The Company of
Biologists

RESEARCH ARTICLE

Sexual cell cycle initiation is regulated by CDK19 and CYC9 in *Tetrahymena thermophila*

Yang Ma[1,2,3], Guanxiong Yan[1,2,3], Xiaojie Han[2,4], Jing Zhang[2], Jie Xiong[2] and Wei Miao[1,2,3,5,*]

ABSTRACT

To investigate the mechanisms underlying initiation of the sexual cell cycle in eukaryotes, we have focused on cyclins and cyclin-dependent kinases (CDKs) in the well-studied model ciliate, *Tetrahymena thermophila*. We identified two genes, *CDK19* and *CYC9*, which are highly co-expressed with the mating-associated factors MTA, MTB and HAP2. Both CDK19 and CYC9 were found to be essential for mating in *T. thermophila*. Subcellular localization experiments suggested that these proteins are located at the oral area, including the conjugation junction area, and that *CDK19* or *CYC9* knockout prevents mating. We found that CDK19 and CYC9 form a complex, and also identified several additional subunits, which may have regulatory or constitutive functions. RNA sequencing analyses and cytological experiments showed that mating is abnormal in both Δ*CDK19* and Δ*CYC9*, mainly at the entry to the co-stimulation stage. These results indicate that the CDK19–CYC9 complex initiates the sexual cell cycle in *T. thermophila*.

KEY WORDS: Cell cycle, Cyclin, Cyclin-dependent kinase, Sexual reproduction, *Tetrahymena thermophila*

INTRODUCTION

In multicellular organisms, most cells divide via the asexual cell cycle, and a specific regulatory mechanism ensures that only a few cells enter the sexual cell cycle. In single-celled organisms, the decision about whether cells enter the asexual or sexual cell cycle is determined by environmental factors and mating partner existence. Thus, controlling when and how to initiate the sexual cell cycle is of universal importance for eukaryotic organisms.

Tetrahymena thermophila is a single-cell model organism with two modes of reproduction: asexual reproduction predominates when the environment contains abundant nutrients, and sexual reproduction predominates in nutrient-deficient environments. Because sexual reproduction can be initiated in these cells by a simple process of switching them from nutrient-rich broth to aqueous buffers, *T. thermophila* serves as a useful model for studying the mechanisms underlying initiation of the sexual cell cycle. *Tetrahymena thermophila* possesses seven mating types (Orias et al., 2017). The sexual cell cycle (conjugation) takes place

[1]State Key Laboratory of Freshwater Ecology and Biotechnology, Wuhan 430072, China. [2]Key Laboratory of Aquatic Biodiversity and Conservation, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China. [3]University of Chinese Academy of Sciences, Beijing 100049, China. [4]College of Fisheries and Life Science, Dalian Ocean University, Dalian 116023, China. [5]CAS Center for Excellence in Animal Evolution and Genetics, Kunming 650223, China.

*Author for correspondence (miaowei@ihb.ac.cn)

Y.M., 0000-0003-0772-6309; G.Y., 0000-0001-6253-5922; X.H., 0000-0003-1255-3430; J.Z., 0000-0002-7343-2184; J.X., 0000-0001-8923-0424; W.M., 0000-0003-3440-8322

between two starved cells of different mating types (Bruns and Brussard, 1974; Bruns and Palestine, 1975), and starts with an initial period of about 30 min for pair formation (called the co-stimulation stage) (Fig. S1). During this period, cells undergo a series of processes to become ready for pairing, including co-stimulation-induced rounding (Fujishima et al., 1993), development of a smooth-surfaced area (Suganuma et al., 1984), tip transformation (Wolfe and Grimes, 1979), and concanavalin A receptor (ConA-R) appearance (Wolfe and Feng, 1988; Wolfe et al., 1986). In addition, once cells enter the co-stimulation stage, they can form pairs immediately (Suganuma et al., 1984).

The cell cycle is regulated by a series of proteins. Among them, cyclin-dependent kinases (CDKs) often play a central role (Morgan, 1995). CDKs are present in all sequenced eukaryotes, and many of their regulatory functions in the cell cycle have been proven to be evolutionarily conserved. CDKs are serine/threonine kinases that usually play an important role in integrating cell signals to regulate gene transcription. The functions of CDKs are modulated by cyclins, CDK inhibitors and CDK-activating kinases (Morgan, 2007; Obaya and Sedivy, 2002).

To determine whether cyclins and CDKs play a role in initiation of the sexual cell cycle, we screened the expression profile and found that a pair of proteins, CDK19 and CYC9 (hereafter denoted CDK19–CYC9), are highly co-expressed with mating type determination genes, *MTA* and *MTB* (Cervantes et al., 2013), and a gene encoding the male-gamete-specific fusion protein, *HAP2* (Cole et al., 2014; Pinello et al., 2017), which are all essential for conjugation in *T. thermophila*. Consistent with this, subcellular localization experiments indicated that both CDK19 and CYC9 localize to the oral area, including the conjugation junction area. Further gene knockout (KO) experiments confirmed that loss of either protein prevents cell pairing. To verify that CDK19–CYC9 complex formation is needed to initiate the sexual cell cycle, immunoprecipitation experiments were conducted at three stages (starvation, co-stimulation and conjugation). Surprisingly, CDK19 was found to interact with CYC9 at all three stages. However, additional regulatory subunits (CKS1 and CIP2) were specifically recruited to the CDK19–CYC9 complex, suggesting that activity of the complex may be controlled by these subunits. Moreover, a WD40 domain-containing protein, CIP1, seems to be a constitutive component of this complex. Lastly, we found that deletion of either the *CDK19* or *CYC9* gene resulted in abnormalities in gene expression and mating behavior at the start of the co-stimulation stage. These results indicate that the CDK19–CYC9 complex is an important regulator of sexual cell cycle initiation in *T. thermophila*.

RESULTS

CDK19 and CYC9 are candidate regulators of mating initiation in *T. thermophila*

Cyclins and CDKs have been identified in all sequenced eukaryotes, and are mainly involved in regulating the cell cycle. We previously

1

showed that *T. thermophila* has a total of 34 cyclins (Yan et al., 2016b), and only four standard CDKs with exact PSTAIRE-like cyclin-binding motifs (Yan et al., 2016a). Because of the divergence between the number of annotated cyclin and CDK genes in this organism, we searched for additional CDK family members and identified a total of 20 CDKs or CDK-like genes in the *T. thermophila* macronuclear genome (Table S1). The phylogeny, domain composition, expression profile and cyclin-binding motifs of all 20 proteins are shown in Fig. 1.

To identify which proteins control initiation of the sexual cell cycle, we inspected their expression profile in *T. thermophila*. In previous studies, three genes have been proven to be essential for mating in *Tetrahymena*: *MTA*, *MTB* and *HAP2* (Cervantes et al., 2013; Cole et al., 2014; Orias, 2014; Pinello et al., 2017). *MTA* and *MTB* appear to be required for cell recognition, while *HAP2* plays a role in stabilizing mating pairs and catalyzing membrane fusion between mating cells. All three genes are silent during the growth stage, but are specifically expressed under starvation and conjugation conditions, with a peak of expression in early conjugation. The particular expression pattern is therefore a probable characteristic to select the sexual cell cycle regulators.

We screened the expression profiles of all CDKs and cyclins and found that only *CDK19* (TTHERM_00339880) and *CYC9* (TTHERM_00940290) are co-expressed with all three genes (Fig. 2). In this case, we focused our studies on these two gene products. Phylogenetic analysis showed that CDK19 is a CDK that has most of the standard CDK characteristics (Fig. S2), except for a slight difference in the cyclin-binding motif. The fact that CDK19 and CYC9 are highly co-expressed with MTA, MTB and HAP2 suggests that these proteins may function in initiating the *T. thermophila* sexual cell cycle.

Both CDK19 and CYC9 are essential for cell pairing in *T. thermophila*

To determine whether CDK19 and CYC9 are involved in cell pairing, we first used immunofluorescence analysis to study the subcellular localization of Hemagglutinin (HA)-tagged proteins. CDK19-HA and CYC9-HA were found to have similar localization patterns (Fig. 3): the fluorescence signal was extremely weak at the starvation stage, but was clearly evident at the co-stimulation stage, and a stronger signal was developed in the oral area (Fig. 3, arrows) in some cells as development proceeded. In conjugating cells, the
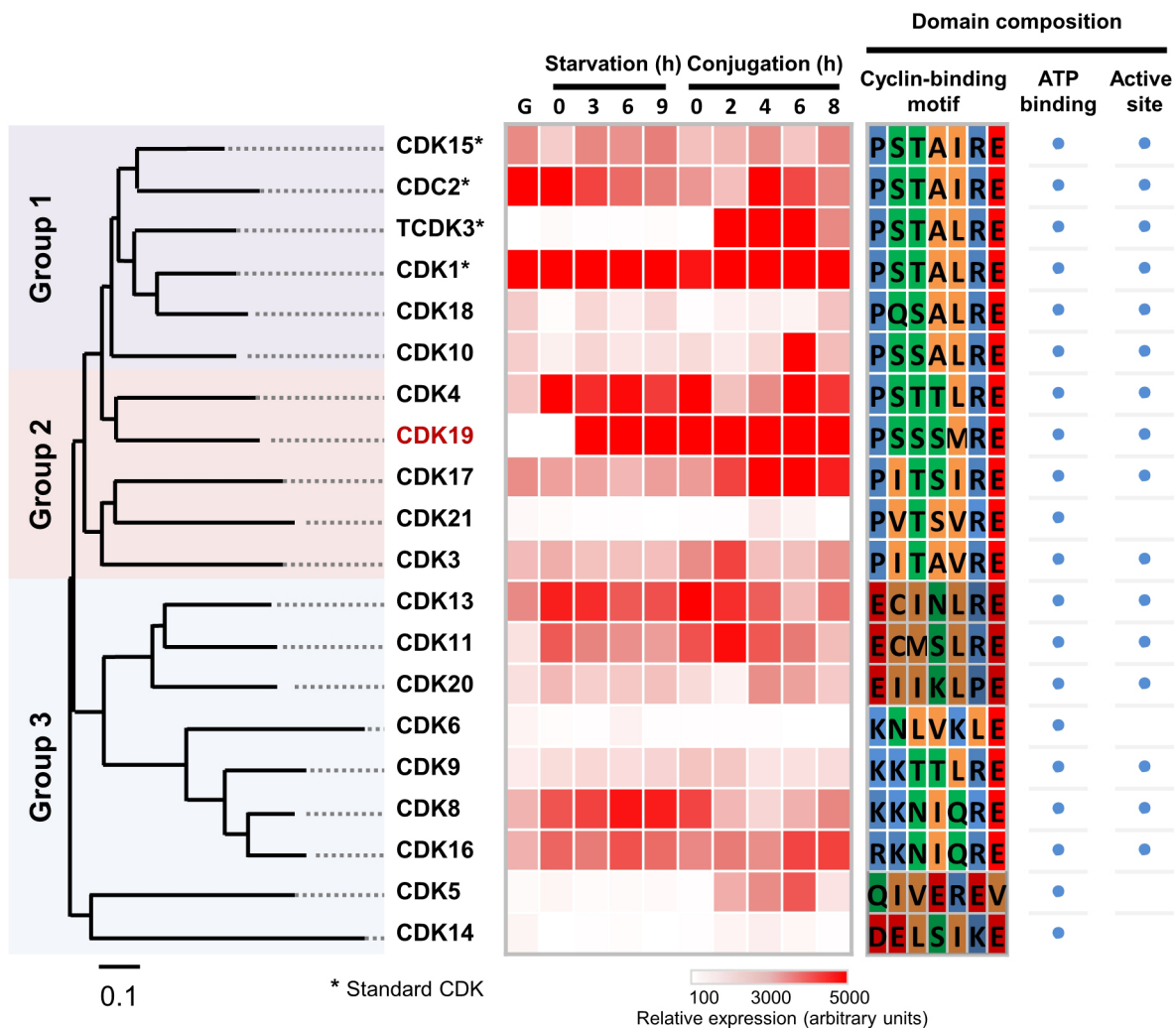

