane domain (blue), and poly-basic region (purple).

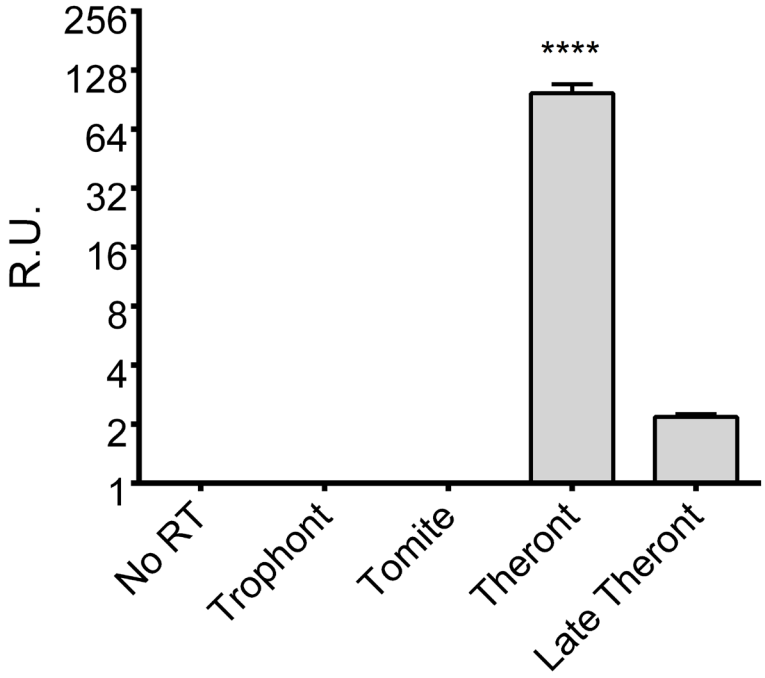
A

```
ATGTTTAAATAGATAAATTTTTTTCTTATTTTAATCAAAACTTTAATCTATCTCATAATTAACTTATCTTCTCTTTTTA
AATGCGAGTTTATATCCAAAACAGAAATTTAAAAATGCTACTAATCTCAGAATTTAAACGCATGTAGTGAAAAGGCAGT
AATAGTACTATCATTAGAAAACAATTAAAGCACTGTTTCAGAAAAAATGACAGCCTAAATAAAATCCGTAAAAATGCCT
GAAGGAAACATGTTACTTAAAAACTATATAGAAATAGAAGTGACGAAATCTCCTGTTGTAGCAGTATACCCATTAAAGT
ATATGAGAGATTACGAAAGTATGCCATAAGAAAAGGTAATATCAAAAAGTGTATTTACATGCTAAGATGGATTTAATGA
GGATTCACCAACATGTGGTTTCTAAAGAGATGAAAAAGGCGAAAAAATATTCGATTCACAAGGTTTTTGCTGTAAATGT
GGAGCAGCTGACTTTTTTGGTTTGGGAAAAGAAGTAATGAGAGGAGTAGATTGCCTTCCATTTAATTTAAATTCAGGTT
CTGCTAGCGCTCATTGTTTACGTTTTCCAGGTAGATGGTATTCTGGCTATGAAATTTTATAATATTATATATATTATGA
AATAAAAGTTGAAGTATACGAACTAGAAGGTAATAATAATAAAAAGAGGAAGCTTAAATACAAGCTAACAACAAGTACT
ACTGATAGAATCAAAAAATCTCCAGACAATAAGTTTCTAGTAAAAATAATCGGTGATTTTTTCCCTACATAACCTCCTC
CAGTATATAATAATGTATATTTAGTACGCCCAACCCCAAACAGGCCTTAAGCTAATAATGAATTAAGAGTAAGAGTATT
AGAAGGAATCTCAAATTGGATGCTAATAGAAAAGAATCAATTTACACTAGATGGAACTGAATGTAACAAAATAGGAGTC
TCTTATGCTGCATTTAGGAGAGAAAATGGATCTTGTAGTAAATAAATAGGCTCTTGTTTAAAAAATCAAATAGAACATT
TCTACCTTAGAGATATTGAAAGAATAAAAAAAGGGTAGCCTACTTAGAATTTACTTTTACCCAAAGGAGACTTTTAAGA
AAGTTGGGATAAATAAAATAATACATAAATGATTCTTTTTATAGAAGGAAGTATGTCTACATTAATTACAATAGAAATG
GATTCCGCGGAAATTTAATTCTTAACAATGTTAGGTTAAGGAAAATTTATTCTTGTAAAAATTAATAATTTTGAAAGTC
ATTCGGGAAGTGGAAAGTTTGAAGCTCATATTTTGAATAAATCAAGTTTCGCAGCAGAATTTAATTTGGGATTTTCTTG
CGATTAAAATGTATTGCCCATTTCTGGATAAAAACTTTTTCTTAATCAAGATTAATTATTTATTTTCAAATCTAGTGTT
AATGTAGTTAGTGATTTAGGAAAAACAAATAATCTTTGTAATGTTACGCTTTCTGATGCAGTTAATAATGTTTTAGATT
TTGCCTAAATTACGTTCAATACAACTGATGTTGTTAGAATTAGTCCTCAAGGTAATGGAACTTATTATAATGAAAATAA
TAGTACTTTGAAAAAACCTTTAATAGAGGTTACTTGTAACTAAAAATGTCCTGACTTTTGGGATATTTTTTGTCATTTT
AGTACTAAGTGTTTAAATAATGGTTTTAAAACATTAGGAACTGGCTTAGGAATTTTAGTTATCTTTTTAGAATTATTCG
ATGTTGTTGCTCTTTTTGTTGTTCTTCATAACCTTCAGAGGAAGAAGAAGAATAATAATAAAAAAAAATGA
```

B

```
Im   MFKQINFFLILIKTLIYLIINLSSLFKCEFISKTEIQKCYQSQN-LNACSEKAVIVLSLE 59
Tt   ----MKFLAFGLIYFHFCILN-----RCEYITSSTIQKCYNSSNEPNNCSQKAVIVLSLE 51
         ::*: : :  : : *:*     :**:*:.: *****:*.*  * **:*********

Im   NNQSTVSEKMTAQIKSVKMPEGNMLLKNYIEIEVTKSPVVAVYPLKYMRDYESMPQEKVI 119
Tt   NGQIANTEQVVATLNQLSDSGVNKQLQNSFIFEVTKSPVTALFPLIYLQDFNSQPLEQVI 111
     *.* : :*::.* ::.:. .  *  *:* : :*******.*::** *::*::* * *:**

Im   SKSVFTCQDGFNEDSPTCGFQRDEKGEKIFDSQGFCCKCGAADFFGLGKEVMRGVDCLPF 179
Tt   ATTLFSCKDGFYDSSPTCKFQYDSKGQKILDSQGYCCYCSLSDILGMGNDLSRGKVCYAL 171
     :.::*:*:*** :.**** ** *.**:**:****:** *. :*::*:*::: **  * .:

Im   NLNSGSASAHCLRFPGRWYSGYEILQYYIYYEIKVEVYELEGNNNKKRKLKYKLTTSTTD 239
Tt   NLGAGSATAHCLKFSPLWYSAFKIQQYQLYFEVNINIYTVDSQNQK--NLKQTLKLSTSN 229
     **.:***:****:*.  ***.::* ** :*:*::::::* ::.:*:*  :** .*. **::

Im   RIKKSPDNKFLVKIIGDFFPTQPPPVYNNVYLVRPTPNRPQANNELRVRVLEGISNWMLI 299
Tt   PTMKSSDNSTISKIIGTFTPTQPPADLSSYYLVKPSFPATDP------RVLQGISSWMFV 283
        **.**. : **** * *****.  .. ***:*:   .:.      ***:***.**::

Im   EKNQFTLDGTECNKIGVSYAAFRRENGSCSKQIGSCLKNQIEHFYLRDIERIKKGQPTQN 359
Tt   DKTMFTLDGTQCNKIGVSYSGFRQQSSSCSQPVGSCLQNQLENLYQSDLILLSQNKQPKY 343
     :*. ******:********:.**::..***: :****:**:*::*  *:  :.:.: .:

Im   LLLPKGDFQESWDKQNNTQMILFIEGSMSTLITIEMDSAEIQFLTMLGQGKFILVKINNF 419
Tt   LLESQGNFNQ-VQFQGQTILQQGLSGSASTLITIEIDAAQIKFVTNLGIGCISQCSINNF 402
     ** .:*:*::  : *.:* :   :.** *******:*:*:*:*:* ** * :   .****

Im   ESHSGSGKFEAHILNKSSFAAEFNLGFSCDQNVLPISGQKLFLNQDQLFIFKSSVNVVSD 479
Tt   ESHSGNGKLVALVQNQGNYSAEFVLGFNCSSNVQPIQGQKLFLTANQLYNFNCSVSVNSD 462
     *****.**: * : *:..::*** ***.*..** **.******. :**: *:.**.* **

Im   LGKTNNLCNVTLSDAVNNVLDFAQITFNTTDVVRISPQGNGTYYNENNSTLKKPLIEVTC 539
Tt   ISAINNNCTINLYDAIGNQLDSKNILFNTTSTNHTSNQGNNTGQQQSSQEYKS---SQSC 519
     :.  ** *.:.* **:.* **  :* ****.. : * ***.*  ::...  *.   . :*

Im   NQKCPDFWDIFCHFSTKCLNNGFKTLG--TGLG---ILVIFLEL---FDVVALFVVLHNL 591
Tt   SDKCSSFWSFWCYFSAGCIKEAFKSIASIAGVASALALVIFLAKNGYLVPIIRFLCCCCC 579
     .:**..**.::*:**: *:::.**::.  :*:.    *****     :  :  *:

Im   QRKKKNNNKKK---------------------------------------------------- 602
Tt   KSKKKENEKNKDKTDKKSIQESCSYDRSCCSHSISQSYQVENKNKYKRSKIQRSFSSESC 639
     : ***:*:*:*
```

Figure 5.3 Stage specific expression of *I. multifiliis HAP2*. Quantitative reverse-transcriptase PCR results of samples taken from different stages of the *I. multifiliis* life cycle. The graph shows the average transcript level (R.U.) from four independent trophont collections and two independent tomite collections. Transcript levels from two independent theront collections are shown separately. Each RNA isolation was performed on total parasites collected *en masse* from 4-5 juvenile catfish. Total RNA was purified and 1µg was used as template for cDNA generation. Biological samples were run in technical triplicates for qPCR. For the theront samples, bars represent the mean level of HAP2 transcript (+/- s.d) from the triplicate runs compared to an *in vitro* transcribed control *HAP2* standard curve. A significant difference between one of the theront stage samples and all other stages was found by a one-way Kruskal Wallis test with Dunn's multiple comparisons post test at (H_5=551 and **** = P<0.0001).


256
128
64
32
16
8
4
2
1
R.U.
****
No RT
Trophont
Tomite
Theront
Late Theront

RT-PCR shows evidence of alternative splicing

As a routine check for reverse transcription in the expression studies described above, we subjected samples that had been generated in the presence or absence of reverse transcriptase to PCR amplification and agarose gel electrophoresis. We were surprised to find not one, but multiple bands indicating differently-sized transcriptional products (Figure 5.4A). The PCR primers used for this amplification were expected to produce a segment of the *Im HAP2* coding sequence 701bp long, spanning the second two introns. However, larger amplicons were observed on the gel just above and below 800bp. To discern what accounted for the differences in size observed in these products, we excised and extracted DNA from gel slices, TOPO-cloned the purified PCR products, and sent the resulting plasmid clones for sequencing. Sequencing showed that the differences observed in the sizes of these cDNA products were due to intron retention, which appeared to occur more frequently in trophont stage parasites (Figure 5.4B). The observed pattern of intron retention at different stages of the parasite life cycle is suggestive of a regulatory role for the processing of Im HAP2 transcripts.

The *I. multifillis* HAP2 ectodomain does not function in *T. thermophila* sexual cell fusion.

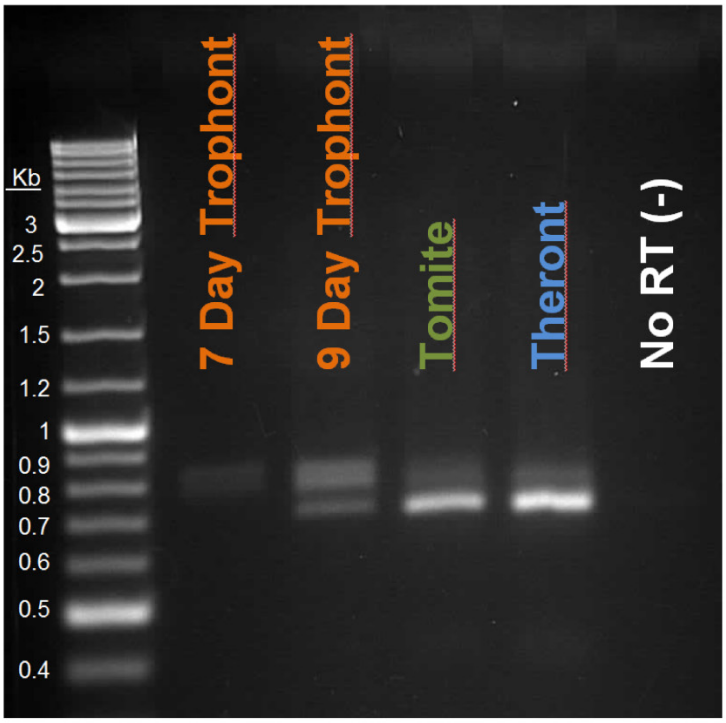
Previous results have indicated that the Im HAP2 protein is predicted to adopt an overall fold that is similar to that of class II viral fusogens (Chapter three). In Figure 5.5A we report a partial Im HAP2 ectodomain structural prediction made by the template-based structural modeling platform Phyre2. Structural homology to the Dengue virus envelope glycoprotein (PDB ID: 3UAJ)[39] was detected at 95.2% confidence (17% sequence identity) to a 196 amino acid stretch (equivalent to 27% coverage) of the Im HAP2 ectodomain, just proximal to the HAP2-GCS1 domain (specifically, residues #103 to 298 of the Im HAP2 protein sequence). The predicted model captures domains typical of class II fusion proteins, namely portions of domain I (red) and domain II (yellow), with the latter containing the predicted Im HAP2 fusion loop (circled). Alignment of the two Phyre2-predicted HAP2 structures from *T. thermophila* and *I.*

multifiliis demonstrated some degree of structural similarity (Z score = 8.4)[40,41] despite the differences in primary sequence of the proteins and the fact that the models were constructed based on different template structures (data not shown). Because of this structural similarity, we decided to test whether the Im HAP2 ectodomain was sufficient to functionally replace *Tetrahymena's* ectodomain in sexual cell fusion events in the latter species.

An interspecific chimeric *HAP2* gene construct was made containing the coding sequence for the *Im HAP2* ectodomain with a *T. thermophila HAP2* transmembrane and cytosolic domain and a C' terminal GFP tag, and introduced into the endogenous macronuclear *HAP2* locus of two *T. thermophila HAP2* deletion strains (mating type VI, *ΔHAP2-427* and VII, *ΔHAP2-428*) by homologous recombination (Figure 5.5B). Although GFP fluorescence was not readily observed when strains harboring the chimeric constructs were mated, anti-GFP antibodies demonstrated that the recombinant protein was expressed and correctly localized at the conjugation junction, an area of dense fusion pore formation between the two mating cells (Figure 5.5C). We then tested the fusogenic functionality of *Im HAP2* chimeras in crosses with both WT and *ΔHAP2* deletion mating partners (Figure 5.5D) using a flow cytometry-based cell-cell fusion assay developed for *T. thermophila* described in Chapter 3. Briefly, this assay allows for the sensitive and high-throughput detection of cells that have mutually exchanged fluorescently labeled cytosolic protein as a result of membrane fusion during normal *Tetrahymena* mating. Similar to previous findings in plants where the HAP2 ectodomains of divergent species were not interchangeable[42], we found that the Im HAP2 chimeras were not competent in rescuing *Tetrahymena* sexual cell fusion events.

Figure 5.4 Intron retention in *HAP2* transcripts during early infection. cDNA samples from different parasite stages were PCR amplified with primers spanning the second two introns in the ImHAP2 gene sequence. **(A)** Representative gel showing RT-PCR products from different *I.multifiliis* life stages. Note the different sized bands in the trophont and tomite samples indicating the presence of transcripts of different lengths. In this study, trophonts were collected from the same fish at 7 days and 9 days post-infection. **(B)** A pictoral representation of entire *Im HAP2* coding sequence (top), and the potential amplicons (bottom) generated by PCR with the primers used (arrows). Exons are represented by black rectangles, and excised introns by thin grey crooked lines. Non-excised introns are shown as thick grey rectangles. Shown to the right are their expected sizes of potential amplicons in base pairs (bp), and immediately to the right of that the number of sequenced clones from each parasite sample denoted by a subscript. The code (R=tRophonts, O=tOmites, H=tHeronts) is colored in accordance with the sample identity shown in (A). No sequences were obtained for the potential 743bp amplicon.

A


Kb
3
2.5
2
1.5
1.2
1
0.9
0.8
0.7
0.6
0.5
0.4
7 Day Trophont
9 Day Trophont
Tomite
Theront
No RT (-)

B

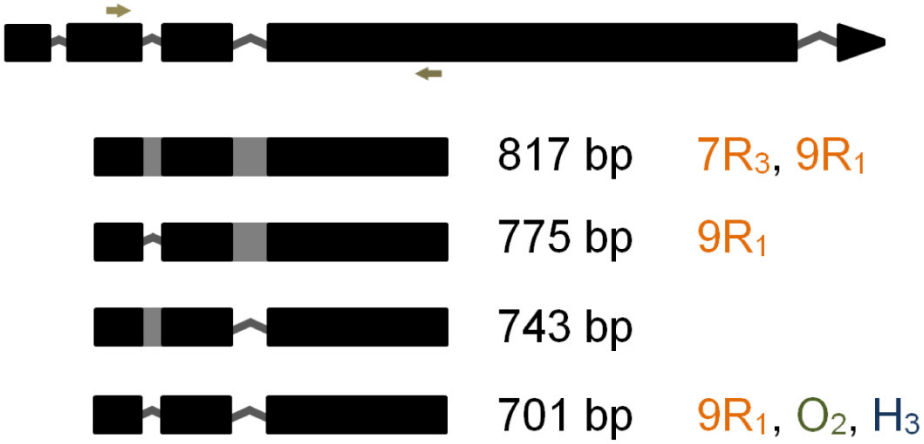
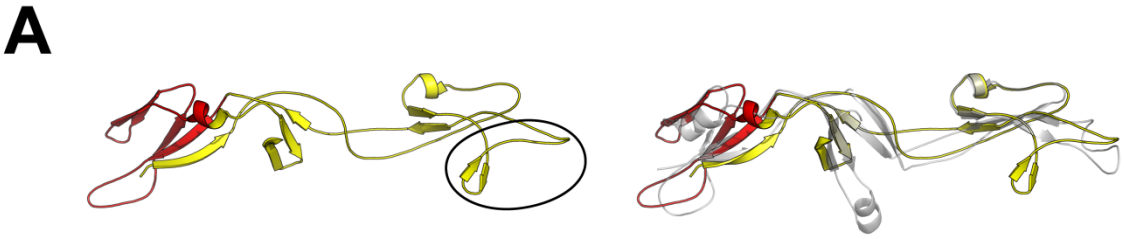
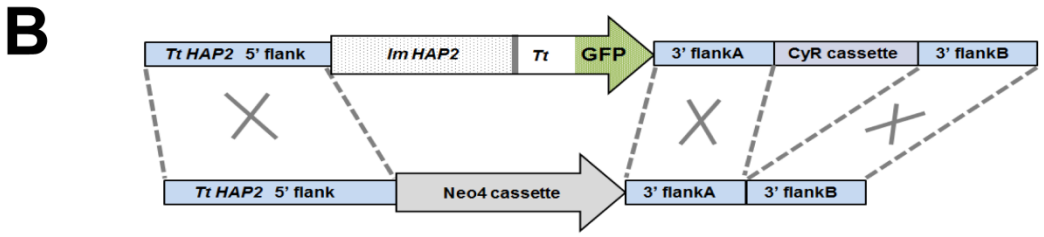
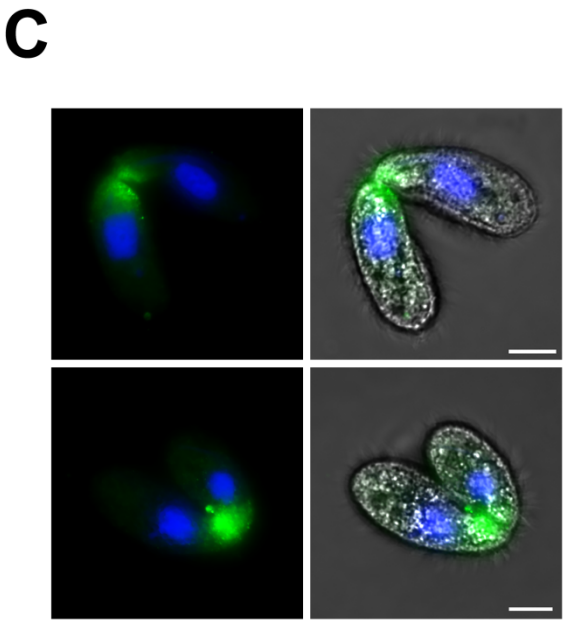
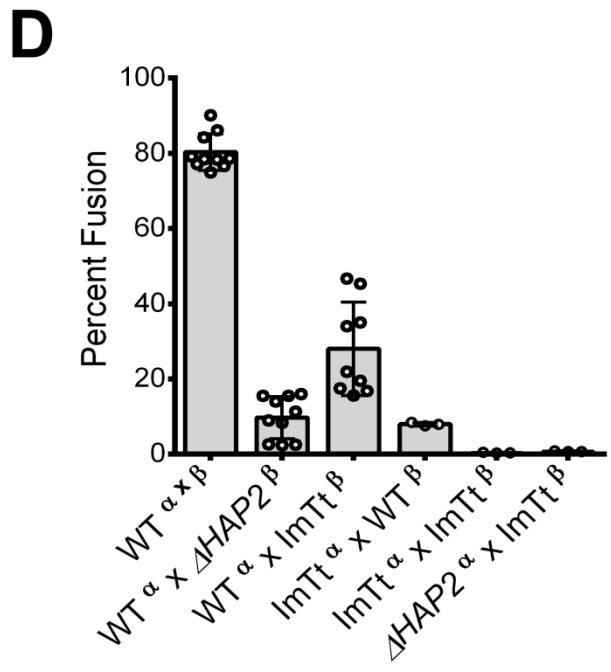

