The characterization of *cnjA*, a *Tetrahymena* gene active only during meiosis

by

Angelika Rosenauer

Department of Microbiology

McGill University

Macdonald Campus

Montreal

Quebec, Canada

A Thesis Submitted to

The Faculty of Graduate Studies and Research

In partial fulfillment of the requirement for the degree of

Master of Science

June, 1993

Short title

The conjugation specific gene, *cnjA*, from *T. thermophila*

ABSTRACT

MSc Angelika Rosenauer Microbiology

The characterization of *cnjA*, a *Tetrahymena* gene active only during meiosis

The nucleotide sequence of the *cnjA* cDNA (formerly pC1) from the ciliated protozoan *Tetrahymena thermophila* was determined. This gene was previously found to be conjugation specific and to peak in expression just prior to or at pachytene in meiotic prophase I. The *cnjA* message is initiated from four transcription start sites, one minor and three major, and encodes a putative protein (CnjA) of 779 amino acids. The protein has a calculated molecular weight of 89.5 kDa and is mainly hydrophilic with an estimated pI of 9.3. CnjA was found to share no sequence similarities with any known protein to date. The gene's coding region demonstrates an unusual codon choice. Flanking regions of the *cnjA* genomic locus were amplified by means of the Inverse PCR method but attempts at subcloning and characterizing its promoter region were unsuccessful.

RÉSUMÉ

MSc Angelika Rosenauer Microbiologie

Caractérisation de *cnjA*, gène de *Tetrahymena* actif

seulement pendant la méiose

La séquence nucléotidique du *cnjA d'ADN-c* (anciennement pC1) d'un protozoaire cilié *Tetrahymena thermophila* a été déterminée. Il a déjà été démontré que ce gène était spécifique de conjugaison et que son expression atteignait son maximum juste avant ou pendant le stade pachytène de la prophase méiotique I. Le message *cnjA* a pour point de départ quatre sites de transcription, un mineur et trois majeurs et encode une protéine hypothétique (*CnjA*) de 779 acides aminés. Selon les calculs, le poids moléculaire de la protéine est de 89,5 kDA; elle est principalement hydrophile avec un pI évalué à 9,3. La *CnjA* n'affiche aucune similarité de séquence avec d'autres protéines connues. La région d'encodage du gène affiche un choix de codon inhabituel. Les régions latérales du locus génomique de *cnjA* ont été amplifiées par la méthode PCR inverse mais les tentatives de sous-clonage et de caractérisation de son site promoteur se sont soldées par un échec.

ACKNOWLEDGEMENTS

I would like to thank a number of people for making this work a pleasant and fun experience.

For technical assistance, I would like to thank Costa Karatzas for the introduction to the IPCR method and for his help with the procedure. To Dr. R. Dhindsa (Department of Biology; McGill University) for kindly allowing me to perform some final experiments in his lab. Special thanks go to my good friend and lab pal Csilla Csank, who taught me the ropes when I first came to the lab and who offered help and support when I needed it most.

To all my lab friends; Brent Wisse for his sometimes unsolicited but always powerfully stimulating comments and for his function as my weight trainer. Frances Taylor, Zeng-Ming Gu and Xiau Peng Feng for their friendship, support and helpfulness. I would also like to thank Alison Jones for her friendship and often invaluable solid and cool-headed advice.

Many thanks go to all the staff and students of the Microbiology Department, especially to Marlene Parkinson, who was always helpful and went out of her way to get things done with a smile and to Emily Magee for her administrative assistance.

Personal thanks are directed to my mother for her superhuman patience and support throughout this period.

My deepest gratitude is bestowed on Duane Martindale who allowed me to work in his wonderful lab. For teaching me to think independently and therefore giving me a solid base on which to build on in the future. His always being available for questions and being, who he is, an overall great person to have known.

TABLE OF CONTENTS

3. MATERIALS AND METHODS

4. RESULTS

LIST OF FIGURES

I. INTRODUCTION

1.1 The organism

The eukaryote *Tetrahymena thermophila* is a unicellular, ciliated protozoan. Ciliates are unique in that they demonstrate nuclear dimorphism by accomodating two functionally distinct nuclei in the same cytoplasm; a small germline micronucleus and a large somatic macronucleus.

The diploid micronucleus (10 bivalent (n=5) chromosomes that appear metacentric at metaphase of meiosis I) (Ray, 1956; Bruns *et al.*, 1983) contains the entire genetic complement of *T. thermophila* and serves as a deposit of nuclear genetic information. This information is passed on to the sexual progeny at the time of conjugation. The germline nucleus is usually transcriptionally silent (Mayo and Orias, 1981) and divides mitotically during the vegetative cell cycle. It takes part in the nuclear events that occur during conjugation and is transcriptionally active only during the brief interval of meiotic prophase I (Sugai and Hiwatashi, 1974; Martindale *et al.*, 1985).

The somatic macronucleus is polyploid (45C in G1), transcriptionally active and responsible for the phenotype of the cell. It divides amitotically (centromeres and metaphase plates are missing) by partitioning approximately equal amounts of chromatin to the two daughter cells (Raikov, 1969; Preer and Preer, 1979). The macronucleus does not contribute DNA to the sexual progeny, but new macronuclei are formed from micronuclei through genetic exchange and recombination at the time of conjugation.

1.1.2 Conjugation

Conjugation represents the sexual stage of the ciliate life cycle. This developmental pathway can be divided into two main micronuclear events. 1) Prezygotic events: cell pairing, meiosis, gametogenesis and fertilization. 2) Postzygotic events: postzygotic mitosis and macronuclear differentiation. *T. thermophila* can be induced to undergo synchronous conjugation (Martindale *et al.*, 1982) by meeting certain requirements. Cells must be mature (approximately 50 fissions) (Nanney, 1974), of different mating type and starved for nutrients (Bruns and Brussard, 1974a).

1.1.3 Prezygotic events

Once starved cells of different mating type are mixed, costimulation occurs (Finley and Bruns, 1980) and conjugation is initiated as pairs form through a temporary oral junction. Meiosis begins when the micronucleus leaves its cuplike pocket on the macronuclear surface and starts to elongate. Meiotic prophase I is characterized by the progressive elongation of the micronucleus resulting in the unique ciliate structure - the micronuclear crescent. At its maximum elongation, the crescent is U-shaped and encompasses an area twice the size of the cell (Sugai and Hiwatashi, 1974; Wolfe *et al.*, 1976). During micronuclear prophase I, the nuclear envelope does not break down and no structures resembling synaptonemal complexes have been observed (see Orias, 1986). Martindale *et al.* (1982) carefully timed the cytological stages of conjugation. Meiotic prophase I was separated into cytological stages I to

VI corresponding to the degree of micronuclear elongation as initially described by Sugai and Hiwatashi (1974). Stage IV represents full crescent elongation and may correspond cytologically to pachytene of meiotic prophase I. Both Sugai and Hiwatashi (1974) and Martindale *et al.* (1985) observed accumulation of newly synthesized RNA over the crescent, peaking at stages III and IV just prior to pachytene. It is also at this time that the transcripts of the early conjugation specific genes (*cnjA* (pC1), *cnjB* (pC2) and *cnjC* (pC7)) peak in expression (Martindale and Bruns, 1983; Martindale *et al.*, 1985). After the crescent shortens, the reductional division of meiosis I and the equational division of meiosis II follow rapidly, generating four haploid meiotic products. One randomly selected product remains functional and positions itself near the oral junction whereas the other three haploid products move posteriorly and disintegrate. The surviving meiotic nucleus divides once mitotically (gametogenesis), forming two haploid nuclei of identical genotypes. These become the migratory and stationary pronuclei. The migratory pronuclei are quickly transferred to the opposite mate where they fuse with the resident stationary pronucleus (cross-fertilization) resulting in one diploid nucleus per cell (reviewed by Orias, 1986).

1.1.4 **Postzygotic development**

The fertilized nucleus divides mitotically twice, generating four diploid products two of which move posteriorly and remain micronuclear in size and DNA content. The other two products take up a more anterior position and become the macronuclear anlagen. During the process of macronuclear differentiation, the anlagen increase in DNA content, undergo DNA rearrangement and DNA loss.

1.1.5 Macronuclear differentiation

The micronuclear chromosomes are fragmented at specific sites (reviewed by Yao, 1989), generating subchromosomal lengths and creating new physical linkages to which telomeres are added (Blackburn and Gall, 1978; reviewed in Blackburn and Karrer, 1986). Coincident with fragmentation, 10-20% of micronuclear DNA is eliminated, decreasing the sequence complexity of the macronucleus from 1.5×10^8 bp to 1.3×10^8 bp (Yao and Gorovsky, 1974). DNA sequences retained by the macronucleus are amplified to a ploidy level of 45 times the haploid DNA content. In the mature macronucleus the single rDNA gene is amplified to 18,000 copies that are arranged on 9,000 identical linear palindromic chromosomes (see Yao, 1986).

The macronuclear and micronuclear anlagen position themselves centrally and one micronucleus is eventually destroyed. The old macronucleus moves to the posterior of the cell and begins to disintegrate. The two cells then separate (exconjugants) and each retains two new macronuclei and one micronucleus. The old macronucleus is eventually destroyed. Refeeding is required for the exconjugants to undergo their first post-zygotic division. The macronucleus does not divide but one whole macronucleus is passed on to the daughter cells (karyonides) whereas the micronucleus divides mitotically once passing one copy to each daughter cell. In the karyonides, the normal genetic complement is restored resulting in four identical genotypes that differ from that of their parents. Although the macronuclear anlagen are of the same genotype initially, macronuclear genotypes of each karyonide may differ due to phenotypic assortment (Sonnenborn, 1975). This phenomenon is due to the amitotic

division of the macronucleus during vegetative growth which results in seemingly unequal segregation of identical sets of chromosomes. The progeny, originally heterozygous for a dominant and recessive allele, will initially express the dominant phenotype but subclones eventually stably express the recessive phenotype. Thus strains may express a particular phenotype while possessing a different micronuclear genotype (Bruns and Brussard, 1974b).

2. LITERATURE REVIEW

Meiosis is a tightly controlled developmental stage and is conserved among sexually reproducing eukaryotic organisms. The process is necessary to assure the reduction of a diploid genome (2N) to a haploid genome (N). Meiosis in most organisms follows a defined program of DNA duplication, pairing and recombination of duplicated chromosomes, orderly segregation of chromosomes (a reductional division) followed by an equational division of the chromosomes that results in the formation of four gametes. The two unicellular eukaryotes, *Saccharomyces cerevisiae* and *T. thermophila,* share some features necessary for the entry into meiosis. Both organisms have to be of different mating type and starved of nutrients. The timing of meiotic events are tightly controlled in both (Martindale *et al.*, 1982; Padmore *et al.*, 1991). *T. thermophila* differs from *S. cerevisiae* in that meiosis is not immediately followed by cell division. In *T. thermophila,* mating induces meiosis and thus meiosis is part of the conjugational process. The meiotic nuclear divisions occur within a single cell (see Orias, 1986). In contrast, conjugation precedes meiosis in *S. cerevisiae* and the two processes are separate. Two hap-

loid cells of opposite mating type fuse (conjugate) in order to form diploid cells that are capable of vegetative growth until starvation triggers meiosis and sporulation (Hartwell, 1974).

The circumstances governing meiosis and sporulation in *S. cerevisiae* are starting to be well understood. A brief account describing early sporulation events in this organism will follow with special emphasis on meiotic prophase I.

2.1 Meiosis in *S. cerevisiae*

There are two levels of control governing the entry of cells into meiosis in *S. cerevisiae;* a mating type control that is primarily effected at the level of transcription and a nutritional control which is determined by the lack of nitrogen and a fermentable carbon source.

2.1.1 Mating type control

In wild type yeast, only *MAT*a/*MAT*α diploid cells can enter meiosis and sporulate under nutrient limitations (Esposito and Klapholz, 1981; reviewed by Malone, 1990). The diploid specific gene product of the *MAT* locus, a1-α2 (a heterodimer) directly represses the transcription of the haploid specific gene *RME1* (negative regulator of meiosis) (Mitchell and Herskowitz, 1986; Dranginis, 1990; Covitz *et al.*, 1991). Inactivation of *RME1* allows the transcription of *IME1* (positive regulator of meiosis), the gene governing the decision to enter meiosis (Kassir *et al.*, 1988). Both *MAT* control and nutritional control converge at *IME1* (Kassir *et*

al., 1988; Smith and Mitchell, 1989; Matsuura *et al.*, 1990). Overexpression of *RME1* does not fully repress sporulation in *MAT*a/*MAT*α diploids and conversely a *rme1* deletion mutation does not support wild type levels of sporulation (Mitchell and Herskowitz, 1986) suggesting another route(s) exists that communicates its signal(s) through the *MAT* locus to *IME1*. A second pathway, also influenced by mating type and nutritional state, seems to act through the gene *IME4* (Shah and Clancy, 1992). This gene was found to be essential for the accumulation of the *IME1* transcript. Additionally, the semidominant mutation *RES1-1* (Kao *et al.*, 1990) supports wild type levels of sporulation when *RME1* is overexpressed in diploid strains under starvation conditions but sporulates poorly in *mat*-insufficient strains. The double mutation *rme1/RES1-1* expresses wild type levels of *IME1* and supports higher levels of sporulation than either mutation alone, indicating that wild type *RES1* acts independently of *RME1*. Neither the function of the *RES1* locus nor its regulation by the *MAT* locus are clear as yet (Kao *et al.*, 1990). The double mutant also expresses wild type levels of *IME1* in the absence of *IME4*, but sporulates poorly, suggesting an additional role for *IME4* during sporulation. Overexpression of *IME4* promotes sporulation in *mat*-insufficient strains whether or not the *rme1* or *RES1-1* bypass mutations are present (Shah and Clancy, 1992), indicating that *IME4* acts in a different pathway from either *RME1* or *RES1*. Furthermore, *IME4* is essential for high levels of *IME1* accumulation and sporulation since *ime4* null mutants do not accumulate sufficient *IME1* and are unable to sporulate (Shah and Clancy, 1992). *RME1* and *IME4* seem to occupy antagonistic pathways. *RME1* suppresses *IME1* expression in haploid cells whereas *IME4* activates *IME1* under sporulation conditions. Both pathways appear to be leaky. *RME1* and *IME4* are both expressed at low levels in diploid cells but the

IME4 transcript is greatly induced under starvation conditions (Mitchell and Herskowitz, 1986; Shah and Clancy, 1992). It has been suggested by Shah and Clancy (1992) that together both genes confer absolute cell type specificity on sporulation.

One of the earliest events following the starvation of a/α cells is the accumulation of the *IME1* RNA (Kassir *et al.*, 1988; Smith and Mitchell, 1989; Mitchell *et al.*, 1990). *IME1* has been shown to permit genetic recombination through the activation of early sporulation specific (ss) genes (Kassir *et al.*, 1988; Smith and Mitchell, 1989) such as *IME2*, *SPO11*, *SPO13*, *HOP1* (Smith *et al.*, 1990). Expression of *IME1* from a GAL1 promoter enabled the moderate accumulation of the *IME2*, *SPO11*, *SPO13* and *HOP1* transcripts in nutrient rich medium but the transcripts of the middle to late ss genes *SPS1* and *SPS2* (normally induced by *IME1*) did not accumulate until cells were starved (Smith *et al.*, 1990). An *ime1* null mutant blocks the expression of these genes altogether under sporulation conditions (Smith *et al.*, 1990). The *SME1 (IME2)* gene product (another activator of meiosis) was identified as a serine threonine kinase homolog (Yoshida *et al.*, 1990) and is itself an early ss gene subject to *MAT* and nutritional control (Smith and Mitchell, 1989). *IME1* and *IME2* act together to allow efficient sporulation. Increased levels of both transcripts stimulate meiotic recombination in vegetative cells (especially at low glucose levels) but do not promote the meiosis I division or spore formation (Smith and Mitchell, 1989). Overexpression of *IME2* partially suppresses the *ime1* null mutation defect (1% of normal sporulation) and allows for some recombination (Mitchell *et al.*, 1990). Since an *ime1* null mutant demonstrates a more severe sporulation defect than an *ime2* null mutant and *IME2* can only partially

relieve the *ime1* defect, Mitchell *et al.* (1990) suggested that *IME1* may play an additional role during sporulation apart from the activation of *IME2*. In wild type strains, *IME1* expression peaks early in meiosis and then decreases. The *IME1* transcript level remains high throughout meiosis when *IME2* is absent and the induction of *SPO11*, *SPO13*, *SPS1* and *SPS2* is delayed along with a depression in their transcript accumulation (Mitchell *et al.*, 1990). This suggests that *IME2*, apart from activating early ss genes, is also a negative regulator of the *IME1* transcript and involved in the temporal expression of these ss genes.

Another positive regulator of meiosis and spore formation is the gene product of *MCK1* *(IME3)* a serine threonine kinase homolog that was independently characterized by Neigeborn and Mitchell (1991) and Shero and Hieter (1991). Mck1* has been identified to play three roles. In vegetative cells, the kinase governs centromere function during mitosis (Shero and Hieter, 1991) and in meiosis it is responsible for the optimal expression of *IME1* and stimulation of ascospore maturation (Neigeborn and Mitchell, 1991). Overexpression of *IME1* in *mck1* mutants supports high levels of sporulation but spores do not form mature asci, whereas overexpression of *MCK1* greatly accelerates sporulation through *IME1* induction but does not inhibit *RME1* expression in *mat*-insufficient strains. *MCK1* therefore seems to operate downstream of *RME1* in the same pathway as *IME1* and may play a role in the temporal regulation of the sporulation process (Neigeborn and Mitchell, 1991).

* when referring to the protein products of the genes discussed, the first letter will be written in uppercase followed by lowercase lettering.

Strich *et al.* (1989) isolated mutations in five complementation groups, called *ume1-ume5* (unscheduled meiotic gene expression) that allowed expression of the early ss genes *SPO13*, *SPO11* and *SPO16* during vegetative growth but did not affect expression of later ss genes like *SPO12* and *SPS2*. The wild type alleles of the *UME* genes seem to express negative regulators of meiosis that act downstream of cell type and nutritional control; either downstream of *IME1* or in parallel. It has been suggested that *IME1* and/or *IME2* act to activate early ss genes through the negative regulation of the *UME* genes (Srich *et al.*, 1989).

2.1.2 Nutritional control: the Ras/cAMP pathway in *S. cerevisiae*

The Ras/cAMP pathway functions during the vegetative growth of *S. cerevisiae* in rich medium. Genes that are involved in the nutritional regulation of meiosis are also involved in mitotic growth. The Ras proteins in the pathway interact in a complex manner with components of signal transduction (reviewed by Broach, 1991).