Fig. 1. Phylogenetic tree of 20 CDK genes in *T. thermophila*, with the domain composition and expression profile of the encoded proteins. The 20 CDK genes could be classified into three groups: group 1 contains six conserved CDKs; the CDKs in group 2 are somehow different in the cyclin-binding motif; and group 3 contains CDK-like proteins. The ATP-binding domain could be identified in all 20 CDKs by InterProScan or CDD/SPARCLE-NCBI, but the active site could only be identified in 16 CDKs. CDK19 is shown in red.
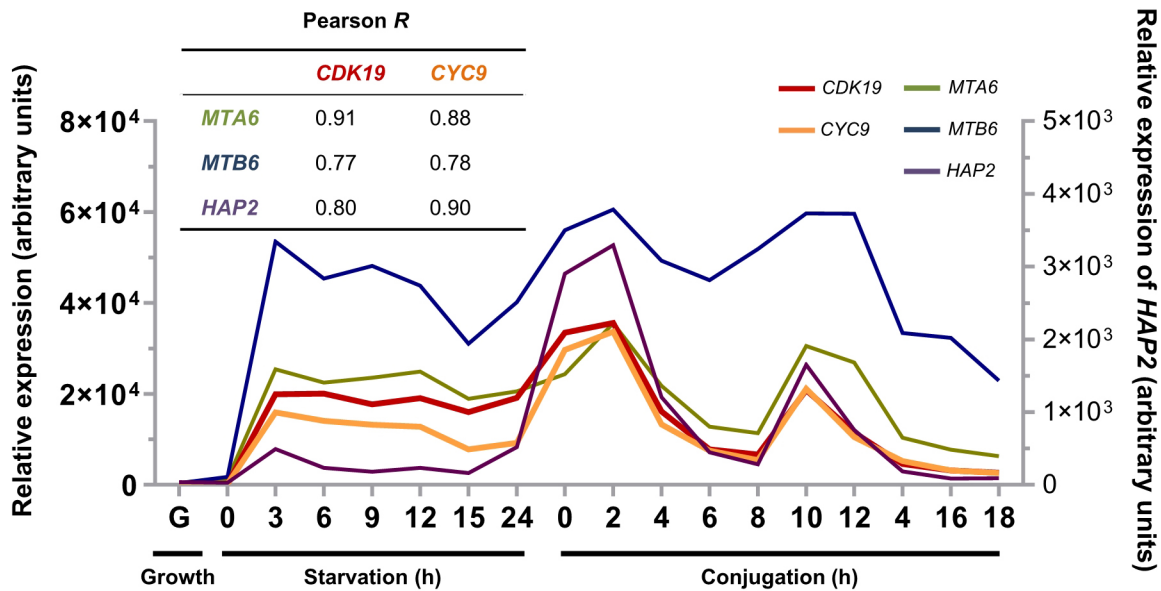
Fig. 2. Expression profiles of *CDK19*, *CYC9*, *MTA*, *MTB* and *HAP2*. Expression levels of *CDK19*, *CYC9*, *MTA* and *MTB* are indicated on the left ordinate, and expression level of *HAP2* is indicated on the right ordinate. The table in the top left corner shows the Pearson correlation (Pearson *R*) coefficients for specific gene pairs.

fluorescence signals in the oral area were brightest, including the junction (mating area). In addition, the weak signal in starving cells may result from low protein levels during the starvation stage, and the strong signal in the co-stimulation and conjugation stages may be due to the increased protein content, as shown by western blotting (Fig. S3). In addition, the high levels of expression at the co-stimulation and conjugation stages suggest that CDK19 and CYC9 function during the co-stimulation stage.

The correct subcellular localization of CDKs and cyclins are of great importance for their function (Obaya and Sedivy, 2002). Both CDK19 and CYC9 show a tendency to localize to the oral area, an important site in cell pair formation and thereby for initiating the sexual cell cycle. Moreover, CDK19 also localized to the macronucleus and micronucleus, while CYC9 did not (Fig. 3), indicating that CDK19 may be imported into the nuclei to regulate transcription.

To verify that CDK19 and CYC9 regulate the sexual cell cycle in *T. thermophila*, we knocked out *CDK19* and *CYC9* separately. Polymerase chain reaction (PCR) analysis confirmed that each gene had been completely knocked out (Fig. S4A), and RNA sequencing (RNA-Seq) analysis verified the absence of transcripts in each KO strain (Fig. S4B). We next studied the progress of the sexual cell cycle in both Δ*CDK19* (*CDK19* knockout strains) and Δ*CYC9* (*CYC9* knockout strains). Similar to wild-type cells, in both KO strains, cell fission occurred normally in culture medium and survived without proliferating under starvation conditions. However, a distinct phenotype was observed during the conjugation stage: Δ*CDK19* and Δ*CYC9* cells were unable to pair. Irrespective of whether they were mixed with KO cells or wild-type (WT) cells, no pairs were formed (Table 1). This suggested that both Δ*CDK19* and Δ*CYC9* cells had lost the ability to mate. Therefore, we conclude that both CDK19 and CYC9 are essential for mating.

Identification of the CDK19–CYC9 complex and additional subunits

All known CDKs are activated by cyclin binding. Our results showed that both CDK19 and CYC9 are essential for mating, are highly co-expressed, and have similar subcellular localization patterns. It was therefore possible that these proteins functioned together in a complex or that they independently formed a complex with another CDK or cyclin. To distinguish between these possibilities, we performed immunoprecipitation-coupled mass spectrometry (MS) experiments. Samples of both CDK19-HA-expressing cells and CYC9-HA-expressing cells were taken under starvation, co-stimulation and conjugation conditions. CDK19 and CYC9 were found to interact with each other at all three stages of the sexual cell cycle (Fig. 4A; Tables S2–S13), illustrating that CDK19 does form a complex with CYC9, and that this interaction appeared even at the starvation stage.

We also found that three additional proteins, CIP1, CKS1 and CIP2, were all interacted with CDK19 and CYC9, respectively (Fig. 4A). It is still unclear what the components are in each individual complex. For example, during conjugation stage, one complex may contain CIP1 and another may contain CIP2.

All three proteins are highly co-expressed with CDK19 and CYC9 (Fig. 4B). The first one, CIP1 (CDK/cyclin interacting protein 1; TTHERM_00891190), was identified in all hemagglutinin (HA)-expressing cell samples from three stages. CIP1 contains a WD40 domain, one of the most abundant protein interaction domains that coordinate multi-protein complex assemblies, where the repeating units serve as a rigid scaffold for protein–protein or protein–DNA interactions (Jain and Pandey, 2018; Xu and Min, 2011). The second protein, CKS1 (CDK regulatory subunit 1; TTHERM_00794670), was identified at both the co-stimulation and conjugation stages. CKS are small, evolutionarily conserved proteins that have little effect on the conformation of the CDK catalytic site, but instead provide an accessory binding site with phosphorylated residues (Harper, 2001; Morgan, 2007). It is possible that CKS1 may modulate the function of the CDK19–CYC9 complex. The last protein, CIP2 (CDK/cyclin interacting protein 2; TTHERM_00449610), is a hypothetical protein. It was detected only at the conjugation stage, suggesting that it might specifically regulate the CDK19–CYC9 complex under conjugation conditions.
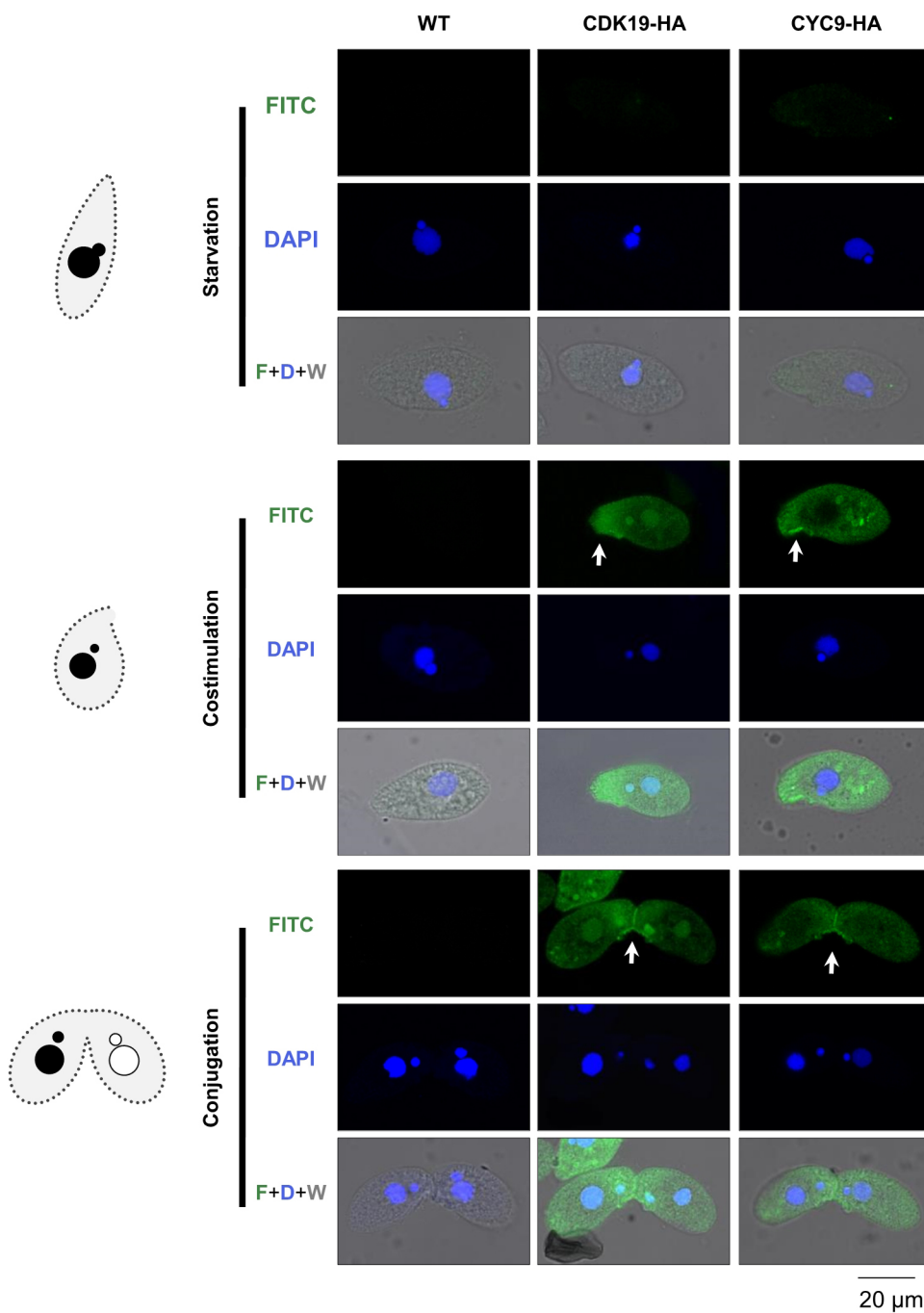
Fig. 3. Subcellular localization of CDK19-HA and CYC9-HA. CDK19-HA and CYC9-HA were visualized by immunofluorescence staining. Cell samples were taken from the starvation, co-stimulation and conjugation stages. FITC immunostaining (green) shows CDK19 and CYC9 localization, while DAPI staining (blue) shows the cell nucleus. White arrows indicate the oral area. CDK19 and CYC9 are almost undetectable at the starvation stage, but strong and specific staining is seen at the co-stimulation and conjugation stages, especially at the oral area, including the junction area. F+D+W denotes merging of FITC, DAPI and white light.