817 bp
7R3, 9R1
775 bp
9R1
743 bp
701 bp
9R1, O2, H3

Figure 5.5 Functional testing of an interspecific *I. multifillis* HAP2 chimera.
(A) The PHYRE2 structural prediction engine was used to render a partial-predicted 3D model of Im HAP2 (left). The location of the fusion loop is indicated by a black circle. An alignment of this prediction to the *T. thermophila* HAP2 protein (grey) is shown to the right. **(B)** A schematic outlining the approach used for generation of the Im HAP2 chimeric *T. thermophila* cell lines. The top cartoon shows the gene construct used for transformation encoding an open reading frame containing the *Im HAP2* ectodomain, *T. thermophila* HAP2 transmembrane and endodomain, and a GFP tag (large arrow), as well as the 5' and 3' *HAP2* flanking regions (light blue rectangles) used for homologous recombination (indicated by dashed grey lines and X), and a cycloheximide (CyR) resistance cassette used in selection. The bottom cartoon depicts the endogenous *HAP2* locus of the *ΔHAP2* deletion cell lines used for this transformation. **(C)** Immunofluorescence images of representative mating pairs of *T. thermophila* that were fixed and stained with anti-GFP antibodies to localize chimeric HAP2, and Hoescht 33258 to stain nuclei. In each case pairs consisted of a HAP2 wild type (WT) cell mated with an Im HAP2 chimeric cell lines of opposite mating type. A fluorescence-only (left) and merged phase-fluorescence images (right) are shown for each mating pair. Scale bar is 10μm. **(D)** A bar chart showing the mean +/- s.d. percent fusion for different crosses between WT, *HAP2* deletion *(ΔHAP2*) and interspecies chimeric HAP2 (ImTt) recombinant cell lines. Open circles represent data from individual mating experiments. On the X-axis, the mating partners in each cross are given the superscripts α and β to denote mating type VII or VI respectively.


A


B
Tt HAP2 5' flank
Im HAP2
Tt
GFP
3' flankA
CyR cassette
3' flankB
Tt HAP2 5' flank
Neo4 cassette
3' flankA
3' flankB


C


D
Percent Fusion
100
80
60
40
20
0
WT α x β
WT α x ΔHAP2 β
WT α x ImTt β
ImTt α x WT β
ImTt α x ImTt β
ΔHAP2 α x ImTt β

Discussion

While many multicellular eukaryotes flagrantly advertise their sexuality, making for readily distinguishable sexual habits, the private lives of most pathogenic and free-living eukaryotic microbes are not so easily observed[7]. The fresh-water fish pathogen, *I. multifiliis,* is one such microbe, with highly motile theronts and an obligate parasitic life style that could involve conjugation events on live fish. These factors cause difficulties in observation and study, making gene expression studies an ideal way to define the details of this parasite's life cycle.

Here we found that *Ichthyophthirius multifiliis* expresses the gamete fusogen *HAP2* in theronts, suggestive of a sexual parasitic stage. The *Im HAP2* start site was mapped using 5' RACE, and *HAP2* RNA expression throughout the life cycle was quantified. We also found evidence for control over *HAP2* transcript maturation through intron retention at the earlier stages in the life cycle, indicating a possible regulatory role for RNA splicing in HAP2 production. Not surpisingly, an Im HAP2 ectodomain chimera failed to rescue sexual cell fusion events in a related, free-living species, *Tetrahymena thermophila*, even though both are predicted to have high structural similarity to Class II viral fusion proteins. This is likely due to amino acid differences between the two orthologs but could also be caused by any differences in the mating behavior of these two species. Together, these results provide additional evidence in support of a sexual life cycle in *I. multifiliis*, and specifically identify theronts as the morphological stage involved in fertilization. These results are in agreement with previous studies examining *I.multifiliis* population genetic structure[9], transcriptional profiling of developmental genes[13], presence of key meiotic genes[16], and cytological findings of parasite cell-cell fusion on infected fish[8]. However, an RNA-sequencing study would be a very valuable step in confirming these accumulated findings.

Now that theront stage parasites are the key suspect for sexual activity, much also remains to be done to detail the species-specific biology involved in this conjugal event. One key question is which cells are the mating partners of sexually-active theronts? The biology of

most related free-living ciliates[11,24,43], would suggest that theronts likely exist in multiple mating types, and transiently pair and fuse to one another just prior to or during the establishment of infection. Another possibility could be that this mating system works more similarly to other anisogamous species in which HAP2 only functions in the male or “minus” gamete. In this case, the expectation would be for “male” theronts to fuse with a yet unknown female partner. While this seems unlikely, there is precedence for anisogamy in some primitive Karyorelictid ciliates[28,44], and previous cytological studies have found evidence of theronts fusing to the morphologically distinct trophonts on live fish[8], a cellular fusion event that could be considered anisogamous, if it is indeed sexual.

Another important question to answer will be whether or not the sexual stage of *I. multifiliis* is an obligate or facultative part of the life cycle. The difficulty in observing a sexual stage in the life cycle may well be because sex is not a necessary step for infection[4,5,7]. If sex is facultative in *Ichthyophthirius*, understanding what triggers it will be important in determining when and why it occurs. Ciliates favored as laboratory models (*Tetrahymena, Paramecium, Oxytricha,* etc.) typically have conjugation events spurred by periods of nutrient starvation and the presence of an opposite, mature mating partner[11] at or around macronuclear G1 phase of the cell cycle[45], but other wild ciliate species have less predictable patterns. For example, the prolific marine planktonic oligotrich *Pelagostrobilidium,* has short (5.6 hr) conjugation events that are apparently synchronized to a diurnal rhythm, with conjugal initiation each daybreak[46]. Complicating matters more, if sex is a necessary step in the course of infection for these protozoans and a suitable partner cannot be found, it’s also possible that theronts could undergo a type of autogamy (or self-fertilization event) that is a regular occurence in other related ciliate species (e.g. *Paramecium*).

The inability to generate transgenic parasites and the nature of the obligate infection cycle make laboratory experiments with *I.multifiliis* difficult. However, outside of expression studies to define a putative sexual life stage, one could envision using existing cell biology tools

in creative ways to study the *I. multifiliis* sexual cell fusion event. For example experiments using the fluorescent nuclear and cytoplasmic markers Edu and CFSE respectively, could be applied to separate theront populations, followed by a secondary mixed infection to examine early trophonts for the presence of both markers in individual cells as an indicator of cell-cell fusion.

Lessons from the sexual cell biology of other ciliates could also be applied to *I. multifiliis* in the examination of morphological features which may be common to a sexual stage. For instance, the fast-swimming cellular behaviour of theront cells is very similar to the fast-swimming *Tetrahymena* cells which develop after extended periods of starvation[47,48]. Also, the oral apparatus of ciliates is commonly absent before and during conjugation[49], and in *I. multifiliis* theronts, the small primitive oral apparatus matures into a large buccal cavity just before or during the invasion of the fish epithelium[1,50,51]. A specialized conjusome organelle forms nearby to the oral region during *Tetrahymena* conjugation[52], but the function an enigmatic theront-specific organelle of Lieberkühn in the same location has yet to be determined[1]. Increases in macronuclear ploidy are also characteristic of ciliate development after conjugation and could be analogous to similar ploidy changes observed in the trophont stage parasites[53]. Finally, the related ciliates *Tetrahymena, Paramecium,* and *Euplotes* all upregulate localized concanavalin A (Con A) receptor proteins on their membranes following stimulation by surface-bound or secreted pheromones in preparation for conjugation[43,54–57]. These Con A-binding proteins have been experimentally shown to play a role in the adhesion of opposite mating partners during conjugation, with ConA treament early in costimulation completely inhibiting the subsequent pairing and mating of *Tetrahymena* cells[56–58]. A study testing the effects of lectin treatment of *Ichthyophthirius* theronts showed modest but significant reductions in trophont development and parasite invasion with lentil, gorse, and wheat germ agglutinin, but the lectin Con A was not tested[59].

Understanding how sexual conjugation fits into the *Ichthyophthirius* life cycle could clearly lead to better treatments for this detrimental pathogenic ciliate. Not only would knowledge of sexual activity be important epidemiologically for tracking infectious outbreaks, but also for the potential utilization of transmission-blocking tools already under development for halting parasite-specific sexual activity. Examples of preventative or therapeutic measures might include the use of fish vaccines targeting HAP2 or other sexually expressed genes[60,61], antimicrobial-resistant-proof small molecule drugs such as atovaquone[62], or bumped kinase inhibitors which block a calcium-dependent kinase signaling pathway widely important for the zygotic development of alveolates[3,63,64]. Finally, HAP2 may provide a window to understanding what other ciliate and non-ciliated unicellular eukaryotes, both free living and pathogenic, have cryptic sexual stages.

Materials and methods

Growth and Collection of Parasites and *T. thermophila*: *I. multifiliis* G15 strain trophonts were collected from 4 infected juvenile channel catfish 7, 8 and 9 days after initial exposure. Trophonts were released from the fish epidermis by gentle rubbing, and allowed to develop in carbon filtered aquarium water at room temperature. Tomite stages were collected at 7 and 16 hrs after harvesting trophonts, and free swimming theronts were collected 18-22 hours following trophont isolation.

Tetrahymena thermophila strains were obtained from the Tetrahymena Stock Center, Cornell University (https://tetrahymena.vet.cornell.edu/) and grown at 30˚C in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 33.3 μM $FeCl_3$) and starved in 10 mM Tris buffer (pH 7.5) as previously described in Chapters 2 - 4. For mating, cells were maintained in Tris buffer for up to 48 hrs, then mixed in equal numbers to a final concentration of 2×10^5 cells/mL.

RNA isolation and cDNA generation: *I. multifiliis* cells were collected as described above, spun down and homogenized in TRIZOL (Thermo Fisher). Total RNA was extracted and quantified using a QuaWell UV spectrophotometer Q3000 (Quawell Technology, Inc.) before being aliquoted and stored at -80°C. Just prior to reverse transcription, 5µg of total RNA was treated with RNAse-free DNase I (Thermo Fisher) and 1ug of this reaction product was used as template for reverse transcription. All cDNAs were synthesized using SuperScript III reverse transcriptase (Thermo Fisher) and the *HAP2* gene-specific reverse primers (Sigma Aldrich) indicated below. For the qPCR quantitation standard, 1µg of SP6 (Promega) *in vitro* transcribed *I. multifiliis* HAP2 RNA was used to generate 1°- strand cDNA, which was then serially diluted to the estimated equivalent of 100fg of starting material. A diagnostic RT-PCR of cDNA generated from the synthetic RNA products was performed on both these positive (synthetic RNA) as well as negative (noRT) controls (see Figure 5.4). A full list of PCR primers and their experimental uses is provided in the table below.

5' Rapid Amplification of cDNA Ends (5'-RACE): *I. multifiliis* HAP2 theront RNA was reverse transcribed using a PNK (New England BioLabs) phosphorylated gene-specific IchHAP2GSP1rev primer and treated with RNAseH (Thermo Fisher). Reaction products were then purified on a Qiagen PCR clean-up column and eluted with 60°C sterile, ddH_2O. 12µL of the resultant purified cDNA was then circularized in an intramolecular ligation reaction mediated by T4 RNA ligase (New England BioLabs) according to an approach developed by Eyal et. al in 1999[34]. The ligation reaction (30µl) was allowed to proceed overnight at room temperature and contained 3µl of 10mM ATP, 5µl of 50% PEG 8000, 4µl ddH_2O, as well as 3µl of both the ligase and the 10x buffer provided (New England BioLabs). A nested inverse PCR (with the primer sets IchHap2for2 / IchHap2GSP3.75rev, followed by IchHap2for3 / IchHap2GSP4rev, Tm=61°C, extension times of 30 and 45 seconds respectively) of single-strand circles using Phusion polymerase (ThermoFisher) with manual hot start allowed for isolation of products slightly larger than the expected size. These blunt-ended PCR products were then cut from the

gel, cloned into a pCR™4Blunt-*TOPO*® plasmid vector (ThermoFisher) and transformed into E. cloni competent cells (Lucigen). The purified plasmid DNA was first restricted with *Eco*RI to identify the transformant clones that contained the 5' RACE PCR product, and then Sanger sequenced at the Cornell Biotechnology Resource Center using an M13rev primer.

Quantitative PCR: RNA was isolated from bulk *I. multifiliis* parasite samples collected from each life cycle stage as described above. cDNA was generated from 1µg of DNase I-treated RNA using the IchHap2RNArev gene-specific primer. The primer set, IchPair1, used for quantitative PCR was designed using PrimerExpress3 software (Thermo Fisher) and detected both mature (spliced) as well as unprocessed messages. 2µl of a 4µM working stock of the forward/reverse primer mix, along with 0.4µl template cDNA was used in a SyberSelect MasterMix (Thermo Fisher) for a total reaction volume of 10µl in a 384 well plate. Reactions were run in triplicate for each sample (both experimental and serially-diluted quantitation standards) on an ABI Viia7 Real-Time thermocycler using fast cycling mode (50°C 2min, 95°C 2min, then 40 cycles of 95°C 1sec, 58°C 30sec). A dissociation curve was performed after the run, and did not contain off-target amplicons. The amount of *HAP2* transcript in samples was determined by comparison of C_ts generated from experimental samples relative to those of the standard curve. All statistical tests were performed using Prism 7 software (GraphPad Inc.). A one-sided non-parametric Kruskal Wallis test with Dunn's multiple comparisons post-test was applied to the qPCR data because low sample sizes did not allow for normality tests. From these tests, there was a significant difference found in the percent fusion for data from the 5 different crosses (H_5= 551.0, P<0.0001). Although both theront samples had elevated *HAP2* transcript levels, significant differences were found only between the first theront sample and the tomite and trophont samples.

Structural homology modeling of *I. multifiliis* HAP2: The Protein Homology/analogY Recognition Engine V2.0 (Phyre2) was queried under "normal" mode (as described in detail in chapter 3) with the newly deduced full-length amino acid sequence of *I. multifiliis HAP2* (as

determined from the newly annotated start codon of the gene IMG5_026080) for predicted homologies to other known crystal structures in the protein data bank.

***T. thermophila* strain generation:** The *ImTt HAP2* ectodomain hybrid::GFP *Tetrahymena* cell lines used in this study were made through modification of existing *ΔHAP2-428* clone 5 and *ΔHAP2-427* clone 6 heterokaryon strains expressing the complementary mating types (VII and VI respectively) described earlier[24,65]. The mutant *ImTt HAP2* Ectodomain hybrid - GFP gene construct was prepared using overlap PCR of products derived from the separate amplifications of both the *I. multifiliis* ectodomain cDNA sequence and the cytosolic cDNA coding region of *T. thermophila.* The reverse and forward primers from each of these respective primer sets contained extended sequence with complementarity to the opposite product, allowing for a third PCR reaction to seal the two amplicons together. After overlap, the *Bcl*I and *Mlu*I restricted PCR product was ligated to a similarly restricted pMTT1-NRK2-GFP plasmid vector[66] for purpose of attaching the GFP coding sequence to the 3' terminus of the gene. Transformed plasmid DNA was then subsequently amplified with primers X and Y in order to attach 5' B*am*HI and 3' K*pn*I sites for subsequent cloning.

The aforementioned PCR product was then gel purified and cloned into a previously constructed "*HAP2* rescue vector" which contained a pHrpl29-B cycloheximide resistance cassette and is described at length in Chapters 2[24] and 3. *Bam*HI and *Kpn*I restriction sites were used for insertion via T4 DNA ligase (New England BioLabs) of the restricted *ImTt HAP2* Ectodomain hybrid -GFP coding sequence (Figure 5.5b) into this vector prior to transformation and amplification in E. cloni 10G competent cells (Lucigen)[24].

Somatic (macronuclear) transformation of linearized plasmid DNA was performed as stated earlier for *T. thermophila*, with *ΔHAP2* target cells first grown to late log phase (~1 × 10^6 cells/mL) in NEFF medium and then starved overnight in 10 mM Tris buffer (pH 7.5) for 24 h at ~2 × 10^5 cells/mL prior to biolistic transformation bombardment[67] using a PDS-1000/He Biolistic Particle Delivery System (Bio-Rad). Positive transformants were selected and pushed

to complete macronuclear replacement by growth in NEFF medium containing 25-50 µg/mL cycloheximide. The transformant clones isolated and used in this study were named ImResc427 clone 5 and ImResc428 clone 1. Other *T. thermophila* strains used were the *HAP2* wild-type cell lines (CU427.4 and CU428.2) and *ΔHAP2-428* and *ΔHAP2-427* clones generated in previous studies[24].

Flow cytometry assays for cell-cell fusion: Complementary mating types of the indicated *T. thermophila* cell lines were grown and placed in starvation medium (10 mM Tris buffer, pH 7.5) for 24 hr and then labeled with amine reactive dyes and subjected to mating and subsequent flow cytometry acquisition as outlined in Chapter 3. Briefly, prior to labeling, 7×10^6 cells were washed once in 0.1 × PBS and resuspended 1 mL of the same buffer. Then, 1mL of 0.1x PBS containing either 20 µM CFSE (Affymetrix eBioscience) or 10 µM CTFR (Life Technologies) was added, and cells were incubated in the dark for 5 min at room temperature (RT), or 15 min at 30˚C for CFSE and CTFR, respectively. NEFF media was immediately added to quench unincorporated label and cells were washed and resuspended in 10 mM Tris (pH 7.5), maintained overnight, then washed again the following day and mated for 16-20 h at 30˚C in petri dishes ($0.5–2 \times 10^6$ total cells/dish). Following mating, exconjugant cells were centrifuged (350-400 × g), fixed with IC Fixation buffer (Affymetrix eBioscience) and resuspended in 1 × PBS containing 0.3% BSA prior to acquisition on a BD FACSCanto™ II Flow Cytometer.

A minimum of 30,000 events were acquired for each mating reaction and data were analyzed using FlowJo software (FlowJo LLC). The criterion for inclusion in the functional analyses was a pairing frequency of >60% in mating cultures 3 hr post mixing. Sample sizes for the crosses tested (Figure 5.5C) are listed in parentheses below with the total number of biological replicates over all experiments (i.e., the total number of individual matings) listed first, followed by the total number of independent experiments performed for each cross: $WT^{\alpha \times \beta}$(10,

3); WT$^{\alpha}$ x Δ*HAP2*$^{\beta}$ (10,3); ImResc427cl.5$^{\beta}$ x WT$^{\alpha}$ (9, 4); ImResc427cl.5$^{\beta}$ x Δ*HAP2*$^{\beta}$ (3,1); ImResc428cl.1$^{\alpha}$ x WT$^{\beta}$ (3,1); ImResc427cl.5$^{\beta}$ x ImResc428cl.1$^{\alpha}$ (3,1).