The membrane bound Ras (GTP-binding/GTP-hydrolyzing protein), encoded by the *RAS1* and *RAS2* loci, transmits a signal to adenylate cyclase (Toda *et al.*, 1985). When activated, Ras is charged with GTP and capable of directly or indirectly stimulating adenylate cyclase activity (encoded for by the *CYR1* locus) (Toda *et al.*, 1985; Kataoka *et al.*, 1985). Binding of GTP to Ras is facilitated by the *CDC25* protein (GTP/GDP exchange factor) (Robinson *et al.*, 1987; Broek *et al.*, 1987). The Ras activity is down-regulated by the product of the *IRA* genes (*IRA1*, *IRA2*) (Tanaka *et al.*, 1989; 1990a) which encode GTPase acti-

vating proteins that are structurally and functionally homologous to mammalian GAP proteins (Tanaka *et al.*, 1990b; 1991). Adenylate cyclase amplifies the signal by increasing the intracellular cAMP level and cAMP acts directly to stimulate the activity of the cAMP-dependent protein kinase A (Toda *et al.*, 1985). The protein kinase exists as a heterotetramer made up of two regulatory subunits (encoded by the *BCY1* locus) (*Toda et al.*, 1987a) and two catalytic subunits (encoded by three loci: *TPK1, TPK2, TPK3*) (Toda *et al.*, 1987b). Binding of cAMP to the regulatory subunits results in the release of the catalytic subunits which are then free to phosphorylate their various cellular targets. High kinase activity stimulates the breakdown of storage carbohydrates, activates the glycolytic pathway, induces a large number of growth specific genes and down-regulates the pathway itself (Broach, 1991). Low kinase activity stimulates the accumulation of cabohydrates, activates the gluconeogenesis pathway, decreases transcription of growth specific genes and induces various stress related genes (Broach, 1991).

Mutations affecting the pathway result in two distinct phenotypes. 1) If the pathway is hyperactivated (i.e. $RAS^{val\text{-}19}$, *ira, bcy1*), cells are capable of normal mitotic growth but exhibit abnormal phenotypes. These mutants show increased sensitivity to heat shock and starvation and diploid strains are unable to sporulate (Broek *et al.*, 1987; Toda *et al.*, 1985; Tanaka *et al.*, 1990; Matsumoto *et al.*, 1982). 2) If the pathway is depressed (i.e. *ras2, cyr1, cdc25*), cells arrest unbudded in G1 of the cell cycle. Diploid strains homozygous for these mutations leave the mitotic cycle and undergo meiosis in rich medium (Toda *et al.*, 1985; Matsumoto *et al.*, 1982; Robinson *et al.*, 1987). The general trend seems

to indicate that increased levels of cAMP prevent the entry into meiosis whereas low cAMP levels allow diploid strains to sporulate

2.1.3 Early events in meiotic prophase I: Meiosis at the level of the chromosome in *S. cerevisiae*

Meiotic prophase I can be divided into six cytological stages: preleptotene, leptotene, zygotene, pachytene, diplotene and metaphase I (Swanson *et al.*, 1981). Each stage is characterized by the degree of chromatin condensation and observable cytological events. Padmore *et al* (1991) have timed and examined landmarks of cytological events during prophase I in the rapidly sporulating strain SK1. The first meiotic observable event is the completion of bulk DNA replication, doubling the DNA content of the cell (preleptotene). In wild type yeast, this process is completed within three hours of meiotic initiation (Padmore *et al* 1991) The chromatin is diffuse during bulk DNA replication and becomes progressively more condensed until it is tightly packed at pachytene (Dresser and Giroux, 1988). During leptotene, the chromosomes are long and slender with some degree of contraction. This stage is initiated approximately 1.3 hours after bulk DNA replication and lasts for about 30 minutes (Padmore *et al.*, 1991). Here the lateral elements of the synaptonemal complex (SC) develop and can be observed as unpaired, short axial core segments that eventually form the lateral elements of the SC (Alani *et al.*, 1990). In wild type yeast, the first physically observable event of recombination is the appearance of site-specific double-stranded breaks (DSBs) as determined by the analysis of the meiotic recombination hot spots, *HIS4-LEU2* (Cao *et al.*, 1990) and *ARG4* (Nicolas *et al* 1989; Sun *et al.*, 1991). The DSBs are transient and can occur prior to or simultaneous

[redacted]

[redacted]

[redacted]

to proper orientation with the meiotic spindle at metaphase I (Baker *et al*., 1976). Only 15 minutes elapse between the dissociation of the SC and the formation of the meiotic spindle. These steps occur sequentially (Padmore *et al*., 1991).

2.1.4 Gene expression during meiotic prophase I in *S. cerevisiae*

There are two classes of genes recruited during meiotic prophase I in *S cerevisiae*. Class I consists of genes that are important for both mitotic and meiotic events. Class II is made up of genes that are required for recombination and chromosome segregation in meiotic cells and play no detectable role in mitosis.

Class I

The *RAD50-57* series of radiation repair genes were originally isolated as mutations that make the cell sensitive to ionizing radiation and methyl methanesulfonate (MMS) treatment (Resnick, 1987). This sensitivity is due to the inability of the mutants to undergo the DSB repair that would normally occur after UV damage (Resnick and Martin, 1976). Diploid lesions in any of these genes affect either the sporulation rate or the production of viable progeny (Game, 1983).

The *RAD50* gene product (Alani *et al*., 1990; Cao *et al*., 1990) is required for recombination and chromosome synapsis. Two types of mutations are identified. The *rad50* deletion mutation blocks meiotic recombination and SC formation at a very early step. These mutants form only short to

medium length axial core elements that fail to synapse, are defective in inter- and intrachromosomal recombination and do not form DSBs. The *rad50S* mutation (which contains a point mutation that maps within the ATP-binding site) blocks recombination and SC formation at an intermediate step. The block occurs at the connection between the 3' single-stranded tail generation and tripartite SC formation. These mutants form DSBs at the normal time but do not generate 3' single-stranded tails. DSBs accumulate stably and long axial core elements are formed with rare tripartite SC. All *rad50* mutations block the formation of tripartite SC but permit the formation of axial core elements. In wild type yeast both processes are coupled (Padmore *et al.*, 1991).

The *RAD51* (Shinohara *et al.*, 1992; Aboussekhra *et al.*, 1992; Basile *et al.*, 1992) and *DMC1* (a meiosis specific gene of class II) (Bishop *et al.*, 1992) gene products share sequence homology with the RecA protein of *Escherichia coli*. RecA catalyzes rapid DNA homology searches resulting in strand exchange between homologous DNAs. These complexes form stable, joined molecules that are necessary for recombination (Conley and West, 1989). Deletion mutations in *rad51*, *rad52*, *rad54* and *dmc1* share similar phenotypes (Game and Mortimer, 1974; Bishop *et al.*, 1992; Shinohara *et al.*, 1992). These include the accumulation of meiosis specific DSBs containing intact 3' single-stranded tails with shorter and more heterogenous 5' ends than are found in the wild type. The DSBs appear to the same extent as in wild type but accumulate to a higher degree and persist (Cao *et al.*, 1990; Bishop *et al.*, 1992). These mutants show a decrease in reciprocal recombination with mature recombinant molecules appearing at normal times but with greatly reduced frequency (Bishop *et al.*, 1992). They also fail to form normal

SCs and arrest late in meiotic prophase, resulting in low spore formation, with the rare spore formed being inviable. It has been shown that Rad51 binds directly to Rad52, suggesting that both proteins act together in the processing of DSBs to the next intermediate (Sinohara *et al.*, 1992). In mitosis, *rad51*, *rad52* and *rad54* block radiation induced mitotic recombination (Game and Mortimer, 1974) and the wild type gene products are also required for mating type switching which involves site-specific DSBs (Malone and Esposito, 1980).

Two proteins catalyzing an ATP-independent single-strand exchange reaction, STPα and STPß, are expressed in both mitotic and meiotic cells (Sugino *et al.*, 1988; Dykstra *et al.*, 1990). Their level of expression is constant under both physiological conditions. The specific activity of these proteins is greatly induced during meiosis. STPα is encoded by the locus *DST1* (Clark *et al.*, 1991) and has been identified as DNA strand transferase I. This protein contains a zinc finger motif, shows sequence similarity to the TFIIS protein (involved in transcript elongation by RNA polymerase II) and acts in homolog pairing during meiosis. The STPß protein (encoded by the *DST2* locus) has been identified as DNA strand transferase II (Dykstra *et al.*, 1991). In the *dst1* and *dst2* deletion mutants, the induction of DNA strand transfer activity and intrachromosomal recombination are greatly reduced but both mutants produce viable spores. The DNA strand transferase activity seems to be involved in meiotic reciprocal exchange but not in spore formation.

Class II

The *REC102* gene product (Cool and Malone, 1992; Bhargara *et al.*,1992) is necded early in meiosis at approximately the time of genetic recombination. The protein contains leucine zippers at its C-terminal end and may associate with other proteins. The *rec102* deletion mutant displays a decrease in intrachromosomal recombination and its phenotype resembles that of the other *rec* mutants. Entry into meiosis is normal and the majority of cells go through both the meiosis I and II divisions, but both divisions are delayed by 1-2 hours.

The *SPO11* gene product (Atcheson *et al.*, 1987) is required for meiotic recombination and acts at or soon after chromosome pairing. In the *spo11-1* disruption mutant, meiotic recombination and ascospore viability are drastically reduced. This defect results in aneuploidy as a result of random segregation at the reductional division in the absence of exchange. Only 1% of spores are viable and these are nonrecombinant. The *spo11-1* mutant does not block tripartite SC formation and thus displays a recombination defect despite the appearance of normal SC (Klapholz *et al.*, 1985).

The *MER1* protein is required for chromosome pairing and genetic recombination (Engebrecht and Roeder, 1989; 1990). The deduced protein sequence has a limited similarity to calmodulin without the calcium-binding motif. The gene product is essential for wild type levels of gene conversion, reciprocal recombination and SC assembly. The *mer1* mutants accumulate full length but unpaired axial core elements which fail to synapse completely (full synapsis is only 10% of wild type

level). Both gene conversion and reciprocal recombination are reduced 10 times over the wild type levels. The defect is due to a failure in pairing and/or recombination which results in the production of inviable spores.

The *MER2* gene is transcribed both in mitosis and meiosis. The transcript is spliced by Mer1 during meiosis generating a functional Mer2 (Engebrecht *et al.*, 1991). *MER2* was originally identified as a repressor of the *mer1* defect. Overexpression of *MER2* in a *mer1* background restores gene conversion and SC assembly but does not correct for the reciprocal recombination defect and only partially restores spore viability. This suggests that the *MER2* gene product is involved in both gene conversion and SC assembly and that the role of Mer1 is not only to activate Mer2 but may play additional role(s) in meiosis.

The *HOP1* protein is a structural element of the meiotic chromosome and is situated along its length (Hollingsworth *et al.*, 1990). The protein contains a zinc finger motif at its C-terminal end that has been shown to be important for its function. Hop1 promotes pairing and is required for SC formation. The *hop1* mutant does not form SCs, has decreased levels of intrachromosomal recombination and a deficiency in homolog pairing. This defect results in the random segregation of chromosomes at the meiosis I division due to a failure in forming chiasmata (Hollingsworth and Byers, 1989).

The *MRE4* (*MEK1*) protein shares sequence homology to serine threonine kinases and is required early in meiosis (Leem and Ogawa, 1992). In the *mre4* deletion mutant, the frequency of DSBs is decreased to 10-20% of

the wild type level suggesting that protein phosphorylation is involved in early recombinational events. The mutants produce inviable spores due to chromosome nondisjunction at the time of the first meiotic division.

The *MEI4* gene product is involved early on in meiosis and it is needed for the induction of recombination and viable spore production (Menees *et al.*, 1992). The *MEI4* gene contains an intron at its 5' nontranslated region for which splicing is not dependent on Mer1. The *mei4* null mutation, like *mer1* and *rec102*, accumulates full length but unpaired axial core elements that fail to form functional SCs. Meiosis II division is delayed for 1 hour in the *mei4* mutant but spore formation occurs at wild type levels even though gene conversion and reciprocal exchange are completely eliminated. The mutation leads to inviable meiotic progeny.

Both the *SPO12* (Malavasic and Elder, 1990) and *SPO13* (Wang *et al.*, 1987; Buckingham *et al.*, 1990) gene products are required for proper chromosome segregation at the meiosis I division. Mutants in these genes fail in chromo ome disjunction but are able to completely bypass the reductional division. They undergo a single meiosis II-like equational division with two viable diploid spores as products (Klapholz and Esposito, 1980).

The *RED1* gene product is also required for the meiosis I division and for SC assembly (Rockmill and Roeder, 1988: Thompson and Roeder, 1989). The *red1* mutant sporulates efficiently. The spores are inviable, due to aneuploidy resulting in chromosome nondisjunction during the reductional division. The mutants demonstrate normal homolog

pairing and meiotic exchange that occur with wild type kinetics but without forming a SC. Interchromosomal recombination is decreased but the mutants are proficient in intrachromosomal exchange (25% of wild type level) (Rockhill and Roeder, 1990).

2.1.5 Regulation and transcription patterns of early meiotic genes

Two negative regulatory sites in the 5' upstream region of *RME1* (negative regulator of meiosis) have been identified (Covitz *et al.*, 1990). An a1-α2 repressor site that allows direct repression via binding of the a1-α2 heterodimer (Covitz *et al*, 1990; Dranginis, 1990) and an URS site (upstream repressor sequence) with sequence identity to the CAR1 upstream repressor site. The *RME1* protein contains three zinc fingers that are characteristic of nucleic binding proteins, of which at least two were found to be necessary for its function (Covitz *et al.*, 1990). The *IME1* transcript contains an unusually long 5' upstream region allowing for both positive and negative regulation (Granot *et al.*, 1989). An UAS (upstream activating sequence) appears to positivly regulate *IME1* expression, possibly through an inducer and a 0.5 kb upstream region that seems to be a cis-acting negative regulatory element (behaves like a URS sequence), capable of binding *RME1* (Granot *et al.*, 1989).

The early meiotic genes are all regulated at the level of transcription. They are either not expressed during vegetative growth or are only detectable at low levels, whereas their transcripts are greatly induced during meiosis. Several of these genes are co-regulated. A URS1 (ubi-

quitous mitotic repressor element) consensus sequence can be found 100-200 bases upstream of the translation initiation codon of *SPO13, SPO16, HOP1, RED1, MRE4, MEI4, MER1, DMC1* and *SME2(IME2)* (Wang *et al.*, 1987; Malavasic and Elder, 1990; Hollingsworth *et al.*, 1990; Rockmill and Roeder, 1988; Leem and Ogawa, 1992; Menees *et al.*, 1992; Engebrecht and Roeder, 1990; Bishop *et al.*, 1992 and Yoshida *et al.*, 1990). This sequence is also found within the *SPO11* coding region (Buckingham *et al.*, 1990; Atcheson *et al.*, 1987) but is absent altogether in *REC102* (Buckingham *et al.*, 1990; Cool and Malone, 1992). Interestingly, two URS1 consensus sequences are found upstream of the genes *MRE4* and *SME1(IME2)*, both of which encode putative kinases. The role of the URS1 sequence in regulating meiotic expression has been shown. A point mutation within this region in the *SPO13* gene allowed the lacZ fusion protein to be expressed during vegetative growth (Buckingham *et al.*, 1990).

Many of these early meiosis specific genes have multiple sized transcripts. Two transcripts are detected for *RED1, SPO13, HOP1* and *RAD50* in which the shorter transcript is the more abundant (Thompson and Roeder, 1989; Wang *et al.*, 1987; Hollingsworth *et al.*, 1990; Farnet *et al.*, 1988). *MRE4* encodes four transcripts that are the result of differential termination within the gene's 3' nontranslated region (Leem and Ogawa, 1992) and the same holds true for *MER1* which is encoded by three transcripts (Engebrecht and Roeder, 1990). In either case, one transcript is the more abundant. The expression patterns for *SPO11, SPO13, DMC1, MRE4, HOP1 , MER1, RED1*and *REC102* are similar to that of *IME2* (Mitchell *et al.*, 1990; Atcheson *et al.*, 1987; Buckingham *et al.*, 1990; Bishop *et al.*, 1992; Leem and Ogawa, 1992; Hollingsworth *et al.*,

1990; Engebrecht and Roeder, 1990; Thompson and Roeder, 1989; Cool and Malone, 1992; Yoshida *et al.*, 1990; Smith and Mitchell, 1989). These transcripts are specifically induced within 1.5-2.0 hours after diploid cells are placed into sporulation medium and peak in expression at 4 hours. This time corresponds to early-mid prophase when the SC is in its earliest stage of formation and recombination related DSBs occur (Padmore *et al.*, 1991). Interestingly, *SPO11* is required for recombination and *SPO13* is necessary for chromosome segregation yet both genes show indistinguishable transcription patterns (Klapholz *et al.*, 1985; Wang *et al.*, 1987). In contrast *SPO12* which is also required for chromosome segregation and is thought to act at the same time as *SPO13*, is induced 2-3 hours later (Klapholz and Esposito, 1980; Malavasic and Elder, 1990). This suggests that the detectable mutant phenotypes do not always coincide with the induction of their transcripts. The genes essential for pairing and recombination, *MER1, RED1, SPO13, SPO11* and *HOP1* accumulate their transcripts with similar kinetics (Engebrecht and Roeder, 1990; Thompson and Roeder, 1989; Wang *et al.*, 1987; Atcheson *et al.*, 1987; Hollingsworth *et al.*, 1990). Expression patterns of the *RED1* transcript and the *red1::lacZ* fusion protein are identical (Thompson and Roeder, 1989; Menees *et al.*, 1992) and the same holds true for the *MER1* and its lacZ fusion protein (Engebrecht and Roeder, 1990). Furthermore, the *red1::lacZ, mer1::lacZ* and *mei4::lacZ* fusion proteins peak in synthesis at the same time (Engebrecht and Roeder, 1990; Menees *et al.*, 1992) and the smaller transcript of *MER2* predominates in meiosis when *MER1* is maximally transcribed (Engebrecht *et al.*, 1991). The *MER4* transcript is induced one hour after cells are placed into sporulation medium and peaks slightly before *SPO13* (Menees *et al.*, 1992). The *RAD51* transcript

is induced immediately upon transfer to sporulation medium with a maximum induction just 1.5 hours before *SPO13* (Shinohara *et al.*, 1992).