These three proteins were detected by CDK19 and CYC9, and all were highly co-expressed. To investigate whether these three proteins are related to cell mating, we knocked out *CIP1*, *CKS1* and *CIP2*, respectively (Fig. S4C). We found that knockout *CKS1* or *CIP2* cells cannot mate (Table 1). These results indicate that these two proteins are indeed involved in cell mating. In addition, cells became selfer when they lose CIP1, which means when Δ*CIP1* cells were transferred to starvation medium, the strain itself formed pairs

Table 1. Percentage pairing of WT cells, Δ*CDK19*, Δ*CYC9*, Δ*CKS1* and Δ*CIP2* cells

	WT×WT	Δ*CDK19*		Δ*CYC9*		Δ*CKS1*		Δ*CIP2*	
		Δ×Δ	Δ×WT	Δ×Δ	Δ×WT	Δ×Δ	Δ×WT	Δ×Δ	Δ×WT
2 h	86	0	0	0	0	0	0	0	0
4 h	86	0	0	0	0	0	0	0	0
6 h	91	0	0	0	0	0	0	0	0

All values are given as percentages. Mating ability of knockout cells is detected by mixing them with cells of another mating type. Regardless of whether they were mixed with WT or KO cells, no mating pairs were seen in the Δ*CDK19*, Δ*CYC9*, Δ*CKS1* and Δ*CIP2* experimental groups.

without the addition of another mating type cell; this indicates that CIP1 may play an inhibition role in cell recognition.

Recent gene orthology studies (Xiong et al., 2019; Yang et al., 2019) showed that these genes have a one-to-one ortholog in all *Tetrahymena* species (except for *CIP2*, which is absent in *T. shanghaiensis* and *T. paravorax*; Fig. 4C). This result suggests that the organization of the CDK19–CYC9 complex is relatively evolutionarily conserved, at least in *Tetrahymena*.

The CDK19–CYC9 complex has important functions in cell co-stimulation

Our results revealed that CDK19 and CYC9 form a complex from the starvation stage onwards, and may therefore regulate the sexual cell cycle from then on. However, we observed no significant difference between the WT and two KO strains at starvation stage (e.g. cessation of fission, normal cell size, and nuclei formation). As shown in Figs 2 and 3 and Fig. S3, the expression of both *CDK19* and *CYC9* reached an obvious peak at co-stimulation and the early conjugation stages, similar to other known essential mating genes (*MTA*, *MTB* and *HAP2*), suggesting that the complex plays a particular role in cell mating.

To determine which stage(s) is controlled by this complex, we first screened for differentially expressed genes (DEGs) between KO and WT strains (DEGs-KO/WT) using the following criteria: a fold change in expression of >2 and fragments per kilobase of exon per million fragments mapped (FPKM) of >10 (see Materials and Methods). As shown in Fig. 5A, more genes are differentially expressed at the co-stimulation stage than under starvation, with more variation in expression at the co-stimulation stage. We also noted that many genes had a $\log_2$ (fold change) value of >16 during co-stimulation, but this difference in expression was rarely seen at the starvation stage.

To determine whether DEGs-KO/WT regulate changes in the cell cycle stage, we next screened for genes that were differentially expressed from growth to starvation conditions (DEGs-G/S) and from the starvation to co-stimulation stages (DEGs-S/Co) of WT cells, and then determined the number of DEGs in each category that were shared with the DEGs-KO/WT group (Fig. 5C,D) to find out the genes with opposite-changed expression in KO cells compared with the normal WT cells. The analysis revealed that during the co-stimulation stage, the intersection of DEGs-KO/WT and DEGs-S/Co is great (1100 down-regulated and 1496 up-regulated genes); while only a few DEGs-KO/WT of starvation were shared with the DEGs-G/S group (307 down-regulated and 90 up-regulated genes). Moreover, the intersection of Δ*CDK19* and Δ*CYC9* DEGs is obviously greater at the co-stimulation stage than that under starvation conditions. Therefore, the transcriptome results indicate that the CDK19–CYC9 complex mainly functions during the co-stimulation stage.

Further gene ontology (GO) enrichment analysis of these overlapping genes at co-stimulation stage was performed. The result shows that overlapping genes in all three sets were mainly involved in protein synthesis, transcription regulation, cytoskeleton and signal conduction (Fig. 5E; Tables S14 and S15). This result is consistent with previous reports that during the co-stimulation stage, cells would undergo a transient increase in the synthesis of some proteins, such as p80 (Garfinkel and Wolfe, 1981). The altered-regulated genes related to protein synthesis may include these proteins. There is also a wave of RNA synthesis during the co-stimulation stage (Allewell and Wolfe, 1977; Ron and Horovitz, 1977). Genes related to transcription regulation were also altered in expression in Δ*CDK19* and Δ*CYC9* cells. These genes are

enriched in cytoskeleton and actin-related proteins, and changes may be due to the failure in co-stimulation-induced rounding (Fujishima et al., 1993) and tip transformation (Wolfe and Grimes, 1979), as these events require the cytoskeleton to be reorganized. Therefore, the GO results also indicated that the altered-regulated genes in Δ*CDK19* and Δ*CYC9* are important for mating initiation.

In addition, we also collected the genes which were reported to associate with cell pairing and early conjugation, and their expression in knockout strains was analyzed (Fig. 5B). The expression of *MTA* and *MTB* genes in knockout strains is similar to that in WT cells. The inability to mate of these two knockout strains is not caused by lack of MTA and MTB. However, *HAP2* and *ZFR1* (which are important for tight pair formation) and *CYC2* (which is essential for meiosis initiation) have higher expression in WT-co compared with WT-S, but are still low in Δ*CDK19*-co and Δ*CYC9*-co. Deficiency of these genes confirmed that mating is not induced in knockout strains.

Δ*CDK19* and Δ*CYC9* cells are disabled to finish the co-stimulation events

Transcriptome results indicated that the altered-regulated genes in Δ*CDK19* and Δ*CYC9* are important for mating initiation. Co-stimulation events were assessed to confirm whether the co-stimulation stage is normal in the Δ*CDK19* and Δ*CYC9* strains. *Tetrahymena thermophila* is known to require several events to occur for pairing (Fig. S1). As the two mating type cells are mixed, the first event to occur is co-stimulation-induced rounding (Fujishima et al., 1993): in this process, cells become rounder than in starvation conditions. Next, in tip transformation (Wolfe and Grimes, 1979), the shape of the cell tip becomes blunter and the transformed area forms a smooth-surfaced area (Suganuma et al., 1984) as various anterior cell structures are lost. Simultaneously, ConA-R appears at the head of cell (Wolfe and Feng, 1988; Wolfe et al., 1986). Once cells accomplish these events, pair formation is rapidly initiated. Therefore, we used tip transformation and ConA-R expression as indicators of the co-stimulation stage because of their high reproducibility and high similarity between cell lines.

Δ*CDK19*, Δ*CYC9* and WT cells were treated by another mating type cells of WT in a ratio of 9:1 (for example, 90% Δ*CDK19*-427 was mixed with 10% WT-428 cells). After 1 h of treatment, tip transformation and ConA-R events were detected. Fig. 6A shows that during the co-stimulation stage, the tip of most WT cells became blunter, with a certain curvature, which could be easily separated from starvation cells. In Δ*CDK19* and Δ*CYC9* cells, the tip showed no significant change. It seems that neither Δ*CDK19* nor Δ*CYC9* can complete tip transformation, which may correspond to altered regulation of the cytoskeleton in both Δ*CDK19* and Δ*CYC9* cells, consistent with the GO results (Fig. 5D; Tables S14 and S15). Similarly, ConA-R expression analysis (Fig. 6B) revealed that in Δ*CDK19* and Δ*CYC9* cells, the fluorescence signal at the anterior of the cells was not as strong as in the WT cells, (although very few cells with little signal can be detected) indicated the Δ*CDK19* and Δ*CYC9* cells cannot complete ConA-R re-localization. These results confirm that cells cannot fully enter the co-stimulation stage without CDK19 or CYC9.

Δ*CDK19* and Δ*CYC9* cells stimulate WT cells to undergo incomplete co-stimulation

The *T. thermophila* mating process requires communication to occur between both cells (Miyake, 1978). Considering the results
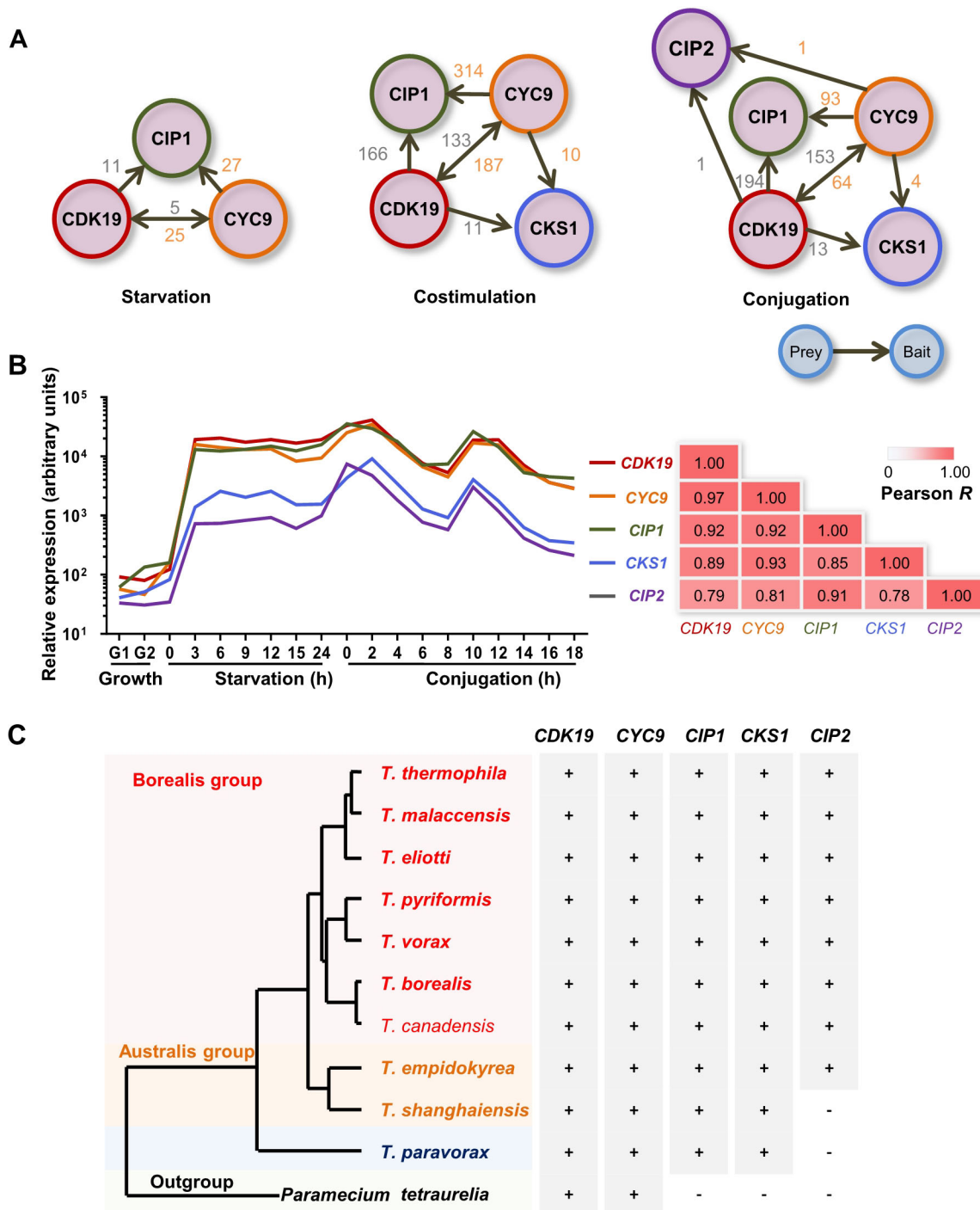
Fig. 4. Identification of the CDK19–CYC9 complex and its interacting proteins. (A) Immunoprecipitation-coupled MS identified interactors of CDK19-HA or CYC9-HA under starvation, co-stimulation and conjugation conditions. Prey means the HA-expressing protein, and Bait means interacting proteins detected by immunoprecipitation of HA-expressing protein. The gray numbers between prey and bait denote the PSMs (peptide spectrum matches) of proteins detected by CDK19-HA-expressing strains, and the orange numbers denote PSMs detected by CYC9-HA-expressing cells. (B) Expression levels of CDK19-HA, CYC9-HA and three interacting proteins under starvation, co-stimulation and conjugation conditions. Pearson correlation (Pearson R) coefficients indicate that the five genes are highly co-expressed. (C) Ortholog analysis of the five genes in sibling species of ciliates. +, homologous gene present; –, no homologous gene.