Table S-5.1 PCR Primers:

Primer Name[a]	Sequence 5'→ 3'	Primer Use[b]
ICHTtHap2EctoHybrev	GCTACTCCTGCAATTGAAGCAATG CCAGTTCCTAATGTTTTAAAACC	*ImTt* HAP2::GFP Construct
TtIchHap2HybtranmemFor	GGTTTTAAAACATTAGGAACTGGCA TTGCTTCAATTGCAGGAGTAGC	*ImTt* HAP2::GFP Construct
NEWICHHap2BclIfor	GATTACTGATCATGTTTAAATAGAT AAATTTTTTTCTTATTTTAATC	*ImTt* HAP2::GFP Construct
NEW3'GFPHap2MluIrev	TATACGACGCGTCTTTCAATTAGTA GATAGAGAGGAGATGTTG	*ImTt* HAP2::GFP Construct
ICHHap2BamHIF	GATTACGGATCCATGTTTAAATAGA TAAATTTTTTTCTTATTTTAATC	*ImTt* HAP2::GFP Construct
ICHHap2KpnIR	atatacGGTACCTCACCCCATGATtttgt atagttc	*ImTt* HAP2::GFP Construct
IchHap2Sp6RNAFor	GATTTAGGTGACACTATAGAAATGT TTAAATAGATAAATTTTTTTCTTATT TTAATC	*in vitro* transcription
IchHap2RNArev	GCCAGTTCCTAATGTTTTAAAACC	qPCR
IchPair1 For	ATATTCGATTCACAAGGTTTTTGCT	qPCR
IchPair1 Rev	TCTCATTACTTCTTTTCCCAAACCA	qPCR
IchHap2For1	GTGACGAAATCTCCTGTTGTAGC	RT-PCR
IchHap2GSP2rev	CCATTTTCTCTCCTAAATGCAG	RT-PCR
IchHAP2GSP1rev	GGAATCCATTTCTATTGTAATTAAT G	5'RACE
IchHap2for2	GCTAGCGCTCATTGTTTACG	5'RACE
IchHap2GSP3.75rev	CAGCTGCTCCACATTTACAG	5'RACE
IchHap2for3	CCGCGGAAATTTAATTCTTAAC	5'RACE
IchHap2GSP4rev	CCACATGTTGGTGAATCCTC	5'RACE
M13Rev	CAGGAAACAGCTATGAC	Sequencing

*All primers used in this project were supplied by Sigma-Aldrich, stored as 100 µM stock solutions, and diluted to 10µM working solutions prior to use in PCR reactions. Tms were determined using Modified Breslauer's thermodynamics, dH and dS parameters as recommended by the manufacturer.

REFERENCES

1. Dickerson, H.W. (2006). 4. Ichthyophthirius multifiliis and Cryptocaryon irritans (Phylum Ciliophora). In Fish diseases and disorders, P. T. K. Woo and J. F. A.-P.-2006 Leatherland, eds. (Wallingford, UK: CABI Pub.), pp. 116–153. Available at: http://newcatalog.library.cornell.edu/catalog/5919538 [Accessed October 28, 2016].

2. Martins, M., Moraes, F., Fujimoto, R., Onaka, E., Nomura, D., Silva, CAH, and Schalch, SHC (2000). Parasitic infections in cultivated freshwater fishes a survey of diagnosticated cases from 1993 to 1998. Rev. Bras. Parasitol. Veterinária *9.1*, 23–28.

3. Coyne, R.S., Hannick, L., Shanmugam, D., Hostetler, J.B., Brami, D., Joardar, V.S., Johnson, J., Radune, D., Singh, I., Badger, J.H., *et al.* (2011). Comparative genomics of the pathogenic ciliate Ichthyophthirius multifiliis, its free-living relatives and a host species provide insights into adoption of a parasitic lifestyle and prospects for disease control. Genome Biol. *12*, R100.

4. Heitman, J. (2006). Sexual Reproduction and the Evolution of Microbial Pathogens. Curr. Biol. *16*, R711–R725.

5. Heitman, J. (2010). Evolution of Eukaryotic Microbial Pathogens via Covert Sexual Reproduction. Cell Host Microbe *8*, 86–99.

6. Speijer, D., Lukeš, J., and Eliáš, M. (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc. Natl. Acad. Sci. *112*, 8827–8834.

7. Dunthorn, M., and Katz, L.A. (2010). Secretive ciliates and putative asexuality in microbial eukaryotes. Trends Microbiol. *18*, 183–188.

8. Matthews, R.A., Matthews, B.F., and Ekless, L.M. (1996). Ichthyophthirius multifiliis: observations on the life-cycle and indications of a possible sexual phase. Folia Parasitol. (Praha) *43*, 203.

9. MacColl, E., Therkelsen, M.D., Sherpa, T., Ellerbrock, H., Johnston, L.A., Jariwala, R.H., Chang, W., Gurtowski, J., Schatz, M.C., Mozammal 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.

10. Doerder, F.P. (2014). Abandoning sex: multiple origins of asexuality in the ciliate Tetrahymena. BMC Evol. Biol. *14*, 112.

11. Orias, E. (1986). Ciliate Conjugation. In The Molecular Biology of Ciliated Protozoa, Gall, J.G., ed. (New York, NY: Academic Press), pp. 45–94.

12. Nanney, D.L. (1974). Aging and long-term temporal regulation in ciliated protozoa. A critical review. Mech. Ageing Dev. *3*, 81–105.

13. Cassidy-Hanley, D.M., Cordonnier-Pratt, M.-M., Pratt, L.H., Devine, C., Mozammal Hossain, M., Dickerson, H.W., and Clark, T.G. (2011). Transcriptional profiling of stage specific gene expression in the parasitic ciliate Ichthyophthirius multifiliis. Mol. Biochem. Parasitol. *178*, 29–39.

14. Abernathy, J.W., Xu, P., Li, P., Xu, D.-H., Kucuktas, H., Klesius, P., Arias, C., and Liu, Z. (2007). Generation and analysis of expressed sequence tags from the ciliate protozoan parasite Ichthyophthirius multifiliis. BMC Genomics *8*, 176.

15. Abernathy, J., Xu, D.-H., Peatman, E., Kucuktas, H., Klesius, P., and Liu, Z. (2011). Gene expression profiling of a fish parasite Ichthyophthirius multifiliis: Insights into development and senescence-associated avirulence. Comp. Biochem. Physiol. Part D Genomics Proteomics *6*, 382–392.

16. Chi, J., Mahé, F., Loidl, J., Logsdon, J., and Dunthorn, M. (2014). Meiosis Gene Inventory of Four Ciliates Reveals the Prevalence of a Synaptonemal Complex-Independent Crossover Pathway. Mol. Biol. Evol. *31*, 660–672.

17. Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). Generative Cell Specific 1 is essential for angiosperm fertilization. Nat. Cell Biol. *8*, 64–71.

18. Kawai-Toyooka, H., Mori, T., Hamaji, T., Suzuki, M., Olson, B.J.S.C., Uemura, T., Ueda, T., Nakano, A., Toyoda, A., Fujiyama, A., *et al.* (2014). Sex-Specific Posttranslational Regulation of the Gamete Fusogen GCS1 in the Isogamous Volvocine Alga Gonium pectorale. Eukaryot. Cell *13*, 648–656.

19. Liu, Y., Tewari, R., Ning, J., Blagborough, A.M., Garbom, S., Pei, J., Grishin, N.V., Steele, R.E., Sinden, R.E., Snell, W.J., *et al.* (2008). The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in Chlamydomonas and Plasmodium gametes. Genes Dev. *22*, 1051–1068.

20. Steele, R.E., and Dana, C.E. (2009). Evolutionary history of the HAP2/GCS1 gene and sexual reproduction in metazoans. PloS One *4*, e7680.

21. Ebchuqin, E., Yokota, N., Yamada, L., Yasuoka, Y., Akasaka, M., Arakawa, M., Deguchi, R., Mori, T., and Sawada, H. (2014). Evidence for participation of GCS1 in fertilization of the

starlet sea anemone Nematostella vectensis: Implication of a common mechanism of sperm–egg fusion in plants and animals. Biochem. Biophys. Res. Commun. *451*, 522–528.

22. Garcia, V.E. (2012). A Generative Cell Specific 1 Ortholog in Drosophila melanogaster. Available at: https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/20270/Garcia_washington_0250O_10191.pdf?sequence=1 [Accessed January 19, 2016].

23. Okamoto, M., Yamada, L., Fujisaki, Y., Bloomfield, G., Yoshida, K., Kuwayama, H., Sawada, H., Mori, T., and Urushihara, H. (2016). Two HAP2-GCS1 homologs responsible for gamete interactions in the cellular slime mold with multiple mating types: Implication for common mechanisms of sexual reproduction shared by plants and protozoa and for male-female differentiation. Dev. Biol. *415*, 6–13.

24. Cole, E.S., Cassidy-Hanley, D., Pinello, J.F., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Jr, T.G., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. Curr. Biol. CB *24*, 2168–2173.

25. Hirai, M., Arai, M., Mori, T., Miyagishima, S.Y., Kawai, S., Kita, K., Kuroiwa, T., Terenius, O., and Matsuoka, H. (2008). Male fertility of malaria parasites is determined by GCS1, a plant-type reproduction factor. Curr. Biol. CB *18*, 607–613.

26. Wong, J.L., and Johnson, M.A. (2010). Is HAP2-GCS1 an ancestral gamete fusogen? Trends Cell Biol. *20*, 134–141.

27. Wong, J.L., Leydon, A.R., and Johnson, M.A. (2010). HAP2(GCS1)-dependent gamete fusion requires a positively charged carboxy-terminal domain. PLoS Genet. *6*, e1000882.

28. Orias, E. (2014). Membrane Fusion: HAP2 Protein on a Short Leash. Curr. Biol. *24*, R831–R833.

29. Schurko, A.M., and Logsdon, J.M. (2008). Using a meiosis detection toolkit to investigate ancient asexual “scandals” and the evolution of sex. BioEssays *30*, 579–589.

30. IchDB | Ichthyophthirius Genome Database Wiki Available at: http://ich.ciliate.org/index.php/home/welcome/ [Accessed October 28, 2016].

31. Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., *et al.* (2015). CDD: NCBI’s conserved domain database. Nucleic Acids Res. *43*, D222-226.

32. Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., *et al.* (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*, D279–D285.

33. Liu, X., and Gorovsky, M.A. (1993). Mapping the 5’ and 3’ ends of Tetrahymena thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). Nucleic Acids Res. *21*, 4954–4960.

34. Eyal, Y., Neumann, H., Or, E., and Frydman, A. (1999). Inverse Single Strand RACE: An Adapter-Independent Method of 5’ RACE. BioTechniques *27*, 656–658.

35. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., *et al.* (2014). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539–539.

36. Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., and Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res. *43*, W580–W584.

37. Liu, Y., Pei, J., Grishin, N., and Snell, W.J. (2015). The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in Chlamydomonas and regulates the fusion reaction. Dev. Camb. Engl. *142*, 962–971.

38. Mori, T., Hirai, M., Kuroiwa, T., and Miyagishima, S.Y. (2010). The functional domain of GCS1-based gamete fusion resides in the amino terminus in plant and parasite species. PloS One *5*, e15957.

39. Cockburn, J.J., Navarro Sanchez, M.E., Goncalvez, A.P., Zaitseva, E., Stura, E.A., Kikuti, C.M., Duquerroy, S., Dussart, P., Chernomordik, L.V., Lai, C.-J., *et al.* (2012). Structural insights into the neutralization mechanism of a higher primate antibody against dengue virus. EMBO J. *31*, 767–779.

40. Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. *38*, W545–W549.

41. Holm, L., and Laakso, L.M. (2016). Dali server update. Nucleic Acids Res., gkw357.

42. Wong, J.L., Leydon, A.R., and Johnson, M.A. (2010). HAP2(GCS1)-dependent gamete fusion requires a positively charged carboxy-terminal domain. PLoS Genet. *6*, e1000882.

43. Cole, E.S. (2016). Cell-Cell Interactions Leading to Establishment of a Mating Junction in Tetrahymena and Paramecium, Two “Contact-Mediated” Mating Systems. In Biocommunication of Ciliates, G. Witzany and M. Nowacki, eds. (Springer International Publishing), pp. 195–220. Available at: http://link.springer.com/chapter/10.1007/978-3-319-32211-7_12 [Accessed September 12, 2016].

44. Kovaleva, V.G. The pronuclei of the lower ciliate Tracheloraphis totevi (Karyorelictida). *134*, 367–377.

45. Wolfe, J. (1976). G1 arrest and the division/conjugation decision in Tetrahymena. Dev. Biol. *54*, 116–126.

46. Ota, T., and Taniguchi, A. (2003). Conjugation in the marine aloricate oligotrich Pelagostrobilidium (Ciliophora: Oligotrichia). Eur. J. Protistol. *39*, 149–160.

47. Nelsen, E.M. (1978). Transformation in Tetrahymena thermophila. Development of an inducible phenotype. Dev. Biol. *66*, 17–31.

48. Nelsen, E.M., and Debault, L.E. (1978). Transformation in Tetrahymena pyriformis: description of an inducible phenotype. J. Protozool. *25*, 113–119.

49. Cole, E.S. (2000). The Tetrahymena Conjugation Junction. Madame Curie Biosci. Database. Available at: http://www.ncbi.nlm.nih.gov/books/NBK6002/.

50. MacLennon, R.F. (1935). Dedifferentiation and redifferentiation in Ichthyophthirius. I. Arch. Protozool. *86*, 191–210.

51. Canella, M.F., and Rocchi-Canella, I. (1976). Biologie des Ophryoglenina (ciliés hyménostermes, histophages). Ann. Univ. Ferrara *NS Section III*, 1–510.

52. Janetopoulos, C., Cole, E., Smothers, J.F., Allis, C.D., and Aufderheide, K.J. (1999). The conjusome: a novel structure in Tetrahymena found only during sexual reorganization. J Cell Sci *112 (Pt 7)*, 1003–1011.

53. Uspenskaja, A.V., and Ovchinnikova, L.P. (1966). Quantitative changes of dna and rna during the life cycle of Ichthyophthirius multifiliis. Acta Protozool. Available at: https://eurekamag.com/research/023/445/023445654.php [Accessed October 28, 2016].

54. Tsukii, Y., and Hiwatashi, K. (1978). Inhibition of early events of sexual processes in Paramecium by concanavalin A. J. Exp. Zool. *205*, 439–445.

55. Luekken, W.W., Breer, H., and Hartkkemeyer, M. (1981). Local and Temporal Pattern of Con A-Binding Site Aggregation During Conjugation in Euplotes vannus (Ciliophora, Hypotrichida)1. J. Protozool. *28*, 414–417.

56. Pagliaro, L., and Wolfe, J. (1987). Concanavalin A binding induces association of possible mating-type receptors with the cytoskeleton in Tetrahymena. Exp. Cell Res. *168*, 138–152.

57. Wolfe, J., and Feng, S. (1988). Concanavalin A receptor “tipping” in Tetrahymena and its relationship to cell adhesion during conjugation. Development *102*, 699–708.

58. Frisch, A., and Loyter, A. (1977). Inhibition of conjugation in Tetrahymena pyriformis by ConA. Exp. Cell Res. *110*, 337–346.

59. Xu, D.H., Klesius, P.H., and Shoemaker, C.A. (2001). Effect of lectins on the invasion of Ichthyophthirius theront to channel catfish tissue. Dis. Aquat. Organ. *45*, 115–120.

60. Blagborough, A.M., and Sinden, R.E. (2009). Plasmodium berghei HAP2 induces strong malaria transmission-blocking immunity in vivo and in vitro. Vaccine *27*, 5187–5194.

61. Hirai, M., and Mori, T. (2010). Fertilization is a novel attacking site for the transmission blocking of malaria parasites. Acta Trop. *114*, 157–161.

62. Goodman, C.D., Siregar, J.E., Mollard, V., Vega-Rodríguez, J., Syafruddin, D., Matsuoka, H., Matsuzaki, M., Toyama, T., Sturm, A., Cozijnsen, A., *et al.* (2016). Parasites resistant to the antimalarial atovaquone fail to transmit by mosquitoes. Science *352*, 349–353.

63. Ojo, K.K., Pfander, C., Mueller, N.R., Burstroem, C., Larson, E.T., Bryan, C.M., Fox, A.M.W., Reid, M.C., Johnson, S.M., Murphy, R.C., *et al.* (2012). Transmission of malaria to mosquitoes blocked by bumped kinase inhibitors. J. Clin. Invest. *122*, 2301–2305.

64. Keyloun, K.R., Reid, M.C., Choi, R., Song, Y., Fox, A.M.W., Hillesland, H.K., Zhang, Z., Vidadala, R., Merritt, E.A., Lau, A.O.T., *et al.* (2014). The gatekeeper residue and beyond: homologous calcium-dependent protein kinases as drug development targets for veterinarian Apicomplexa parasites. Parasitology *141*, 1499–1509.

65. Bruns, Peter J, and Brussard, Trudy B. (1974). Positive selection for mating with functional heterokaryons in Tetrahymena pyriformis. Genetics *78*, 831–841.

66. Gaertig, J., Thatcher, T.H., Gu, L., and Gorovsky, M.A. (1994). Electroporation-mediated replacement of a positively and negatively selectable beta-tubulin gene in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. *91*, 4549–4553.

67. Cassidy-Hanley, D., Bowen, J., Lee, J.H., Cole, E., VerPlank, L.A., Gaertig, J., Gorovsky, M.A., and Bruns, P.J. (1997). Germline and somatic transformation of mating Tetrahymena thermophila by particle bombardment. Genetics *146*, 135–147.

Chapter six

Sex without cell fusion: a case for autogamy in the ciliate *Tetrahymena thermophila*

Abstract

The ability of individual cells to form new macronuclei from meiotic germline micronuclei in the absence of sexual cell fusion - referred to as autogamy – has been well described in various ciliate species. Nevertheless, despite years of study, the phenomenon has yet to be reported in *Tetrahymena thermophila.* A recent attempt in our laboratory to generate macronuclear transformants via introduction of foreign genes into the *MTT5* gene locus suggest that, under certain conditions, autogamy may occur in *T. thermophila* as well. We report these short observational findings both to highlight the unexpected (and often unpublished) variation in transformation efficiency of biolistic-mediated homologous recombination into different *Tetrahymena* macronuclear loci, as well as to encourage further hypothesis-based testing in model ciliate species for identification of the environmental and genetic factors which influence the developmental decision between autogamy and sexual fertilization.

Introduction

Parthenogenesis is a form of asexual reproduction that does not require cellular fusion between heterologous gamete cells, instead orchestrating development of a new organism from only one parent's germline genome. A spectrum of single and multi-cellular organisms obligately or facultatively reproduce in this way, but the factors which allow for the adoption of this reproductive pathway are not well understood[1].

Autogamy and cytogamy are two forms of parthenogenesis that are common in single-celled ciliated protozoans[2,3]. Broadly, both terms refer to the development of a new transcriptionally-active somatic macronucleus from the silent meiotic germline micronucleus of a single parental cell, however the cellular circumstances surrounding these two processes are quite different. Cytogamy is often seen as a conjugal mistake during ciliate sex, where cells of opposing mating types pair and undergo meiosis, but then a failure of pronuclear exchange leads to the zygotic product of only one parental cell's pronuclei being used as a template for development[4]. In autogamy, however, the parent cell initiates meiosis and macronuclear development independent of the presence of a mating partner, as an adaptive reaction to certain intrinsic and environmental stimuli[5–8]. But despite the uncertainty around how this is accomplished, there are clear reasons why facultative autogamy would be beneficial to ciliates. Autogamy, like parthenogenesis or fertilization, is a process allowing for genome self-renewal through corrective meiotic cross-over events (during gametogenesis) which can ultimately increase the chances of advantageous adaptation to changing environments.

Normal ciliate sexual conjugation involves the transient pairing and exchange genetic material between mating cells, which then develop into their own progeny. Because of this fact, it is clear that both cells involved have all the molecular machinery necessary to make a new organism. What is unclear is how some ciliate species have individual control over the initiation of development, while others do not. In ciliates with obligate sexuality, there is likely a key intercellular "communication" event during mating that incites the developmental cascade,

whereas in autogamous ciliates, this signal must be controlled by a single parent. Being semelparous in nature, individual ciliates are only afforded one reproductive episode per somatic lifetime. If an obligately sexual cell is starving but without a suitable mating partner, they will lose their chance to reproduce and die. In situations like this, incest is preferable to extinction, making the ability to self-initiate development particularly advantageous for continued individual survival under circumstances where a mate might be difficult to find[1,5]. Parthenogenetic development holds another benefit in ciliates, as the transcriptionally silent germline nucleus might house a treasure-trove of accumulated beneficial mutations allowing for the potential future success of daughter cells[9].

Yet, despite the benefits of facultative autogamy, it's frequency varies widely among different ciliate species[5]. As a phylum, ciliates contain both obligately sexual, facultatively autogamous, and obligately asexual species. Autogamy is known to occur naturally and with regularity in ciliates of the *Paramecium aurelia* complex[7], especially with increased clonal age or low physical interactions between mating types during starvation[5,7,10]. This has been an especially useful genetic tool for *Paramecium* researchers to create whole genome homozygote strains making it possible to perform mass-screening for recessive mutations[2,6]. Other ciliates, like *Tetrahymena thermophila*, however, are obligately sexual and autogamy has never been observed. This makes the generation of whole genome homozygote cell lines much more difficult, involving either conjugal tricks employing mating partners without functional micronuclei (star strains) or hyperosmotic shock to increase the frequency of cytogamy[2,4,11–13]. A valuable tool for researchers using *Tetrahymena thermophila* might be a protocol that allows for the generation whole-genome homozygote strains without such complicated cellular manipulations.

In this study, we accidently observed what appear to be chance occurrences of autogamy in the model ciliated protist *Tetrahymena thermophila* while attempting to introduce gene products into what we discovered to be a rather intractable macronuclear locus.