2.1.6 Epistasis groups

There are roughly two classes of mutants acting on early meiotic events (Malone and Esposito, 1981): 1) Those mutants that can be rescued by the *spo13* mutation (bypassing the reductional division) (Klapholz and Esposito, 1980). These include; *hop1* (Hollingsworth *et al.*, 1990), *red1* (Rockmill and Roeder, 1988), *mei4* (Menees and Roeder, 1989), *rec102* (Cool and Malone, 1992), *mre4* (Leem and Ogawa, 1991), *spo11* (Klapholz *et al.*, 1985), *mer1* (Engebrecht and Roeder, 1989) and *rad50* (Malone, 1983) and consist of genes that may act before the formation of DSBs. 2) Those mutants that are not rescued by *spo13*. These include; *rad50S* (Alani *et al.*, 1990; Cao *et al.*, 1990), *rad51* (Game, 1983; Shinohara *et al.*, 1992), *rad52* (Malone, 1983) and *dst2* (Dykstra *et al.*, 1991). In these mutations, DSBs are more processed or accumulate over wild type levels (*dst2*, not determined). Their gene products may be involved in steps after DSB formation. Malone *et al.* (1991) propose that mutations affecting an early block in recombination (before strand exchange) are rescued by the *spo13* mutation whereas mutations blocking later recombination steps may generate intermediates that prevent the equational segregation.

It has been demonstrated that recombination and homolog pairing are essential for proper chromosome segregation at the meiosis I division (Baker *et al.*, 1976) and it has been suggested by Carpenter (1987) that recombination is necessary for chromosome synapsis to occur. Mutations

defective in recombination or pairing often lead to chromosome nondisjunction (i.e. *rad50, mer1, spo11-1, mre4, hop1* and *red1*) (Malone, 1983; Engebrecht and Roeder, 1989; Klapholz *et al.*, 1985; Leem and Ogawa, 1991; Hollingsworth *et al.*, 1990; Rockmill and Roeder, 1989).

The *spo11-1* mutation severely restricts intra- and interchromosomal recombination but forms normal SC although the mutant may be unable to resolve the SC due to the lack of recombination (Klapholz *et al.*, 1985). Cao *et al.* (1990) have shown that *SPO11* is required for the generation of DSBs by *RAD50* and the *spo11-1* mutant also rescues *rad51* and *rad57* mutants in a *spo13* mutant background (Malone and Esposito, 1981) indicating that *SPO11* acts before or at the formation of DSBs and that the gene product is epistatic to the wild type *RAD* products. On the other hand, the *red1* mutation blocks the production of SC but the mutant exhibits 25% of wild type intrachromosomal recombination, though these mutants also fail at chromosome disjunction (Thompson and Roeder, 1989; Rockmill and Roeder, 1990). These results would suggest that the SC is not absolutely necessary for recombination.

HOP1 has been shown to be epistatic to *RED1* and acts in the same pathway, since a *red1/hop1* mutant has the same phenotype as a *hop1* double mutant whereas *MER1* appears to act in a different pathway from both genes (Rockmill and Roeder, 1990). *MEI4* has also been shown to be epistatic to the *rad52* mutation in a *spo13* mutant background (Menees *et al.*, 1992).

2.2 Meiosis in *T. thermophila*

2.2.1 Meiotic prophase I in *T. thermophila*

The kinetics and precise cytological stages of conjugation have been carefully timed by Martindale *et al.* (1982). Meiotic prophase I was divided into six cytologically defined stages corresponding to the degree of micronuclear elongation (Sugai and Hiwatashi, 1974, Martindale *et al.*, 1982). Meiosis begins soon after cells pair. Stage I of prophase I represents the beginning of micronuclear elongation and occupies approximately 43 minutes followed by stages II and III that take up about 56 minutes. Stage IV represents the maximum micronuclear elongation (full crescent) and may correspond cytologically to pachytene of meiotic prophase I. The end of pachytene to anaphase I (stages V to VI) occupies approximately 23 minutes. These studies demonstrated that meiosis as well as conjugation in *T thermophila* follows a precisely defined program.

2.2.2 Transcription and protein synthesis patterns during meiotic prophase I

Martindale *et al.* (1985) also examined mRNA transcription and protein synthesis patterns during conjugation. During prophase I the induction of RNA synthesis (as measured by the rate of ^{3}H-uridine incorporation) occurs within the first half hour upon cell mixing. This time corresponds to costimulation and probably represents mainly rRNA synthesis. The levels decrease steadily until a base level is reached between 5 to 6 hours into conjugation. Protein synthesis (measured as the rate of ^{3}H

[redacted]

[redacted]

[redacted]

tion (Martindale *et al.*, 1985; unpublished data).

2.2.5 Study rationale

This work describes the characterization and analysis of the *cnjA* cDNA sequence in an attempt to gain information about the structure and function of the *cnjA* protein. Others, Dr. Allis (Syracuse University) and Dr. Gorovsky (University of Rochester) and their groups have utilized *cnjA* as a standard for stage-specific gene expression during early conjugation (Stargell *et al.*, 1990; Dedon *et al.*, 1991) since the *cnjA* message represents the most abundant transcript of the conjugation specific genes isolated (Martindale and Bruns, 1983; Martindale *et al.*, 1985). This feature promises to make *cnjA* a desirable model for stage-specific gene expression in *T. thermophila*. It was therefore of particular interest to isolate the gene's promoter region so it could be examined for possible conjugation controlling elements, when compared to similar regions in *cnjB* and *cnjC*, just as similar sequence elements have been identified in the early meiotic genes in yeast (Buckingham *et al.*, 1990).

3. MATERIALS AND METHODS

3.1 Screening of a *T. thermophila* early conjugation cDNA library

The poly $(A)^+$ RNA for the construction of the cDNA library was isolated from *T. thermophila* early conjugating cells (2-5 hours) by D. Martindale and F. Taylor. The cDNA library (Uni-Zap^{TM} XR) was then constructed by Stratagene using the Lambda ZAP II vector which allows the excision of the phagemid (pBluescript, SK-) in the XLI-Blue host strain with the aid of the helper phage R408 (Stratagene). The cDNA was directionally inserted into the EcoRI site (representing the 5' end of the gene) and the XhoI site (representing the 3' end). Screening of the cDNA library was performed according to Stratagene's instructions.

3.1.1 Preparation of host cells for infection

The XLI-Blue host was selected for the F' episome prior to infection by growth on LB media (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 15 g/L agar) containing 12.5 µg per ml tetracycline. To prepare for infection by lambda, cells were incubated overnight at 37°C (standing) in LB liquid medium supplemented with 0.2% maltose and 10 mM $MgSO_4$. The overnight culture was then pelleted at 1,000 x g for 10 minutes at 4°C and resuspended in 10 mM $MgSO_4$ to a final OD_{600} of 0.5.

3.1.2 **Infection**

(Quantities mentioned below apply to plates of 86 mm diameter). The cDNA library phage stock was diluted by a factor of 10^4 in SM buffer (10 mM NaCl, 1 mM $MgSO_4$, 50 mM Tris·Cl pH 7.5 and 10% gelatin) and 10 µl (approximately 1.0 x 10^5 plaque forming units (PFU) per ml) of this dilution was incubated with 200 µl of the above host cell preparation for 15 minutes at 37°C. Subsequently, 3 ml of top agarose (0.7% agarose in NZY medium: 0.5% NaCl, 0.2% $MgSO_4$, 0.5% yeast extract, 1% NZ amine casein hydrolysate) was added at 48°C, plated on NZY solid media (1.5% agar) and incubated at 37°C for an average of 7 to 8 hours. Approximately 1,000 PFU were obtained per plate. The plates were placed at 4°C for at least 1 hour prior to plaque lifts.

3.1.3 **Plaque lifts**

Plaques were then transferred to nitrocellulose filters (Schleicher and Schuell) as specified by Stratagene. The filters remained in contact with the plaques for 2 minutes and in subsequent duplicate lifts, the contact time was 7 minutes. Nucleic acids were denatured by submerging the filters in 0.5 M NaOH, 1.5 M NaCl for 2 minutes, neutralized in 0.5 M Tris·Cl pH 8.0, 1.5 M NaCl for 5 minutes and rinsed for 30 seconds in 2 x SSC (for 1 x SSC: 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0). The nucleic acids were fixed to the filters by baking for 2 hours at 80°C under vacuum.

3.1.4 Hybridization with double-stranded probes

The method followed was that of Thomas (1980). Prehybridizations and hybridizations were performed under stringent conditions in 50% formamide (Sigma), 5 x SSC, 1 x Denhardt's, 0.2% SDS and 250 mg per ml denatured herring sperm DNA (Boehringer Mannheim). The filters were prehybridized on average for 4 hours at 42°C and hybridizations were allowed to proceed for 18 hours at the same temperature. DNA fragments were cut from 0.8% agarose gels run in TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA) and Gene-Cleaned (Bio/Can) according to the manufacturer's specifications. Nick translations (BRL, Nick Translation Kit) were performed according to the Kit's instructions with 40 µCi of alpha-^{32}P dCTP (ICN, 3000 Ci per mmol). Removal of unincorporated dCTP occured via a Sephadex G50 (Pharmacia) spin column according to Sambrook *et al.* (1989). The DNA fragment was usually nick translated to a specific activity of 2.0 x 10^8 counts per minute (cpm) per µg of DNA. The double-stranded probe was denatured by boiling for 5 minutes and 1.0 x 10^6 cpm were added per ml of the hybridization mixture. Four washes were performed at room temperature in 2 x SSC, 0.1% SDS for 15 minutes each. The final wash consisted of 0.1 x SSC, 0.1% SDS for 30 to 60 minutes at 55°C to 60°C.

3.1.5 End-labelling and hybridization conditions for oligonucleotides

The end-labelling protocol was followed according to Sambrook *et al.* (1989). A 22-mer primer (100 ng) was end-labelled using 200 µCi of gamma-^{32}P dATP (ICN, 4000 Ci per mmol) and 8 units of T_4

Polynucleotide Kinase (Pharmacia) in a buffer consisting of 0.5 M Tris·Cl pH 7.5, 0.1 M $MgCl_2$, 50 mM DTT, 1 mM spermidine 1 mM EDTA. The reaction was incubated for 45 minutes at 37°C. Unincorporated dATP was removed by passing the reaction through a Sephadex G10 (Pharmacia) spin column. The primer was end-labelled to a specific activity of 2.3 x 10^9 cpm per µg of DNA. Protocols for prehybridization and hybridization were performed according to Stratagene. Prehybridization took place in 6 x SSC, 5 x Denhardt's, 500 µg per ml denatured herring sperm DNA and 20 mM NaH_2PO_4 pH 7.0 for 4 hours at 42°C. The hybridization was performed in 6 x SSC, 20 mM NaH_2PO_4 pH 7.0, 500 µg per ml denatured herring sperm DNA and 0.4% SDS for 18 hours at the above temperature. A total of 2.0 x 10^6 cpm per ml of hybridization mixture was added. Washes were performed in 6 x SSC, 0.1% SDS at 60°C (2°C below the TM of the hybrid) approximately 2 to 4 times for 15 minutes until the background was low.

3.1.6 In vivo excision

The protocol was performed according to Stratagene. Positive plaques were cored with a pasteur pipet and resuspended in 500 µl of SM buffer and 20 µl chloroform. The phage were allowed to diffuse out of the agarose for 2 hours at room temperature and then stored at 4°C. XLI-Blue were prepared for infection as previously described, except the cells were resuspended to a final OD_{600} of 1.0. A total of 200 µl of the host cells were simultaneously infected with 200 µl (>1.0 x 10^5 PFU) of the desired purified phage stock and 1 µl (approximately 1.0 x 10^8 PFU) of the supplied helper phage R408 (Stratagene). The culture was incubated for 15 minutes at 37°C (standing) after which 5 ml of 2 x YT

media (1.6% tryptone, 1.0% yeast extract, 1.0% NaCl) was added and incubation allowed to resume for 3 hours at 37°C with shaking (250 rpm). Aliquots (1.5 ml) were taken and heated to 70°C for 20 minutes. The cells were then pelleted by microcentrifugation at 4,000 x g for 5 minutes at room temperature and 1 ml of the supernatant, containing the pBluescript phagemid, recovered. XLI-Blue host cells (200 µl) at an OD600 of 1.0 were transformed with 200 µl of the supernatant by incubation at 37°C for 15 minutes. Transformants (100 µl) containing the rescued phagemids were plated on LB solid media (1.5% agar) supplemented with 100 µg per ml ampicillin. The plates were then incubated at 37°C overnight.

3.1.7 **Plasmid DNA isolation**

All plasmids were isolated according to the alkaline lysis mini-prep method of Birnboim (1983). A single colony of the culture of interest was inoculated into 5 ml of TB (1.2% tryptone, 2.4% yeast extract, 0.4% glycerol, 17mM KH_2PO_4, 72 mM K_2PO_4 and 100 µg per ml ampicillin) (Sambrook *et al.*, 1989) and incubated at 37°C, shaking (250 rpm), overnight. The above culture (1.4 ml) was then pelleted in a microcentrifuge at 4,000 x g for 2 minutes and resuspended in GTE (50mM glucose, 25 mM Tris·Cl pH 8.0 and 10 mM EDTA). Lysozyme (10 mg per ml in GTE) was added and the mixture allowed to incubate for 5 minutes at room temperature. The cells were lysed in 0.2 M NaOH and 1% SDS. Cell debris, proteins and chromosomal DNA were precipitated on ice for 5 minutes by the addition of 5 M potassium acetate in glacial acetic acid. After the removal of the precipitate by microcentrifuga-

tion (23,000 x g) for 5 minutes, the aqueous phase was recovered and extracted once with an equal volume of phenol/chloroform/isoamyl-alcohol (25:24:1) followed by an additional chloroform/isoamyl-alcohol (24:1) extraction. The plasmid DNA was then precipitated in 1/10 volume of 3M sodium acetate pH 5.3 and 2.5 volumes 95% ethanol for 1 hour at room temperature. The DNA was recovered by microcentrifugation (23,000 x g) for 30 minutes at 4°C and the pellet washed in 70% ethanol and dried. The pellet was then resuspended in TE8 (10 mM Tris·Cl pH 8.0, 1mM EDTA) to which RNase A (Boehringer Mannheim) was added to a final concentration of 0.1 µg per ml and the mixture incubated at 37°C for 30 minutes. The solution was made to 20% polyethylene glycol 8000, 2.5 M NaCl (PEG/NaCl) and incubated for at least 1 hour on ice. The plasmid DNA was pelleted by microcentrifugation (23,000 x g) for 15 minutes at 4°C, washed with 70% ethanol and dried. The pellet was then resuspended in TE8. An average yield of 15 µg per 1.4 ml of culture was usually obtained.

3.2 Preparation for sequencing

3.2.1 Generation of nested deletions

Unidirectional deletions were generated via the digestion method of Henikoff (1984). A total of 5 µg of each plasmid was first linearized with KpnI (Pharmacia) or SphI (Promega), leaving a 3' overhang to protect the universal sequencing primer. The linearized plasmids were then digested with BamHI (Pharmacia) or XhoI (Pharmacia) according to the construct. Deletions were generated by exonuclease III (Boehringer Mannheim) or by the Erase-a-Base system provided in Kit form

by Promega. After the restriction enzyme digest, the plasmids were washed with phenol/chloroform and ethanol precipitated as described previously. The pellets were taken up in 1 x exonuclease buffer (66 mM Tris·Cl pH 8.0, 0.66 mM $MgCl_2$) and the mixture prewarmed to the reaction temperature (37°C). Four hundred units of exonuclease III was then added and aliquots were taken every 30 seconds for 10 minutes. Each aliquot (240 ng DNA) was placed into a pre-chilled S1 nuclease mix (0.25 M NaCl, 30 mM potassium acetate pH 4.5, 1 mM $ZnSO_4$, 5% glycerol and 67 Vogt units per ml S1 nuclease; Promega). The aliquots were then placed at room temperature for 30 minutes followed by heat inactivation of the enzymes at 70°C for 10 minutes. Aliquots from appropriate time points were chosen by running 70 ng of each sample on a 0.8% agarose gel. Chosen samples were then blunt-ended with the Klenow fragment of *E. coli* DNA polymerase I (Promega). The reaction buffer contained 20 mM Tris·Cl pH 8.0, 7 mM $MgCl_2$, 0.125 mM of each dNTP and 10 units per ml Klenow) and incubation occurred at 37°C for 3 minutes. Ligations were performed according to King and Blakesley (1986) in 50 mM Tris·Cl pH 7.6, 10 mM $MgCl_2$, 5% PEG 8000, 1 mM ATP, 1 mM DTT and 1 unit T_4 DNA ligase (Pharmacia) for 4 hours at room temperature. The ligated deletions were then transformed into the *E. coli* strain 71-18 made competent by the calcium chloride method (Pharmacia, from the manual for the M13 cloning/ sequencing system) and plated on LB solid medium (1.5% agar) containing 100 µg per ml ampicillin.

3.2.2 Generation of single-stranded DNA

The method was performed according to the IBI protocol for single-stranded DNA generation of pIBI31 derivatives. The pIBI30 and pIBI31 vector series (IBI) and pBluescript (Stratagene) harbor an f1 origin of replication that allows the generation of single-stranded forms of these plasmids. A single colony was inoculated into 2 ml of TB liquid medium (100 µg per ml ampicillin) and grown shaking overnight at 37°C. The overnight culture was then diluted at 1:50 in TB (100 µg per ml ampicillin), incubated under the same conditions for an additional 30 minutes and followed by infection with 10 µl (approximately 1.0×10^{11} PFU) of M13KO7 helper phage (IBI). After further incubation for 30 minutes, kanamycin was added to a final concentration of 70 µg per ml and incubation resumed for 6 hours. The host cells (1.2 ml) were pelleted by microcentrifugation (23,000 x g) for 5 minutes and the phage precipitated from 1 ml of supernatant in PEG/NaCL (20% polyethelene glycol 8,000; 2.5 M NaCl) for 1 hour on ice. The phage was recovered by microcentrifugation (23,000 x g) for 30 minutes, washed in 70% ethanol, dried and resuspended in TE8. The coat proteins were removed by a phenol/chloroform extraction and the single-stranded DNA ethanol precipitated (as previously described). The dried pellet was resuspended in distilled water. An average of 8 µg of single-stranded DNA was recovered per 1.2 ml of culture.

3.2.3 Sequencing

Clones were sequenced according to a modification (Pharmacia) of the chain termination method by Sanger *et al.* (1977) using a T7 DNA polymerase sequencing Kit (Pharmacia). Sequencing reactions were performed with the double- or single-stranded forms of the plasmids.