described above, it is possible that $\Delta CDK19$ and $\Delta CYC9$ cells cannot be stimulated to advance through the co-stimulation stage even when they receive the mating signal from WT cells. Nevertheless, it was of interest to ask whether the opposite is also true, that is, could the $\Delta CDK19$ and $\Delta CYC9$ cells impart a signal to WT cells that would allow them to enter co-stimulation?

After mixing WT cells with $\Delta CDK19$ or $\Delta CYC9$ cells, we inspected both tip transformation and ConA-R events. WT cells did not go through both tip transformation and ConA-R events after mixing with $\Delta CDK19$ or $\Delta CYC9$ cells (Fig. 7A,B). However, we found that $\Delta CDK19$- or $\Delta CYC9$-stimulated WT cells (Fig. 7C,D, blue line) form pairs more rapidly than unstimulated
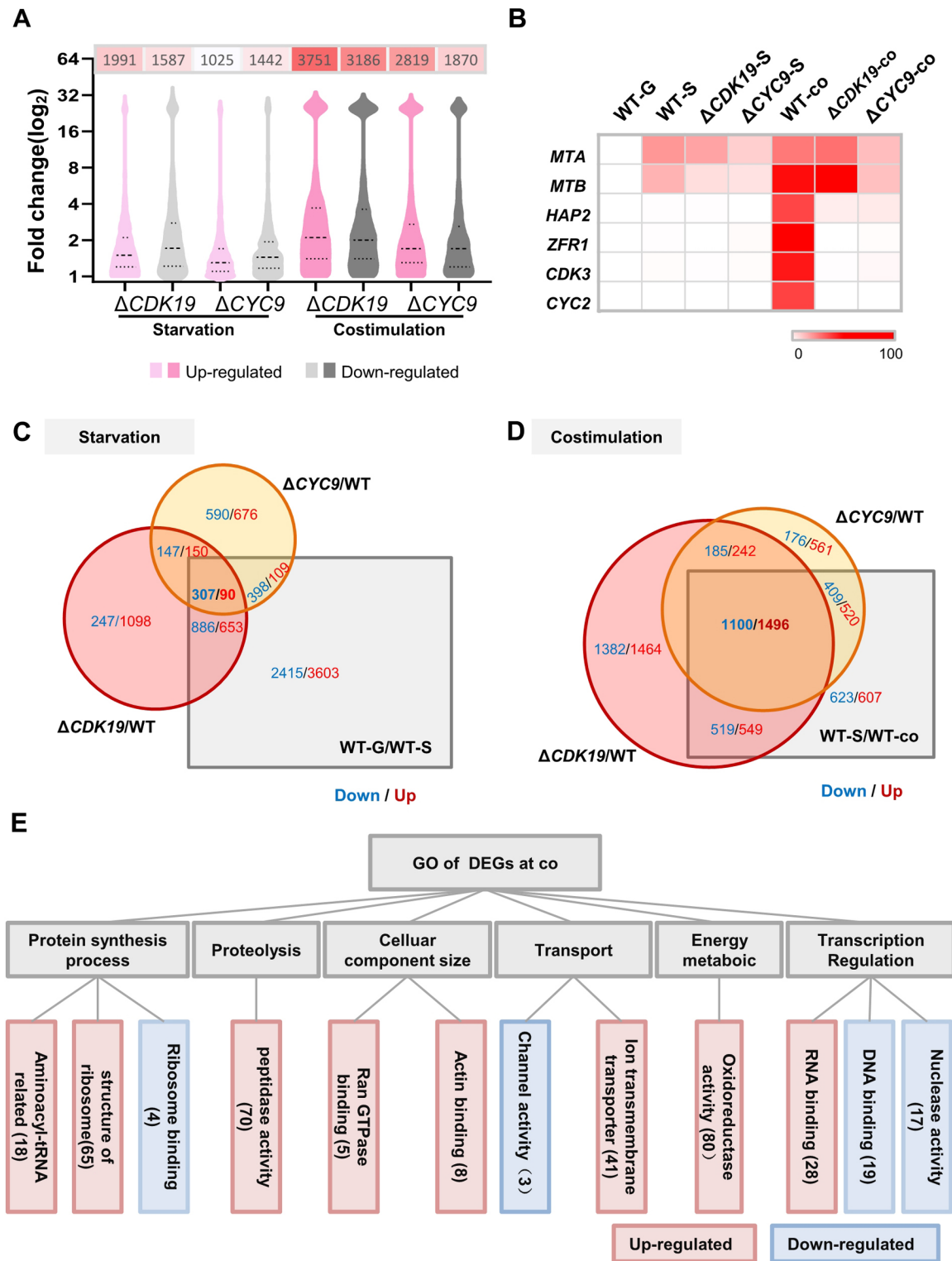
Fig. 5. CDK19 and CYC9 function during the co-stimulation stage. (A) DEGs for Δ*CDK19* and Δ*CYC9* during starvation and co-stimulation. (B) Fragments per kilobase of exon per million fragments mapped (FPKM) of mating-related genes in WT, Δ*CDK19* and Δ*CYC9*Δ*CDK19* and Δ*CYC9* of starvation and co-stimulation stage. (C,D) Venn diagrams showing overlapping DEGs of KO cells and WT cells under starvation and co-stimulation conditions. (E) GO enrichment analysis of overlapping genes at co-stimulation stage; co, co-stimulation.

cells (Fig. 7C,D, black line), although it is still slower than co-stimulation cells (Fig. 7C,D, red line). This finding suggests that both Δ*CDK19* and Δ*CYC9* cells are still capable of stimulating WT cells to proceed developmentally to some extent, but cannot fully initiate the WT to complete all the co-stimulation events. Thus Δ*CDK19* and Δ*CYC9* can stimulate WT cells to undergo incomplete co-stimulation.
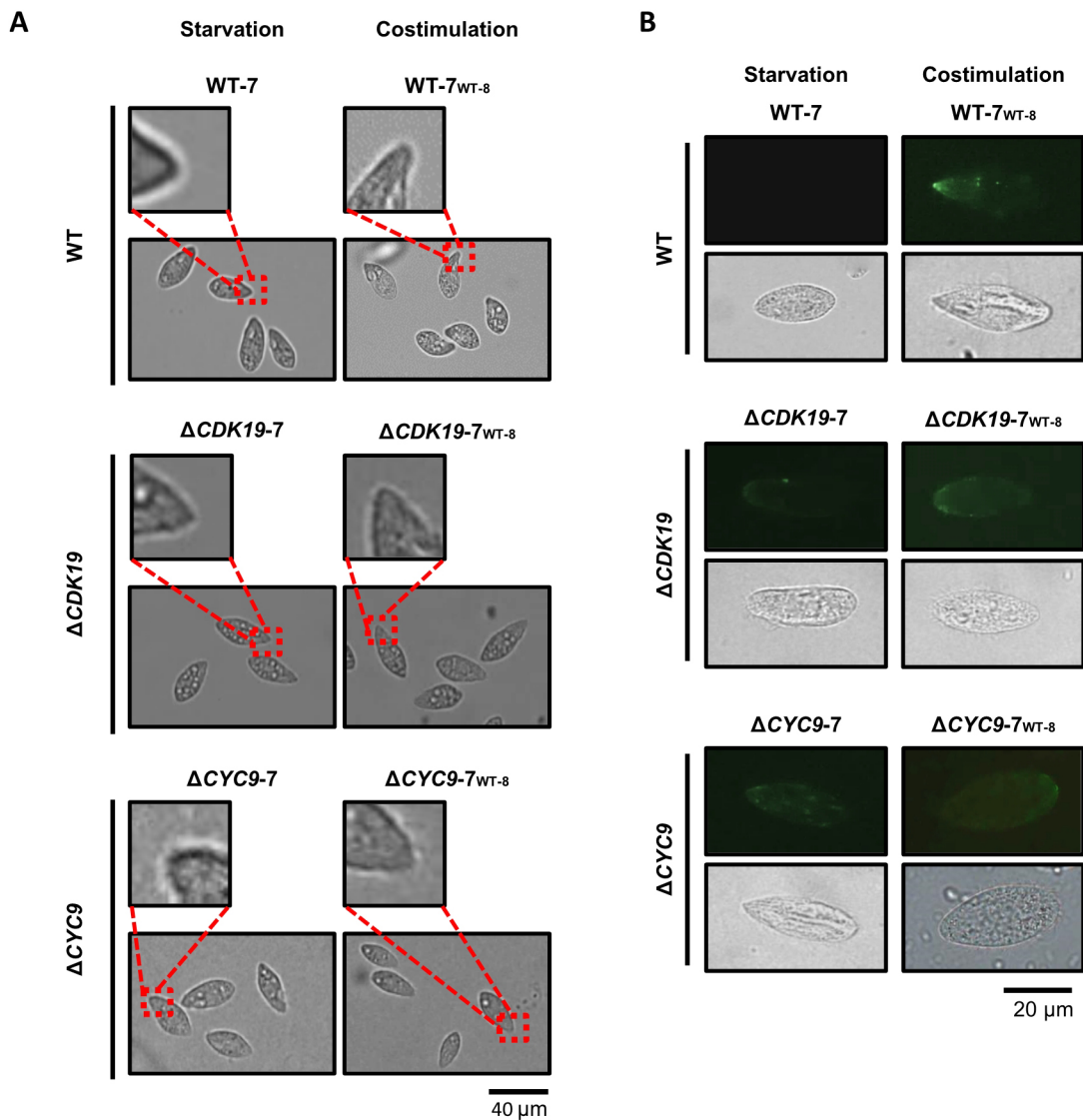
Fig. 6. ΔCDK19 and ΔCYC9 cells are disabled to finish the co-stimulation events. (A) Tip transformation in WT, ΔCDK19 and ΔCYC9 cells at starvation and co-stimulation stages (n=3). (B) ConA-R expression in WT, ΔCDK19 and ΔCYC9 cells (n=3). WT-7$_{WT-8}$ in panels A and B mean that 90% WT-7 and 10% WT-8 were mixed to acquire the stimulated WT-7 cells. The same holds for ΔCDK19-7$_{WT-8}$ and ΔCYC9-7$_{WT-8}$. '7' denotes the CU427 strain, and '8' denotes the CU428 strain.