Specifically, we used a gene construct carrying a cycloheximide (Cy) resistance cassette in an effort to express mutant forms of the presumptive gamete fusogen, HAP2, in a *HAP2* deletion CU427 cell line. These cells were mating type VI and carried a paromomycin (neo) resistance marker at the deleted *HAP2* locus of their macronuclei. After biolistic bombardment and selection in increasing concentrations of Cy, individual drug resistant clones were isolated. Surprisingly, PCR results revealed that these cell lines had not undergone the typical genetic transformation. Instead, the Cy-resistant clones isolated were wild type at the *MTT5* locus (no incorporation of the transgene at the target locus) and had gained sensitivity to paromomycin (lost their parental resistance marker). Additionally, while parental CU427 cell line was uniformly mating type VI, the majority of "pseudotransformants" had switched their mating type. Because CU427 is a heterokaryon strain which contains a Cy resistance marker in its germline micronucleus, the most likely explanation for these findings is that these "pseudotransformants" were actually the products of autogamy, and had gone through a round of new macronuclear development at some point during the transformation process. Routine microscopic observations before biolistic bombardment revealed no evidence for pairing (that is, "selfing," or other strain contamination). Because cadmium was added to the medium to "loosen up" the *MTT5* locus during introduction of the selection cassette, we surmise that autogamy may have occurred in limited numbers of cells in response to heavy-metal stress.

From a practical standpoint, these results highlight the unexpected (and often unpublished) variation in transformation efficiency of biolistic-mediated homologous recombination into different *Tetrahymena* macronuclear loci, and open the possibility that additional tools could be developed to facilitate in the generation of whole-genome homozygote strains in *Tetrahymena thermophila.* From a more basic perspective, these findings display the potential utility of single-celled parthenogens in answering basic research questions regarding the environmental and genetic factors controlling the cellular choice between parthenogenesis and sexual fertilization in the initiation of development.

Results

Attempted transformation of the *MTT5* locus

The original goal of this project was to introduce altered versions of the gamete fusion protein HAP2's coding sequence into the macronuclear *MTT5* locus for stable, cadmium-inducible over-expression during conjugation. We considered this to be a fairly straightforward endeavor for several reasons including that another metallothionein locus (*MTT1*) is routinely used as a destination locus for gene introduction[14–16], the *MTT5* promoter is commonly used to initiate high levels of transcription for heterologous gene expression[14,17], and constructs with the *MTT5* promoter and termination flanking regions already exist and are used frequently in our lab as shuttle vectors into larger plasmids. The MTT5 vector used in this study (Figure 6.1A) was designed to replace the *MTT5* gene's open reading frame with that of a gene sequence encoding a *HAP2* cDNA isoform, an *MTT1* termination region, and a cycloheximide resistance cassette (Cy-r) (for identification of transformant cell lines). The target cell line used for biolistic transformation was a predominantly *ΔHAP2*-427 cell line created in previous work (see Chapter 2 and 3) in which a paromomycin resistance cassette completely replaced the HAP2 coding sequence at the endogenous macronuclear locus[18]. This heterokaryon cell line had a different genetic identity in its transcriptionally expressed Mac than in its transcriptionally silent Mic. Phenotypically, it was mating type VI, paromomycin resistance and cycloheximide sensitive(Mac), but was homozygous for cycloheximide resistance in its Mic (a marker which had been useful for the previous genetic analyses done in Chapter 2). Due to the relative lack of good alternative selectable markers in *Tetrahymena*[19], and the fact we were only interested in somatic modifications to the Mac in the proposed study, we aimed to simply use the cycloheximide resistance cassette to introduce modified versions of *HAP2* (Figure 6.1A) into the phenotypically cycloheximide-sensitive Mac genome, then isolate resistant transformants.

To reduce any possible heterochromicity that could impede homologous recombination at the *MTT5* locus it is normal to add 1µg/mL $CdCl_2$ to the recovery growth medium after biolistic

transformation. After a couple of initial unsuccessful transformation attempts, we also began adding 0.1µg/mL $CdCl_2$ to the starvation medium before transformation as well. As is done routinely before transformation of *Tetrahymena thermophila*, the starving cells were examined microscopically (Figure 6.1C) and found to be slender, single cells, with one Mic and one Mac - although there were rather unusual but rare instances of what appeared to be two Mac, one Mic cells (Figure 6.1D).

Great difficulty was experienced when attempting to transform these cells at the *MTT5* Mac locus (Figure 6.1E). From 55 individual transformation attempts a total of only 156 cycloheximide-resistant cells were isolated, of which, 39 clones were saved for further experimentation. This was a much lower frequency than expected for normal biolistic transformation in *Tetrahymena thermophila*[20], and there was also much variability observed in the number of transformants isolated in each shot, with four shots on 2/19/2015 yielding 135 of the total Cy-resistant cells, whereas other successful shots only had Cy-resistant clones obtained in the single digits. Furthermore, the Cy-resistant clones recovered exhibited abnormal conjugal behavior just after shooting, including extremely low-to-no pairing in preliminary mating experiments. It did not appear that the low efficiency of "transformation" was in anyway due to the shooting conditions, as other transformations were being successfully performed at this time in the lab into other Mac loci with different transformation vectors. Low transformation success also did not appear to be an effect of the HAP2 gene in the vector or the target cell line used, as attempted shots of two other different gene products (the *I. multifiliis* I-antigen and *Brambleberry*) into the *MTT5* locus of different target cell lines using the same *MTT5* flanking regions gave similar low-to-no "transformant"-yielding results (data not shown).

Macronuclear identity of the pseudotransformant clones

To ensure that the few Cy-resistant clones we isolated were true transformants which had incorporated the *HAP2* cDNA at their MTT5 locus, we extracted genomic DNA from 7 of the 39 saved "transformants" as well as four control strains (*ΔHAP2*-CU428, *ΔHAP2*-CU427, WT

CU428, and *HAP2* cDNA Rescue) which all contained the normal endogenous *MTT5* locus and performed PCR using primers flanking the *MTT5* locus (Figure 6.2A). Using these primers, the *MTT5* locus in wild-type cells yields a 914bp PCR product, which was noticeably unchanged in all 7 Cy-resistant clones tested, indicating a lack of transformation at this locus. Primers were also used to detect the entire open reading frame of *HAP2*, in case homologous recombination had mistakenly occurred at a different locus (Figure 6.2B). If recombination had not occurred at all we expected there to be no product at all as the target cells we had used for shooting were deleted for HAP2 in their Mac. Surprisingly, there was a PCR product for *HAP2* in these Cy-resistant cells, but only at the size of the genomic 2.8kb version of *HAP2* (note the HAP2 cDNA Rescue control of a shorter size, 2.2kb). These results confirmed that these Cy-resistant clones were not the true products of successful biolistic transformation as they lacked the incorporation of HAP2 cDNA at the mtt5 locus. Furthermore, Cy-resistance phenotype in the absence of successful transformation in these heterokaryon cells was indicative of obtaining expression of both Cy-resistance and a genomic *HAP2* from their genetically-distinct Mic through a new round of macronuclear development. Therefore, the most plausible initial explanation for these results was fertilization-induced Micronuclear development, whereby a new expressed macronucleus is generated from the Mic. However, since there was not a complementary mating type present in these transformation experiments, and *T. thermophila* has an obligate sexual life cycle, where the presence of autogamy has never been reported, this explanation seemed unlikely.

To further explore this possibility, we looked at other phenotypic macronuclear markers of these cells to see if they were consistent with this hypothesis. As the target cell line shot was originally mating type VI, and paromomycin-resistant, if these pseudotransformants were the results of a new-round of Mic development they would have lost both markers, instead adopting the expression of only the traits present in the target cell's germinal Micronucleus. After fertilization in *Tetrahymena thermophila,* adolescent cells do not inherit, but rather randomly select their mating type in a period of clonal immaturity where they are incapable of forming

pairs and mating[21,22]. Specifically, since the development of this new mating type in *T. thermophila* requires the formation of a new somatic MAC, the expression of a different mating type in the pseudotransformants would be particularly indicative of a new round of macronuclear development.

Several weeks after isolation of the pseudotransformants we performed mating-type tests on 32 out of 39 of the saved clones (Figure 6.2C-F). Mating type tests are based on the ability of any nutritionally-starved *T. thermophila* clone to form pairs and mate with any of the other six mating types but not its own, allowing for identification of a clone's mating type by its proclivity to pair with the other mating types[5]. To do these tests *en masse* we first starved cells of a mating type panel (mating types I –VI), and pseudotransformant clones to be tested in separate plates (Figure C,D), then mixed the contents of these plates together for mating in a separate 96 well plate (Figure 6.1E) so that every pseudotransformant clone was tested in a mating with each of the seven mating types. Mating-pair formation was monitored three hours after mixing cells together for these experimental clones as well as control cell lines of known mating type (HAResc427cl7 [MT=VI] and ΔHAP2CU428cl5[MT=VII]), and the target cells shot (ΔHAP2CU427cl6 [MT=VI]). We did find that the pseudotransformant clones tested expressed all different mating types rather than the mating type VI of the cell line shot, with a tendency towards MT IV, which is more commonly selected under the temperature-controlled settings typical of laboratory settings (personal communication, Dr. Donna Cassidy-Hanley). Additionally, all clones tested showed a single mating type, a somewhat unusual outcome for a typical *Tetrahymena* conjugation event, which generally yields four karyonid cells after exconjugant cell divisions that can all independently adopt a different mating type[21]. Likewise, when 12 of these pseudotransformant clones were tested for their growth and resistance to the drug paromomycin (Figure 6.2G-H) they died, while control strains and the target cells shot (which are known to be resistant) thrived in the presence of the drug. Together with the genetic identity of *MTT5* and *HAP2* locus in the Mac in these pseudotransformants, the phenotypic

mating type and drug-resistance changes in these cell lines were all consistent with a fertilization-induced round of macronuclear development.

Micronuclear identity of the pseudotransformant clones

While it was clear from the macronuclear phenotype that the pseudotransformant clones obtained from attempted transformation of the *MTT5* locus were the result of a new Mac being generated from the Mic of this cell line, it was not certain what circumstances had led to this unexpected development. We predicted three distinct scenarios that could have led to Mac development in this case, and conceived a series of genetic crosses that might enable us to discern the parentage of these unusual clones (Figure 6.3).

It was possible that Mac development could have been initiated by either one, autogamy, two, "selfing" in which *Tetrahymena* of similar mating types pair, or three, low level contamination with cells of a different mating type / genetic background. Selfing seemed unlikely, as these strains did not exhibit that tendency in any of the previous mating type assays described, plus no previous reported studies could be found correlating higher incidences of cytogamy with selfing in *T. thermophila*[23,24]. However, although careful microscopic examination of starving cells before transformation did not lead to any observation of cellular pairing, the extremely low occurrences of these pseudotransformants meant that low level contamination or "selfing" could have been missed due to their low frequency.

To rule out contamination as a possible cause of these events, it was necessary to determine the Mic genotype of the pseudotransformant clones. Since the new Mac and Mic of progeny cells is the zygotic product of the two *Tetrahymena* parent cell's Mics, we predicted two possible genetic identities of the pseudotransformant strain's Mics. One, was that they were homozygous for Cy-resistance (Cy-r/Cy-r, Figure 6.3A) and were the products of either autogamy, or maybe selfing (i.e. nuclear fusion of the Mic products from a single parent). The second possibility was that they were heterozygous for Cy-resistance (Cy-r/Cy-s, Figure 6.3B) and had been the unobserved result of a fertilization event with a contaminating cell line.

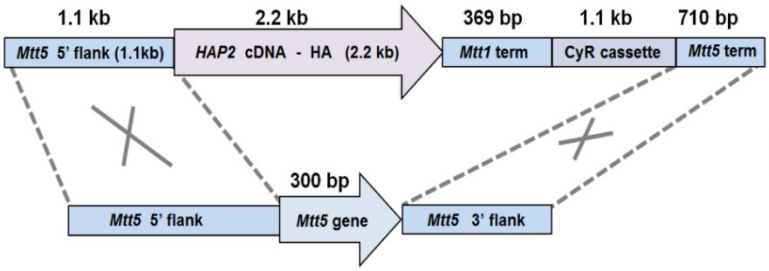
Because the diploid germline Micronucleus of the *Tetrahymena* is transcriptionally silent, the only way to find out whether these pseudotransformant strain's Micronuclei were heterozygous or homozygous for cycloheximide resistance was through a genetic analysis of their progeny cells (when their Mic would be brought into expression in the form of a new transcriptionally-active Mac). Since the Mac is polyploid and divides amitotically, any genetic resistance in the Mac, whether originally from a heterozygous or homozygous Micronucleus is expressed.

Therefore, depending on the two predicted Mic genotypes, we then expected two possible different frequencies of Cy-resistance in progeny generated by any given pseudotransformant outcross, either 100% Cy-r or 50% Cy-r (Figure 6.3, bottom). If the pseudotransformants were the result of a fertilization event with a contaminating strain, then whether outcrossed with either the CU428.2 cell line (which contain a Mic 6-methylpurine resistance [mp-r] marker) or a B2086.2 cell line, we could then predict 50% of all true progeny cells would be Cy-r. To determine the drug resistance frequency in progeny, individual mating pairs from these crosses were isolated and their offspring were grown in 96 well plates and surveyed for both immaturity (a non-pairing phenotype which is a hallmark of true *Tetrahymena* progeny) and their drug resistance phenotypes.

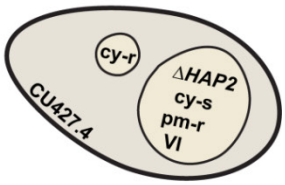
Eight pseudotransformant strains were tested in these outcrosses with the cell lines CU428.2, B2086.2, or both. In the case of B2086.2 outcrosses, cellular immaturity was used as the predominant marker of true progeny cells, whereas in CU428.2 outcrosses (which had Mic mp-r marker), 6-methylpurine and immaturity were both used as markers of true progeny. We found that in both outcrosses, all the true progeny isolated were 100% Cy-r, thus ruling out the possibility of fertilization with a contaminating strain (that didn't have Cy-r in its Mic) as the source of the pseudotransformants (Figure 6.4A-D).

Figure 6.1 Results of attempted *MTT5* locus transformation
(A) The transformation vector designed for targeting the *HAP2* cDNA into the *MTT5* locus (top) is shown with its cycloheximide resistance cassette (CyR) placed downstream of HAP2 before the *MTT5* 3' flanking region. Below, a model of the endogenous *MTT5* Macronuclear locus is shown with dashed lines and X's indicated expected regions of homologous recombination. **(B)** A model of the target cell line used. This strain was a somatically modified version of the CU427.4 (mating type VI) cell line in which the endogenous *HAP2* gene was deleted through replacement with a neomycin cassette conferring resistance to the drug paromomycin. These cells were also functional heterokaryons carrying Cy resistance (Cy-r) in their transcriptionally silent Micronuclei, but maintaining Cy sensitivity (Cy-s) in their expressed macronuclei. **(C-D)** Fluorescence and phase Microscopy showing the nuclear morphology of the starved, target cells described in (B). In both (C) and (D) DAPI nuclear fluorescence (left) show the positioning of micro- and macronuclei, where as phase images show the locations of individual cells. Note the lack of cell pairing. Red arrows in (D) point to cells containing two macronuclei. Scale bars are 20μm. **(E)** A table detailing the results of several biolistic transformation attempts into the *MTT5* locus with the transformation vector described in (A). # of "Transformants" heading indicates the number of cycloheximide resistant progeny clones resulting from shots on the dates given.

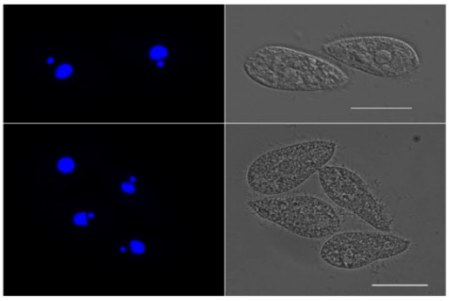
A



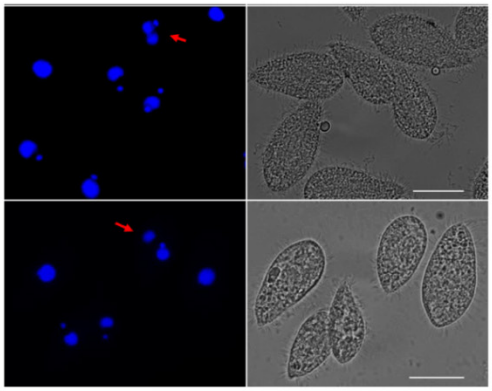
B



C



D

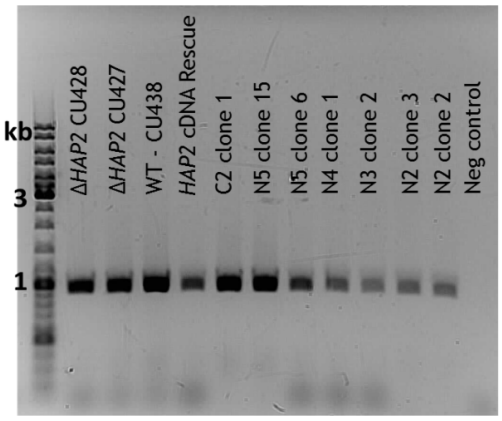


E

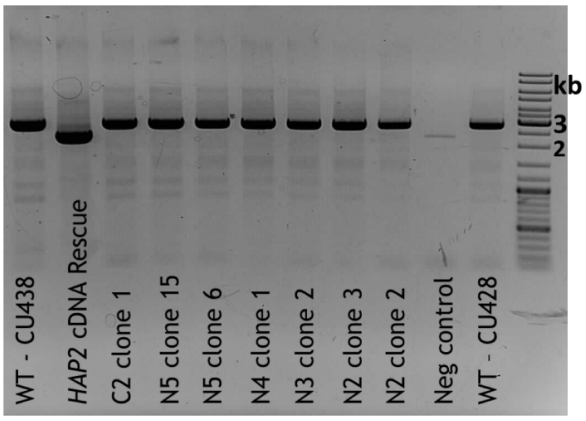
Date	# of Shots	# of "Transformants"	Average # of "Transformants" per Million Cells Shot
1/16/2015	7	0	0.00
2/4/2015	4	8	0.20
2/11/2015	4	3	0.08
2/12/2015	6	1	0.02
2/19/2015	4	135	3.38
2/20/2015	4	2	0.05
2/26/2015	4	3	0.08
3/5/2015	4	1	0.03
3/10/2015	6	0	0.00
3/18/2015	6	0	0.00
3/26/2015	4	0	0.00
Total	55	156	0.347 ± 1.0

Figure 6.2 Macronuclear identity of pseudotransformants.
(A) A DNA gel electrophoresis showing PCR results from amplification across the *MTT5* locus of genomic DNA isolated from seven of the Cy-resistant, "pseuodotransformant" clones, as designated by an alpha-numeric labels (clones were from different shot dates in Figure 6.1E). The endogenous *MTT5* gene yields a 914bp product. The expected size for the PCR product of a true transformant incorporating the *HAP2* gene at this site would be 2,882bp. **(B)** A DNA gel electrophoresis showing PCR results from amplification across the *HAP2* locus in the same seven clones as in (A). The target cells were expected to have a HAP2 deletion. Instead, a PCR product corresponding to the genomic *HAP2* (2.8kb) was present in all clones. In both (A) and (B), DNA extracted from WT (CU438.1 & CU428), *HAP2* cDNA Rescue (2.2kb), and *ΔHAP2* strains were used as controls. **(C - F)** A diagram outlining the experimental set-up and mating type test results for 32 of the "pseudotransformant" clones. Two master 96 well plates were created where in one plate (C), each of the seven *Tetrahymena thermophila* mating types was grown in a separate row and in the second plate, (D) each "pseudotransformant" clone was grown in a separate column. Once grown to log phase (~1 x 10^6 cells/mL), both plates were replicated into separate plates containing 10mM Tris starvation media (equivalent to C & D). After one day of starvation, starving clones were replicated into the same plate (E) containing starvation media to initiate the mating type test. Negative controls are highlighted in blue and purple and positive control clones of known mating type are in blue text (D). The mating type test results for each clone are wells highlighted in yellow. (F) Tabular results of the mating type tests as described in (C-E) for 32 of the "pseudotransformant" clones identified. **(G-H)** A schematic and table showing the results of cellular growth tests of 12 "pseudotransformant" clones to the drug paromomycin. (G) schematic showing that pseudotransformant clones originally growing in Cy-containing media were transferred to a separate tube containing normal growth media supplemented with paromomycin (Pm). (H) After four days of incubation, the Pm growth-tolerance results for each clone were tallied. Positive (Pm-resistant cell lines *ΔHAP2* CU428 and *HAP2* cDNA Rescue) and negative (CU427.4) control cell lines in blue text showed the expected phenotype.