Approximately 3.2 to 4.2 µg of double-stranded plasmid DNA and 2.0 µg of single-stranded DNA were used per sequencing reaction. The double-stranded DNA (isolated as previously described) was denatured in 1/10 volume of 2 M NaOH for 5 minutes at room temperature, neutralized in 5 M ammonium acetate pH 7.5 and ethanol precipitated prior to primer annealing (as recomended by Pharmacia). Either 0.2 µM of the universal (Pharmacia) or reverse sequencing primer (Boehringer Mannheim) were employed for the double-stranded sequencing reactions and 0.1 µM of the primers were used for the single-stranded reactions (as specified by the Kit's instructions). Primer annealing and sequencing reactions were also performed according to the Kit's instructions using 10 µCi alpha-^{35}S dATP (Amersham, 1000 Ci/mmol) per reaction.

3.3 Conjugation

Conjugation was performed between *T. thermophila* strain B derivatives CU428 (mating type VII) and CU427 (mating type VI). These strains were kindly provided by Peter Bruns (Cornell University).

The experimental procedure was carried out as previously described by Martindale and Bruns, (1982). CU428 (doubling time of 2 hours and 45 minutes) and CU427 (doubling time of 2 hours and 15 minutes) were grown exponentially (134 rpm at 30°C) to a final cell density of 2.0 x 10^5 cells per ml in NEFF medium (0.25% proteose peptone, 0.25% yeast extract, 0.5% glucose, 0.3 mM $MgSO_4$, 17 mM $CaCl_2$ and 33 μM ferric citrate). The cell density was monitored via Coulter Counter (Coulter Electronics). Cells were then pelleted by low speed centrifugation (100 x g) for 2 minutes in a table top centrifuge and resuspended in 10 mM Tris·Cl pH 7.4 (starvation buffer). This was repeated once more with the cells being resuspended to a final density of 3.0 x 10^5 cells per ml. The two different mating types were starved independently at 30°C without shaking for 18 hours, then mixed at equal cell numbers and allowed to conjugate. The number of conjugating pairs were monitored microscopically. Pairs were observed 1 hour after mixing and 86% of the cells were paired by 5 hours.

3.3.1 RNA isolation

Total RNA was isolated as described by Martindale and Bruns (1983) from the above culture at time 0 after mixing the starved cultures of different mating type and subsequently at time 5, 6 and 7 hours after mixing. For each sample, 10 ml of culture corresponding to approximately 2 x 10^6 cells were pelleted in a table top centrifuge at 1,000 x g for 5 minutes. One ml of RNA lysis buffer (0.1 M NaCl, 10 mM EDTA, 10 mM Tris·Cl pH 7.4, 1% SDS and 5 mg per ml heparin) was added to the cell pellet and incubated for 10 minutes at room temperature. Cell debris and proteins were extracted with phenol/chloroform followed

by two more chloroform extractions (as described previously). The RNA was ethanol precipitated under standard conditions and finally resuspended in diethylpyrocarbonate (DEPC, Sigma) treated distilled water. The total RNA recovered varied between 160 and 190 µg per time point.

3.3.2 Northern analysis

In order to determine the integrity of the isolated RNA and to chose the time point with the most abundant transcript of *cajA* for the primer extension experiment, a Northern analysis was performed according to Sambrook *et al.* (1989). Five µg of total RNA from each sample and the RNA standard (BRL) were denatured in 5.4 µl deionized 6M glyoxal (Sigma) and 16 µl dimethylsulfoxide (DMSO, Sigma) for 1 hour at 50°C. The samples were run on a 1.2% agarose gel at 4 volts per cm (V/cm) for 4 hours in 10 mM Na_2PO_4 pH 7.0 buffer. Upon completion, the RNA was stained by placing the gel in distilled water containing 1 µg per ml ethidium bromide for 45 minutes. The RNA was directly transferred to a Zeta probe membrane (BioRad) without any prior treatment of the gel. The transfer took place overnight in 10 x SSC according to the manufacturer's instructions. Prehybridization and hybridization were performed under stringent conditions as previously described. The 2.42 kb insert of pAR103 was isolated and nick translated to a specific activity of 2.75×10^8 cpm per µg of DNA as previously described and 2.0×10^6 cpm were added per ml of hybridization mixture.

[illegible] pellet and the pellet taken up in [illegible] buffer (Boehringer Mannheim DNA sequencing kit).

DNA sequencing

[illegible] sequencing the 3' end of the [illegible] was carried out with a Boehringer Mannheim DNA sequencing kit by following the manufacturer's instructions. Annealing of primer and template was performed with 1.5 µg of total DNA prepared from the 2 days-conjugation time point (Fig. 3a and 3b) 0.5 µg of pUC19 DNA contained in 5 µl sterile water and 0.5 µg of the [illegible] primer respectively in a 10 µl reaction volume. The mixture was heated to 100°C for 1 minute followed by annealing at 50°C for 5 minutes. Preparation of the dATP/dTP labelling mixes was performed according to the kit's instructions and the sequencing reactions carried out with [illegible] (Amersham 1000 Ci/mmol). The nucleotide mixes contained at final concentrations 0.5 mM dGTP, 0.05 mM dTTP, 0.5 mM dCTP, 0.05 mM dATP, 1 mM dGTP, 0.05 mM dTTP, 1 mM dCTP, 0.05 mM dATP, and 1 mM dCTP, 0.05 mM dGTP. The concentration of the remaining dNTPs per nucleotide mix was kept constant at 0.5 mM. Sequencing reactions consisted of the annealed primer and template, 4.5 µl reaction buffer (as provided by the primer manufacturer), 1.5 µg of Actinomycin D (Boehringer Mannheim), 10 µCi [illegible] and 20 units of AMV reverse transcriptase (Pharmacia) in a final volume of 15.5 µl. Four µl of this mixture was added to 3 µl of the appropriate dNTP/ddNTP mixes and incubated for 30 minutes at 42°C. The reaction was stopped by ad-

ding 7 µl formamide gel loading buffer (Boehringer Mannheim). The second attempt utilized the same procedure as described above, except that primer annealing was performed as specified for the primer extension protocol. The reactions were run on a 6% acrylamide sequencing gel.

3.4 Computer analysis

The *cnjA* cDNA sequence was analyzed on a Macintosh Plus computer using the DNA Inspector IIe Sequence Analysis programs (Textco) and the Pustell programs (version 2.03) as well as MacVector sequence analysis programs (International Biotechnologies Inc.) (version 3.5). Homology searches were performed by screening the National Biological Research Foundation (NBRF) database (release 19.0) and the Swiss Protein database (version 22.0) with FASTP or FASTA (Lipman and Pearson, 1985; Pearson and Lipman, 1988).

3.5 Screening of *T. thermophila* genomic libraries

Three independent genomic libraries were screened for the *cnjA* genomic sequence.

Library I

A limited amount of a micronuclear specific *T. thermophila* genomic library was available that had been donated by Kathleen Karrer (Rogers and Karrer, 1989). The library was cloned into the SalI site of pUC18 with an average insert size of 5.4 kb. Colonies were obtain-

ed by transforming DH5α (provided by Karl Tibelius, McGill University) made competent by the calcium chloride method (Pharmacia, from the manual for the M13 cloning/sequencing system). Colony lifts were performed according to Sambrook *et al.* (1989). The transformants were transferred (2 minutes) from a master plate (LB, 1.5% agar, 0.002% X-Gal, 80 mM IPTG, 100 μg per ml ampicillin) to sterile nitrocellulose filters (Schleicher and Schuell) and the filters placed, colony side up, on LB solid media (1.5% agar, 100 μg per ml ampicillin). The plates containing the filters were then incubated for 3 hours at 37°C and the master plates for 24 hours at the same temperature. Cell lysis and DNA denaturation occurred when filters were placed for 5 minutes on a 3 MM Whatman paper presoaked in 0.5 M NaOH, followed by two successive neutralization steps; first, in 1 M Tris·Cl pH 7.5 and second in 0.5 M Tris·Cl, 1.5 M NaCl for 5 minutes each. The filters were then rinsed in 2 x SSC for 5 minutes and baked at 80°C under vacuum for 2 hours. Cell debris was removed by soaking the filters in 6 x SSC for 5 minutes followed by a 2 hour wash at 42°C in 50 mM Tris·Cl pH 8.0, 1 M NaCl and 0.1% SDS. Colonies were probed with the 1.84 kb insert of pAR100 (purified and nick translated as previously stated) and hybridizations were performed under standard conditions.

Library II and III

Two amplified *T. thermophila* genomic libraries were kindly provided by Gary Bannon (University of Arkansas); a macronuclear-specific (II) and a micronuclear-specific (III) library. These libraries were cloned into the BamHI site of Lambda DASH (Stratagene) with an average insert size of 20 kb. Host cells, *E. coli* LE392 (Stratagene), were prepared

for infection by lambda as previously described for Lambda Zap II, except that the cells were resuspended to a final OD_{600} of 2.0 in 10 mM $MgSO_4$ (Sambrook *et al.*, 1989). Infections and plaque lifts were carried out under the same conditions as described for the cDNA library screening. The hybridizations were performed under stringent conditions as mentioned earlier using the 2.42 kb insert of pAR103 as a probe (purified and nick translated as mentioned earlier).

3.6 Generation of subgenomic libraries

3.6.1 Genomic DNA isolation

This method was adapted from Howard and Blackburn (1985) by D. Martindale. Genomic DNA was isolated from either *T. thermophila* strain CU427 or CU428. Cells were grown to late log phase (134 rpm, 30°C) in NEFF medium to a final density of 2.0 x 10^5 cells per ml. Approximately 4.0 x 10^6 cells (20 ml) of the culture were pelleted for 5 minutes in a table top centrifuge at 100 x g and washed in 10 ml of 10 mM Tris·Cl pH 8.0. The cells were lysed in 2 ml NDS (18 mM Tris·Cl pH 8.0, 0.5 M EDTA and 1% SDS) by incubation at 65°C for 45 minutes. Cell debris and proteins were removed via a phenol/chloroform extraction and the aqueous phase precipitated in ethanol as previously described. The DNA was collected by centrifugation at 10,000 x g for 30 minute at 4°C, washed with 70% ethanol, dried and resuspended in TE8. RNase A (Boehringer Mannheim) was then added to a final concentration of 50 µg per ml and the mixture incubated at 37°C for 10 minutes. The reaction volume was then adjusted to 0.1% SDS and proteinase K (Boehringer Mannheim) at 100 µg per ml added.

Following incubation for 2 hours at 37°C, the mixture was extracted with phenol/chloroform and the DNA ethanol precipitated. The dried pellet was resuspended in TE8. An average yield of 1 mg of genomic DNA was usually obtained per 4.0 x 10^6 cells.

3.6.2 **Size-selected genomic library**

Three attempts were made to generate restriction fragment size-selected genomic libraries containing the *cnjA* genomic locus. The restriction fragments chosen for subcloning (see Fig. 6a) were: 1) a 2.9 kb PstI fragment consisting of the *cnjA* 5' upstream region, 2) a 3.5 kb HhaI fragment containing the entire *cnjA* coding region along with 5' and 3' noncoding sequences and 3) a 1.0 kb TaqI fragment comprising part of the 5' coding region along with 5' upstream sequences. For each subgenomic library, a minimum of 50 µg genomic DNA was digested with PstI (Boehringer Mannheim), HhaI (Pharmacia) (or its isoschizomer CfoI; Boehringer Mannheim) or TaqI (Pharmacia). The completion of the digest was verified by the method of Southern (1975) as recommended by the BioRad protocol for Zeta probe membranes. Two µg of DNA per restriction digest were run on a 0.7% agarose gel in TBE buffer (50 mM Tris·Cl pH 8.3, 100 mM boric acid , 10 mM EDTA) at 4 V/ cm for 3 hours. The DNA was denatured by submerging the gel in 0.5 M NaOH, 1M NaCl for 30 minutes at room temperature with agitation and neutralized in 0.5 M Tris·Cl pH 7.4,1.5 M NaCl for an additional 30 minutes. Transfer to the Zeta probe membrane (BioRad) was performed in 10 x SSC for 18 hours. The membrane was then briefly washed in 2 x SSC and baked for 1 hour at 80°C. The hybridization was performed under stringent conditions

as previously described using the 2.42 kb insert of pAR103 as a probe (purified and nick translated as stated before). The remaining portions of the restriction digests were electrophoresed on a 0.7% agarose gel in TAE and the desired fragments excised. Collected size fragments usually encompassed a window of 2.6-3.2 kb for PstI, 3.2-3.9 kb for HhaI (CfoI) and 0.7-1.3 kb for TaqI. These fragments were then Gene-Cleaned (BioCan), according to the manufacturers instructions. Both the HhaI (CfoI) and TaqI fragments were blunt-ended by incubation in 1 x Klenow buffer (20 mM Tris·Cl pH 8.0, 7 mM $MgCl_2$, 0.125 mM of each dNTP) and 10 units per ml Klenow (Promega) for 30 minutes at 37°C. The blunt-ended fragments were ligated to pIB131 digested with HincII (Pharmacia) and the PstI fragments into pIB131 digested with PstI (Pharmacia). Both vectors had been dephosphorylated according to Sambrook *et al.* (1989), excised from a 0.8% agarose gel and Gene-Cleaned prior to ligation. Ligations were performed according to King and Blakesley (1986) using 1 unit of T_4 DNA ligase (Promega) for the blunt-end ligations and 0.1 units for the sticky-end ligation. The ligated products were then transformed into *E. coli* DH5α made competent via the Hanahan method (Hanahan, 1983). Approximately 10^9 transformants per μg of DNA was obtained with the control plasmid pIB131.

3.7 **Inverse PCR (IPCR)**

To obtain the 5' and 3' flanking regions of the *cnjA* locus, genomic DNA was amplified via the IPCR method according to Ochman *et al.* (1988).

Primers:

a) The 22-mer *cnjA* 5' primer was previously described. It allows the outward extension by Taq DNA polymerase toward the 5' end of the gene (Fig. 2 and 6b).

b) The 21-mer *cnjA* 3' primer, also purchased from the Regional DNA Synthesis Laboratory at the University of Calgary, corresponds to the sequence 5'CGATGACACTTTGATCCAAGC3'. This oligonucletide is located at positions +2277 to +2297 in the cDNA and encodes the *cnjA* coding strand. It allows the outward extension by Taq DNA polymerase towards the gene's 3' end (Fig. 2 and 6b).

Preparation for amplification

Circularization of the genomic DNA fragments was performed according to the method of Collins and Weissman (1984) as specified by Ochman *et al.* (1988). Five µg of *T. thermophila* CU427 genomic DNA was digested to completion with HhaI (Pharmacia) (verified by Southern hybridization, see subgenomic libraries). The restriction fragments were extracted with phenol/chloroform and ethanol precipitated as described earlier and the dried pellet taken up in TE8. The HhaI fragments were blunt-ended with Klenow (Promega) (see subgenomic libraries). For circularization, 0.1 µg of the blunt fragments were diluted to a final concentration of 0.5 µg per ml in T_4 DNA ligase mix (50mM Tris·Cl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1mM ATP and 25 µg per ml BSA and 1 unit per µl T_4 DNA ligase;

all components were purchased from New England Biolabs) and allowed to ligate for 16 hours at 12°C. The ligated products were phenol/chloroform extracted and ethanol precipitated (as described previously).

3.7.1 IPCR reaction

Conditions for the IPCR amplification were carried out according Saiki *et al.* (1988). The Taq DNA polymerase amplification took place in a 100 µl reaction volume containing 0.1 µg of the HhaI genomic ligation mixture in 50 mM KCl, 10 mM Tris·Cl pH 8.4, 2.5 mM MgCl2, 1 µM of each primer (*cnjA* 5' primer and *cnjA* 3' primer), 200 µM of each dNTP pH 7.5 (Pharmacia) and 2.5 units Taq DNA polymerase (Pharmacia). The amplification was performed on a Perkin-Elmer-Cetus Instrument thermocycler under the following conditions: an initial denaturation step at 94°C for 10 minutes followed by 35 cycles at a) 94°C for 1 minute, b) primer annealing at 54°C for 2 minutes and c) extension by Taq DNA polymerase at 72°C for 3 minutes. A final extension step consisted of 10 minutes at 72°C. The cDNA controls were prepared and run under the same conditions: 1) the 2.42 kb insert of pAR103, the 2.47 kb insert of pAR105 and the 2.40 kb insert of pAR106. The PCR product was treated with proteinase K according to Sambrook *et al.* (1989) by adjusting the reaction volume to 5 mM EDTA, 0.5% SDS and adding Proteinase K (Boehringer Mannheim) to a final concentration of 100 µg per ml. The mixture was incubated at 56°C for 30 minutes, phenol/chloroform extracted and ethanol precipitated as previously described, except that the sodium acetate solution was at pH 7.0.

The DNA was then blunt-ended with Klenow (Pharmacia) as already described for subgenomic libraries.

3.7.2 Southern analysis of the IPCR product

Conditions for Southern analysis were the same as previously described. One tenth volume of the amplified IPCR products were loaded on a 1.2 % agarose gel and run at 4 V/cm for 1.5 hours in TBE buffer. Ten µg of *T. thermophila* genomic DNA was digested with HhaI, PstI, Hind III or TaqI (all enzymes from Pharmacia) and 5 µg per digest were loaded in duplicate and run on a 0.7% agarose gel in TBE buffer at 4 V/cm for 6 hours. The DNA was denatured and transferred to a Zeta probe membrane (BioRad) as described earlier. The 1.4 kb IPCR product and the 2.47 kb insert of pAR105 were gel purified, Gene-Cleaned (BioCan) and 200 ng of each were nick translated (BRL, Nick Translation Kit) to a specific activity of 1.95×10^8 and 5.0×10^8 cpm per µg of DNA respectivly. Hybridizations were performed under standard conditions and 1.0×10^6 cpm per ml of hybridization solution were added per filter.

3.7.3 Cloning

The purified PCR product was blunt-ended with Klenow (as described above) and phosphorylated according to Sambrook *et al.* (1989). The volume was adjusted to 10 mM Tris·Cl pH 7.5, 7mM $MgCl_2$, 0.1 mM EDTA, 5 mM DTT, 10 mM ATP and 30 units of T_4 Polynucleotide Kinase (Pharmacia). The mixture was incubated at 37°C for 30 minutes, phenol/chloroform extracted and ethanol precipitated as already de-

scribed. The recovered PCR product was ligated to 0.5 µg pIBI30 digested with HincII according to King and Blakesley (1986). The vector had been dephosphorylated and gel purified as previously stated. The ligated products were transformed into *E. coli* DH5α and Epicurian coli Sure^{TM} (capable of accepting repeats; kindly donated by Dr. Kuhnlein, McGill University) that had been made competent by the Hanahan method (Hanahan, 1983). Colony hybridzations were performed as previously described using the 1.4 kb IPCR product as a probe (specific activity of 2.3 x 10^8cpm per µg of DNA) (purified and nick translated as mentioned earlier).