DISCUSSION

Initiation of the sexual cell cycle is important for organisms. In this paper, we found two proteins: CDK19 and CYC9, which are important for initiating the sexual cell cycle of *T. thermophila*. Either knockout CDK19 or CYC9 prevents cells from completing co-stimulation (Fig. 6A,B). The absence of the CDK19 or CYC9 gene products leads to a change in expression of many genes (Fig. 5) that are presumably important for control of the sexual cycle as well.

According to Fig. 1, the 20 CDKs or CDK-like proteins could be classified into three groups based on the phylogenetic analysis. The characteristics of group 1 proteins are almost exactly consistent with known CDKs: they contain an ATP-binding domain, a serine/threonine protein kinase active site, and a relatively well-conserved PSTAIRE cyclin-binding domain. All group 2 CDKs have some variations in the cyclin-binding domain; CDK19 belongs to this group. Group 3 contains CDK-like proteins, with differences in size and domain composition, and

significant differences in cyclin-binding domain compared with classic CDKs.

CDK19–CYC9 was identified through immunoprecipitation mass spectrometry; however, the number of differentially expressed genes was significantly different in ΔCDK19 and ΔCYC9 strains (Fig. 5A–C). This may be due to the different protein localization of CDK19 and CYC9: CDK19 bound specifically to CYC9, and only CDK19 localized to the nucleus (whereas CYC9 did not; Fig. 3). This result indicates that a pool of CDK19 molecules does not bind to CYC9. In addition, the function of CDK/cyclin complexes may be regulated by their translocation to the correct site by the cyclin subunit (Loog and Morgan, 2005; Morgan, 2007). Therefore, CDK19 must be able to translocate into the nucleus either alone, or in combination with other proteins, rather than via CYC9.

Co-stimulation is a crucial stage for cell pairing. Our results show that CDK19 and CYC9 play important roles in this stage. The incomplete co-stimulation events in ΔCDK19 and ΔCYC9 (Figs 6 and 7) indicate that both tip transformation and ConA-R are

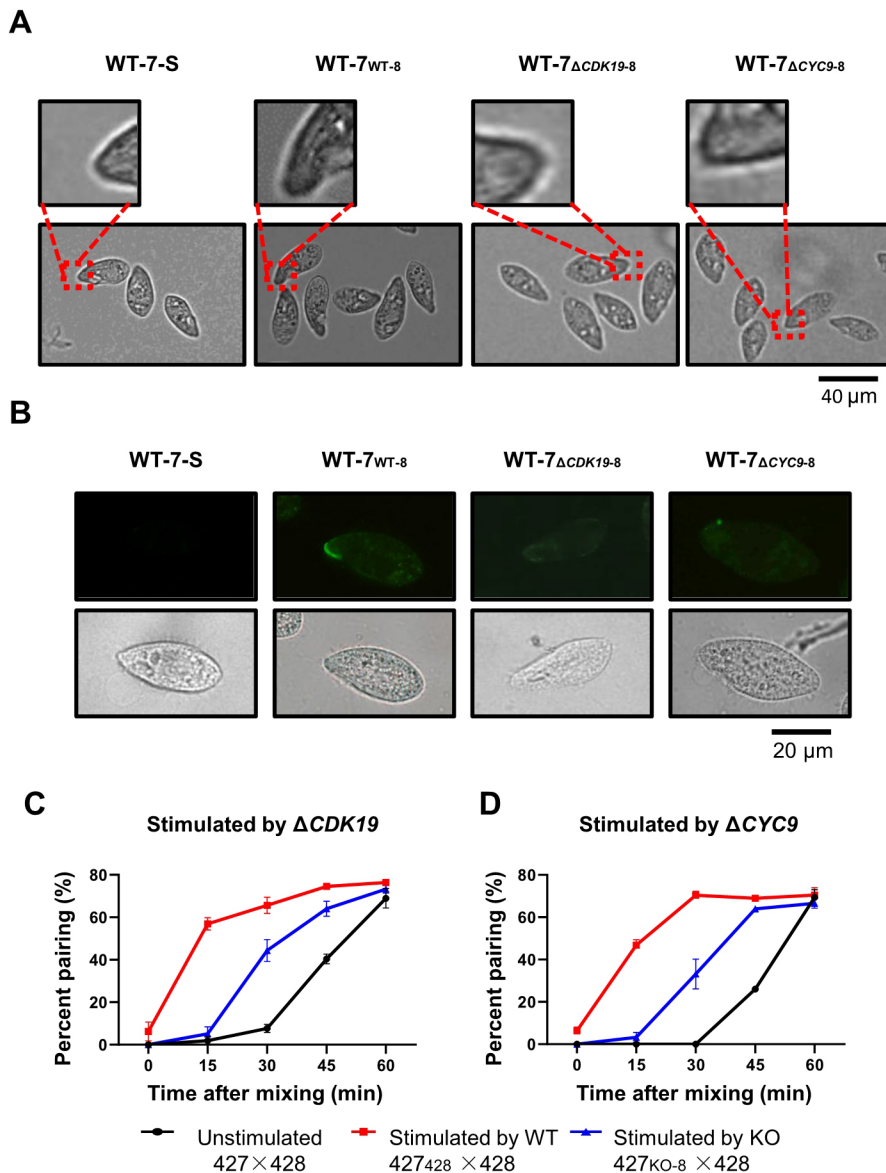**Fig. 7. WT cells are semi-stimulated by Δ*CDK19* and Δ*CYC9* cells.** (A) Tip transformation in WT cells mixed with WT (i.e. co-stimulation), Δ*CDK19* or Δ*CYC9* of another mating type (*n*=3). Only the tip of costimulated cells will bend. (B) ConA-R expression in WT cells mixed with WT, Δ*CDK19* or Δ*CYC9* of another mating type (*n*=3). None of Δ*CDK19* and Δ*CYC9* cells can stimulate ConA-R immunofluorescence in WT cells. WT-7$_{WT-8}$ in both panels A and B mean that 90% WT-7 cells were mixed with 10% WT-8 cells, same as WT-7$_{ΔCDK19-8}$, WT-7$_{ΔCYC9-8}$ and others. '7' denotes the CU427 strain, and '8' denotes the CU428 strain. (C,D) Percentage pairing of WT cells stimulated with WT, Δ*CDK19* or Δ*CYC9* cells of another mating type. Δ*CDK19* or Δ*CYC9* stimulated WT cells (blue line) more quickly than unstimulated cells (black line), but still more slowly than co-stimulation cells (red line). In panels C and D, more than 100 cells/pairs were counted, *n*=3; data are expressed as means± s.e.m.

necessary for mating, and CDK19 and CYC9 regulate these events. Δ*CDK19* or Δ*CYC9* cells still can stimulate the WT cell into an incomplete co-stimulation stage, which suggests that co-stimulation events (including some unknown events that prepare for mating) are not regulated simultaneously, but instead, might be sequential or independent of each other. CDK19–CYC9 regulates some events but not others (which may be the reason why pairing rate for WT cells mixed with Δ*CDK19* or Δ*CYC9* cells is slower than for WT cells mixed with WT cells of a different mating type). In addition, co-stimulation seems to be coincident in both mating cells. At least for the Δ*CDK19* and Δ*CYC9* strains, the events that they cannot complete are the same as the events they cannot make the WT cells do, indicating that communication between the two cells continues downstream of CDK19–CYC9 to ensure synchronization of the two cells.

In addition, as the combined transcriptional and cell cycle data show that the CDK19–CYC9 complex mainly functions in cell co-stimulation, the specific recruitment of CKS1 at the co-stimulation stage might be important for the function of the complex. Knockout CKS1 cells, unable to mate, indicate its importance for cell pairing.

Our results show that CDK19–CYC9 is an important regulatory protein complex for initiating the sexual cell cycle in *T. thermophila*. The complex influences the expression of many proteins at the co-stimulation stage (which in turn regulates co-stimulation events) and induces cells to initiate pairing. The function of the CDK19–CYC9 complex might be regulated by specific subunits, indicating a special regulatory mechanism for this CDK–cyclin complex.

MATERIALS AND METHODS
CDK identification and bioinformatics analysis
All *T. thermophila* CDKs were identified using KinBase (http://www.kinase.com/web/current/kinbase/browser/SpeciesID/10092), along with previously identified CDK genes in the Tetrahymena Genome Database (http://ciliate.org/index.php/home/welcome) (Stover et al., 2006), and the annotated domain composition was verified by InterProScan (http://www.ebi.ac.uk/interpro/) (Mitchell et al., 2015). A total of 20 proteins were selected for phylogenetic analysis, starting with multiple sequence alignment with ClustalW (Larkin et al., 2007). The alignment was used to construct a phylogenetic tree using MEGA6 with the NJ (neighbor-joining) method and 1500 bootstrap replicates (Tamura et al., 2013). All microarray data were based on TetraFGD (Xiong et al., 2013) (http://tfgd.ihb.ac.cn/).

InterProScan and CDD/SPARCLE-NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer et al., 2017) were used to determine the domain composition of all CDKs and the PSTAIRE-like sequence was used to search for cyclin-binding motifs. The multiple sequence alignment between CDK19 and other CDKs was visualized by ESPript 3 (Robert and Gouet, 2014).

Cell culture, starvation and conjugation
Tetrahymena thermophila WT strains CU427 (mating type VI) and CU428 (mating type VII) (*Tetrahymena* Stock Center; http://tetrahymena.vet.cornell.edu/) were used as controls. Cells were grown in Super Proteose Peptone (SPP) medium (1% proteose peptone, 0.2% glucose, 0.1% yeast extract, 0.003% Sequestrene) (Orias et al., 1999). Cells (at ∼2×10^5 cells/ml) were starved in 10 mM Tris-HCl (pH 7.4) for 16 h. For co-stimulation experiments, two different mating types cells were mixed at a ratio of 9:1 (Suganuma et al., 1984) for 1 h; for conjugation induction, starved cells were mixed in equal proportions.

Macronuclear gene KO and construction of HA-labeled strains
To construct KO strains, one DNA fragment (∼1 kb) upstream of the open reading frame and one downstream fragment (∼1 kb) were amplified using the primers listed in Table S16. Using fusion PCR, the two fragments were joined to NEO4 cassette, which contains a neomycin resistance gene driven by a Cd^{2+}-inducible *MTT1* metallothionein promoter. To construct the KO plasmid, the fusion fragment was then cloned into the pBlueScript SK (+) vector. The KO construct was obtained by restriction endonuclease digestion and introduced into starved WT CU427 and CU428 cells by biolistic transformation (Cassidy-Hanley et al., 1997) to obtain a KO strain of two mating types. Transformants were cultured in SPP medium containing decreasing CdCl$_2$ concentrations (from 1 to 0.05 μg/ml) and increasing paromomycin concentrations (from 0.12 to 40 mg/ml) until all WT chromosomes in the macronucleus were entirely replaced by KO chromosomes.