A



B



C

MATING-Type Panel Plate:

	1	2	3	4	5	6	7	8	9	10	11	12
A	BLANK ---------->											
B	**(I)** SB3539 ---------->											
C	**(II)** B2086.2.14 ---------->											
D	**(III)** A*III ---------->											
E	**(IV)** CU438.1 ---------->											
F	**(V)** A*V ---------->											
G	**(VI)** B*VI ---------->											
H	**(VII)** B*VII ---------->											

D

Clones to be Tested Plate:

	1	2	3	4	5	6	7	8	9	10	11	12
A	C2 clone 1 ---------->	C2 clone 2 ---------->	C1 clone 1 ---------->	C1 clone 4 ---------->	C1 clone 5 ---------->	C1 clone 9 ---------->	C1 clone 17 ---------->	C1 clone 20 ---------->	BLANK ---------->	HAResc427 clone 7 ---->	HAP2Δ CU427 cl.6 ---->	HAP2Δ CU428 cl.5 ---->
B												
C												
D												
E												
F												
G												
H												

E

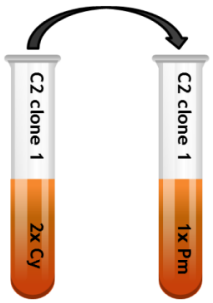
Experimental Plate:

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	M	M	M	M	M	M	M	M		M	M	M
C	M	M	M	M	M	M	M	M		M	M	M
D	M	M	M	M	M	M	M	M		M	M	M
E	IV	IV	M	IV	M	IV	M	IV		M	M	M
F	M	M	M	M	M	M	M	M		M	M	M
G	M	M	M	M	M	M	M	M		VI	VI	M
H	M	M	VII	M	VII	M	VII	M		M	M	VII

KEY

	starving alone control, clone to test
	Blank well, pipetting control
	starving alone control, mating type panel
	Mating type selected
M	Pairs observed in these wells

G



F

Mating Type Chosen	No. of clones with this mating type/ total clones tested	Percent of Clones selecting this Mating type
I	0/32	0%
II	5/32	16%
III	0/32	0%
IV	22/32	69%
V	0/32	0%
VI	1/32	3%
VII	4/32	13%

H

clone tested	Paromomycin Resistant?
C2 clone 1	no
C2 clone 2	no
C1 clone 1	no
C1 clone 5	no
C1 clone 17	no
N5 clone 15	no
N5 clone 6	no
N4 clone 1	no
N3 clone 2	no
N2 clone 2	no
N2 clone 3	no
HA clone 9	no
ΔHAP2 CU428	yes
HAP2 cDNA Rescue	yes
CU427.4	no

Other important mating characteristics of the pseudotransformant strains were also found through doing these outcrosses. Several pseudotransformant strains, especially the clones HA-9, C2-1, and C2-2, had lower than expected pair survivorship during isolation (meaning cell pairs died rather than completing or aborting conjugation) (Figure 6.4B). Of the remaining pairs that survived mating, there was also wide variability in the number of pairs leading to "back-outs", which is a *Tetrahymena*-specific mating phenomenon where parental cells abort sexual conjugation for unknown reasons, choosing instead to keep their old parental macronuclei. "Back-outs" were scored based on their continued cellular maturity (ability to pair), and for CU428.2 crosses, the lack of 6-methylpurine resistance. In a typical wild-type cross, less than 10% of pairs isolated lead to such back-out events[18], but with pseudotransformant strains, the percentage of back-outs was as high as 80% (Figure 6.4C). Crosses with the highest percentage of "back-outs" were those cell lines mated to the B2086.2 cells, as well as the HA-9 pseudotransformants. Lastly, although all viable true progeny were also Cy-resistant in these pseudotransformant outcrosses, the percentage of true progeny obtained from each individual cross was also widely variable (Figure 6.4D). These results show not only that the pseudotransformant strains likely resulted from autogamy, but that these strains also have very abnormal behavior and low fertility in subsequent mating events.

Testing the induction conditions for autogamy

Due to the variable efficiency of pseudotransformant generation by the biolistic transformations performed in Figure 6.1, it was unclear what conditions were necessary to stimulate the autogamous-like events that were occurring in these cells. It was also unsure whether low unobserved levels of selfing may be triggering cytogamy in this cell line leading to these results. Fortunately, it has been previously reported that high concentrations of Tris-HCl starvation medium, such as 60mM, can be used to block pairing and mating events in *T. thermophila*[25]. Therefore, we expected that if we starved cells in 60mM TrisHCl, we would not obtain any pseudotransformants. At the same time, we were unsure of whether the biolistic

transformation event itself was playing any role in stimulating Mic development, or if it was simply the metabolic cellular transition between growth, starvation, and then growth again that was having this effect. We also wanted to test other heterokaryon cell lines (cells that have silent drug resistance markers in their Mics but drug sensitivity in their Mac) to see if this was a cell line intrinsic problem or something more universal to other inbred lab strains of *T. thermophila*.

We performed a series of experiments in which different environmental and cellular conditions were altered to see if we could define the conditions necessary to spark this abnormal autogamy in *Tetrahymena thermophila*. We found that in the majority of conditions tested, there were very few-to-no pseudotransformants generated (Table 6.5). Those tests which did produce pseudotransformants though, did provide valuable insight.

In the first of such conditions, we got one pseudotransformant clone from starvation in 0.1µg/mL $CdCl_2$- containing 60mM Tris medium. This implied that selfing is likely not a cause of these developmental events because *Tetrahymena* cells are known not to pair in 60mM Tris. It also suggested that the cellular stress of biolistic transformation or presence of the DNA construct we were using for shooting was not important for initiation of these events. The only other pseudotransformant clones we obtained from these experiments were from a biolistic shot of CU427.4 (instead of *ΔHAP2*-427 target cells that were predominantly used in noticing this effect). Because *ΔHAP2*-427 was made originally from CU427.4, and because tests of other heterokaryon cell lines (CU438.1 and *ΔHAP2*-428) did not generate any pseudotransformants, this result suggests that the lack of *HAP2* in the original strain in which this effect was noticed is ultimately not a factor and instead there may be some intrinsic propensity of the CU427.4 cell line to experience these autogamy-like events. That being said, the other heterokaryon cell lines were not tested as rigorously as the CU427 cell line, and with the extremely low frequency of these events it is likely more tests need to be done to support this notion. The same goes

with our tests of another starvation media, Dryl's, which also did not produce any pseudotransformants although it is known to enhance the number of cytogamonts produced[2].

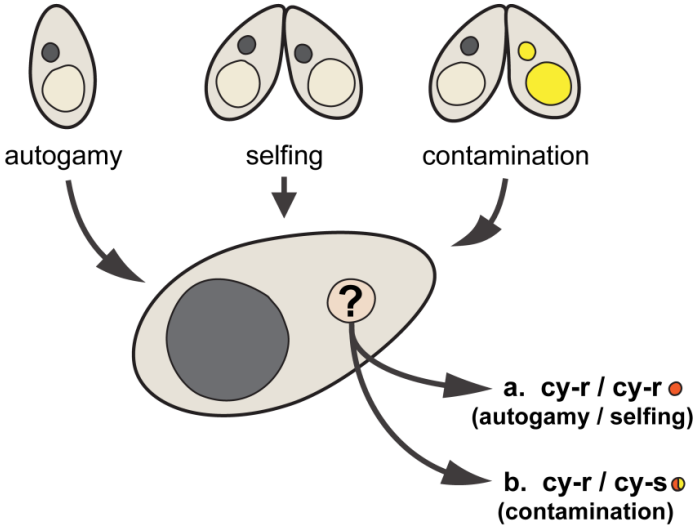
One untested factor that was different between these induction tests and the shots in Figure 6.1 that remains to be tested is the presence of high concentration of $CdCl_2$ in the growth media after starvation. In all earlier shots, cells were recovered after starvation in 1μg/mL $CdCl_2$ containing Neff, whereas in these tests cells were always transferred to Neff without $CdCl_2$ after starvation treatment. It is also possible that the potential hyperosmotic shock experienced by cells upon transfer to growth media could be important. We suspect that future tests of high $CdCl_2$ concentrations during starvation and growth, or and hyperosmotic shock to the cells could yield informative results.

Discussion

Although autogamy has yet to be described for *Tetrahymena thermophila*, in attempting to transform cells at the intractable *MTT5* locus, we may have stumbled on a set of conditions that give rise to low frequencies of autogamy in this organism. *Tetrahymena* is well-known among the ciliates for its genetically tractable macronuclear genome, allowing researchers the ability to generate stable transformants with relative ease[26]. The above experiments, however, show that not every Mac target is equally acquiescent to transformation. DNA constructs generated for the targeted replacement of *MTT5* failed to generate a single Mac transformant in this locus despite dozens of transformation attempts, while contemporaneously, the same target cells were easily transformed at a different locus (*HAP2*). Heterochromicity at the *MTT5* locus may play a role in its obstinate nature, but the addition of cadmium to the media before and after shooting did not increase transformation efficiency. Instead it led to what appeared to be autogamy, *albeit* at low frequency. Anecdotally it has been noted by other researchers that some loci are more difficult to transform than others. It may be useful for *Tetrahymena* researchers to look for a commonality among transformation resistant targets.

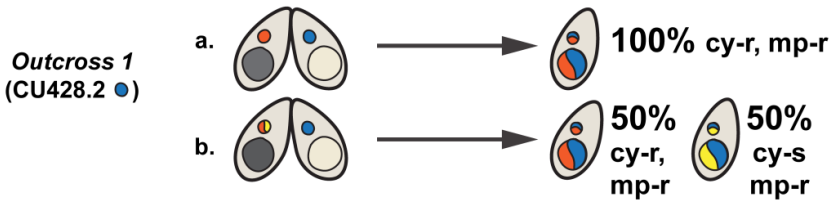
Figure 6.3 Possible micronuclear identities and proposed crosses. A cartoon outlining the potential origins of pseudotransformant strains (top) and their diploid Micronuclear identities for either the case of autogamy/selfing (a) or contamination (b). Below, proposed outcrosses to cell lines of known genetic composition and expected phenotypic results of progeny cells whether (a) or (b) was the genotype of the Mic.

Events that could have led to the pseudotransformants:



Proposed matings:

Expected progeny phenotype:



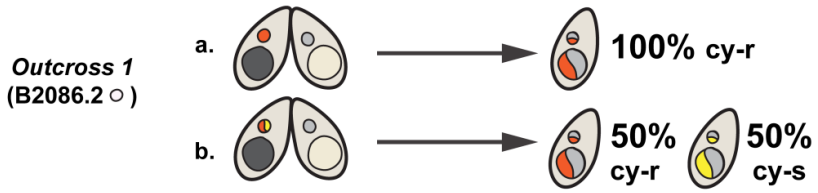
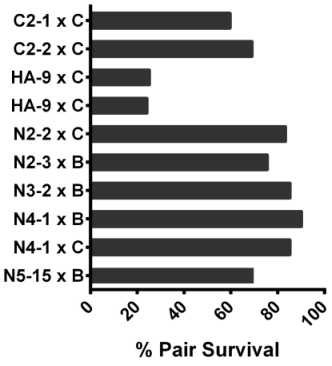


Figure 6.4 Micronuclear identity of the pseudotransformants
The results of progeny genetic analyses from pseudotransformant strains mated with either the cell lines CU428.2 (outcross 1) or B2086 (outcross 2). **(A)** Table of raw results from crosses and individual pair isolations showing the pseudotransformant clone name (far left), mating partner used in each outcross, and the relevant survival and drug resistance percentages (top) and sample sizes (parenthesis, bottom) for each cross tested. **(B-D)** Percentages shown in (A) were graphically plotted. Numbers and bars were color-coded to match in both (A) and (B-D). (B) Percent pair survival, gray. **(C)** % back-outs, green **(D)** % true progeny shown as two bars for each cross. One bar is the total percentage of true progeny as scored by mp-r (blue bars) and/or immaturity (grey bars), and orange is the percentage of Cy-r true progeny (orange bars).

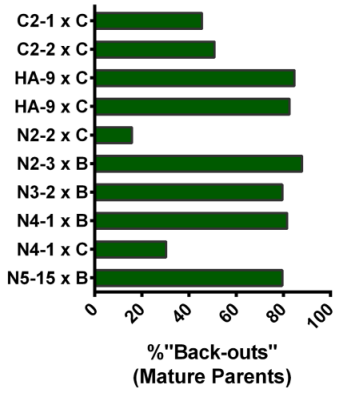
A

pseudo-transformant	mating partner	% survival	*Outcross 1* % mp-r	*Outcross 1* % mp-r & cy-r immature	*Outcross 2* % immature	*Outcross 2* % cy-r immature	% "Back-outs" mature
C2 clone 1	CU428.2	59.8 (55/92)	56.4 (31/55)	50.9 (28/55)	-	-	45.5 (25/55)
C2 clone 2	CU428.2	69.2 (63/91)	50.8 (32/63)	49.2 (31/63)	-	-	50.8 (32/63)
HA clone 9	CU428.2	24.3 (46/189)	4.3 (2/46)	0 (0/46)	-	-	84.7 (39/46)
HA clone 9 (repeat exp.)	CU428.2	25.3 (23/91)	17.3 (4/23)	17.3 (4/23)	-	-	82.6 (19/23)
N2 clone 2	CU428.2	83.5 (76/91)	81.6 (62/76)	80.3 (61/76)	-	-	15.8 (12/76)
N2 clone 3	B2086	75.7 (140/185)	-	-	12.1 (17/140)	11.4 (16/140)	87.9 (123/140)
N3 clone 2	B2086	85.4 (234/274)	-	-	20.5 (48/234)	19.23 (45/234)	79.5 (186/234)
N4 clone 1	B2086	90.4 (255/282)	-	-	18.4 (47/255)	14.5 (37/255)	81.6 (208/255)
N4 clone 1	CU428.2	85.4 (158/185)	67.1 (106/158)	64.6 (102/158)	-	-	30.3 (48/158)
N5 clone 15	B2086	69.9 (190/272)	-	-	20.5 (39/190)	17.4 (33/190)	79.5 (151/190)

B



C



D

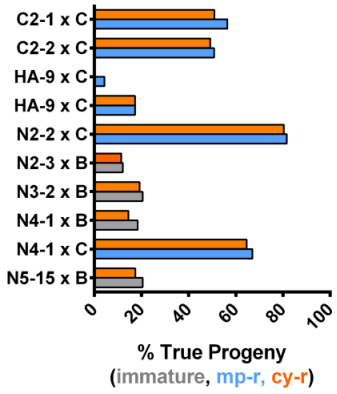


Table 6.5 Testing induction conditions for autogamy.
A table showing the culmulative tests done and conditions used for testing the experimental induction of autogamy. Target heterokaryon cell lines were starved either in the presence or absence of 0.1µg/mL $CdCl_2$ as if for biolistic transformation, and then either shot or not, and placed into fresh Neff recovery growth medium containing penicillin/streptomycin and fungizome for 24 h at 30ºC before the addition of the indicated selection drug. The cell resuspension (~2 x 10^5/mL) was then dispersed into 96 well plates at 100µl increments. Cell survival in a well is indicative of development and survival of just a single cell expressing the resistance marker when fewer than 1/3 of total wells are showing growth.

target cells used	date	shot	starvation buffer	$CdCl_2$ concentration (µg/mL)	selection drug	frequency of drug-resistant wells
ΔHAP2-427	6/09/2015	no	10mM Tris	0.1	cycloheximide	0 / 936
ΔHAP2-427	6/09/2015	no	60mM Tris	0.1	cycloheximide	1 / 1908
ΔHAP2-427	6/17/2015	no	Dryl's	0.1	cycloheximide	0 / 960
ΔHAP2-427	6/17/2015	no	Dryl's	0.0	cycloheximide	0 / 960
ΔHAP2-428	6/17/2015	no	Dryl's	0.1	6-methylpurine	0 / 952
ΔHAP2-428	8/18/2015	yes	10mM Tris	0.1	6-methylpurine	0 / 936
CU427.4	8/18/2015	yes	10mM Tris	0.1	cycloheximide	0 / 848
CU438.1	8/18/2015	yes	10mM Tris	0.1	paromomycin	0 / 456
CU438.1	8/27/2015	yes	10mM Tris	0.1	paromomycin	0 / 904
CU427.4	8/27/2015	yes	10mM Tris	0.1	cycloheximide	0 / 952
ΔHAP2-428	8/27/2015	yes	10mM Tris	0.1	6-methylpurine	0 / 456
ΔHAP2-427	9/03/2015	yes	10mM Tris	0.1	cycloheximide	0 / 952
ΔHAP2-427	9/03/2015	yes	60mM Tris	0.1	cycloheximide	0 / 960
ΔHAP2-428	9/03/2015	yes	10mM Tris	0.1	6-methylpurine	0 / 960
CU427.4	9/03/2015	yes	10mM Tris	0.1	cycloheximide	4 / 960
CU438.1	9/03/2015	yes	10mM Tris	0.1	paromomycin	0 / 960

The Stubborn MTT5 Locus. What factors might make the *MTT5* locus so recalcitrant to transformation? The difficulty in transformation of this locus is difficult to explain given the ease of modification at a routine destination of *Tetrahymena* gene introduction, the *MTT1* locus[14,15]. It is possible that our construct design was for some reason suboptimal, and to this end, testing other constructs targeting *MTT5* with larger flanking regions or with introduction of shorter gene bodies should be tried. After all, *MTT5* is the shortest of the *T. thermophila* metallothioneins[27]. However, the fact that previous studies have seemingly avoided gene introductions to the *MTT5* locus when it made little sense to (e.g. *MTT5* promoter::heterologous gene expression from the *MTT1* and *btu1* loci[14,17]), makes us question whether this effect has perhaps been seen before and just not reported.

Also, while the *MTT* promoters have been extremely useful in *Tetrahymena* for the inducible over-expression of introduced genes, some skepticism for the deletion and/or rampant overexpression of the metallothionein proteins themselves is warranted, especially when these promoters are used in functional studies. The MTT proteins themselves are thought to be multifunctional proteins[28] but the extent of their effects on *Tetrahymena* cell biology and life cycle have not been thoroughly investigated. Interestingly, the *Tetrahymena* Functional Genomics Database[29] and recent literature[27] show all five of *Tetrahymena's* metallothionein proteins' endogenous expression is upregulated upon initial nutritional starvation, then also at late conjugation stages (8-10 h) for *MTT*1, 3 and 5. It is an open question as to whether the results observed by addition of $CdCl_2$ in this study could in anyway be connected to the over-expression of *MTT* during starvation. The popular theory is that the metallothioneins function for the purpose of mitigating toxic doses of heavy metals, and while this does seem to be a property of the metallothioneins[14,30], it is also thought that historically, levels of heavy metals like Cd^{2+} were too low environmentally to have been much involved in the original evolution of *MTT* function[31]. Rather, it is likely that metal detoxification was not the initial function of *MTTs* in the cell, but that they may have originally served as a cellular reservoir for essential

micronutrients such as Zinc or Copper[31], which have been environmentally available and play critical roles in gamete development[32–34]. In line with this theory, the multifunctional metallothioneins could then have served a more homeostatic role under fluctuating intracellular concentrations of these metals, among other roles possibly even helping to finely tune the gene expression of metal-dependent transcription factors by acting as a competitive reservoir. Further work must first be done to determine the reasons behind the *MTT5* locus transformation difficulties, and to uncover what purpose, if any, *MTT* proteins may be serving in *T. thermophila* conjugal events.

How could autogamy happen? This case of autogamy is surprising because usually an initial interaction with a cell of opposite mating type is required for *Tetrahymena thermophila* Mac development. In nature, two different types of parthenogenesis have been described. One rests solely on gamete activation and post-zygotic development without meiosis to produce identical parental clones. The second type generates non-identical meiotic progeny of a single parental gamete and works to restore diploidy either through endoreduplication of chromosomes, or the pronuclear fusion (karyogamy) of homologous meiotic products[1]. In *Paramecium* autogamy[3,8], *Tetrahymena* cytogamy[4], and the facultative parthenogenetic behavior of other ciliates[35], only the meiotic-type of parthenogenesis has been observed.

Short-circuit genomic exclusion: In rare events, *Tetrahymena thermophila* employs a form of cytogamy without karyogamy referred to a “short-circuit genomic exclusion,” in which a *Tetrahymena* cross is performed with a mating partner containing a defective Mic (named a star strain), resulting in a small proportion of cells utilizing a single haploid meiotic pronuclear product for zygotic development[2]. In contrast to normal genomic exclusion events or cytogamy, the results of short-circuit genomic exclusion are usually a genetic dead-end, as it’s thought that a resulting meiotic misfire or half-hazard endoreduplication of the Mic often ends in germinal aneuploidy[2,11]. The forty-five-fold amplification of chromosome number associated with Mac development may rescue these progeny cells, allowing them to survive, but the Mic

defects incurred in this reproductive pathway give this progeny little chance of ever successfully sexually reproducing again[13].

The cells obtained from short-circuit genomic exclusion occur infrequently[13] and have much lower fertility rates[2,36]. This is similar to the phenotype of the autogamonts resulting from this study and we speculate that the Mic events resulting in autogamy may parallel those occurring in the originally described rare short-circuit genomic exclusion events of star matings (except in this case, without the necessity of an opposite mating partner cell). Chromosome preps to identify the possibly aneuploid status of Mic in these strains could be done to test this hypothesis.

Selfing: Although unlikely due to the lack of observed cell pairing, another avenue from which these strains could have been generated is from selfing. Selfing strains are those which erroneously pair intraclonally under starvation conditions. Selfing has been commonly reported in many ciliate species[24,37,38], and can be provoked by a range of genetic and environmental factors[39], including cellular immaturity and decreases in temperature(*Euplotes*[35]), but in many ciliates appears to be positively correlated with the advancing clonal age (higher number of vegetative cellular divisions) of a culture[35,40]. These changes, age-related or otherwise, are thought to be due to variations in gene expression from the Mac[35] or problems with mating type instability or inheritance[41,42].