Another attempt at subcloning the PCR product utilized phosphorylated synthetic EcoRI linkers that were added to the PCR product according to Sambrook *et al.* (1989). A portion of the previously purified PCR product (0.2 µg) was reamplified, treated with Proteinase K, phenol/chloroform extracted and gel purified as previously stated. The PCR product was heated to 68°C for 15 minutes and cooled to 37°C. The DNA was blunt-ended by incubation for 15 minutes at 37°C in a buffer consisting of 90 mM $(NH_4)_2SO_4$, 0.33 M Tris·Cl pH 8.3, 33 mM $MgCl_2$, 50 mM ß-mercaptoethanol, 5 mM each dNTP (Pharmacia) and 2 units of T_4 DNA polymerase (Pharmacia). The reaction was stopped by adding 1 µl of 0.5 M EDTA and extracted with phenol/chloroform. Unincorporated dNTPs were removed by passing the aqueous phase through a Sephadex G50 spin column as mentioned earlier for nick translations. The recovered fraction was then ethanol precipitated and the dried pellet taken up in 10 mM Tris·Cl pH 7.6. The volume was adjusted to 0.05 M Tris·Cl pH 7.6, 10 mM $MgCl_2$, 10 mM DTT and 10 mM ATP to which 1000 ng of a 10-mer phosphorylated EcoRI

linker (Boehringer Mannheim) was added and 100 Weiss units of T_4 DNA ligase (Stratagene). The Linkers and PCR product were allowed to ligate for 18 hours at 16°C. The unincorporated linkers were removed by a Sephacryl S-400 (Stratagene) spin column and the aqueous phase was phenol/chloroform extracted, ethanol precipitated and the dried pellet resuspended in TE8. The resultant product was digested with 5 units EcoRI (Pharmacia) for 2 hours at 37°C, phenol/chloroform extracted, ethanol precipitated and the pellet resuspended in TE8. The recovered DNA was then ligated to 0.5 μg of pBluescript (KS+) (Stratagene) that had been digested with EcoRI (Pharmacia), dephosphorylated, gel purified and Gene-Cleaned as described earlier. The ligation was performed according to King and Blakesley (1986) and used to transform XLI-Blue and Epicurian coli Sure™ made competent by the Hanahan method (Hanahan, 1983). Transformants were screened via colony hybridizations using the 1.4 kb PCR product as a probe (specific activity 3.5×10^8 cpm per μg of DNA) as previously described.

4. RESULTS

4.1. Isolation, cloning and sequencing of *cnjA* cDNA

The largest cDNA clone (pC1-2; 2074 bp), isolated from the initial cDNA library screen (Martindale and Bruns, 1983) and representing *cnjA*, was completely sequenced. This cDNA library had been constructed from poly(A)$^+$ RNA isolated during early conjugation of *T. thermophila*. The cDNA had been 3' end tailed with polydeoxycytidilic acid and cloned into the PstI site of pBR322 tailed at its 3' end with polyguanylic acid (Martindale and Bruns, 1983). The *cnjA* restriction map, relevant constructs and sequencing strategy are outlined in Fig. 1. An internal PstI site served to subclone pC1-2 into the PstI site of pIBI31 resulting in the constructs pAR100 and pAR101. The 234 bp insert of pAR101 contained 23 guanine residues at its 5' end and the 1840 bp insert of pAR100 revealed 22 cytidine residues at its 3' end. Since no poly(A)$^+$ tail was found in the cDNA represented by pC1-2, the original cDNA synthesis probably initiated internally. In order to obtain the missing 3' end sequence, another cDNA library (constructed by Stratagene from poly(A)$^+$ RNA isolated during early conjugation) was screened using the 1840 bp insert of pAR100 as a probe. The resultant isolate pAR102 (2420 bp) contained a poly(A)$^+$ tail of 18 adenine residues but lacked 146 bp of the *cnjA* 5' end sequence previously identified in pC1-2. Due to sequencing difficulties with pAR102, the 2420 bp insert was subcloned into the vector pIBI30 resulting in the plasmid pAR103. To obtain a full length clone, the Stratagene library was re-screened with a 22-mer oligonucleotide (*cnjA* 5' primer) located at position +160 to +180 in the sequence (Fig. 2) and complementary to

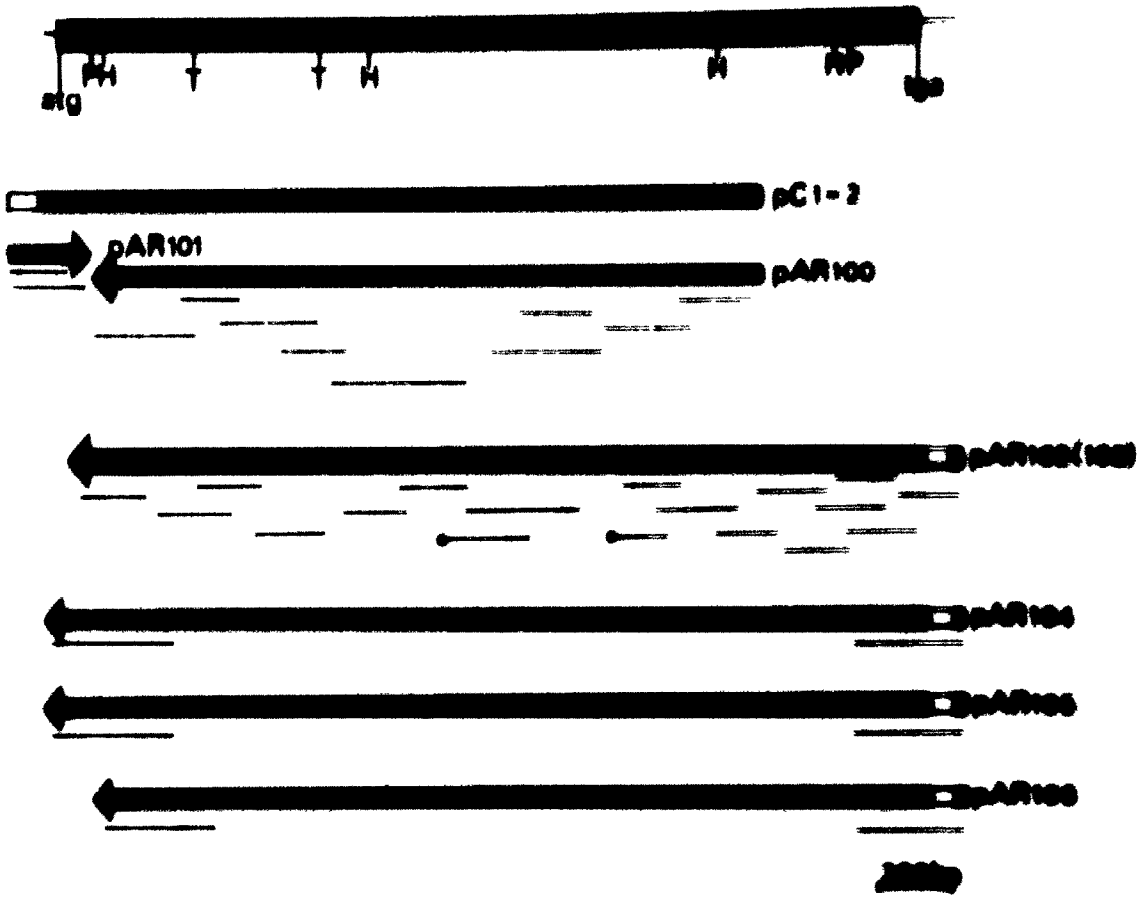

Fig. 1. The restriction map of the sequenced 2494 bp *cnjA* cDNA is shown at the top of the figure with atg and tga delineating the coding region of the gene. Relevant isolates, obtained from two different cDNA libraries (see text), are shown below the restriction map with pC1-2 representing the original clone. The 5' and 3' end repeats are indicated by open bars within their respective cDNAs and subclones or individual isolates are identified with arrows indicating the direction sequenced. Nested deletions are shown as thin lines below their respective cDNAs. These were sequenced using the M13 universal sequencing primer, except where marked with a dot for which the reverse sequencing primer was used. Restriction endonucleases are abbreviated as follows: H, HindIII; P, PstI and T, TaqI.

[redacted]

[redacted]

[redacted]

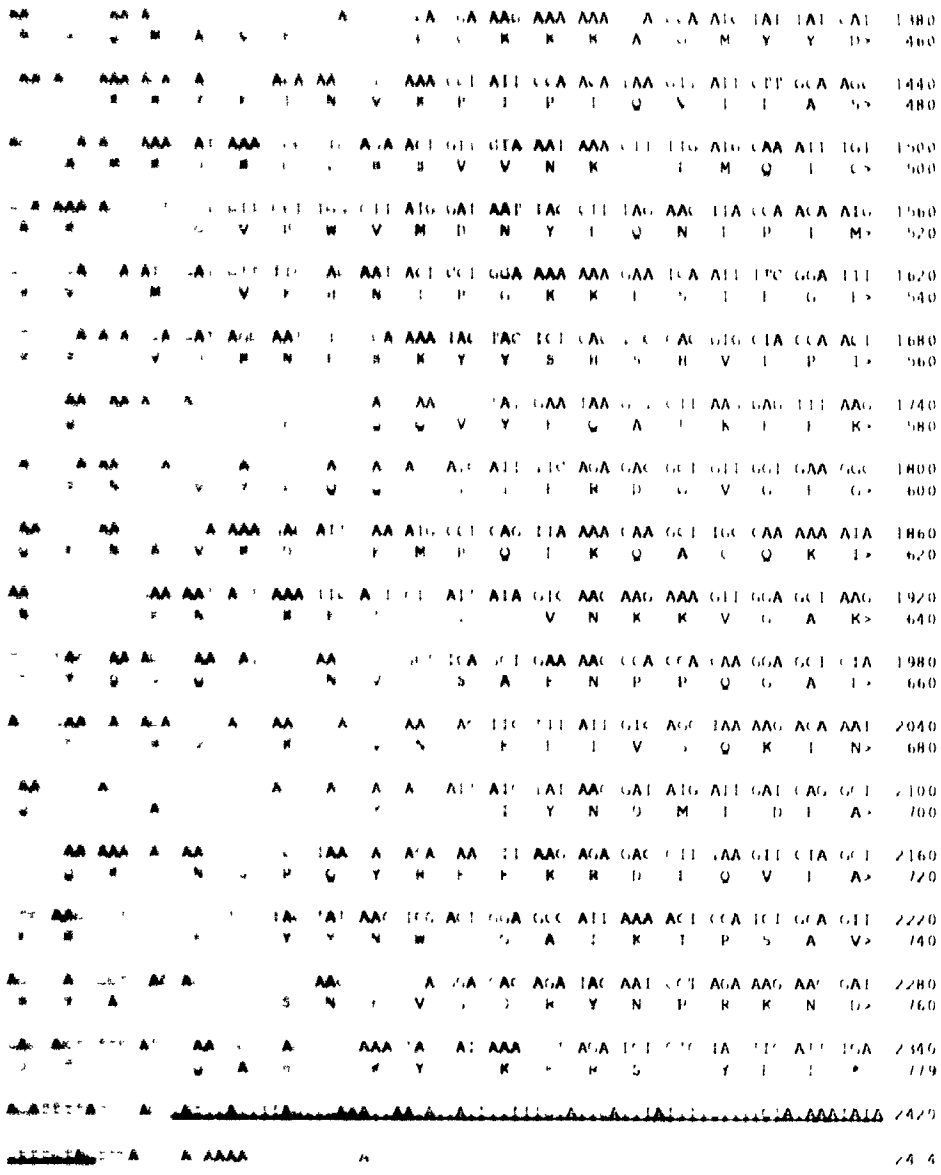

Fig. 2 - continued -

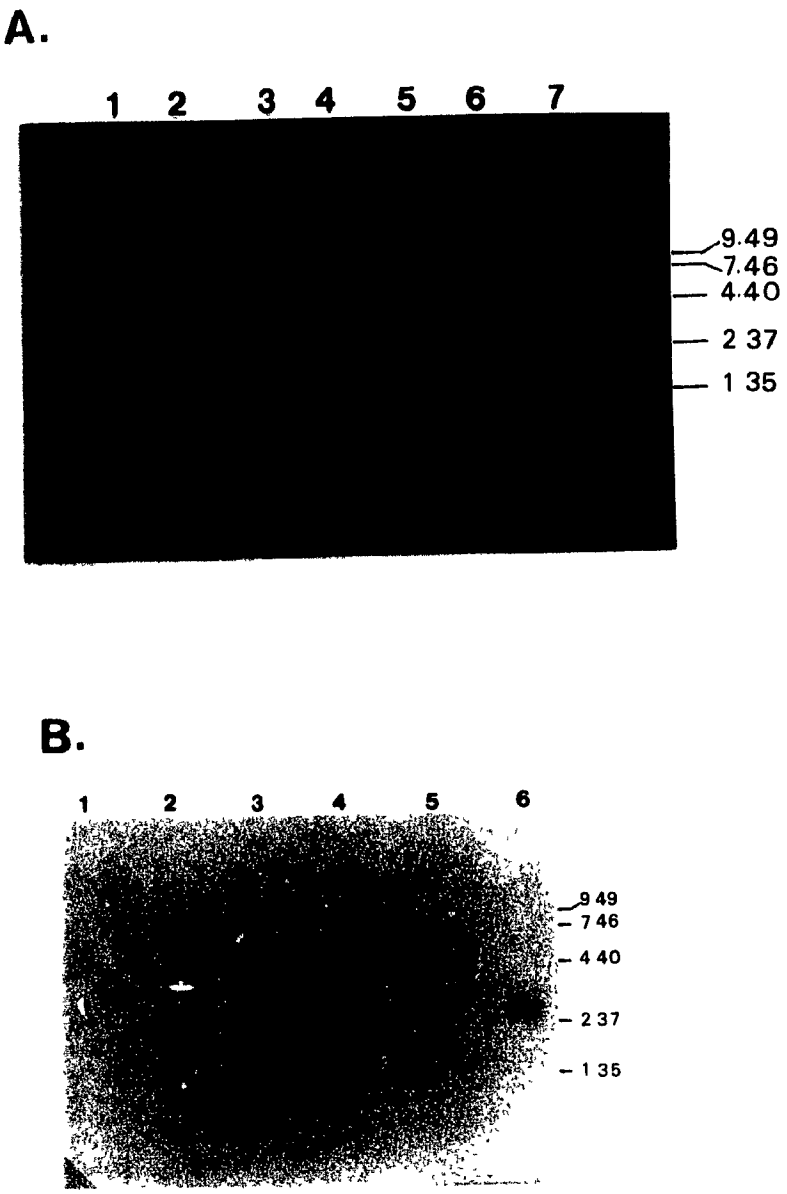

Fig. 3 Northern analysis of total RNA isolated from *T. thermophila* conjugating cells CU428 and CU427. a) Ethidium bromide stained RNA gel; approximately 5 µg of total RNA from each time point and RNA standard were denatured in glyoxal-DMSO and run on a 1.2% agarose gel. Lanes 1 and 2, total RNA isolated from logarithmically grown cells; Lanes 3, 4, 5 and 6 represent 0, 5, 6 and 7 hours into conjugation. b) Resultant autoradiograph of the filter from the above gel probed with the 2.42 kb insert of pAR103.

pC1-2 and the 3' end repeat is found at position +2356 to +2428 (Fig. 1 and 2), 16 bp after the stop codon to 27 bp from the poly(A)$^+$ tail. To verify the reality of the 3' end repeat, three independent isolates (pAR104, pAR105 and pAR106) were sequenced in this area (Fig. 1). These cDNAs all contained the 3' end sequence identified in pAR102 (103) thereby supporting the authenticity of this region. Since repeated attempts to obtain the genomic sequence or an independent cDNA that would confirm the 5' end repeat failed, it was decided to map the 5' end of the *cnjA* message via primer extension.