Strains expressing HA-labeled CDK19 and CYC9 were constructed using a knock-in method: two fragments (∼500 bp and ∼1 kb) downstream of the *CDK19* or *CYC9* gene were amplified and then ligated to NEO4 cassette, after which the fusion fragment and a fragment from the end of the *CDK19* or *CYC9* gene were cloned into the pBlueScript SK (+) backbone. The next steps are the same as used for the KO strains.

Immunofluorescence analysis
For immunofluorescence analysis, 5 ml cell samples (∼2×10^5) from the starvation (HA-expressing-427), co-stimulation (HA-expressing-427 and HA-expressing-428 cells were mixed in a ratio of 9:1 for 1 h), and conjugation stages (HA-expressing-427 and HA-expressing-428 cells were mixed in a ratio of 1:1 for 1 h) were prepared following the methods of Loidl and Scherthan (2004). An aliquot of cell suspension (80 μl) was spread onto a slide and air-dried in a fume hood. Slides were then washed twice with 1× phosphate-buffered saline (PBS) and once with PBS containing 0.05% Triton-X 100 (PBST; 5 min each). Before immunostaining with primary antibodies, non-specific binding sites were blocked by incubating slides in blocking buffer (5% milk powder and 2.25% glycine in 1× PBST), and then washing twice with PBS and once with PBST. The slides were then incubated with mouse anti-HA antibody (1:100 dilution; clone 16B12, Covance, Berkeley, CA, USA). Next, slides were washed again, fluorescein-isothiocyanate (FITC)-labeled goat anti-mouse immunoglobulin (IgG) (H+L, 1:1000 dilution, SA00003-1, Proteintech Group, Chicago, IL, USA) was applied, and slides were mounted and incubated with anti-fading agent for 2 h at room temperature, before being washed again. Finally, the slides were mounted with anti-fading agent (Vector Laboratories, Burlingame, CA, USA) supplemented with 0.5 mg/ml DAPI (4′,6-diamidino-2-phenylindole). Cells were then observed at 40× magnification by laser scanning confocal microscopy (Leica TCS SP8, Leica Microsystems, Mannheim, Germany) using Leica LAS AF lite software (http://downloads.informer.com/leica-las-af-lite/).

Transcriptome analysis
RNA was extracted, sequenced and mapped according to a method described previously (Yan et al., 2016a). Briefly, total RNA was extracted from WT cells at the starvation, co-stimulation and conjugation stages, and from Δ*CDK19* and Δ*CYC9* cells at the starvation (KO cells) and co-stimulation stages (KO cells and WT cells mixed in a ratio of 9:1 for 1 h) using the RNeasy Protect Cell Mini Kit (Qiagen), as previously described (TetraFGD: http://tfgd.ihb.ac.cn/) (Wuitschick et al., 2002). Poly-A tailed mRNA was enriched using Sera-Mag magnetic oligo (dT) beads (GE) and Illumina sequencing libraries were constructed using the manufacturer's recommendations. Paired-end (150 bp×2) sequencing was performed using an Illumina HiSeq4000 sequencer. Most sequencing data have been deposited in NCBI (GSE132677), except the growth sequencing data derived from TetraFGD. Adaptors of raw reads were trimmed with Trim-Galore (version 0.4.0) (Wu et al., 2011) and the reads were mapped to the *T. thermophila* genome (version 2014: http://ciliate.org/index.php/home/downloads) using TopHat (version 2.0.9) (Kim and Salzberg, 2011). Gene expression was quantified to the FPKM value using Cuffdiff (version 2.1.1) (Trapnell et al., 2014). The raw FPKM for each gene in cell samples from the co-stimulation stage were corrected as follows: WT-co-stimulation FPKM=(raw FPKM – 0.2×WT-conjugation)/0.8; and KO-co-stimulation FPKM=(raw FPKM – 0.1×WT-conjugation)/0.9.

DEGs were identified using a cut-off of a 2-fold change between KO and WT cells (DEGs-KO/WT) at each time point, and genes with very low expression (FPKM<10) were excluded to control for false positives. DEGs of Δ*CDK19* and Δ*CYC9* were identified at both the starvation (DEGs-KO/WT-S) and co-stimulation (DEGs-KO/WT-Co) stages. Overlapping DEGs were identified between DEGs-KO/WT-S and DEGs-WT-G/S, and also between DEGs-KO/WT-Co and DEGs-WT-S/Co. Results are presented as a Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/). GO enrichment analyses were carried out using BiNGO (version 3.2.1) (Ashburner et al., 2000; Maere et al., 2005).

Immunoprecipitation-coupled mass spectrometry
Samples from the starvation, co-stimulation and conjugation stages of HA-labeled and WT cells were prepared as described in the 'Immunofluorescence analysis' section, and were lysed using lysis buffer (30 mM Tris-HCl, 20 mM KCl, 2 mM MgCl$_2$, 1 mM phenylmethylsulfonyl fluoride, 1% Triton-X 100) containing complete proteinase inhibitor (Roche Diagnostics, Indianapolis, IN, USA) and ultrasound treatment. The supernatant was incubated with anti-HA agarose beads (Sigma) at 4°C for 2 h. After washing with lysis buffer containing 150 mM NaCl, the HA peptide was added to elute target proteins from the beads. To identify immunoprecipitating proteins, an EASY-nLC chromatography system (Thermo Fisher Scientific) was coupled online to an Orbitrap Elite instrument (Thermo Fisher Scientific) via a Nanospray Flex Ion Source (Thermo Fisher Scientific) in positive ion mode at a spray voltage of 2.2 kV. MS data were acquired in a data-dependent strategy by selecting fragmentation events based on precursor abundance in the survey scan (300–1300 Da). High-resolution mass spectra of the peptide mixture were deconvoluted using Xtract software (Thermo Fisher Scientific). Raw MS data were analyzed using Proteome Discoverer 2.1 software.

Percentage pairing experiments
Experiments were performed as follows: WT CU427 (mating type VI) and WT CU428 (mating type VII) cells were mixed in a 1:1 ratio without pre-treatment. Alternatively, CU427 cells were pre-mixed with CU428, Δ*CDK19*–428 or Δ*CYC9*–428 cells at a ratio of 9:1 for 1 h, and then mixed with 80% CU428 cells to obtain a final ratio of 1:1 of mating type VI and VII. The percentage pairing was then calculated as follows: percentage pairing (%)=(pairs×2)/[single cell+(pairs×2)]×100.

Fluorescein-labeled ConA labeling
Fluorescein ConA labeling was performed as described by Wolfe et al. (1986) with slight modifications. Samples in Fig. 6 were prepared from WT, Δ*CDK19* and Δ*CYC9* cells at the starvation and co-stimulation (cells were mixed with WT cells of another mating type in 9:1 ratio for 1 h) stages. The co-stimulation sample in Fig. 7 was prepared as follows: 90% WT cells were mixed with 10% WT, Δ*CDK19* and Δ*CYC9* of another mating type for 1 h, respectively. All samples were fixed first and washed with 0.1 M PB buffer three times, then cells were incubated with fluorescein-labeled ConA (Vector Laboratories) at a final concentration of 10 μg/ml. They were then

washed a further three times, and cells were immediately observed at 40×
magnification by fluorescence microscopy (Olympus BX51).

Tip transformation

WT, Δ*CDK19* and Δ*CYC9* cells were mixed with WT cells of another
mating type in a ratio of 9:1. After mixing, 20 µl samples of living cells were
taken at the starvation and co-stimulation stages (after mixing for 1 h). They
were observed immediately at 20× magnification under a light microscope
and photographed (Olympus BX51).

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: Y.M., G.Y., J.X., W.M.; Methodology: Y.M.; Software: G.Y., Y.M.;
Validation: Y.M.; Formal analysis: Y.M., G.Y.; Investigation: Y.M., X.H., J.Z.;
Resources: W.M.; Data curation: Y.M., G.Y.; Writing - original draft: Y.M.; Writing -
review & editing: Y.M., G.Y., W.M.; Visualization: Y.M.; Supervision: J.X., W.M.;
Project administration: J.X., W.M.; Funding acquisition: W.M

Data availability
RNA sequencing data have been deposited in Gene Expression Omnibus
(GSE132677), except for the growth sequencing data derived from TetraFGD
(http://tfgd.ihb.ac.cn/).

Supplementary information
Supplementary information available online at
http://jcs.biologists.org/lookup/doi/10.1242/jcs.235721.supplemental

References
Allewell, N. M. and Wolfe, J. (1977). A kinetic analysis of the memory of a
developmental interaction. Mating interactions in *Tetrahymena pyriformis*. *Exp.
Cell Res.* **109**, 15-24. doi:10.1016/0014-4827(77)90039-8

**Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M.,
Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T. et al.** (2000). Gene
ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat.
Genet.* **25**, 25-29. doi:10.1038/75556

Bruns, P. J. and Brussard, T. B. (1974). Pair formation in *Tetrahymena pyriformis*,
an inducible developmental system. *J. Exp. Zool.* **188**, 337-344. doi:10.1002/jez.
1401880309

Bruns, P. J. and Palestine, R. F. (1975). Costimulation in *Tetrahymena pyriformis*: a
developmental interaction between specially prepared cells. *Dev. Biol.* **42**, 75-83.
doi:10.1016/0012-1606(75)90315-2

**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.

**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. doi:10.1371/journal.pbio.1001518

**Cole, E. S., Cassidy-Hanley, D., Fricke Pinello, J., Zeng, H., Hsueh, M., Kolbin,
D., Ozzello, C., Giddings, T., Jr, Winey, M. and Clark, T. G.** (2014). Function of
the male-gamete-specific fusion protein HAP2 in a seven-sexed ciliate. *Curr. Biol.*
24, 2168-2173. doi:10.1016/j.cub.2014.07.064

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. doi:10.1006/dbio.1993.1018

Garfinkel, M. D. and Wolfe, J. (1981). Alterations in gene expression induced by a
specific cell interaction during mating in *Tetrahymena thermophila*. *Exp. Cell Res.*
133, 317-324. doi:10.1016/0014-4827(81)90323-2

Harper, J. W. (2001). Protein destruction: adapting roles for Cks proteins. *Curr. Biol.*
11, R431-R435. doi:10.1016/S0960-9822(01)00253-6

Jain, B. P. and Pandey, S. (2018). WD40 repeat proteins: signalling scaffold with
diverse functions. *Protein J.* **37**, 391-406. doi:10.1007/s10930-018-9785-7

Kim, D. and Salzberg, S. L. (2011). TopHat-Fusion: an algorithm for discovery of
novel fusion transcripts. *Genome Biol.* **12**, R72. doi:10.1186/gb-2011-12-8-r72

**Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A.,
McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al.** (2007).
Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948. doi:10.1093/
bioinformatics/btm404

Loidl, J. and Scherthan, H. (2004). Organization and pairing of meiotic
chromosomes in the ciliate *Tetrahymena thermophila*. *J. Cell Sci.* **117**,
5791-5801. doi:10.1242/jcs.01504

Loog, M. and Morgan, D. O. (2005). Cyclin specificity in the phosphorylation of
cyclin-dependent kinase substrates. *Nature* **434**, 104-108. doi:10.1038/
nature03329

Maere, S., Heymans, K. and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to
assess overrepresentation of gene ontology categories in biological networks.
Bioinformatics **21**, 3448-3449. doi:10.1093/bioinformatics/bti551

**Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F.,
Derbyshire, M. K., Geer, R. C., Gonzales, N. R. et al.** (2017). CDD/SPARCLE:
functional classification of proteins via subfamily domain architectures. *Nucleic
Acids Res.* **45**, D200-D203. doi:10.1093/nar/gkw1129

**Mitchell, A., Chang, H.-Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R.,
McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S. et al.** (2015). The InterPro
protein families database: the classification resource after 15 years. *Nucleic Acids
Res.* **43**, D213-D221. doi:10.1093/nar/gku1243

Miyake, A. (1978). Chapter 3 cell communication, cell union, and initiation of meiosis
in ciliate conjugation. *Curr. Top. Dev. Biol.* **12**, 37-82. doi:10.1016/S0070-
2153(08)60593-1

Morgan, D. O. (2007). *The Cell Cycle: Principles of Control. Primers in Biology*.
London, UK: New Science.