In *T. thermophila*, it has been observed that <10% of wild clones collected exhibit selfing behavior[23,41,43]. Selfing in this ciliate is thought to be caused by a karyonidal mistake causing macronuclei mosaicism and the coincidental failure to establish expression of a single mating type during cell immaturity[21,44]. Although occasional stable selfing subclones are found expressing a single mating-type, experiments showing the induction of selfing behavior through the induced fusion of heterologous macronuclei from cells of opposite mating type would appear to support this hypothesis[24]. Since the parental cell lines CU427, and *ΔHAP2*-427 used in this study were pure for mating type VI, and an autogamont was obtained from a

60mM Tris condition (a starvation buffer shown to inhibit pair formation[25]) this argues against "selfing" as a cause for these autogamy-like events.

Moreover, earlier selfing studies report evidence of a suicidal effect where incestuous pairs die during the course of intraclonal mating events, producing no or very few viable progeny[23,45]. While this could be construed as similar to the low frequency of autogamy events observed here, due to uncertainty regarding the frequency at which cells experiencing autogamy die, it is difficult to make a direct comparison. It is known, however, that those few viable progeny isolated from intraclonal selfing strains were found more likely to be selfers themselves (~14%)[41,42]. However neither the parental cell lines used, nor the autogamy strains isolated in this study exhibited intraclonal pairing behavior under any conditions. Finally, little is known cytologically about the nuclear maneuvers of selfing *T. thermophila*, but preliminary investigations observed wide variations of phenotype depending on the strain tested; ranging from exconjugants with significant abnormalities to those that appear completely normal[23,24]. Difficulty in studying selfing strains has been partially due to their tendency towards poor stability over long term vegetative growth[23,24,42], again, an issue that was not observed here in either the parental or autogamy strains.

Passing the developmental check points: In order for autogamy to occur in *T. thermophila*, cells must be able to bypass the normal developmental checkpoints of sexual conjugation without cues from an opposite mating partner. Failure to pass these checkpoints usually causes the abortion of conjugation and the retention of parental macronuclei ("back-outs"). At least three developmental checkpoints have been defined for *Tetrahymena* conjugation[46], one for prezygotic development which consists of meiotic initiation and is induced by physical intercellular contacts and pairing between opposite mating types, the second for postzygotic development which happens around the third prezygotic division of the gametic pronuclei, and the third is for exconjugant development. Since early mating events are controlled by the parental macronucleus, any changes in gene expression control might

influence this process[47] but it is difficult to envision a change drastic enough to override all three of these checkpoints in sequence.

To this end, we at least know there was not an inherent checkpoint control defect in the of the parental cell lines CU427 and *ΔHAP2*-427 used in this study. A previous study (Chapter 2)[18] examining matings of these same cell lines observed that crosses of membrane-fusogen deficient *HAP2* cell lines still progressed through meiosis despite the lack of cytoplasmic intimacy, accumulating instead at the second checkpoint (just after the third prezygotic division and prior to pronuclear exchange), and leading to high numbers of viable “back-outs” (retained parental macronuclei) in progeny isolations.

While this is reassuring for the results of this study, peripherally, it also raises other questions regarding the specific nature of requirements for passing this second checkpoint, as *ΔHAP2* crosses were not associated with higher levels of cytogamy despite their progression past the third pre-zygotic division. Solving the genetic identity of defect responsible for the *bcd* mutants, which despite pronuclear exchange, fail in karyogamy and post-zygotic development might help illuminate this problem[46]. But as the chemical inhibition of karyogamy by anti-microtubule agents still allows for macronuclear development[48,49], it appears that neither the third pre-zygotic division or nuclear fusion are requirements for passing this checkpoint. It could be that induction of the ubiquitous bidirectional transcription[47] necessary to create the scnRNAs used in sculpting is the new macronuclear anlagen[50] may be the true hallmark of the second post-zygotic developmental checkpoint, perhaps even capable of initiating this developmental program independently of whether or not cells have passed the prior pre-zygotic checkpoints. Testing this explanation may provide some understanding for how autogamonts seem to overcome these developmental checkpoints.

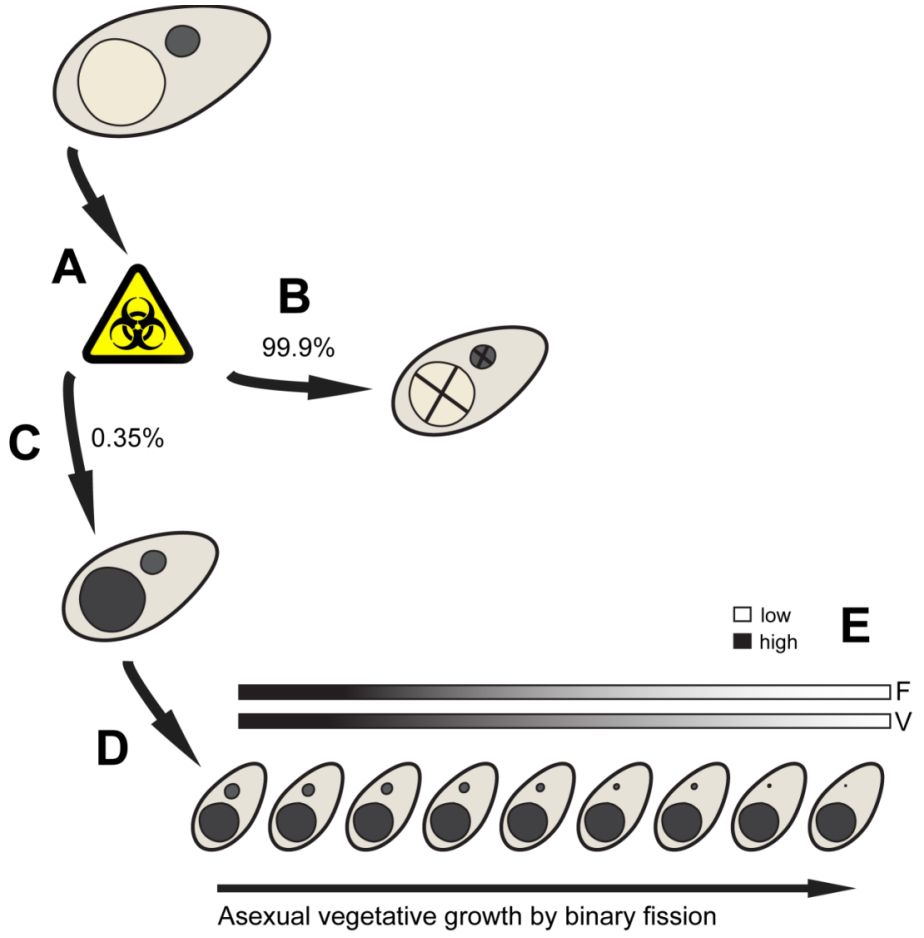
Predicted long-term consequences of autogamy. Since the autogamont strains were observed to be less adept at sexual conjugation (Figure 6.4) this may provide a means for understanding the generation of amicronucleate *Tetrahymena* strains (Figure 6.6). Originally, it

was thought that removal of *Tetrahymena thermophila's* Mic either physically[51] or through germinal senescence[44] would lead exclusively to cellular death[52], but the prevalence and success of amicronucleate *Tetrahymena* species in nature argues otherwise. Amicronucleate strains are purely asexual due to their lack of a Mic but are thought to be originally descended from sexual, micronucleated species which experienced some type of germinal senescence[52,53]. It is unclear how these strains could have initially come into existence, but Doerder in 2014 proposed a model[53] predicting that rearrangements in the mating type locus might lead to a cell's immediate abandonment of sex and eventual loss of the Mic through germinal senescence. Other models propose the lack of mating in these cells is due to some factor causing permanent cellular immaturity. Importantly, any hypothesis regarding the origin of amicronucleate *Tetrahymena sp.* must account for their high frequency in nature as well as their inability to mate[53]. Although these other two hypotheses are alluring, they both rely on defects reproducibly happening at specific Mac loci, which might not be expected to occur at high frequencies in nature. Considering our autogamy mating results (Figure 6.4) we now present an alternative, more gradual, evolutionary model (Figure 6.6) for loss of sex in amicronucleate strains.

The progeny of short circuit genomic exclusion or autogamy appear to ultimately be mistakes in *T. thermophila* sexual conjugation, which occur reproducibly, albeit at low frequencies, through conjugal pathways that haven't been observed in other genera of ciliates. While it's possible most cells might die during induction of autogamy or short-circuit genomic exclusion (Figure 6.6B), the selection on those cell lines that survive the process may be shunted towards an amicronucleate endpoint (Figure 6.6C). The already likely damaged state of Mic health would accelerate germinal aging (Figure 6.6E) in comparison with that seen normal wild-type cells, but in addition to this, the reactions of these cells to subsequent mating events (as observed in Figure 6.4 mating results) could increase selection pressure on the Mac for the eventual avoidance of mating behavior.

Three potential outcomes are possible in the subsequent mating events of autogamont cell lines: (1) Cells that can mate will replenish their potentially damaged micronuclei through either the production of fertile viable progeny, or the uptake of a partner cell's pronucleus without associated development. In both cases this will result in the reconstitution of a Mic and a future fertile *Tetrahymena* cell. (2) Cells will die as a result of mating (lower pair viability in these crosses) and be removed from the gene pool. (3) Cells will only survive conjugation by "backing out" and retaining their parental Mac. Further phenotypic assortment in the Mac followed by several subsequent rounds of mating, will eventually select for earlier "backouts" that can continue vegetatively reproducing, their descendants "avoiding" the risk or conjugation-induced death, and therefore slowly losing the ability to mate. In the end, we hypothesize the potentially faulty round of Mac/Mic development experienced by autogamonts predisposes them to poor viability and fertility outcomes in subsequent crosses, with increased selection pressure for mating avoidance behaviors which first manifest as higher numbers of "back-outs" in crosses (because the other cells in these situations either die or have their fertility restored). The pre-existing poor micronuclear health of these strains (which has absolutely no affect on their overall macronuclear health), coupled with their aforementioned abstinent tendencies and lower viability in crosses allows for a more straightforward path to germinal senescence (and gradual Mic loss) than the singular disruption of any one mating-required factor. Since sexual conjugation is a time costly process for *Tetrahymena*, amicronucleate clones could outcompete sexual clones in natural environments by creating ~256 copies of themselves by asexual division in the time it takes conjugating clones to produce 4 cells[5]. Additionally, the high ploidy level of the *Tetrahymena* Mac is thought to make these cells very resistant to the accumulation of harmful genetic mutations known as Muller's Ratchet[54], allowing asexual clones to be just as evolutionary successful as sexual clones[55]. Autogamont strains saved from this project, including the eldest of these, clone HA-9, could be used in future mating and selection experiments testing this hypothesis.

Figure 6.6 Proposed model for the results of *T. thermophila* autogamy. A proposed sequence of events representing what may occur after *T. thermophila* autogamy. **(A)** A given cell in a clonal population experiences a yet unknown, possibly heavy-metal type environmental stressor and reacts by committing a rare individual cycle of development (autogamy). **(B)** In the majority of cases, either this autogamous development or the otherwise high degree of inflicted cellular stress leads to death. **(C)** In rare circumstances, the cell survives through development of a new Mac and converts back to vegetative growth when conditions improve. The state of micronuclear health in these cells is questionable (see Figure 6.4, low progeny yields), and their likelihood to engage in subsequent successful mating events (Figure 6.4, high numbers of backouts) is poor **(D)** Over time of continued vegetative expansion, the clonal descendants of the autogamont undergo germinal senescence, losing functionality and mass in their micronuclei. **(E)** Owing to these cells reduced ability to form progeny and higher likelihood to backout from mating or death at each following conjugal event, low survivorship (S) and fertility (F) in mating results in accelerated germinal aging of the surviving cells in this lineage, eventually leading to clonal populations of amicronucleate cells.


A
B
99.9%
C
0.35%
low
high
E
F
V
D
Asexual vegetative growth by binary fission

Is Autogamy Useful? Given our current observed results, we conclude that this induction method for *T. thermophila* autogamy, at least for practical purposes, will not be useful for the generation of whole-genome homozygote lab strains because there was variable-to-low fertility in the strains generated, and the specific triggers of autogamy remain poorly-defined. In this regard, specific treatments like trypsin[40,56] or methyl-cellulose treatment[7,57] used in other ciliates to induce autogamy may be able to enhance the efficiency of *T. thermophila* autogamont generation observed here. Until then, however, other methods of genomic exclusion-induced homozygosity through uniparental cytogamy are recommended as they both lead to fertile offspring and are more reproducible and efficient[2].

Much remains to be done to simply define the developmental process occurring in these cells. It is unknown whether the effects observed in these experiments are specific to the CU427 strain background, or whether similar events can be observed in other cell lines. Also, despite similarities with short-circuit genomic exclusion progeny, we still do not definitively know whether meiosis occurs in this clones, or if the offspring are formed through karyogamy or endoreduplication of the Mic, or even what the specific environmental causes and timing are of these events. We speculate that they are somehow related to the environmental stress introduced to cells through addition of either heavy metals, starvation, osmotic shock (from media change), or the timely application of any or all of these three factors.

What is clear is that the *Tetrahymena* genus is unique in containing species adopting all different versions of eukaryotic sexuality with a high degree of long-term evolutionary success. Some species are predisposed towards sexual conjugation between genetically-distinct individuals[58], where others are either facultative[59] or obligate[9] parthenogens, and still others, lack a Mic and have dispensed with sex entirely[52,53]. Continued research in organisms, like ciliates, with such diversity and plasticity in their sexual life cycles, will aid in our future understanding of which environmental and genetic cues are the key in the ecogenetic

maneuvering between autologous and heterologous syngamy and the timely initiation of germline development.

Materials and methods

***Tetrahymena thermophila* strains**. *Tetrahymena* strains were obtained from the Tetrahymena Stock Center, Cornell University. The "pseudotransformants" or autogamonts described were obtained using the strains CU427 and Δ*HAP2*-427 (Figure 6.1B), the panel of cell lines used in mating type tests (Figure 6.1C-F) and subsequent outcrosses (Figure 6.3 and 6.4) are described further in the strain information tables of chapters 2 and 3. Both of the parental cell lines CU427 and Δ*HAP2*-427 used for autogamont generation are heterokaryons, containing genetically distinct germline and somatic nuclei. Briefly, the Δ*HAP2*-427 cell line was created by somatic macronuclear transformation via biolistic bombardment, replacing the endogenous *HAP2* gene with a neomycin selection cassette conferring resistance to paromomycin. Complete knock-outs were made by selecting for resistant clones after gradually increasing drug dosage and verified through PCR.

***Tetrahymena thermophila* culture and mating conditions.** *Tetrahymena thermophila* cells described above were grown at 30°C in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 33.3μM $FeCl_3$) on a platform shaker at ~ 100 rpm. For cell matings, log phase cells were starved at ~2 ×10^5 cells/ml for 12-18 hrs in 10 mM Tris buffer (pH 7.5), and mixed in equal concentration at 30°C. Pseudotransformant cell lines were phenotyped for their drug resistance by transfer of the indicated clones into stock tubes of NEFF medium containing 80 μg/mL of the drug paromomycin.

Procedure for attempted biolistic macronuclear transformation of the *MTT5* locus. For somatic transformation[60,61], target cells were grown to mid-to-late log stage (~6 × 10^5 – 1 × 10^6 cells/ml) in NEFF medium, and starved overnight in 10 mM Tris buffer (pH 7.5) at ~2 × 10^5 cells/ml containing 0.1 μg/mL $CdCl_2$ prior to biolistic transformation. Starvation cultures were

incubated shaking at 30°C overnight followed by 2 h of stationary incubation before either biolistic bombardment (Figure 6.2) or no treatment (Figure 6.6). The DNA construct used in initial transformation attempts (Figure 6.1A) had different *HAP2* cDNA constructs placed at the multiple cloning site (see Chapter 3), while no DNA construct was used in later biolistic bombardment tests (Table 6.5). Following transformation, (or starving as in select conditions of Table 6.5) cells were transferred to NEFF medium containing 250µg/ml Penicillin G, 250 µg/ml Streptomycin, and 1.25µg/ml Fungizone, and grown at 30°C. Cells were rested in this media for 1 day (24 h) at 30°C before addition of 25 µg/mL cycloheximide and dispersal into 96 well plates for drug selection. Wells which showed positive growth in cycloheximide after 5 days were the "pseudotransformants" described and tested in this study.

Genotyping the pseudotransformant clones. Genomic DNA was extracted with phenol:chloroform:iso-Amyl alcohol (25:24:1, VWR), PCR amplified and the resulting products analyzed by agarose gel electrophoresis. All PCR reactions were carried out with Phusion High-Fidelity *Taq* DNA Polymerase (ThermoFisher). PCR primers for amplification of the *HAP2* and *MTT5* loci respectively were:

pTievMtt5Forw 5' CTTGAATACAATCATGAGTTCACC 3';

pTievMtt5Rev 5' TACCCCTATCACCTAGCGAAG 3';

Hap2for 5'ATGAAATTTTTGGCTTTTGGATTGATTTATTTTC 3';

GSP1rev 5' TCATTCAATTAGTAGATAGAGAGGAGATGTTGA 3'

Microscopy. Samples of 0.1 µg/mL $CdCl_2$ starved cells were fixed with DAPI were imaged with a Zeiss Axio Imager M1 Microscope equipped with an AxioCamMR3 camera to capture bright field and fluorescence images of starved cells.

Mating type and cellular immaturity tests. As described and shown in Figure 6.1, pseudotransformant and mating-type panel clones were grown in the columns and rows, respectively of separate 96 well plates containing NEFF medium for two days. These plates were replicated into separate plates of 10mM Tris for one day of starvation, and then replicated

for into the same 96 well 10mM Tris plate to mix cells for mating. Four hours after initially mixing cells together plates were scored for cell pairing between the “pseudotransformants” tested and the cells of mating type I-VII of the panel. Wells in which no pairing was observed were identified as the mating type of the given pseudotransformant tested.

Genetic analysis of pseudotransformants micronuclear genotype and mating ability. As described in Figure 6.3 and 6.4, the micronuclear genotype of a given pseudotransformant cell line was tested through bringing it into somatic expression by a single round of macronuclear development (induced by sexual conjugation). This was done by outcrossing pseudotransformant clones to a different cell line of known genotype, isolating individual mating pairs into drops, and examining the frequencies of drug resistance in progeny cells.

Mass matings of the cell lines of interest were performed, followed by isolation of single persistent mating pairs by mouth-pipetting individual pairs into drops of NEFF starting at 7 h after the mixing of opposite mating types. Typically, at least 96 pairs were isolated from each mating condition and their “synclones” (mixed progeny expanded from both of the resulting exconjugants) were initially incubated in a 30ºC humidified box for 3-4 days, followed by transfer to 96 well NEFF-containing microtiter master plates to allow for expansion of progeny cells. Percent pair survival was based on the recorded growth or death of individual synclones in hanging drops. Drops that merged or dried up or were excluded from final analyses.

As described in detail in Figure 6.3 and the methods section of Chapter 2, outcrosses to the CU428.2 heterokaryon cell line result in the dominant 6-methylpurine resistance of progeny cells, but drug-sensitivity of parental “back-outs”. This is due to the homozygous presence of this silent germline mp-r drug resistance marker in the Mic of CU428.2 parents being brought into somatic expression newly-derived progeny Mac. Because the B2086 parental cells are not heterokaryons and do not carry any Mic drug-resistance alleles, progeny cells were scored on the basis of cellular immaturity (the inability progeny cells to mate and form pairs) in outcrosses to this cell line. Immaturity tests were performed in plate format, similar to the mating type tests

outlined above, but with the exception that only a mating type I strain (SB3539) was utilized as a mating partner instead of the entire mating-type panel. As neither of the two parent cells have the mating type I gene in the mating type cassette of their germline genome[21] they shouldn't be able to make mating type I progeny.

The percent of surviving progeny displaying drug-resistance phenotypes was then tallied by replicating them into 96 well microtiter plates containing NEFF media with either 15µg/ml 6-methyl purine or 25 µg/ml cycloheximide. The growth or death of individual synclone wells was scored as either drug-resistant or not after 4-5 days of incubation with the respective drugs. For CU428 outcrosses, wells were scored as double drug resistance if they consecutively grew in first cycloheximide, then 6-methylpurine – containing medium. For B2086 outcrosses, only drug resistance to cycloheximide was scored.

REFERENCES

1. Schön, I., Martens, K., and Dijk, P. eds. (2009). Lost Sex (Dordrecht: Springer Netherlands) Available at: http://link.springer.com/10.1007/978-90-481-2770-2 [Accessed November 8, 2016].

2. Cole, E.S., and Bruns, P.J. (1992). Uniparental cytogamy: a novel method for bringing micronuclear mutations of Tetrahymena into homozygous macronuclear expression with precocious sexual maturity. Genetics *132*, 1017–131.

3. Yanagi, A. (2010). Timing of commitment to autogamy in Paramecium bursaria. Jpn J Protozool Vol *43*, 95.

4. Orias, E., and Hamilton, E.P. (1979). Cytogamy: An Inducible, Alternate Pathway of Conjugation in TETRAHYMENA THERMOPHILA. Genetics *91*, 657–671.

5. Nanney, D.L. (1980). Experimental ciliatology: an introduction to genetic and developmental analysis in ciliates (New York: Wiley).

6. Tsukii, Y., and Hiwatashi, K. (1979). Artificial induction of autogamy in Paramecium caudatum. Genet. Res. *34*, 163–172.