The transcription start sites were determined by using total RNA from the 5 hour conjugation time point (Fig. 3) and the *cnjA* 5' primer (+160 to +182) (Fig. 2 and 6b). Primer extension (Fig. 4) revealed three major start sites at positions -45, -48 and -54 and one minor start site at position -57 from the proposed translation initiation codon (Fig. 2). This is strong evidence that the 5' portion of the inverted repeat (starting at -113 and ending at -40) is a cloning artifact. Multiple start sites in *T. thermophila,* situated approximately 58 bases from the translation initiation codon (Csank and Martindale, manuscript in preparation), have been previously demonstrated for nuclear genes transcribed by RNA polymerase II such as histone genes *H4-1* and *H4-2* (Horowitz *et al.*, 1987); *hv1* (van Daal *et al.*, 1990); *H2B-1* and *H2B-2* (Nomoto *et al.*, 1987); *HMGB* (Schulman *et al.*, 1991); actin (Cupples and Pearlman, 1986); F-ag (Hummel *et al.*, 1991); *SerH3* (pC6) (Tondravi *et al.*, 1990); *cupC* (pC8) (Csank and Martindale, 1992) and *cnjB* (pC2) (Taylor and Martindale, submitted) as well as for genes transcribed by RNA poly-

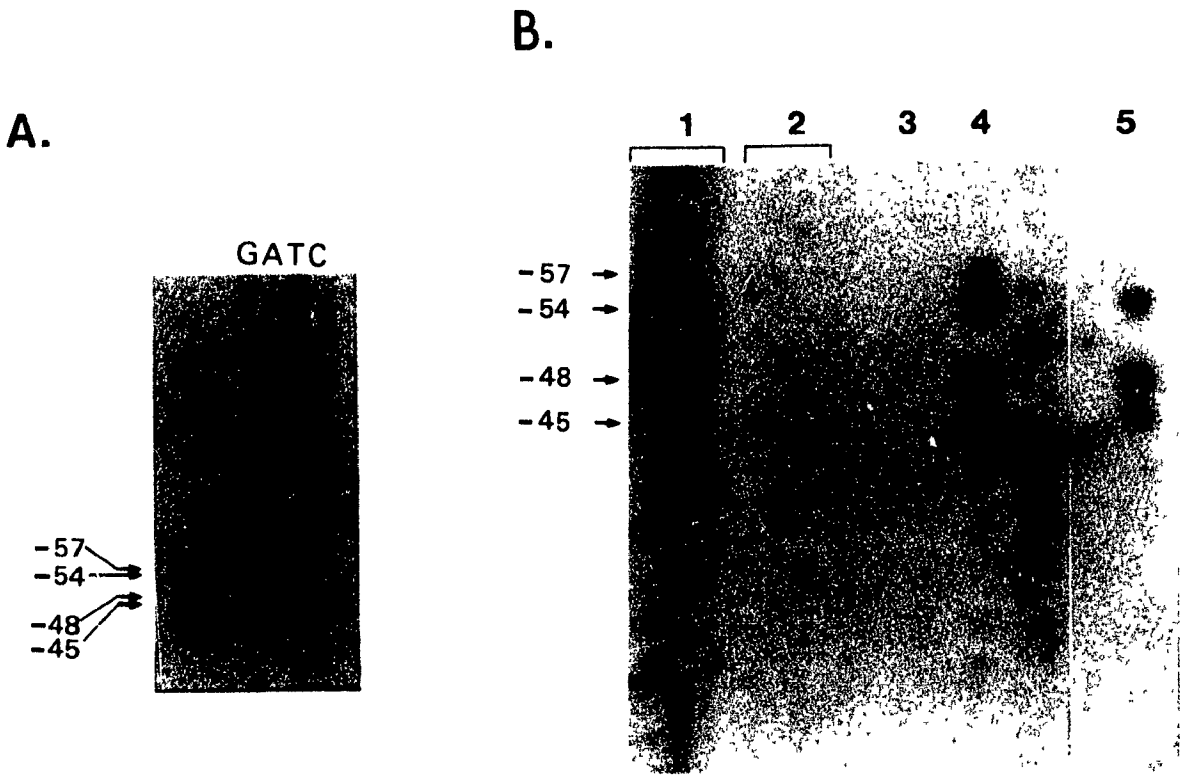

Fig. 4 5' end point analysis of the *cnjA* transcript. Primer extension was performed on 30 µg of total RNA from the 5 hour conjugation time point (Fig.3). The *cnjA* 5' primer (Fig.6b) was used for both the DNA sequencing reactions and primer extension. a) Shows a low resolution 6% acrylamide sequencing gel with lanes GATC denoting the DNA sequencing standard pAR105 (^{35}S-dATP). b) A more resolved sequencing gel with lanes 1 and 2 representing the DNA standard pC1-2 (1, reactions performed with gamma-^{32}P dATP end-labelled primer and 2, a standard ^{35}S-dATP reaction). Lane 3, primer extension performed on log RNA; Lane 4, 5 hour conjugation (exposed 1 week); Lane 5, 5 hour conjugation (exposed 3 days).

merase III, *L21* (Rosendahl *et al.*, 1991) and *L37* (Hansen *et al.*, 1991). Multiple start sites are thought to be due to the AT-richness of the ciliate 5' nontranslated region and a lack of clearly definable TATA boxes (Horowitz *et al.*, 1987; Brunk and Sadler, 1990).

Additional evidence favouring the idea that the 5' end repeat of *cnjA* is an artifact comes from another conjugation specific gene *cnjB* (pC2-1) examined in our laboratory. The original cDNA clone containing sequences from this gene came from the same cDNA library as pC1-2. A large inverted repeat (200 bp) was found, like in *cnjA*, in the 5' and 3' untranslated regions of pC2-1. The genomic sequence of *cnjB* identified the 5' end repeat as an artifact (Taylor and Martindale, submitted). Thus the repeat in *cnjA*, like *cnjB*, may be a possible artifact generated during the construction of the original cDNA library (Fig. 9).

The *cnjA* leader sequence (40 bases) falling outside the repeat region has an A+T content of 82.5% of which 62.5% are adenine residues, whereas the G+C content is only 17.5%. This A-richness (68% A-content, 12 different *T. thermophila* genes examined; Csank and Martindale, manuscript in preparation) is characteristic of the ciliate 5' nontranslated leaders having a mean A+T content of 84.8% (Yamauchi, 1991, 26 different ciliate genes examined). The *cnjA* coding region is 34.2% G+C which is much lower than the mean G+C content (45%) found in highly expressed genes of *T. thermophila* (Martindale, 1989) but compares well with the other conjugation specific genes analyzed (*cnjB*, 32.3% G+C; *cnjC*, 30.5% G+C) (Taylor and Martindale, submitted; Martindale, 1990). The 3' untranslated region of *cnjA* (114 bases) is highly T-rich (60.5%) with 87.7% A+T.

4.1.3 Amino acid sequence of *cnjA*

The long ORF of *cnjA* predicts a protein of 779 amino acids starting with the adenine of ATG at position +1 and ending with the stop codon TGA at position +2338 in the sequence (Fig. 2). The derived protein has a calculated molecular weight of 89.5 kDa and an isoelectric point of 9.3. Charged amino acids make up 25% of the protein's composition of which 14.6% are basic (9.6% lysine, 3.9% arginine, 1.2% histidine) and 10.4% are acidic (5.5% aspartic acid, 4.9% glutamic acid). The predicted protein sequence reveals no internal repeats or unusual amino acid distributions except for the high percentage of lysine residues. There are no extended regions of secondary structure and a hydropathy plot calculated according to Kyte and Doolittle (1982) with a window size of 7 residues, predicts the protein to be essentially hydrophilic (Fig. 5). No hydrophobic leader sequence or membrane spanning regions can be detected.

4.1.4 Codon usage

The Fpr value (frequency of preferred codons), thought to be a reflection of codon choice in ciliates (Martindale, 1989), shows *cnjA* to follow the general trend already observed for the other two conjugation specific genes (*cnjB*, *cnjC*) with similar expression patterns (Taylor and Martindale, submitted; Martindale, 1990) (Table 1). A much lower Fpr value can be observed for these stage specific genes (*cnjA*, 0.60; *cnjB*, 0.52 and *cnjC*, 0.64) than for the Fpr value seen for highly expressed genes (0.90) (Martindale, 1989). As noted by Martindale (1990), the lower Fpr value can be related to the lower G+C content of the genes' coding regions (*cnjA*, 34.2%, *cnjB*, 32.3% and *cnjC*, 30.5%) when compared

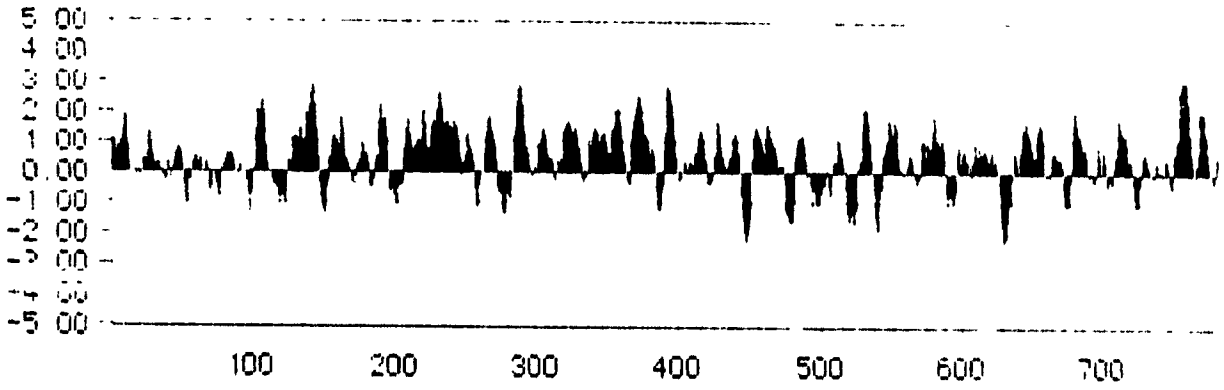

Fig. 5. Hydropathy profile of the *cnjA* cDNA from amino acid 1 to 779. The solid bars above the midline denote hydrophilic residues and bars below predict hydrophobic regions. The profile was determined according to Kyte and Doolittle (1982) with a window size of 7 residues.

Table 1: Codon usage

Amino Acid	codon	cnjA	cnjB	cnjC	TET
Ala	**GCT**	25	29	6	**128**
	GCC	4	3	2	65
	GCA	11	15	2	4
	GCG	0	2	0	0
Arg	CGT	5	1	1	7
	CGC	0	1	1	0
	CGA	0	0	0	0
	CGG	0	0	0	0
	AGA	22	42	11	**130**
	AGG	3	6	0	2
Asn	AAT	36	129	16	17
	AAC	19	57	5	**63**
Asp	**GAT**	31	59	15	**75**
	GAC	12	21	3	33
Cys	TGT	4	23	8	6
	TGC	9	25	5	**24**
Gln	**CAA**	20	34	3	**30**
	CAG	3	16	0	2
	TAA	26	75	8	14
	TAG	6	40	4	5
Glu	**GAA**	30	87	29	**121**
	GAG	8	26	2	6
Gly	**GGT**	15	32	8	**139**
	GGC	2	16	0	8
	GGA	21	48	4	10
	GGG	4	9	0	1
His	CAT	0	19	6	6
	CAC	9	6	3	**40**
Ile	**ATT**	24	65	15	**61**
	ATC	13	22	7	**73**
	ATA	9	24	4	2
Leu	**TTA**	9	30	11	**35**
	TTG	12	17	4	**43**
	CTT	18	24	6	**28**
	CTC	7	10	1	**45**
	CTA	12	16	3	1
	CTG	2	7	0	1

Amino Acid	codon	cnjA	cnjB	cnjC	TET
Lys	AAA	54	109	17	36
	AAG	21	36	9	**220**
Phe	TTT	22	51	9	10
	TTC	19	30	6	**76**
Pro	**CCT**	21	29	9	**48**
	CCC	0	3	0	**59**
	CCA	11	21	1	3
	CCG	0	1	0	2
Ser	**TCT**	14	70	7	**67**
	TCC	2	13	0	**67**
	TCA	11	51	6	13
	TCG	0	3	0	1
	AGT	6	29	1	6
	AGC	8	27	0	4
Thr	**ACT**	23	32	7	**57**
	ACC	7	8	3	**67**
	ACA	13	21	6	5
	ACG	1	5	0	0
Tyr	TAT	18	34	7	21
	TAC	21	19	2	**47**
Val	**GTT**	26	24	8	**67**
	GTC	7	4	3	**76**
	GTA	10	17	6	3
	GTG	7	10	2	0
Met	ATG	21	36	6	51
Trp	TGG	5	29	2	19
total		779	1748	310	2250
%GC		34.2	32.3	30.5	45.2
Fpr		0.60	0.52	0.64	0.90

A comparison of codon usage in conjugation specific genes *cnjA*, *cnjB*, *cnjC* and highly expressed *T. thermophila* genes (TET). Class I codons are written in bold letters, class II codons are underlined and Fpr is the frequency of preferred codons as determined by Martindale (1989). TET (Martindale, 1989), *cnjB* (Taylor and Martindale, submitted) and *cnjC* (Martindale, 1990).

to highly expressed genes (greater than 40% G+C) indicating a tendency towards more AT-rich codons. An example of this tendency is seen in the codon choices for proline (CCT, CCC, CCA and CCG). In highly expressed genes, CCT and CCC are the most commonly used, whereas *cnjA* uses only CCT and CCA.

4.1.5 Comparison of the *cnjA* protein with known protein sequences

The deduced protein (CnjA) was compared with proteins in the National Biological Research Foundation (NBRF, release 19.0) and the Swiss Protein (release, 22.0) databases. No significant similarities were found. An amino acid comparison was made between CnjA and the following published sequences of proteins encoded by meiosis-specific or meiosis-related genes: Spo11 (Atcheson *et al.*, 1987); Spo12 and Spo16 (Malavasic and Elder, 1990); Spo13 (Buckingham *et al.*, 1990); Spo7 (Whyte *et al.*, 1990); Red1 (Thompson and Roeder, 1989); Hop1 (Hollingsworth *et al.*, 1990); Mer1 (Engebrecht and Roeder, 1990); Mer2 (Engebrecht *et al.*, 1991); Mre4 (Leem and Ogawa, 1992); Mei4 (Menees *et al.*, 1992); Rme1 (Covitz *et al.*, 1991); Ime1 (Smith *et al.*, 1990); Sme1 (Yoshida *et al.*,1990); Mck1 (Neigeborn and Mitchell, 1991); Hpr1 (Aguilera and Klein, 1992); Rec102 (Cool and Malone, 1991); Dmc1 (Bishop *et al.*,1992); Rec1 (Holden *et al.*, 1989); Rad3 (Harosh and Deschavanne, 1991); Rad4 (Choi *et al.*, 1990); Rad5 (Johnson *et al.*, 1992); Rad9 (Murray *et al.*, 1991); Rad14 (Bankman *et al.*, 1992); Rad16 (Bang *et al.*, 1992); Rad21 (Birkenbihl and Subramani, 1992); Rad25 (Park *et al.*, 1992); Rad51 (Shinohara *et al.*, 1991); Rad54 (Emery *et al.*, 1991); Rad57 (Krans and Mortimer, 1991); RhoNUC (Chow *et al.*, 1992); STPα (Clark *et al.*,1991);

STPß (Dykstra *et al.*, 1991); Sepl (Tishkoff *et al.*, 1991); Iral (Tanaka *et al.*, 1989); Ira2 (Tanaka *et al.*, 1990); Srv2 (Fedor-Chaiken *et al.*, 1990); Cap (Field *et al.*, 1990); Tfsl (Robinson and Tatchell, 1991); as well as *Schizosaccharomyces pombe* Ranl (DeVoti *et al.*, 1991); Mei2 (Watanabe *et al.*, 1988); Mei3 (McLeod *et al.*, 1987); the *E.coli* RuvAB (Benson *et al.*, 1988); RuvC (Takahagi *et al.*, 1991) and *Drosophila melanogaster* BJ1 (Frasch, 1991) and Rrpl (Sander *et al.*, 1991) without detecting any sequence similarities with *cnjA*.

4.1.6 Consensus sequences

The putative *cnjA* protein sequence was searched for several consensus sequences. No homologies were found to motifs characteristic of transcriptional regulators, such as zinc fingers (Evans and Hollenberg, 1988); helix-turn-nelix (Dodd and Egan, 1987); homeodomain (Scott *et al.*, 1989); helix-loop-helix (Davis *et al.*, 1987); leucine zippers (Landschulz *et al.*, 1988) and MCM1/SRF (Passmore *et al.*, 1991). In addition, we did not find calcium-binding sites (Baum *et al.*, 1986); GTP/GDP-binding sites (Dever *et al.*, 1987); ATP-binding sites (Walker *et al.*, 1985); NTP-binding (Gorbalenya et al., 1988; Hodgman, 1988) or nuclear localization signals (Roberts, 1989; Garcia-Bustos *et al.*, 1991). CnjA was also searched without success for motifs such as SPKK and (R)GRP(K) that have been identified in chromosomal proteins that bind to the minor groove of DNA in AT-rich regions (Churchill and Suzuki, 1989) and the HMG motif (Schulman *et al.*, 1991). The deduced protein was further compared with sequence patterns found in proteins of known function by scanning the Prosite database (version 9.0) but no matches were found. Two sites similar to the glycosylation consensus sequences NXT and NXS (Hunt and

Dayhoff,1970) are located starting with the amino acids at positions 547 and 729 in the protein.

The cnjA genomic locus

Since the meiosis-specific *cnjA* is under tight transcriptional control (Martindale *et al.*, 1985; Sturgell *et al.*, 1990) and is the most abundantly transcribed of the conjugation specific genes (Martindale and Bruns, 1983), it was of considerable interest to obtain the promoter region of this gene. Three independent *T. thermophila* genomic libraries were screened without success.

4.2. Screening *T. thermophila* genomic libraries

Library I: A limited amount of a micronuclear specific genomic library was available that had been cloned into the SalI site of pUC18 with an average insert size of 5.4 kb (kindly donated by Kathleen Karrer, Marquette University). For a 99% probability of finding the *cnjA* sequence once per haploid genome, a total of 1.3×10^5 colonies needed to be screened. Only 2.1×10^4 colonies were recovered after transformation into *E. coli* DH5α of which 60% did not contain recombinant DNA (determined by color selection) reducing the actual number screened to 8.4×10^3 colonies.

Library II and III: Both a macronuclear-specific (library II) and a micronuclear-specific (library III) *T. thermophila* genomic library were made available by Gary Bannon (University of Arkansas). The libraries

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

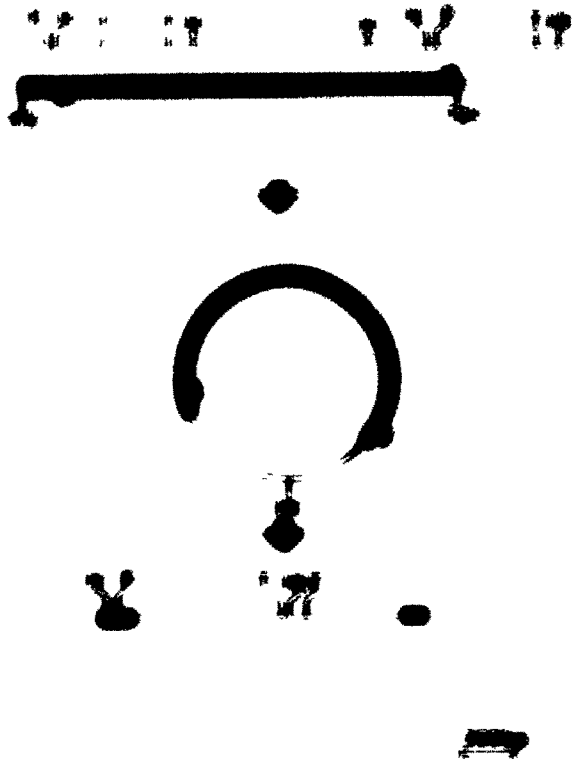

Fig. 1. The restriction map of the *cryA* genomic locus. The solid bar below indicates the *cryA* ORF along with the 5′ and 3′ flanking regions contained within the genomic MboI fragment. Primers are represented by arrows (not drawn to scale) indicating the direction of amplification. A schematic representation of the RPCR method is shown below. Primers hybridized close to the 5′ and 3′ ends of the *cryA* coding region allow the entire *cryA* sequence to be amplified by the DNA polymerase. The amplified PCR product is shown below with the MboI site recommended originally. Restriction enzymes are MboI, H HindIII, P PstI and T TaqI

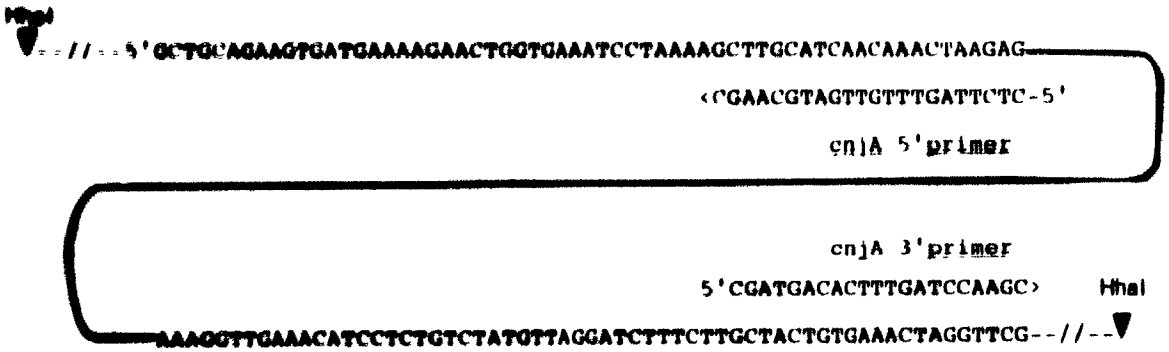

Fig. 4b. **The partial nucleotide sequence of the *cnjA* cDNA is shown along with the location and orientation of each oligonucleotide primer. The *cnjA* 5' primer is located at position +160 to +182 in the sequence and is complementary to the coding strand. The *cnjA* 3' primer is located at position +2277 to +2297 in the sequence and encodes the coding strand**

for this purpose. The anticipated product (1430 bp) would contain part of the 5' coding and upstream region (790 bp) and part of the 3' coding and downstream sequences (640 bp) of the *cnjA* locus. *T. thermophila* genomic DNA was completely digested with HhaI and the fragments ligated to favour monomeric circles (Collins and Weissman, 1984). Primers (Fig. 6b) were designed to allow the outward extension by Taq DNA polymerase from the coding region thereby amplifying the circularized portion of the unknown sequence.