Morgan, D. O. (1995). Principles of CDK regulation. *Nature* **374**, 131-134. doi:10.
1038/374131a0

Obaya, A. J. and Sedivy, J. M. (2002). Regulation of cyclin-Cdk activity in
mammalian cells. *Cell. Mol. Life Sci.* **59**, 126-142. doi:10.1007/s00018-002-
8410-1

Orias, E. (2014). Membrane fusion: HAP2 protein on a short leash. *Curr. Biol.* **24**,
R831-R833. doi:10.1016/j.cub.2014.08.004

Orias, E., Hamilton, E. P. and Orias, J. D. (1999). Chapter 4 tetrahymena as a
laboratory organism: useful strains, cell culture, and cell line maintenance. *Method
Cell Biol.* **62**, 189-211. doi:10.1016/S0091-679X(08)61530-7

Orias, E., Singh, D. P. and Meyer, E. (2017). Genetics and epigenetics of mating
type determination in *Paramecium* and *Tetrahymena*. *Annu. Rev. Microbiol.* **71**,
133-156. doi:10.1146/annurev-micro-090816-093342

**Pinello, J. F., Lai, A. L., Millet, J. K., Cassidy-Hanley, D., Freed, J. H. and Clark,
T. G.** (2017). Structure-function studies link class II viral fusogens with the
ancestral gamete fusion protein HAP2. *Curr. Biol.* **27**, 651-660. doi:10.1016/j.cub.
2017.01.049

Robert, X. and Gouet, P. (2014). Deciphering key features in protein structures with
the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324. doi:10.1093/nar/
gku316

Ron, A. and Horovitz, O. (1977). Increased RNA synthesis during preconjugation
and its effect on pair formation in *Tetrahymena*. *Experientia* **33**, 1146-1149.
doi:10.1007/BF01922294

**Stover, N. A., Krieger, C. J., Binkley, G., Dong, Q., Fisk, D. G., Nash, R.,
Sethuraman, A., Weng, S. and Cherry, J. M.** (2006). Tetrahymena Genome
Database (TGD): a new genomic resource for *Tetrahymena thermophila* research.
Nucleic Acids Res. **34**, D500-D503. doi:10.1093/nar/gkj054

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.* **33**, 10-18. doi:10.1093/oxfordjournals.
jmicro.a050431

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6:
molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**,
2725-2729. doi:10.1093/molbev/mst197

**Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H.,
Salzberg, S. L., Rinn, J. L. and Pachter, L.** (2014). Erratum: corrigendum:
differential gene and transcript expression analysis of RNA-seq experiments
with TopHat and Cufflinks. *Nat. Protoc.* **9**, 2513-2513. doi:10.1038/nprot1014-
2513a

Wolfe, J. and Feng, S. (1988). Concanavalin-a receptor tipping in tetrahymena and
its relationship to cell-adhesion during conjugation. *Development* **102**, 699-708.

Wolfe, J. and Grimes, G. W. (1979). Tip transformation in *Tetrahymena*: a
morphogenetic response to interactions between mating types. *J. Eukaryot.
Microbiol.* **26**, 82-89. doi:10.1111/j.1550-7408.1979.tb02737.x

Wolfe, J., Pagliaro, L. and Fortune, H. (1986). Coordination of concanavalin-a-
receptor distribution and surface differentiation in *Tetrahymena*. *Differentiation* **31**,
1-9. doi:10.1111/j.1432-0436.1986.tb00375.x

Wu, Z., Wang, X. and Zhang, X. (2011). Using non-uniform read distribution models
to improve isoform expression inference in RNA-Seq. *Bioinformatics* **27**, 502-508.
doi:10.1093/bioinformatics/btq696

Wuitschick, J. D., Gershan, J. A., Lochowicz, A. J., Li, S. and Karrer, K. M.
(2002). A novel family of mobile genetic elements is limited to the germline

genome in *Tetrahymena thermophila*. *Nucleic Acids Res.* **30**, 2524-2537. doi:10.1093/nar/30.11.2524

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* **2013**, bat008. doi:10.1093/database/bat008

Xiong, J., Yang, W., Chen, K., Jiang, C., Ma, Y., Chai, X., Yan, G., Wang, G., Yuan, D., Liu, Y. et al. (2019). Hidden genomic evolution in a morphospecies – the landscape of rapidly evolving genes in *Tetrahymena*. *PLoS Biol.* **17**, e3000294. doi:10.1371/journal.pbio.3000294

Xu, C. and Min, J. (2011). Structure and function of WD40 domain proteins. *Protein Cell* **2**, 202-214. doi:10.1007/s13238-011-1018-1

Yan, G.-X., Zhang, J., Shodhan, A., Tian, M. and Miao, W. (2016a). Cdk3, a conjugation-specific cyclin-dependent kinase, is essential for the initiation of meiosis in *Tetrahymena thermophila*. *Cell Cycle* **15**, 2506-2514. doi:10.1080/15384101.2016.1207838

Yan, G. X., Dang, H., Tian, M., Zhang, J., Shodhan, A., Ning, Y.-Z., Xiong, J. and Miao, W. (2016b). Cyc17, a meiosis-specific cyclin, is essential for anaphase initiation and chromosome segregation in *Tetrahymena thermophila*. *Cell Cycle* **15**, 1855-1864. doi:10.1080/15384101.2016.1188238

Yang, W., Jiang, C., Zhu, Y., Chen, K., Wang, G., Yuan, D., Miao, W. and Xiong, J. (2019). *Tetrahymena* Comparative Genomics Database (TCGD): a community resource for *Tetrahymena*. *Database* **2019**, baz029. doi:10.1093/database/baz029

Fig. S1

Costimulation (~30 min)

Cell recognition   **Pair formation**

Starvation

Pronuclei mitosis

Meiosis

Reciprocal fertilization

Sexual

Postzygotic mitosis

Asexual

Rounding   ConA-R appear   Pairing
Smooth-surfaced area
Tip-transformation

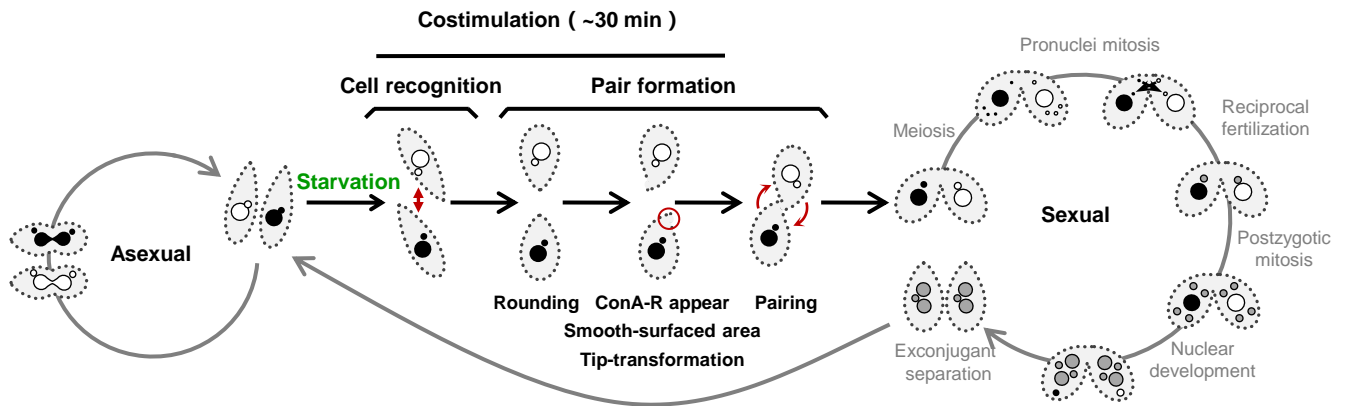Exconjugant separation

Nuclear development

Fig. S1. Life cycle of *T. thermophila*. The schematic diagram includes the asexual (vegetative growth period), sexual cell cycle (conjugation) and most importantly, the costimualtion stage, during which stage cells will undergo a series of events to prepare for pairing. *T. thermophila* enter sexual cell cycle after starvation induction. During costimulation stage, cells first undergo costimulation-induced rounding, then form a smooth-surfaced area at the in front of oral, tip transformation, and concanavalin A receptor (ConA-R) appear. After that, cells form pairs and initate meiosis.
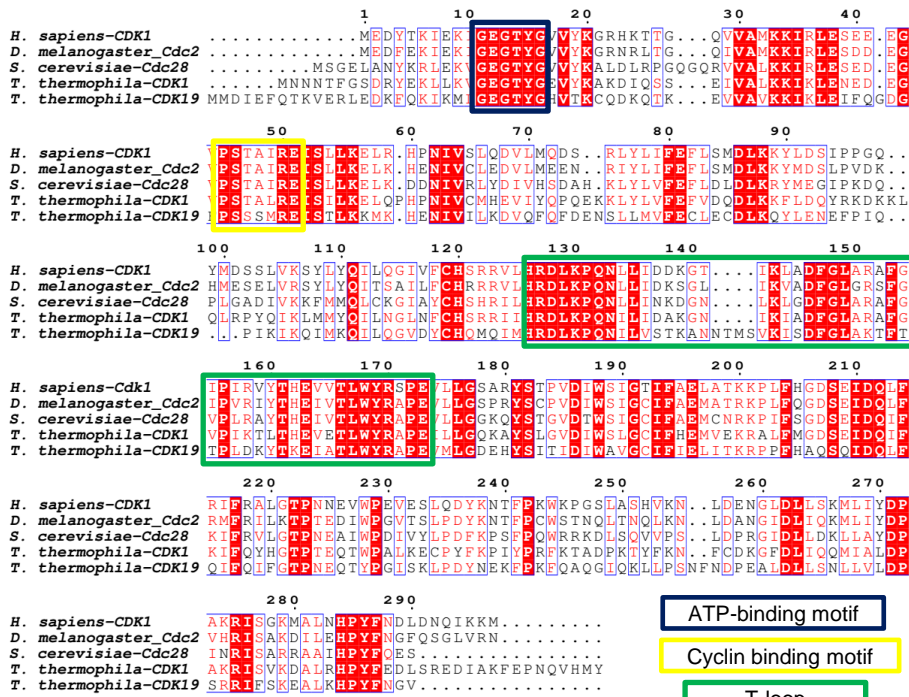
Fig. S2