7. Yanagi, A., and Haga, N. (1996). A simple method of induction of autogamy by methyl cellulose in Paramecium caudatum. Eur. J. Protistol. *32*, 183–186.

8. Berger, J.D. (1986). Autogamy in Paramecium cell cycle stage-specific commitment to meiosis. Exp. Cell Res. *166*, 475–485.

9. Kaczanowski, A., Brunk, C.F., and Kazubski, S.L. (2016). Cohesion of Clonal Life History, Senescence and Rejuvenation Induced by Autogamy of the Histophagous Ciliate Tetrahymena rostrata. Protist *167*, 490–510.

10. Orias, E. (1986). Ciliate Conjugation. In The Molecular Biology of Ciliated Protozoa, Gall, J.G., ed. (New York, NY: Academic Press), pp. 45–94.

11. Sanford, Y.M., and Orias, E. (1981). Phenylketonuric Tetrahymena: phenylalanine hydroxylase mutants and other tyrosine auxotrophs. Proc. Natl. Acad. Sci. U. S. A. *78*, 7614–7618.

12. Allen, S.L., File, S.K., and Koch, S.L. (1967). Genomic exclusion in tetrahymena. Genetics *55*, 823–837.

13. Bruns, P.J., Brussard, T.B., and Kavka, A.B. (1976). Isolation of homozygous mutants after induced self-fertilization in Tetrahymena. Proc. Natl. Acad. Sci. U. S. A. *73*, 3243–3247.

14. Amaro, F., Turkewitz, A.P., Martín-González, A., and Gutiérrez, J.C. (2014). Functional GFP-metallothionein fusion protein from Tetrahymena thermophila: a potential whole-cell biosensor for monitoring heavy metal pollution and a cell model to study metallothionein overproduction effects. BioMetals *27*, 195–205.

15. Shang, Y., Song, X., Bowen, J., Corstanje, R., Gao, Y., Gaertig, J., and Gorovsky, M.A. (2002). A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. *99*, 3734–3739.

16. Busch, C.J.-L., Vogt, A., and Mochizuki, K. (2010). Establishment of a Cre/loxP recombination system for N-terminal epitope tagging of genes in Tetrahymena. BMC Microbiol. *10*, 191.

17. Formigari, A., Boldrin, F., Santovito, G., Cassidy-Hanley, D., Clark, T.G., and Piccinni, E. (2010). Functional Characterization of the 5′-upstream Region of MTT5 Metallothionein Gene from Tetrahymena thermophila. Protist *161*, 71–77.

18. Cole, E.S., Cassidy-Hanley, D., Pinello, J.F., Zeng, H., Hsueh, M., Kolbin, D., Ozzello, C., Jr, T.G., Winey, M., and Clark, T.G. (2014). Function of the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. Curr. Biol. CB *24*, 2168–2173.

19. Iwamoto, M., Mori, C., Hiraoka, Y., and Haraguchi, T. (2014). Puromycin resistance gene as an effective selection marker for ciliate Tetrahymena. Gene *534*, 249–255.

20. Cassidy-Hanley, D., Bowen, J., Lee, J.H., Cole, E., VerPlank, L.A., Gaertig, J., Gorovsky, M.A., and Bruns, P.J. (1997). Germline and Somatic Transformation of Mating Tetrahymena thermophila by Particle Bombardment. Genetics *146*, 135–147.

21. Cervantes, M.D., Hamilton, E.P., Xiong, J., Lawson, M.J., Yuan, D., Hadjithomas, M., Miao, W., and Orias, E. (2013). Selecting one of several mating types through gene segment joining and deletion in Tetrahymena thermophila. PLoS Biol. *11*, e1001518.

22. Rogers, M.B., and Karrer, K.M. (1985). Adolescence in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. *82*, 436–439.

23. Simon, E.M., and Meyer, E.B. (1992). Suicide is not the inevitable outcome of “perpetual” selfing in tetrahymenines collected from natural habitats. Dev. Genet. *13*, 47–52.

24. Allen, S.L., and Nanney, D.L. (1958). An Analysis of Nuclear Differentiation in the Selfers of Tetrahymena. Am. Nat. *92*, 139–160.

25. Wellnitz, W.R., and Bruns, P.J. (1979). The Pre-Pairing Events in Tetrahymena thermophila Analysis of Blocks imposed by High Concentrations of Tris-Cl. Exp. Cell Res. *119*, 175.

26. Ruehle, M.D., Orias, E., and Pearson, C.G. (2016). Tetrahymena as a Unicellular Model Eukaryote: Genetic and Genomic Tools. Genetics *203*, 649–665.

27. de Francisco, P., Melgar, L.M., Díaz, S., Martín-González, A., and Gutiérrez, J.C. (2016). The Tetrahymena metallothionein gene family: twenty-one new cDNAs, molecular characterization, phylogenetic study and comparative analysis of the gene expression under different abiotic stressors. BMC Genomics *17*, 346.

28. Coyle, P., Philcox, J.C., Carey, L.C., and Rofe, A.M. (2002). Metallothionein: the multipurpose protein. Cell. Mol. Life Sci. CMLS *59*, 627–647.

29. Xiong, J., Lu, Y., Feng, J., Yuan, D., Tian, M., Chang, Y., Fu, C., Wang, G., Zeng, H., and Miao, W. (2013). Tetrahymena functional genomics database (TetraFGD): an integrated resource for Tetrahymena functional genomics. Database J. Biol. Databases Curation *2013*, bat008.

30. Wang, Q., Xu, J., Chai, B., Liang, A., and Wang, W. (2011). Functional comparison of metallothioneins MTT1 and MTT2 from Tetrahymena thermophila. Arch. Biochem. Biophys. *509*, 170–176.

31. Palmiter, R.D. (1998). The elusive function of metallothioneins. Proc. Natl. Acad. Sci. *95*, 8428–8430.

32. Duncan, F.E., Que, E.L., Zhang, N., Feinberg, E.C., O’Halloran, T.V., and Woodruff, T.K. (2016). The zinc spark is an inorganic signature of human egg activation. Sci. Rep. *6*, 24737.

33. Tvrda, E., Peer, R., Sikka, S.C., and Agarwal, A. (2015). Iron and copper in male reproduction: a double-edged sword. J. Assist. Reprod. Genet. *32*, 3–16.

34. Hester, J., Hanna-Rose, W., and Diaz, F. (2016). Zinc deficiency reduces fertility in C. elegans hermaphrodites and disrupts oogenesis and meiotic progression. Comp. Biochem. Physiol. Toxicol. Pharmacol. CBP.

35. Heckmann, K. (1967). Age-dependent intraclonal conjugation in Euplotes crassus. J. Exp. Zool. *165*, 269–277.

36. Orias, E., Hamilton, E.P., and Flacks, M. (1979). Osmotic shock prevents nuclear exchange and produces whole-genome homozygotes in conjugating Tetrahymena. Science *203*, 660–663.

37. Webb, T.L., and Francis, D. (1969). Mating Types in Stentor coeruleus. J. Protozool. *16*, 758–763.

38. Corliss, J.O. (1972). Tetrahymena and Some Thoughts on the Evolutionary Origin of Endoparasitism. Trans. Am. Microsc. Soc. *91*, 566–573.

39. Watanabe, T. (1990). The role of ciliary surfaces in mating in Paramecium. In Ciliary and Flagellar Membranes (Springer), pp. 149–171. Available at: http://link.springer.com/chapter/10.1007/978-1-4613-0515-6_6 [Accessed November 8, 2016].

40. Cronkite, D.L. (1974). Genetics of Chemical Induction of Conjugation in Paramecium Aurelia. Genetics *76*, 703–714.

41. Nanney, D.L., and Caughey, P.A. (1955). An Unstable Nuclear Condition in Tetrahymena Pyriformis. Genetics *40*, 388–398.

42. Simon, E.M., and Orias, E. (1987). Genetic instability in the mating type system of Tetrahymena pigmentosa. Genetics *117*, 437–449.

43. Simon, E.M., Meyer, E.B., and Preparata, R.M. (1985). New wild Tetrahymena from Southeast Asia, China, and North America, including T. malaccensis, T. asiatica, T. nanneyi, T. caudata, and T. silvana n. spp. J. Protozool. *32*, 183–189.

44. Simon, E.M., and Nanney, D.L. (1979). Germinal aging in Tetrahymena thermophila. Mech. Ageing Dev. *11*, 253–268.

45. Nanney, D.L. (1953). Nucleo-cytoplasmic interaction during conjugation in tetrahymena. Biol. Bull. *105*, 133–148.

46. Cole, E.S., and Soelter, T.A. (1997). A Mutational Analysis of Conjugation inTetrahymena thermophila. Dev. Biol. *189*, 233–245.

47. Ward, J.G., and Herrick, G. (1996). Effects of the transcription inhibitor actinomycin D on postzygotic development of Tetrahymena thermophila conjugants. Dev. Biol. *173*, 174–184.

48. Hamilton, E.P., Suhr-Jessen, P.B., and Orias, E. (1988). Pronuclear fusion failure: an alternate conjugational pathway in Tetrahymena thermophila, induced by vinblastine. Genetics *118*, 627–636.

49. Kaczanowski, A., Ramel, M., Kaczanowska, J., and Wheatley, D. (1991). Macronuclear differentiation in conjugating pairs of Tetrahymena treated with the antitubulin drug nocodazole. Exp. Cell Res. *195*, 330–337.

50. Chalker, D.L., Meyer, E., and Mochizuki, K. (2013). Epigenetics of Ciliates. Cold Spring Harb. Perspect. Biol. *5*, a017764–a017764.

51. Nanney, D.L. (1957). Inbreeding Degeneration in Tetrahymena. Genetics *42*, 137–146.

52. Kaney, A.R., and Speare, V.J. (1983). An amicronucleate mutant of Tetrahymena thermophila. Exp. Cell Res. *143*, 461–467.

53. Doerder, F.P. (2014). Abandoning sex: multiple origins of asexuality in the ciliate Tetrahymena. BMC Evol. Biol. *14*, 112.

54. Muller, H.J. (1964). THE RELATION OF RECOMBINATION TO MUTATIONAL ADVANCE. Mutat. Res. *106*, 2–9.

55. Zufall, R.A. (2016). Mating Systems and Reproductive Strategies in Tetrahymena. In Biocommunication of Ciliates, G. Witzany and M. Nowacki, eds. (Springer International Publishing), pp. 195–220. Available at: http://link.springer.com/chapter/10.1007/978-3-319-32211-7_12 [Accessed September 12, 2016].

56. Mikami, K., and Koizumi, S. (1979). Induction of autogamy by treatment with trypsin in Paramecium caudatum. J. Cell Sci. *35*, 177–184.

57. Miyake, A. (1968). Induction of conjugation by chemical agents in Paramecium. J. Exp. Zool. *167*, 359–379.

58. Cole, E.S. (2006). The Tetrahymena Conjugation Junction. In Cell-Cell Channels, F. Baluska, D. Volkmann, and P. W. Barlow, eds. (New York, NY: Springer New York), pp. 39–62. Available at: http://dx.doi.org/10.1007/978-0-387-46957-7_3.

59. Lynn, D.H., and Doerder, F.P. (2012). The Life and Times of Tetrahymena. In Methods in Cell Biology (Elsevier), pp. 9–27. Available at: http://linkinghub.elsevier.com/retrieve/pii/B9780123859679000025 [Accessed April 15, 2016].

60. Bruns, P.J., and Cassidy-Hanley, D. (2000). Biolistic transformation of macro- and micronuclei. Methods Cell Biol. *62*, 501–512.

61. Seashell Technology - Carrier Particle Protocols for Plasmid DNA Available at: http://www.seashelltech.com/protocols.shtml [Accessed October 24, 2016].

Chapter seven

Summary, future directions, and evolutionary implications

A. Summary

In the work presented in this dissertation we used the fresh-water ciliated protist *Tetrahymena thermophila,* a model organism with an easily inducible sexual cycle, to study the function of a conserved gamete membrane protein in cell-cell fusion during sexual fertilization. In *T. thermophila*, we found that HAP2/GCS1 localized to the site of membrane fusion (the conjugation junction), and unlike other species where function is limited to only one gamete, was instead expressed during conjugation in all seven mating types. It was possible for HAP2 to induce low levels of fusion when expressed from only one cell of a mating pair, but function was necessary on both partners for efficient membrane fusion and fertilization to take place. Next, we made the surprisingly finding that this gamete membrane fusogen adopts a predicted structure and function similar to that of class II viral fusogens. In these studies we were able to identify and functionally dissect the *T. thermophila* HAP2 fusion loop, testing both it's necessity in *Tetrahymena* cell-cell fusion using a newly developed flow-cytometry based sexual cell fusion assay, and sufficiency to interact with model membranes and induce lipid mixing using biophysical methods.

During the course of this work we examined another conjugation-specific *Tetrahymena* protein, ZFR1, which although showing localization patterns and fertility defects similar to HAP2, was not involved in either HAP2 trafficking to the conjugation junction, or *Tetrahymena* membrane fusion. Rather, preliminary data suggests ZFR1 helps to reshape the cellular environment during progeny development. We also studied *HAP2* expression as a potential marker for sexual reproduction in the obligate parasitic ciliate of fish, *Ichthyophthirius multifiliis,* where the infective free-swimming theront stage was found to have the highest expression of HAP2 transcripts. Finally while cell-cell fusion is a primary characteristic of eukaryotic sex, parthenogenetic eukaryotes have lost this ability. It is thought that *T. thermophila* is an obligate sexually reproducing species, but through an unexpected experimental result, we uncovered evidence that a type of parthenogenesis known as 'autogamy' may occur in this ciliate.

B. Future Directions

B1. Steps to defining HAP2's fusogenic sufficiency and trigger

Since it is now known that HAP2 is structurally related to CII viral fusogens and has a functional fusion loop in *Tetrahymena, Chlamydomonas,* and *Arabidopsis,* it is reasonable to apply techniques commonly used in the study of viral fusogens to probe for HAP2's trigger and mechanism of action. In these studies, HAP2's fusogenic sufficiency in cell-cell and virus-cell fusion should be tested and HAP2's potential interaction partners should be identified. Because recognition and adhesion are the first steps in allowing two membranes to come close enough together for fusion to occur, and since HAP2 is only involved in fusion and separate species-specific proteins mediate membrane adhesion, appropriate intercellular membrane recognition and adhesion are likely the key rate-limiting steps to fusion in HAP2 fusogenic sufficiency assays using heterologous cells. Taking this into consideration, there are then two directions that can be taken with these types of future experiments. One, the species-specific proteins important for gamete recognition and adhesion can be identified and reconstituted along with HAP2 into heterologous fusion assays, or two, membranes can be tethered together by known adhesion proteins that are not species-specific. In the latter method, it is known that with one class I viral fusion protein, namely hemagglutinin (HA), that substituting the normal sialic acid adhesion with a streptavidin-biotin-mediated membrane adhesion still allows for HA-initiated fusion[1]. Other additional factors, such as actin polymerization foci near the sites of membrane fusion could also promote adhesion as well, and may even provide a supportive force for fusion to occur[2].

The literature regarding the adhesion and triggering of class II (CII) viral fusogens will likely also give us insight into HAP2-mediated gamete membrane fusion that will be useful in designing fusogenic sufficiency assays. While many viral fusion proteins wrap adhesion and fusion functions into the same protein, some CII fusogens from alphaviruses encode a separate fusion and adhesion protein[3]. With the Dengue viral E glycoprotein, however, adhesion and

fusion functions are both accomplished by the same envelope protein which uses a myriad of different proteins including DC-Sign, heparan-sulfated glycoproteins, and HSP70 [4] for adhesion to host cells. It is not known whether a cellular receptor on the opposite gamete membrane exists for HAP2, but data from both CII viral and developmental fusogens suggests there could actually be a receptor on target female gametes, even if this interaction no longer functions in cell-cell adhesion. Other potential functions of a putative female gamete receptor could be in the proper orientation or oligomerization of HAP2 monomers before or during fusion.

We also gain understanding regarding the possible triggering mechanisms of HAP2 from information already known about class II viral fusogens. A fusion trigger is an environmental factor that stimulates either the timely membrane insertion or conformational changes of a fusogen to allow membrane merger. As reviewed in Chapter 1, it is known that there are different triggers necessary for different class II viral fusogens, including low pH, enzymatic cleavage of chaperone proteins, calcium[5], and lipid content[6] which could be systematically tested for their capacity to promote HAP2-mediated fusion. A mild hypotonic shock in some systems has also been shown to increase the number of cell-cell fusion events[7]. It is thought that this type of shock increases the lateral membrane tension applied to both hemifused and non-expanding fusion pores, causing an increase in the number of full fusion events[8]. In *Tetrahymena,* sexual cell fusion events often occur in hypotonic media (10mM Tris in the lab or water in the wild). Furthermore, a brief hyperosmotic shock after meiosis has been shown to block pronuclear exchange in *Tetrahymena* (possibly through blocking membrane pore formation or expansion), creating whole genome homozygotes [9]. For other HAP2-mediated cell-cell fusion events, it may be possible that specific ion channels have evolved along with this ancient cellular fusion system to temporarily decrease the osmolality of the extracellular microenvironment proximal to cellular fusion pores, permitting their eventual expansion. While this might be intuitively possible, (as the cell's themselves would control their own fusion "trigger" event) no specific tests of this hypothesis have been attempted.

Finally, in defining HAP2's sufficiency and trigger it may also be interesting to test the potential sufficiency of cognate viral fusion proteins in mediating a developmental cell-cell fusion event. We already know that developmental fusogens (EFF-1 / AFF-1) are capable of rescuing viral envelope fusion to host cells[10], but we still don't know if the reverse is possible – a viral envelope protein functionally substituting in a developmental cell-cell fusion event. One could envision experiments replacing *HAP2* in simple model organisms (like *Tetrahymena or Chlamydomonas)* with a full length CII viral fusogen to see if it is capable of rescuing sexual cell fusion events, or even expressing HAP2 on pseudotyped viruses and applying these particles to activated gametes lacking the HAP2 gene in order to see if "viral" infection alone can stimulate the fusion of already adhered mating partners.

While these are interesting future questions, at least in initial tests of pseudotyped viruses expressing either HAP2 or VSV-G on their surface, fusion was not enhanced over *Δenv* particle controls in sexual cell fusion assays using *Tetrahymena* (preliminary data not shown). Instead, oddly, it appeared that the mere presence of the concentrated pseudotyped particles themselves (either decorated with a fusion protein or not), caused a slight increase in the total amount of *Tetrahymena* cells capable of sexual cell fusion compared to matings treated with the same volume of the Tris buffer not containing pseudotyped particles. The relevance of this slight increase is not yet understood, but it could be that minor changes to the lipid or cholesterol content of the media might increase the overall ability of these cells to fuse during sex.

B2. HAP2's fusogenic mechanism: cis or trans?

Another key next step in this work will be uncovering the precise mechanisms of HAP2-mediated membrane fusion. As discussed in Chapter 1 and Figure 1.4, there are two alternate mechanisms proposed as to how HAP2 may function in its merger of cellular membranes during sex. The big difference between these two models is that certain developmental fusogens (like EFF-1) do not have a fusion loop, and are necessary on both cell membranes for fusion to occur.

Thus, in this model, it is thought that interactions in trans between monomers on opposite membranes results in the formation of the pre-fusion trimers, which when followed by conformational changes of those trimers, draws membranes together for fusion. In viral fusogens that adopt a CII structure, however, these proteins are known to act unilaterally (in cis), with monomers on the same viral envelope membrane in order to form pre-fusion trimers under specific environmental triggering conditions and eventually fuse their membranes with those of the target host cells.

While initially it might seem that these two mechanisms are quite distinct from one another, it's possible that they are not entirely mutually exclusive, especially in the *Tetrahymena* system, where it is known that both mating partners express HAP2 (with a functional fusion loop) on their membranes at the sites of fusion between mating cells. Thus, it's possible in *Tetrahymena* HAP2, that a hybrid version of these two models is employed, utilizing both insertion of a fusion loop and trans-oligomerization to accomplish a productive and highly regulated sexual fusion between the two mating cells. But in this case, I think there is a slight distinction between those trans-oligomerization events suggested in the original EFF-1 model, and those employed by *Tetrahymena* HAP2. Instead of individual monomers forming trimers between the two target bilayers, due to the presence of a known functional fusion loop in *Tetrahymena* HAP2, I instead think it might be more likely that prefusion trimers form first on a single membrane, insert their fusion loops into the opposite lipid bilayer, and then form trans-trimer interactions (with HAP2 prefusion trimers on the opposite membrane before conformational changes occur) to help in more efficiently zippering the two membranes together. Certainly, more work needs to be done to identify the extent to which HAP2 may utilize a "cis" or potentially "trans" action during sexual cell fusion events, but at least initially, from the presence of HAP2 on only the male gamete in the majority of species studied so far, coupled with the presence of a functional fusion loop in HAP2 orthologs, it appears that the most logical

conclusion so far from the existing data would be that HAP2 functions via the viral “cis” model of CII fusion protein action.