4.2.3 Southern analysis of the IPCR product

The PCR product is shown in Fig. 7 along with the cDNA controls employed. A band of the expected size (1.4 kb) was observed as the major product of the amplification. To verify if the PCR product was indeed the desired one, Southern analysis was performed on *T. thermophila* genomic DNA using the 1.4 kb PCR product as a probe (Fig. 8). The PCR product hybridized strongly and specifically to the 3.5 kb HhaI fragment as expected (Fig. 8, lane 1). The most prominent band in the PstI digest (Fig. 8, lane 2) is seen to be the 2.9 kb fragment representing the 5' upstream region of *cnjA* whereas the 2.1 kb core fragment, sharing only 100 bp with the expected PCR product, appears as a weakly hybridizing band. A large (9 kb) PstI fragment, weakly hybridizing to the probe, may represent the next PstI site at the gene's 3' end. The strongest signal in the HindIII digest (Fig. 8, lane 3) is seen to be the 1.7 kb fragment containing the 5' upstream region. A weaker signal from a 2.7 kb band may represent the 3' HindIII fragment and the larger weaker band incomplete cutting. The TaqI digest (Fig. 8, lane 4), as expected, shows a strong signal to the 1.0 kb 5' flanking region and

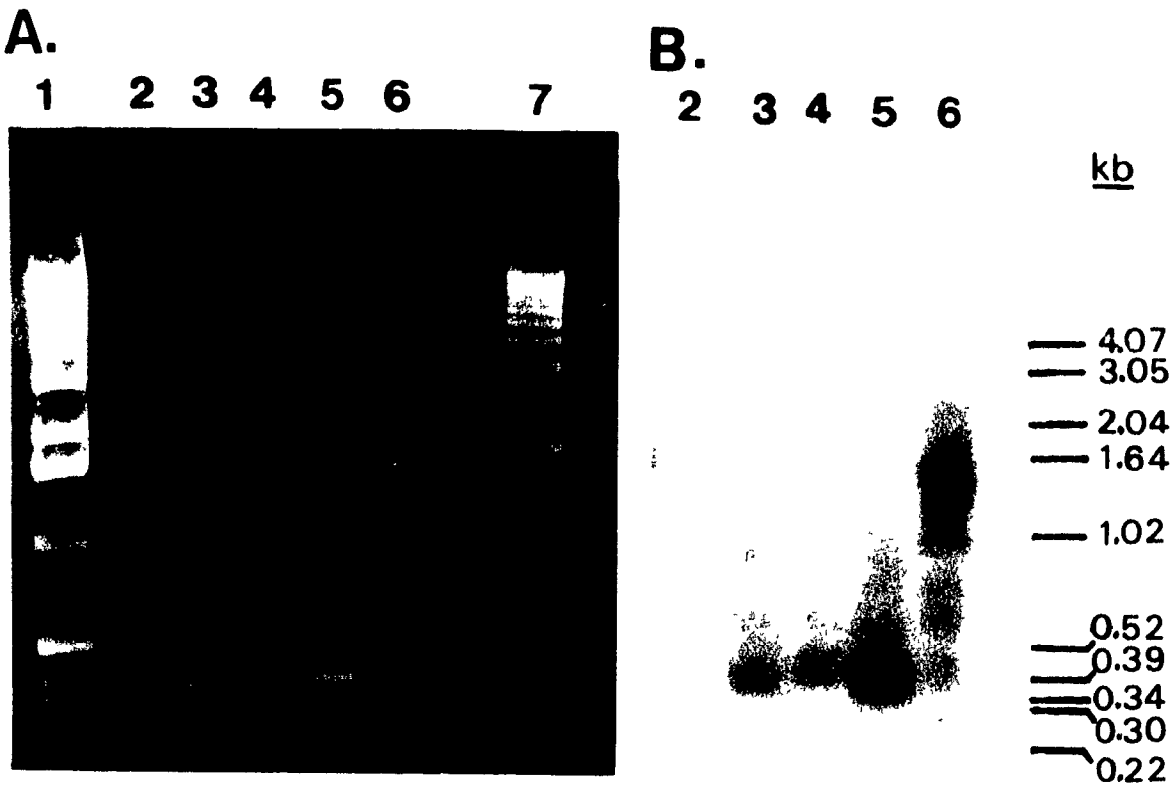

Fig. 7. Result of the IPCR amplification. a) Ethidium bromide stained gel; one tenth volume of each reaction was run on a 1.2% agarose gel. Lanes 1 and 7 show the 1 kb ladder standard; Lane 2, no DNA; Lanes 3 to 5 represent cDNA controls treated in the same way as the genomic fragments. Lane 3, the 2.42 kb insert of pAR103 with an expected PCR product of 340 bp; lane 4, the 2.47 kb insert of pAR105 with an expected product of 390 bp; lane 5, the 2.40 kb insert of pAR106 with an expected product of 350 bp and lane 6, amplified *T. thermophila* genomic HhaI fragments with an expected product of 1.43 kb. b) Resultant autoradiograph of the filter from the above gel probed with the 2.42 kb insert of pAR103.

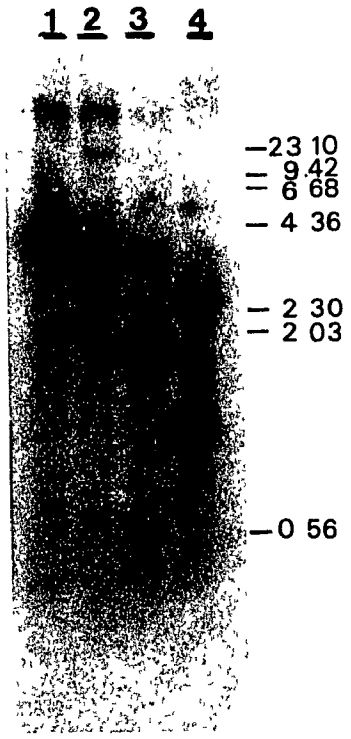

Fig. 8. Southern analysis of *T. thermophila* genomic DNA digested with: lane 1, HhaI; lane 2, PstI; lane 3, HindIII and lane 4, TaqI. Five μg of each digest were loaded per lane and probed with the 1.4 kb PCR product.

a slightly weaker signal to the 2.3 kb fragment containing both the core and 3' flanking regions of *cnjA*.

The Southern analysis confirmed the 1.4 kb amplification product to be the product anticipated. Repeated attempts at subcloning this PCR product have failed (see Materials and Methods). This region therefore remains as yet uncharacterized.

5. DISCUSSION

Previous studies have shown *cnjA* to be stage-specific, highly transcribed (Martindale and Bruns, 1983), under tight transcriptional control (Martindale *et al.*, 1985; Stargell *et al.*, 1990) and peaking in expression just prior to or at pachytene in micronuclear prophase I (Martindale *et al.*, 1985). The *cnjA* cDNA was sequenced in the hope of gaining an insight into the gene's function during early meiosis.

A single long ORF encodes a putative protein of 779 amino acids (89.5 kDa) that is hydrophilic along its length (Fig. 5) and basic in nature (pI 9.3). The deduced protein shows no sequence similarity to any protein in the NBRF or Swiss Protein databases. The size of the ORF (2337 bp) is consistent with the estimated size of the *cnjA* mRNA (2.65 kb) seen in Northern blots (Fig. 3). The flanking 5' and 3' nontranslated regions are highly AT-rich (82.5% and 87.7% respectivly) when compared to the coding region (65.8% A+T). This AT-richness is characteristic of ciliate intergenic regions (>75%) (Horowitz *et al.*, 1987; Brunk and Sadler, 1990). Transcription initiation in *cnjA* is also consistent with that observed in other *T. thermophila* nuclear mRNAs. Csank and Martindale (manuscript in preparation) found an average of 4 transcription start sites when *T. thermophila* genes coding for 12 different nuclear mRNAs were examined. These start sites are located approximately 58 nts from the translation initiation codon with a mean A-content of 68% in the region spanning the most proximal transcription start site to the start codon. Four transcription start sites were identified in the *cnjA* mRNA (1 minor and 3 major) (Fig. 4) with the most proximal located 45 nts from the proposed translation start. This region contains 62.5% A-residues (5 bases not determined) (Fig. 2).

A 73 bp inverted sequence was found at the extreme 5' end (at position -113 to -40) of the *cnjA* cDNA with respect to a 3' end sequence (located at +2356 to +2428) (Fig. 2). The 3' end repeat was verified by sequencing independently derived *cnjA* cDNA clones (Fig. 1) and proved to be a genuine component of the *cnjA* 3' nontranslated region. There are two factors that support the notion that the 5' end repeat found in the original *cnjA cDNA clone* (pC1-2) is an artifact; 1) the transcription start sites fall near the 3' end of this repeat region (Fig. 2) and 2) evidence from *cnjB* (pC2-1), where a 200 bp 5' end repeat was absent from the genomic sequence and therefore a cloning artifact (Taylor and Martindale, submitted) introduced during the construction of the same cDNA library from which pC1-2 was isolated. This evidence would lead us to propose that conditions for the original cDNA synthesis were such as to favour a snap-back phenomenon of the reverse transcriptase (Fig. 9). The coding regions of both *cnjA* and *cnjB* contain frequent stretches of A-residues. Since the cDNA library was primed with oligo (dT), the synthesis most likely initiated internally. This is supported by the fact that neither cDNA (pC1-2 nor pC2-1) contained a poly(A)$^+$ tail. We propose that the reverse transcriptase copied a portion of the coding region and part of the 5' nontranslated region. The new strand then melted off the template, looped back and hybridized to an A-rich island in the 3' nontranslated region of the mRNA where the reverse transcriptase reinitiated and continued to elongate the new strand. Since the 3' flanking region is rich in T-residues, the copied sequences would be A-rich and resemble the 5' leader sequences of ciliate mRNAs.

In addition to a similar expression pattern (Martindale *et al.*, 1985), certain common features are shared among the three conjugation specific

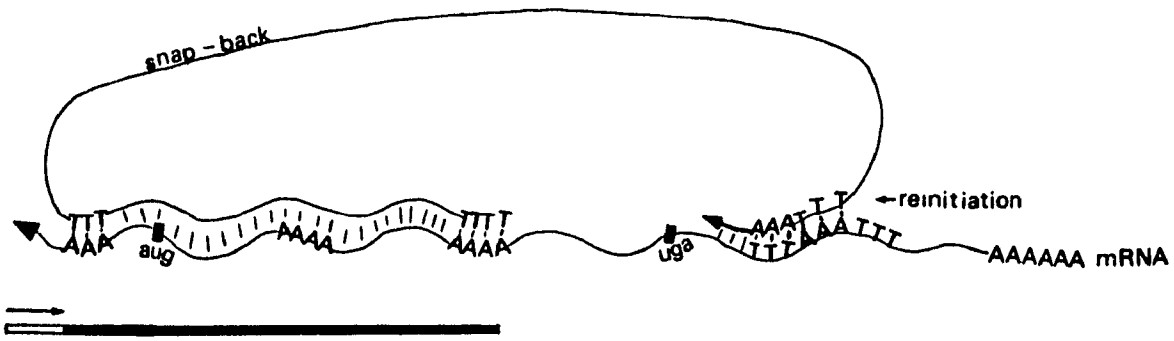

Fig. 9. A proposed model for the generation of the 5' end inverted repeat during the original cDNA synthesis. Conditions may have been such as to allow the reverse transcriptase to melt off the template, after having copied part of the 5' leader sequence, loop back and hybridize to an A-stretch in the 3' nontranslated region. Following reinitiation, the enzyme continued to elongate the new strand for a short distance and finally to terminate just prior to the translation stop codon. Since the copied 3' nontranslated region is T-rich, the resultant duplicated sequence would be A-rich and found in an inverted manner at the extreme 5' end of the cDNA.

genes (*cnjA, cnjB* and *cnjC*). The lower G+C content in the genes' coding region is reflected by a lower Fpr value (Martindale, 1989) when compared to highly expressed *T. thermophila* genes, indicating a tendency toward more AT-rich codons. This may reflect that optimal codon usage is not necessary for stage-specific genes. It has been shown that codon preference is strongly correlated with the abundance of isoaccepting tRNAs in yeast and *E. coli* (Ikemura, 1985) but little is known about tRNA abundance in *T. thermophila*. Martindale (1989) suggests that the conjugation specific genes (expressed only for a brief period) may not be under the same selection pressure as constitutively expressed genes, since the latter need to optimize translation time during vegetative growth and therefore select against minor tRNAs. Recent evidence from yeast data suggests that codon choice may also reflect mRNA stability. Surosky and Esposito (1992) found that the early meiotic transcripts of *SPO11, SPO13* and *SPO16* are highly unstable and contain a large percentage of rare codons in their coding region. The mRNA of *SPO13* was greatly stabilized when nonsense or frameshift mutations were introduced. Hoekema *et al.* (1987) also demonstrated that the *PGK1* mRNA became three times less stable when the coding region, containing mostly optimal codons, was changed to rare codons without changing the protein sequence. Choice of AT-rich codons versus optimal codon usage in the *T. thermophila* conjugation specific genes may also reflect mRNA stability and therefore influence, at least in part, their temporal expression.

Cis-acting sequence elements that positivly or negativly control early meiotic transcripts in yeast have been documented (Buckingham *et al.*, 1990; Vershon *et al.*, 1992) but have not yet been identified in *T. thermophila*. The best candidate for identifying such elements is *cnjA*, since

it constitutes the most abundant transcript of the three conjugation specific genes examined (Martindale and Bruns, 1983). Unfortunately, all attempts at isolating the gene's promoter region have failed so far.

The nucleotide sequence surrounding the translation initiation codon for both *cnjA* and *cnjB* seem to deviate from the consensus established for ciliate mRNAs by Brunk and Sadler (1990) and Yamauchi (1991). They determined the consensus sequence AAA<u>ATG</u>$^{A}_{G}{}^{C}_{G}{}^{T}_{A}$ (61 and 26 ciliate genes examined, respectively) in which A-residues at position -1 to -3 are found 70-90% of the time and the majority of genes contain a guanine at position +4. Of the three conjugation specific genes, only *cnjC* (TAA<u>ATG</u> GAA) follows the consensus sequence closely whereas *cnjA* (TCC<u>ATG</u>TCT) conforms the least and *cnjB* (CCT<u>ATG</u>GAG) only retains the purine at postion +4. Yamauchi (1991) also examined the 13 bases of the 5' leader preceding the translation start. He found that an A-residue frequency of 49-65% and a G+C content of 2-11% was the norm. Table 2 shows a subpopulation of genes with their nucleotide composition in the -13 window to demonstrate differences and similarities between the conjugation specific genes and highly expressed genes.

The consensus sequence established for the translation initiation in ciliate mRNAs was compiled from a small sample size (few ciliate genes have been sequenced so far) and is highly biased toward constitutively expressed genes, such as genes coding for actin, tubulin, histone and ribosomal proteins. Even though it has been shown that a preference for a purine at position -3 is demonstrated for mRNAs among eukaryotic groups as a whole (Cavener and Ray, 1991), one may wonder at the significance of A-residues at positions -1 to -3 in an organism that

Table 2: Nucleotide composition in the -13 window of ciliate nuclear mRNAs.

		nt composition (%)				
organism	gene	A	G	C	G+C	ref.
T.thermophila	*cnjA*	46.2	0	23.0	23.0	-
	cnjB	7.7	15.4	23.0	38.4	Taylor and Martindale, submitted
	cnjC	54.0	7.7	0	7.7	Martindale, 1990
	ileA	46.2	15.4	7.7	23.1	Csank and Martindale, 1992
	HI-4	69.2	0	15.4	15.4	Horowitz *et al.*, 1987
	HII-4	46.2	7.7	23.0	30.7	" "
	SerH3	84.6	0	7.7	7.7	Tondravi *et al.*, 1990
	HMGB	61.5	0	15.4	15.4	Schulman *et al.*, 1991
T.pyriformis	ubiquitin	69.2	30.8	0	30.8	Neves *et al.*, 1991
E.crassa	actin	53.8	0	15.4	15.4	Harper and Jahn, 1989

is notoriously AT-rich in its flanking regions. The significance of the novel nucleotide composition surrounding the translation initiation codon in *cnjA* and *cnjB* is unclear as yet and must await the analysis of more *T. thermophila* genes.

Very little sequence similarity was seen among the sequenced conjugation specific genes *cnjA*, *cnjB* and *cnjC*. A stretch of 7 nucleotides (5'AACATCC3') in the 5' nontranslated leader is shared among *cnjA* (-7 to -1) and *cnjB* (-27 to -21) just upstream of the ATG but was not detected in the 5' leaders of other *T. thermophila* genes (including *cnjC*). A small amino acid similarity is found at the N-terminal of both the putative proteins, CnjA (119-KTVFAKIKL) and CnjB (229-KDGTNKIKL). The significance, if any, of these sequence similarities must await further analysis.