Fig. S2. Multiple sequence alignment of *T. thermophila CDK1* and *CDK19*, *H. sapiens CDK*1, *S. cerevisiae CDC28* and *D. melanogaster CDC2*.
Three important characteristics of CDKs: ATP-binding motif, Cyclin binding motif and T-loop are shown in deep blue, yellow and green box respectively.
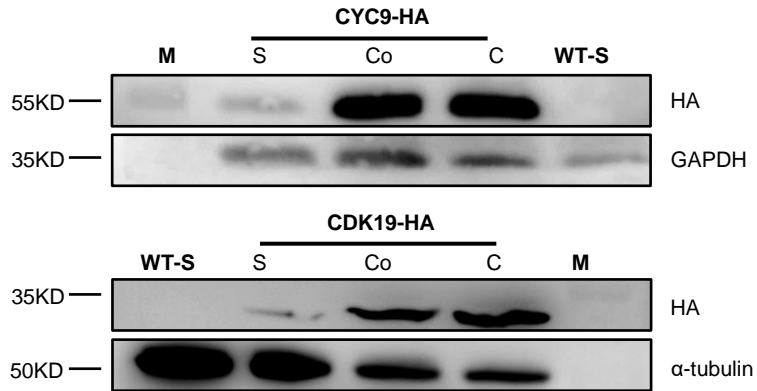
Fig. S3

Fig. S3. CDK19-HA and CYC9-HA expression at starvation, costimulation and conjugation stages. M: marker; WT-S: WT cells from starvation stage; S: starvation; Co: costimulation; C: conjugation;
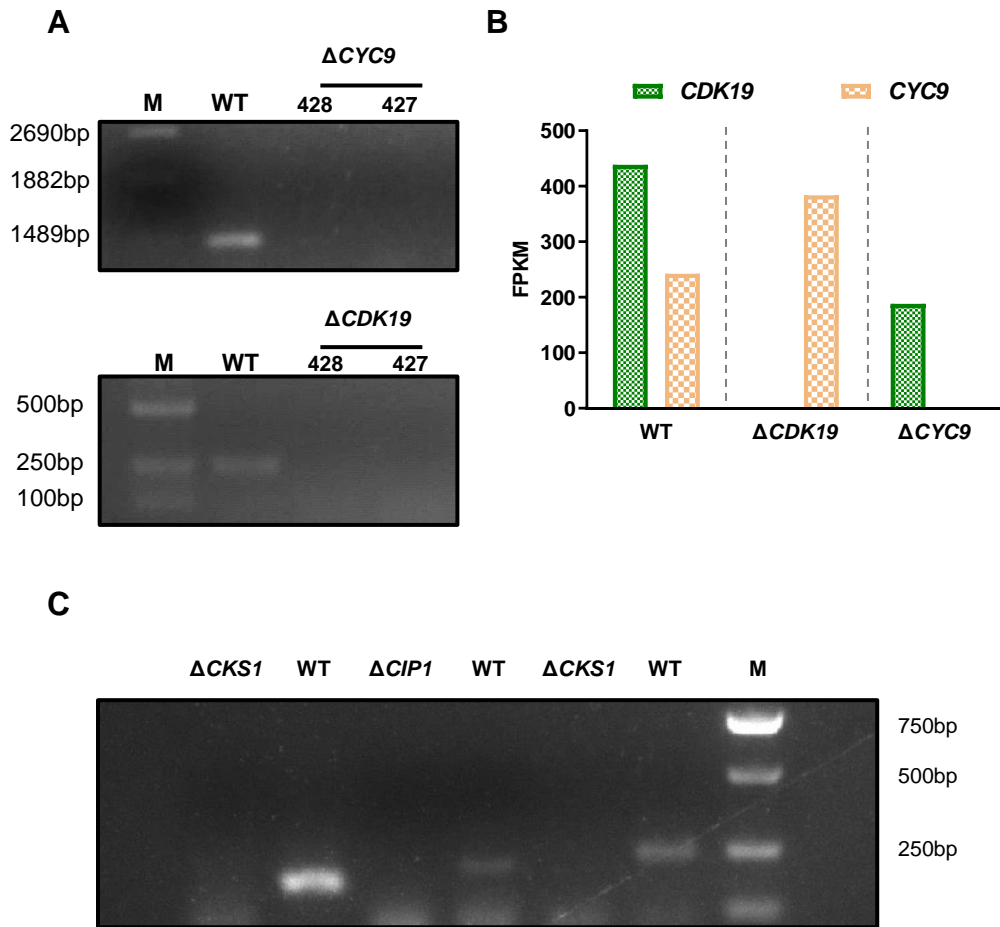
Fig. S4

A

B

C

Fig. S4. Verification of CDK19 and CYC9 gene knock out.(A) PCR analysis ΔCDK19 and ΔCYC9 using primers targeting gene coding sequence. Absence of a brand indicates gene KO. (B) RNA-seq shows no CDK19 expression in ΔCDK19 cells and no CYC9 expression in ΔCYC9 cells. 427, CU427; 428, CU428 strains. M: marker. (C)PCR analysis Δ*CKS1*, Δ*CIP1* and Δ*CIP2* using primers targeting gene coding sequence.

Table S1 Gene IDs and names of all CDKs or CDK-likes.

Gene ID	Gene name
TTHERM_00411810	Cdk1
TTHERM_00318700	Cdk10
TTHERM_00394590	Cdk14
TTHERM_01035490	Cdk15
TTHERM_00624260	Cdk16
TTHERM_00066860	Cdk17
TTHERM_00784290	Cdk18
TTHERM_00339880	Cdk19
TTHERM_01207660	CDC2
TTHERM_00011670	Cdk3
TTHERM_00286770	Cdk4
TTHERM_01080590	Cdk5
TTHERM_00717780	Cdk8
TTHERM_00185770	Cdk9
TTHERM_00483640	TCdk3
TTHERM_01094780	CDK21 (new named)
TTHERM_00267860	Cdk13
TTHERM_00576780	Cdk11
TTHERM_01347900	Cdk20
TTHERM_000837979	CDK6 (misannotation before)

Table S2 IP-MS results. (CDK19-HA Conjugation)

Click here to Download Table S2

Table S3 IP-MS results. (CDK19-HA Costimulation)

Click here to Download Table S3

Table S4 IP-MS results. (CDK19-HA Starvation)

Click here to Download Table S4

Table S5 IP-MS results. (CYC9-HA Conjugation)

Click here to Download Table S5

Table S6 IP-MS results. (CYC9-HA Costimulation)

Click here to Download Table S6

Table S7 IP-MS results. (CYC9-HA Starvation)

Click here to Download Table S7

Table S8 IP-MS results. (WT Conjugation-1)

Click here to Download Table S8

Table S9 IP-MS results. (WT Conjugation-2)

Click here to Download Table S9

Table S10 IP-MS results. (WT Costimulation-1)

Click here to Download Table S10

Table S11 IP-MS results. (WT Costimulation-2)

Click here to Download Table S11

Table S12 IP-MS results. (WT Starvation-1)

Click here to Download Table S12

Table S13 IP-MS results. (WT Starvation-2)

Click here to Download Table S13

Table S14 GO enrichment results of overlapping DEGs in costimulation stage. (down-regulated)

Click here to Download Table S14

Table S15 GO enrichment results of overlapping DEGs in costimulation stage. (up-regulated)

Click here to Download Table S15

Tabel S16 primers used in this work.

Primer name	Sequence (5' - 3')
CDK19-5F-Not1	AGTTCTAGAGCGGCCGCATAGCAAGAAATTTAAAGAGGA
CDK19-5R-Neo4	GTCAGGTGCCTGGTACCCTTTATCTATACTACCCAATATTTCA
CDK19-3F-Neo4	CTGACGTCGCACCATCCCTTAATATAATATGCATTAATGAGAA
CDK19-3R-Not1	ACCGCGGTGGCGGCCGCACTTGGTCAGGTAGTCTATTTTA
CYC9-5F-Neo4	CTGACGTCGCACCATCCCAAATCAAGTGAAACTTTAATGTATC
CYC9-5R-NotI	ACCGCGGTGGCGGCCGCTTGACAGACTCAGCTACTTATTTAT
CYC9-3F-NotI	AGTTCTAGAGCGGCCGCTTATTTGGTGTTAACTGGTTG
CYC9-3R-Neo4	GTCAGGTGCCTGGTACCCAATCAATATGAATATGAATGAGC
CDK19-CF-Not1	AGTTCTAGAGCGGCCGCTTAAGTTAGAGATTTTTTAAGGTGA
CDK19-CR-HA	GTCGGGGACGTCGTAGGGGTATCCATTAATCAAACACCATTAAAGTAAG
CDK19-3NF-HA	CCCTACGACGTCCCCGACTACGCCATAATATGCATTAATGAGAATATATTC
CDK19-3NR-Neo4	GTCAGGTGCCTGGTACCCACAATTATATTGATTATGTTTCTAGC
CDK19-3-TF-Neo4	CTGACCGTCGCACCATCCCACATTTCAAATTAAAATTCAGTAAG
CDK19-3-TR-Not1	ACCGCGGTGGCGGCCGCGTCTTAAGAAGTTTTTCCATTTATT
CYC9-TF-Not1	AGTTCTAGAGCGGCCGCAAATAACAATTCTTTAAACATAGGA
CYC9-TR-Neo4	CTGACCGTCGCACCATCCCATGAGGATTTAAAGTATAAACTAATTT
CYC9-NF-Neo4	GTCAGGTGCCTGGTACCCTTTAATCCTGAAATTAAAATATCAC
CYC9-NR-HA	CCCTACGACGTCCCCGACTACGCCCTTATTTGTCTAATCAATATGAATATG
CYC9-CF-HA	GTCGGGGACGTCGTAGGGGTATCCCTTTTTTGATTTTGGATTAGAGAGA
CYC9-CR-Not1	ACCGCGGTGGCGGCCGCTGATTGAAGTTTTTGGTAATTACAA
CDK19-check-F	CACTATTCAATTACAATTGATATTT
CDK19-check-R	ATCATTAAAATTAGAAGGTAGAAGT
CYC9-check-F	ATGAACACATTACTTCAAA
CYC9-check-R	AAGAACTTAAAGATTGGATTCACTA
CIP1-5F-Xba1	GCTCTAGAGCTAAATAAACTAATAGATTTTGGTTT
C1P1-5R-BamH1	CGGGATCCCGTAAAGTGTGATTAATTTTATTTTGT
CIP1-3F-Sal1	ACGCGTCGACGTCGGCCATAGCGGCCGCGGAATTTTCTCATTTTCATCTATAAAATT
CIP1-3R-Kpn1	GGGGTACCCCGGAACACATTAAAACAGATAGAAGT
CKS1-5F-Xba1	GCTCTAGAGCCTCCAGACATAATTATGTATTTTT
CKS1-5R-BamH1	CGGGATCCCGTTGTGATAAAATAGAAATAAAGAGT
CKS1-3F-Sal1	ACGCGTCGACGTCGGCCATAGCGGCCGCGGAAATCCTTAAATTATTCAACTAACTAA
CKS1-3R-Kpn1	GGGGTACCCCAAACAGTTATCTTTCAAATGATAAA
CIP2-5F-Xba1	GCTCTAGAGCAAGAATTTTCATTTATAATACAAGA
CIP2-5R-BamH1	CGGGATCCCGAGGATTTTTATTATAATATTGAGTG
CIP2-3F-SAL1	ACGCGTCGACGTCGGCCATAGCGGCCGCGGAATTTAGTTATCATTTACTTCCACATA
CIP2-3R-Kpn1	GGGGTACCCCTTCAAAGTAAATTGTTTTAGACTTA
CKS1-check-F	CTAAAGAAATCGAATACTCTGACAC
CKS1-check-R	ATAATGAATGGTTCTGGTTTATAGA
CIP1-check-F	TAACTAATAGACTGAGATTGGTTAA
CIP1-check-R	TGTATCTTAGAAATATAGTGTGGGT
CIP2-check-F	ACTATAAATGTCAGTTAATCCTACC
CIP2-check-R	CCTTATCATTATTCTTAGGAGTTGA