Central to any discussion of these different mechanisms is the extent to which the different facets of fusogen structure contribute to the ultimate result of membrane fusion. The two main different characteristics of fusogens thought to be important for function are the fusion peptide’s insertion into target bilayers, and the dramatic conformational changes which occur in fusion protein oligomers that are expected to draw the two target bilayers together. Both models discussed above expect conformational changes to be fundamental, but membrane insertion of just the fusion peptide alone is known to initiate lipid mixing events between liposomes (see Chapter 3) through a yet unknown pathway, implying that just fusion peptide insertion is also equally sufficient for fusion pore formation. These data suggest that fusion peptide insertion alone is not simply serving a role as a membrane anchor in the target bilayer (a holdfast for later conformational changes), but may actually instead be more important in imparting a physical disturbance in the target bilayer, possibly aiding as a starting point for initiating hemifusion intermediate structures between the two membranes. Understanding and being able to measure how different CII fusogens fusion peptide’s structures change upon membrane insertion, cluster with prefusion trimer formation, and the extent of the physical stress or even membrane damage caused by peptide insertion may help us better delineate which of these two essential characteristics of CII fusion proteins contributes most significantly in these alternate models of fusion. Furthermore, examining the pre- and post-fusion structures of different HAP2 proteins and their respective fusion loops may also help to explain how come the HAP2-domain has come to be the most sequence-conserved part of this protein.

Structural similarities existing between EFF-1, CII viral fusogens, and HAP2 clearly point to these proteins having diverged from a common ancestor, but somehow these similar structures have ended up relying on the two characteristics of fusion protein action (fusion loop insertion and conformational changes) to different extents in their final respective mechanisms

of action. So far, emerging evidence suggests that careful specialization and regulation are hallmarks of cell-cell fusion processes (to limit fusion protein expression or action either to specific cells or temporally during development), where viral fusogens are more generalists when it comes to fusion (so as to ensure transmission to as many cell types as possible during infection). Broadly speaking, it is tempting to speculate that these different biological imperatives have informed the evolution of CII fusion protein mechanistic action in the viral versus developmental settings.

B3. Clinical applications for blocking the sexual transmission of parasites

Eukaryotic microbial pathogens such as *Plasmodium,* the causative agent of malaria, or others including *Trypanosomes, Leishmania, Eimeria, Babesiosis, Toxoplasma* etc. are continued threats to the health of humans and/or agricultural animals worldwide. Many of the parasitic organisms responsible for these devastating diseases have a well defined, obligatory sexual stage that, if blocked, would halt their transmission and arrest the continued spread of disease. So far, however, vaccines against many of these pathogens have proved elusive.

Since HAP2 is a highly conserved protein necessary for the sexual development of many of these parasitic protists, it represents a good potential target upon which new vaccines or therapies against these diseases could be developed. There is already some interest in HAP2 as a candidate target for transmission-blocking vaccines against malaria [11,12], but some reticence exists as there are no commercially available transmission blocking vaccines in existence. As a prevention method, transmission blocking vaccines are a relatively new idea and have yet to be thoroughly tested for efficacy or efficiency. Ideally, a model parasitic infection, perhaps of some agricultural relevance, would be a more appropriate stepping stone to demonstrate the value of these vaccines before trials in a major human pathogen like malaria.

Now that it is known that HAP2 is structurally similar to CII viral fusion proteins, the epitopes known to be important for conferring protection to other CII viral fusogens could be used to inform us as to which parts of the HAP2 protein might be the most integral for use in a

subunit vaccine. This is important, as often it is difficult to produce correctly-folded parasitic proteins easily given the standard protein expression platforms, and so being able to synthesize just the key epitope necessary for a protective response might be extremely helpful in generating effective vaccines. Since many of the CII viral envelope protein's fusion loops are the targets of neutralizing antibodies, this fusion peptide sequence might represent a good starting point for the development of a transmission-blocking vaccine targeting HAP2. Learning from successful vaccines that have been developed in the past which targeted CII viral fusogens will also be helpful here[13,14]. Alternative to the vaccination approach, small molecules inhibiting the conformational changes of these CII fusogens or even lipid bilayer intermediates in the membrane fusion process could also be tested to block transmission.

B4. Identifying developmental cell-cell fusogens through structural homology: TULP4

Based on primary sequence analysis, HAP2 does not have an identifiable vertebrate ortholog and the gene products controlling gamete membrane fusion in humans and other vertebrates are still unknown. However, given the finding that HAP2 and CII viral fusogens are structurally similar, using structural homology modeling techniques may provide avenues towards the discovery of more, yet unidentified developmental fusion proteins that also adopt protein folds analogous to those of known viral fusogens. Towards this end, I have recently identified a mammalian protein, Tubby-Like Protein 4 (TULP4), as having features reminiscent of both HAP2 and CII viral fusion proteins. Because TULP4 is also expressed in the testis, and is encoded in a region of the mouse genome that is intimately linked with male fertility, we believe that this protein is a strong candidate for the elusive sperm-egg fusogen of mammals.

The similarities that led to the identification of the TULP4 as a candidate gamete fusogen were uncovered using the homology modeling tool "BackPHYRE" [15], which allows a researcher to use a known structure from the protein data bank as query to search all predicted open-reading frames within genomes of interest in order to identify putative protein products that may have similar structures. I choose the CII viral fusogen structure with the highest degree of

similarity to HAP2 (the Dengue virus envelope protein structure; PDB ID: 1UZG, chain A [16] and used it to search several vertebrate genomes, including human and mouse, for proteins that would be predicted to adopt a similar fold. The top hit in both the human and mouse genomes was *TULP4* (88.9% and 78.0% confidence respectively).

At first glance, TULP4 is an unlikely contender as a potential membrane fusogen. Despite the high confidence match to Class II viral fusion proteins based on BackPHYRE analysis, TULP4 is not predicted to contain a signal peptide or a transmembrane domain (normally present in other membrane fusion proteins). However, there is some disagreement between the various prediction software, as recently I found that despite the lack of a signal peptide, several LOCATE [17] predictors do place TULP4 on the plasma membrane.

Regardless, other attributes of TULP4 are highly suggestive of it's potential fusogenic role in the male gamete fertility. For instance, the promoter region of the *TULP4* gene is located within a 225kb stretch of the mouse *t* haplotype known to cause complete male sterility [18,19]. Furthermore, TULP4 orthologs in many mammals are highly conserved (like HAP2 in other more basal eukaryotes) suggesting an important shared biological function for the gene product. Correspondingly, both the gene and protein are expressed in the testis based on searches of data available from the EMBL-EBI expression atlas [20] and MOPED [21] databases. Finally, despite the absence of readily identifiable signal peptides, tubby domain-containing proteins have been localized to both intracellular and extracellular cellular compartments and are known to be secreted [22]. Secretion, in this case, is hypothesized to depend on the ability of these proteins to bind to phosphatidylinositol 4,5-bisphosphate. Indeed, just the tubby domain itself has been recognized to having a 3D fold similar to that of phospholipid scramblases, proteins that rearrange lipid orientations between inner & outer leaflets of the bilayer[23]. This raises the intriguing possibility that TULP4, acting as a membrane scramblase, could catalyze a "flipped-out" membrane fusion event between male and female gametes following its regulated expression in sperm.

Given these suggestive characteristics, we have initiated experiments to test the possibility that TULP4 could function as the mammalian sperm-egg fusion protein through the construction of genetic modifications to the *TULP4* gene in mice using CRISPR/Cas9. Once complete gene knockouts have been made, we will set up breeding cages to see if these male knockout animals are infertile. If this is the case, further experiments could be done via *in vitro* fertilization experiments, which could potentially show that the fertility defect in these mice might be rescued by intracytoplasmic injection of knockout sperm into WT female eggs. These experiments could be similar to those that have already been performed for other *t haplotype* mice[24]. Likewise, other methods will also be useful in testing this hypothesis if initial mouse breeding experiments result in infertility. Such experiments might include (1) the heterologous expression of the TULP4 protein in cells (potentially causing membrane fusion and syncytia formation), (2) obtaining an actual crystal structure of the protein to examine the true extent of structural homology that might exist between this protein and viral fusogens, or even (3) using biophysical fusion peptide assays to see if the TULP4 fusion loop peptide has the capacity to interact with model membranes. These are only a short list of the many very exciting projects that might ensue should this hypothesis bear out. If infertility does not result in male mice lacking *TULP4*, I believe the structural homology modeling approach is still a logical method to pursue in the difficult endeavor to find these still undiscovered developmental fusogens, and continued improvements to these modeling approaches may lead to yet new candidate proteins to be tested.

C. Evolutionary Implications

Together the results presented in this dissertation, including predictive structural modeling and biophysical tests of the HAP2 fusion loop, show that at least in *Tetrahymena* HAP2 may both look and act like a viral fusogen. But what does this mean for eukaryotic sex? These data give us insight about two major evolutionary aspects regarding eukaryotic sex.

First, this work brings new insight into the evolution of the gamete-specific functions during sexual conjugation. It is widely thought that isogamous species (with gametes of similar size and function) were among the earliest eukaryotes to have sex[25]. But how then did most of the species on earth then end up anisogamous (with gametes that are different in size and function)? What evolutionary challenge was met, and perhaps continues to be met, that limits HAP2 to just one gamete in the majority of species on earth? We aren't able to answer these questions yet, but we do know that this work uncovers the evolutionary adaptability of this gamete membrane fusion protein, which preserves its fusogenic-function in an organism that has dispensed with the production of male gamete cells[26]. And so, *Tetrahymena* has offered us a glimpse into how some of the earliest eukaryotes may have initially solved the problem of cell-cell fusion during sex, thus opening up the possibility that HAP2 functioned in fusion long before it ever became gamete-specific.

Second, and perhaps most importantly, the mere existence of sex has long troubled evolutionary biologists[27,28]. Why should sex exist, when it is expected that natural selection would instead favor an organism that can produce offspring all by itself (using a process called parthenogenesis). Finding a mate, producing males, creating showy traits – only to pass along the genetic equivalent of half of yourself to the next generation are all quite costly activities. So what is the biological benefit of sex and why is it so prevalent? Given the evidence existing, I propose we should now expect that selection favoring sex in eukaryotic populations acts at two different levels; selection favoring the maintenance of sex, and selection favoring the origin of sexual activity in eukaryotes.

While existing work has already shown that sex is maintained in eukaryotic populations by endowing species with an enhanced ability to adapt to ever-changing environments through increased horizontal spread of positive mutations throughout the population (Figure 7.1A)[29,30], this mode of maintenance is through group selection, making it very difficult to understand how sex would have been selected for initially, in the first cells ever to try it. Dr. Donal Hickey's

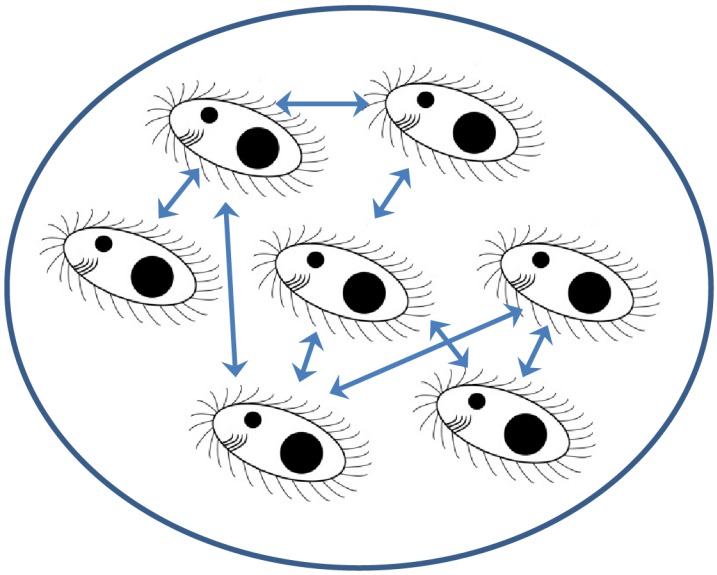
theory (described in detail in Chapter 3, and [31,32]) suggests that the origin of eukaryotic sex was driven by parasitic, semi-autonomous genetic elements, such as viruses or transposons, which promoted the initial sexual cell fusion events in order to promote their own intergenomic transmission between early eukaryotic cells (Figure 7.1B). Our data showing structural and functional similarities between HAP2 and viral fusogens could be viewed as early evidence supporting this infectious theory on the origin of sex.

These evolutionary observations are consistent with the findings of Dr. Yorgo Modis regarding the nature of relationships between of CII viral fusogens. Dr. Modis found in 2014 that the conservation of solved class II viral fusogen structures doesn't line up with their overall viral phylogenetic relatedness[33], a result that predicts the genetic elements encoding class II viral fusogen structures had independent mobility between different genetic systems.

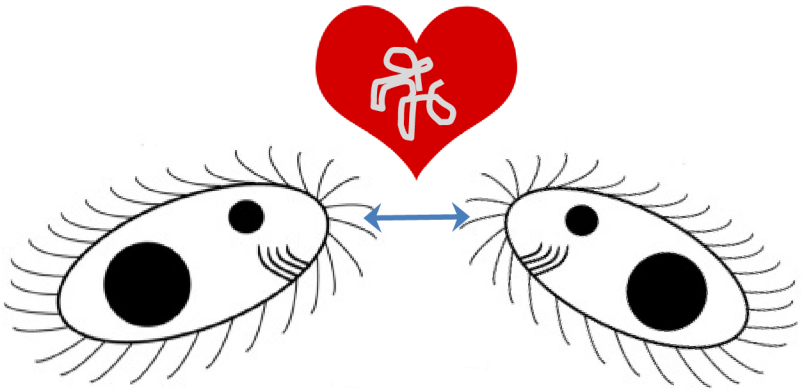
Subsequently, the necessity of cell-cell fusion as a primary step in eukaryotic sex, coupled with the wide prevalence and conservation of HAP2 orthologs in basal lineages of eukaryotes would expect HAP2 to have had a key role in the origin of eukaryotic sex. These insights will likely begin to reshape our understanding of how profoundly horizontal gene transfer events between viruses and cells may have ultimately shaped the course of eukaryotic evolution.

Figure 7.1 Selection for sex. A cartoon illustrating the two levels at which the selective forces work on sex in eukaryotic species **(A)** Sex is maintained in eukaryotes by the increased gene flow (arrows) it provides between members of a given population (circled). **(B)** Sex began in eukaryotes by selection on a selfish genetic element, which initiated the first cell-cell fusion event, thus favoring its own transmission between early eukaryotic cells.

A



B



REFERENCES

1. Schoen, P., Leserman, L., and Wilschut, J. (1996). Fusion of reconstituted influenza virus envelopes with liposomes mediated by streptavidin/biotin interactions. FEBS Lett. *390*, 315–318.

2. Shilagardi, K., Li, S., Luo, F., Marikar, F., Duan, R., Jin, P., Kim, J.H., Murnen, K., and Chen, E.H. (2013). Actin-propelled Invasive Membrane Protrusions Promote Fusogenic Protein Engagement During Cell-Cell Fusion. Science *340*, 359–363.

3. Leung, J.Y.-S., Ng, M.M.-L., and Chu, J.J.H. (2011). Replication of Alphaviruses: A Review on the Entry Process of Alphaviruses into Cells. Adv. Virol. *2011*, e249640.

4. Smit, J.M., Moesker, B., Rodenhuis-Zybert, I., and Wilschut, J. (2011). Flavivirus cell entry and membrane fusion. Viruses *3*, 160–171.

5. Dubé, M., Rey, F.A., and Kielian, M. (2014). Rubella virus: first calcium-requiring viral fusion protein. PLoS Pathog. *10*, e1004530.

6. Zaitseva, E., Yang, S.T., Melikov, K., Pourmal, S., and Chernomordik, L.V. (2010). Dengue virus ensures its fusion in late endosomes using compartment-specific lipids. PLoS Pathog. *6*, e1001131.

7. Podbilewicz, B., Leikina, E., Sapir, A., Valansi, C., Suissa, M., Shemer, G., and Chernomordik, L.V. (2006). The C. elegans developmental fusogen EFF-1 mediates homotypic fusion in heterologous cells and in vivo. Dev. Cell *11*, 471–481.

8. Chernomordik, L.V., and Kozlov, M.M. (2003). Protein-lipid interplay in fusion and fission of biological membranes. Annu. Rev. Biochem. *72*, 175–207.

9. Orias, E., Hamilton, E.P., and Flacks, M. (1979). Osmotic shock prevents nuclear exchange and produces whole-genome homozygotes in conjugating Tetrahymena. Science *203*, 660–663.

10. Avinoam, O., Fridman, K., Valansi, C., Abutbul, I., Zeev-Ben-Mordehai, T., Maurer, U.E., Sapir, A., Danino, D., Grunewald, K., White, J.M., *et al.* (2011). Conserved eukaryotic fusogens can fuse viral envelopes to cells. Science *332*, 589–592.

11. Blagborough, A.M., and Sinden, R.E. (2009). Plasmodium berghei HAP2 induces strong malaria transmission-blocking immunity in vivo and in vitro. Vaccine *27*, 5187–5194.

12. Hirai, M., Arai, M., Mori, T., Miyagishima, S.Y., Kawai, S., Kita, K., Kuroiwa, T., Terenius, O., and Matsuoka, H. (2008). Male fertility of malaria parasites is determined by GCS1, a plant-type reproduction factor. Curr. Biol. CB *18*, 607–613.

13. Pulendran, B. (2009). Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology. Nat. Rev. Immunol. *9*, 741–747.

14. Frierson, J.G. (2010). The Yellow Fever Vaccine: A History. Yale J. Biol. Med. *83*, 77–85.

15. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

16. Modis, Y., Ogata, S., Clements, D., and Harrison, S.C. (2005). Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. J. Virol. *79*, 1223–1231.

17. Mammalian Protein Localization Database Available at: http://locate.imb.uq.edu.au/cgi-bin/show_slpdata.cgi?entry=6116448 [Accessed January 9, 2017].

18. Schimenti, J.C., Reynolds, J.L., and Planchart, A. (2005). Mutations in Serac1 or Synj2 cause proximal t haplotype-mediated male mouse sterility but not transmission ratio distortion. Proc. Natl. Acad. Sci. U. S. A. *102*, 3342–3347.

19. Planchart, A., and Schimenti, J.C. (2001). Experimental and computational approaches yield a high-resolution, 1-Mb physical map of the region harboring the mouse t haplotype sterility factor, tcs1. Mamm. Genome Off. J. Int. Mamm. Genome Soc. *12*, 668–670.

20. Petryszak, R., Burdett, T., Fiorelli, B., Fonseca, N.A., Gonzalez-Porta, M., Hastings, E., Huber, W., Jupp, S., Keays, M., Kryvych, N., *et al.* (2014). Expression Atlas update–a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. Nucleic Acids Res. *42*, D926-32.

21. Kolker, E., Higdon, R., Haynes, W., Welch, D., Broomall, W., Lancet, D., Stanberry, L., and Kolker, N. (2012). MOPED: Model Organism Protein Expression Database. Nucleic Acids Res. *40*, D1093-9.

22. Caberoy, N.B., and Li, W. (2009). Unconventional secretion of tubby and tubby-like protein 1. FEBS Lett. *583*, 3057–3062.

23. Bateman, A., Finn, R.D., Sims, P.J., Wiedmer, T., Biegert, A., and Soding, J. (2009). Phospholipid scramblases and Tubby-like proteins belong to a new superfamily of membrane tethered transcription factors. Bioinforma. Oxf. Engl. *25*, 159–162.

24. Kuretake, S., Maleszewski, M., Tokumasu, A., Fujimoto, H., and Yanagimachi, R. (1996). Inadequate function of sterile tw5/tw32 spermatozoa overcome by intracytoplasmic sperm injection. Mol. Reprod. Dev. *44*, 230–233.

25. Togashi, T., and Cox, P.A. eds. (2011). The evolution of anisogamy: A fundamental phenomenon underlying sexual selection. Ch.4 Nucleo-cytoplasmic conflict and the evolution of gamete dimorphism. In The evolution of anisogamy: a fundamental phenomenon underlying sexual selection (New York: Cambridge University Press), pp. 111–130. Available at: http://assets.cambridge.org/97805218/80954/cover/9780521880954.jpg.

26. Orias, E. (2014). Membrane Fusion: HAP2 Protein on a Short Leash. Curr. Biol. *24*, R831–R833.

27. Schön, I., Martens, K., and Dijk, P. eds. (2009). Lost Sex (Dordrecht: Springer Netherlands) Available at: http://link.springer.com/10.1007/978-90-481-2770-2 [Accessed November 8, 2016].

28. Smith, J.M. (1978). The Evolution of Sex (Oxford, Greate Britain: Cambridge University Press).

29. Goddard, M.R. (2016). Molecular evolution: Sex accelerates adaptation. Nature *531*, 176–177.

30. McDonald, M.J., Rice, D.P., and Desai, M.M. (2016). Sex speeds adaptation by altering the dynamics of molecular evolution. Nature *531*, 233–236.

31. Hickey, D.A. (1982). Selfish DNA: A Sexually-Transmitted Nuclear Parasite. Genetics *101*, 519–531.

32. Hickey, Donal A., and Rose, Michael R. (1988). The role of gene transfer in the evolution of eukaryotic sex. In The Evolution of sex: an examination of current ideas, R. E. Michod and B. R. Levin, eds. (Sunderland, Mass: Sinauer Associates), pp. 161–175.

33. Modis, Y. (2014). Relating structure to evolution in class II viral membrane fusion proteins. Curr. Opin. Virol. *5*, 34–41.