What role does *cnjA* play during meiosis in *T. thermophila*? The transcriptional product of *cnjA* constitutes approximately 0.8% of the total mRNA during the height of its expression (Martindale and Bruns, 1983) which occurs just prior to or at pachytene (Martindale *et al.* 1985). If the protein is also highly expressed and translation follows a similar pattern as transcription, one may speculate at a possible structural role for CnjA during chromosome pairing and/or recombination. Though no obvious DNA binding motifs have been detected in the protein sequence, its inherent basic nature or as yet undetermined motifs may allow DNA binding. Suhr-Jessen (1984) identified two proteins (84 and 86) during the crescent stage of micronuclear prophase I that are highly expressed (2-5 hours into conjugation), are basic (pI 8.5) and have a molecular weight of 90 and 78 kDa respectively. Martindale *et al.* (1985) also show

[illegible]

[illegible]

[illegible]

experimentally whether CnjA is phosphorylated and what effect phosphorylation has on its function.

It has been estimated that between 50-200 genes are involved during meiosis in yeast (Esposito and Klapholz, 1981). Only a fraction of these genes have been sequenced so far and most display no sequence similarity to any known proteins. An interesting feature of the sequenced early meiotic genes is that most are hydrophilic (either basic or acidic) or contain stretches of polar or acidic residues mainly at their C-termini. Cross-hybridization between *cnjA* and yeast genomic DNA has shown no sequence conservation at the DNA level (Martindale *et al*, unpublished data.

Future Prospects

Since the protein sequence of *cnjA* did not provide any clue as to its possible function during early meiosis in *T thermophila*, future areas of study should address the following questions.

1. Do cis-acting elements also function in *T thermophila* in controlling the expression of early meiotic transcripts as they do in yeast? Since *cnjA* is highly transcribed during its brief period of expression, the gene represents an ideal candidate for answering this question. Another attempt at isolating its promoter region could utilize the IPCR method with primers that contain restriction enzyme sites in order to facilitate subcloning.

2. To identify the native protein, antibodies need to be raised against CnjA. This would allow the purification and characterization of the native protein on which enzymatic studies can be performed in order to determine its properties. There are different routes available for raising antibodies to a ciliate gene, either via a synthetic peptide (Harlow and Lane, 1988), expressing a truncated form of the protein in *E. coli* (Studier *et al.*, 1990) or expressing the complete protein in an *E. coli* suppressor strain for the UAA and UAG tRNAs (developed by Cohen *et al.*, 1990).

3. Where does cnjA localize? Is the protein able to enter the micronucleus and does it associate with the meiotic chromosomes? An in situ-hybridization with fluorescently tagged antibodies would allow the localization of CnjA during conjugation.

4. Is *cnjA* as efficiently translated as it is transcribed? Attempts can be made to translate the *cnjA* message (either native or expressed from a plasmid) in a cell free extract of *T. thermophila*

5. Is CnjA phosphorylated? Isolation and purification of the native protein via antibodies would allow selective in vitro phosphorylation studies.

6. Is CnjA essential for meiosis in *T. thermophila* ? Recently, a *T. thermophila* transformation system has been developed by Yao and Yao (1991). Transformation occured by replacement of the host sequences (ribosomal protein L29) via a nonrandom integration. This system could be adapted to determine the role of *cnjA* during early conjugation in

a similar manner as studies are performed with yeast through gene replacement or knockout experiments (Rothstein, 1983).

6. Contribution to knowledge

This study represents one of the few stage specific genes characterized in *T. thermophila* thereby enriching the sequence data for ciliates. It was shown that all three conjugation specific genes demonstrate an unusual codon choice. Though the protein sequence does not as yet identify the gene's function during early meiosis, future comparisons with meiotic proteins isolated from other eukaryotic organisms may shed light on the conservation or divergence of this ancient process between ciliates and their eukaryotic counterparts.

7. REFERENCES

Aboussekhra, A., R. Chanet, A. Adjiri and F. Fabre. 1992. Semidominant suppressors of Srs2 helicase mutations of *Saccharomyces cerevisiae* map in the *RAD51* gene, whose sequence predicts a protein with similarities to prokaryotic RecA proteins. Mol. Cell. Biol. 12: 3224-3234

Aguilera, A. and H. L. Klein. 1990. *HPR1*, a novel yeast gene that prevents intrachromosomal excision recombination, shows carboxy-terminal homology to the *Saccharomyces cerevisiae TOP1* gene. Mol. Cell. Biol. 10 1439-1451.

Alani, E., R. Padmore and N. Kleckner. 1990. Analysis of wild-type and *rad50* mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. Cell 61: 419-436.

Atcheson, C. L., B. DiDomenico, S. Frackman, R. E. Esposito and R. T. Elder. 1987. Isolation, DNA sequence, and regulation of a meiosis-specific eukaryotic recombination gene. Proc. Natl. Acad Sci. USA 84: 8035-8039.

Baker, B. S., A. T. C. Carpenter, M. S. Esposito, R. E. Esposito and L. Sadler. 1976. The genetic control of meiosis. Annu. Rev. Genet. 10: 53-134.

Bang, D., R. Verhage, N. Goosen, J. Brouwer and P. van de Putte. 1992. Molecular cloning of *RAD16*, a gene involved in differential repair in *Saccharomyces cerevisiae* Nucleic Acids Res. 20: 3925-3931.

Bankman, M., L. Prakash and S. Prakash. 1992. Yeast *RAD14* and human xeroderma pigmentosum group A DNA-repair genes encode homologous proteins. Nature 355: 555-558.

Basile, G., M. Aker and R. K. Mortimer. 1992. Nucleotide sequence and transcriptional regulation of the yeast recombinational repair gene *RAD51*. Mol. Cell. Biol. 12. 3235-3246.

Baum, P., C. Furlong and B. Byers. 1986. Yeast gene required for spindle pole body duplication: homology of its product with Ca^{2+} binding protein. Proc. Natl. Acad. Sci. USA 83: 5512-5516.

Benson, F. E., G. T. Illing, G. J. Sharples and R. G. Lloyd. 1988 Nucleotide sequencing of the *ruv* region of *Escherichia coli* K-12 reveals a LexA regulated operon encoding two genes. Nucleic Acids Res. 16: 1541-1549.

Bhargara, J., J. Engebrecht and G. S. Roeder. 1992. The *rec102* mutant of yeast is defective in meiotic recombination and chromosome synapsis. Genetics 130: 59-69.

Birkenbihl, R. P. and S. Subramani. 1992. Cloning and characterization of *rad21* an essential gene of *Schizosaccharomyces pombe* involved in DNA double-strand-break repair. Nucleic Acids Res. 20: 6605-6611.

Birnboim, H. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. 7: 2141-2147.

Bishop, D. K., D. Park, L. Xu and N. Kleckner. 1992. *DMC1*: a meiosis-specific yeast homolog of *E coli* recA required for recombination, synaptonemal complex formation, and cell cycle progression. Cell 69: 439-456.

Blackburn, E. H. and J. G. Gall. 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA gene in *Tetrahymena*. J. Mol. Biol. 120: 33-53

Blackburn, E. H. and K. M. Karrer. 1986. Genomic reorganization in ciliated protozoa. Annu. Rev. Genet. 20: 501-521.

Boorstein, W. R. and E. A. Craig. 1989. Primer extension analysis of RNA. Methods Enzymol. 180: 347-369.

Broach, J. R. 1991. *RAS* genes in *Saccharomyces cerevisiae* signal transduction in search of a pathway. Trends in Genetics 7: 28-32.

Broek, D., T. Toda, T. Michaeli, L. Levin. C. Birchmeier, M. Zoller, S. Powers and M. Wigler. 1987. The *S cerevisiae CDC25* gene product regulates the RAS/adenylate cyclase pathway. Cell 48: 789-799.

Brunk, C. F. and L. A. Sadler. 1990. Characterization of the promoter region in *Tetrahymena* genes. Nucleic Acids Res. 18: 323-329.

Bruns, P. J. and T. B. Brussard. 1974a. Pair formation in *Tetrahymena pyriformis*, an inducible developmental system J. Exp. Zool. 188: 337-344.

Bruns, P. J. and T. B. Brussard. 1974b. Positive selection for mating with functional heterokaryons in *Tetrahymena pyriformis*. Genetics 78 831-841.

Bruns, P. J., T. B. Brussard and E. V. Merriam. 1983. Nullisomic *Tetrahymena* II. A set of nullisomics define the germinal chromosomes. Genetics 104: 257-270.

Buckingham, L. E., H.-T. Wang, R. T. Elder, R. M. McCarroll, M. R. Slater and R. E. Esposito. 1990. Nucleotide sequence and promoter analysis of *SPO13*, a meiosis-specific gene of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 87: 9406-9410.

Cao, L., E. Alani and N. Kleckner. 1990. A pathway of generation and processing of double-strand breaks during meiotic recombination in *S. cerevisiae* Cell 61. 1089-1101.

Carpenter, A. T. C. 1975. Electron microscopy of meiosis in *Drosophila melanogaster* females II. The recombination nodule - a recombination-associated structure at pachytene? Proc. Natl Acad. Sci. USA 72: 3186-3189

Carpenter, A. T. C. 1979. Synaptonemal complex and recombination nodules in wild-type *Drosophila melanogaster* females. Genetics 92: 511-541.

Carpenter, A. T. C. 1987. Gene conversion, recombination nodules and the initiation of meiotic synapsis. Bio Essays 6: 232-236.

Cavener, D. R. and S. C. Ray. 1991. Eukaryotic start and stop translation sites. Nucleic Acids Res. 19: 3185-3192.

Choi, I. S., J. B. Kim and S. D. Park. 1990. Nucleotide sequence of *RAD4* gene of *Saccharomyces cerevisiae* that can be propagated in *Escherichia coli* without inactivation. Nucleic Acids Res. 18: 7137.

Chow, T. Y.-K., E. L. Perkins and M. A. Resnick. 1992. Yeast *RNC1* encodes a chimeric protein, RhoNUC, with a human rho motif and deoxyribonuclease activity. Nucleic Acids Res. 20: 5215-5221.

Churchill, M. E. A. and A. A. Travers. 1991. Protein motifs that recognize structural features of DNA. Trends in Biochem. 16: 92-97.

Clark, A. B., C. C. Dykstra and A. Sugino. 1991. Isolation, DNA sequence, and regulation of a *Saccharomyces cerevisiae* gene that encodes DNA strand transfer protein α. Mol. Cell. Biol 11: 2576-2582.

Cohen, J., P. Dupuis and B. Vigues. 1990. Expression of a ciliate in *Escherichia coli* using a suppressor tRNA to read the UAA and UAG glutamine codons. J. Mol. Biol. 216: 189-194.

Collins, F. S. and S. M. Weissman. 1984. Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. Proc. Natl. Acad. Sci. USA 81: 6812-6816.

Conley, E. C. and S. C. West. 1989. Homologous pairing and the formation of nascent synaptic intermediates between regions of duplex DNA by RecA protein. Cell 56: 987-995.

Cool, M. and R. E. Malone. 1992. Molecular and genetic analysis of the yeast early meiotic recombination genes *REC102* and *REC107/MER2*. Mol. Cell. Biol. 12: 1248-1256.

Covitz, P. A., I. Herskowitz and A. P. Mitchell. 1991. The yeast *RME1* gene encodes a putative zinc finger protein that is directly repressed by a1-α2. Genes and Devel. 5: 1982-1989.

Csank, C. and D. W. Martindale. 1992. Isoleucyl-tRNA synthetase from the ciliated protozoan *Tetrahymena thermophila* J. Biol. Chem. 267: 4592-4599.

Csank, C. and D. W. Martindale. A compilation of *Tetrahymena thermophila* transcription start site sequences. Manuscript in preparation.

Cupples, C. G. and R. E. Pearlman. 1986. Isolation and characterization of the actin gene from *Tetrahymena thermophila* Proc. Natl. Acad. Sci. USA 83: 5160-5164.

van Daal, A., E. M. White, S. C. R. Elgin and M. A. Gorovsky. 1990. Conservation of intron position indicates separation of major and variant H2As is an early event in the evolution of eukaryotes. J. Mol. Evol. 30: 449-455.

Davis, R. L., H. Weintraub and A. B. Lasser. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell 51: 987-1000.

Dedon, P. C., J. A. Soults, D. C. Allis and M. A. Gorovsky. 1991. Formaldehyde cross-linking and immunoprecipitation demonstrate developmental changes in H1 association with transcriptionally active genes. Mol Cell. Biol. 11: 1729-1733.

Dever, T. E., M. J. Glynias and W. C. Merrick. 1987. GTP-binding domain: three consensus sequence elements with distinct spacing. Proc. Natl. Acad. Sci. USA 84: 1814-1818.

DeVoti, J., G. Seydoux, D. Beach and M. McLeod. 1991. Interaction between ran1+ protein kinase and cAMP dependent protein kinase as negative regulators of fission yeast meiosis. EMBO J. 10: 3759-3768.

Dodd, I. B., and J. B. Egan. 1987. Systematic method for the detection of potential lambda Cro-like DNA-binding regions in proteins. J. Mol. Biol. 194: 557-564.

Dresser, M. E. and C. N. Giroux. 1988. Meiotic chromosome behavior in spread preparations of yeast. J. Cell. Biol. 106: 567-573.

Dranginis, A. M. 1990. Binding of yeast a1 and α2 as a heterodimer to the operator DNA of a haploid-specific gene. Nature 347: 682-685.

Dykstra, C. C., R. K. Hamatake and A. Sugino. 1990. DNA strand transfer protein ß from yeast mitotic cells differs from strand transfer protein α from meiotic cells. J. Biol. Chem. 265: 10968-10973.

Dykstra, C. C., K. Kitada, A. B. Clark, R. K. Hamatake and A. Sugino. 1991. Cloning and characterization of *DST2*, the gene for DNA strand transfer protein ß from *Saccharomyces cerevisiae*. Mol. Cell. Biol. 11: 2583-2592.

Emery, H. S., D. Schild, D. E. Kellogg and R. K. Mortimer. 1991. Sequence of *RAD54*, a *Saccharomyces cerevisiae* gene involved in recombination and repair. Gene 104: 103-106.

Engebrecht, J. and G. S. Roeder. 1989. Yeast *mer1* mutants display reduced levels of meiotic recombination. Genetics 121: 237-247.

Engebrecht, J. and G. S. Roeder. 1990. *MER1*, a yeast gene required for chromosome pairing and genetic recombination, is induced in meiosis. Mol. Cell. Biol. 10: 2379-2389.

Engebrecht, J., K. Voelkel Meiman and G. S. Roeder. 1991. Meiosis-specific RNA splicing in Yeast Cell 66: 1257-1268.

Esposito, R. E. and S. Klapholz. 1981. Meiosis and ascospore development. In: J. N. Strathern, E. W. Jones and J. R. Broach (eds.) The molecular biology of the yeast *Saccharomyces*, life cycle and inheritance. Cold Spring Harbor Lab., Cold Spring Harbor, New York, pp. 211-287.

Evans, R. M. and S. M. Hollenberg. 1988. Zinc fingers: guilt by association. Cell 52: 1-3.

Farnet, C., R. Padmore, L. Cao, E. Alani and N. Kleckner. 1988. The *RAD50* gene of *S cerevisiae*. UCLA Symp. Mol. Cell. Biol. New. Ser. 83: 201-215.

Fedor-Chaiken, M., R. J. Deschenes, J. R. Broach. 1990. *SRV2*, a gene required for *RAS* activation of adenylate cyclase in yeast. Cell 61: 329-340.

Field, J., A. Vojtek, R. Ballester, G. Bolger, J. Colicelli, K. Ferguson, J. Gerst, T. Kataoka, T. Michaeli, S. Powers, M. Riggs, L. Rodgers, I. Wieland, B. Wheland and M. Wigler. 1990. Cloning and characterization of *CAP*, the *S. cerevisiae* gene encoding the 70 kd adenylyl cyclase-associated protein. Cell 61: 319-327.

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

Frasch, M. 1991. The maternally expressed *Drosophila* gene encoding the chromatin-binding protein BJ1 is a homolog of the vertebrate gene regulator of chromatin condensation, *RCC1* EMBO J. 10: 1225-1236.

Game, J. C. and R. K. Mortimer. 1974. A genetic study of X-ray sensitive mutants in yeast. Mutat. Res. 24: 281-292.

Game, J. C. 1983. Radiation-sensitive mutants and repair in yeast. In: Yeast genetics: fundamental and applied aspects, J. F. T. Spencer, D. H. Spencer and A. R. W. Smith (eds.) Springer Verlag, New York, pp. 109-137.

Garcia-Bustos, J., J. Heitman and M. N. Hall. 1991. Nuclear protein localization. Biochim. Biophys. Acta 1071: 83-101.

Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko and V. M. Blinov. 1988 A conserved NTP-motif in putative helicases. Nature 333: 22.

Granot, D., J. P. Margolskee and G. Simchen. 1989. A long region upstream of the *IME1* gene regulates meiosis in yeast. Mol. Gen. Genet. 218: 308-314.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166: 557-580.

Hanks, S. K. and A. M. Quinn. 1991. Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. Meth. Enzymol. 200: 38-78.

Hansen, T. S., P. H. Andraesen, H. Dreisig, P. Hojup, H. Nielsen, J. Engberg and K. Kristiansen. 1991. *Tetrahymena thermophila* acidic ribosomal protein L37 contains an archaebacterial type C-terminus. Gene 105: 113-150.

Hanyu, N., Y. Kuchino. S. Nishimura and H. Beier. 1986. Dramatic events in ciliate evolution: alteration of UAA and UAG termination codons to glutamine codons due to anticodon muta tions in two *Tetrahymena* $tRNA_s^{Gln}$. EMBO J. 5: 1307-1311.

Harlow, E. and D. Lane. 1988. Antibodies, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Harosh, I. and P. Deschavanne. 1991. The *RAD3* gene is a member of the DEAH family RNA helicase-like protein. Nucleic Acids Res. 19: 6331.

Harper, D. S. and C. Jahn. 1989. Differential use of termination codons in ciliated protozoa. Proc. Natl. Acad. Sci. USA 86: 3252-3256.

Hartwell, L. H. 1974. *Saccharomyces cerevisiae* cell cycle. Bact. Rev. 38: 165-198.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28: 351-359.

Henderson, A. J. 1970. The time and place of meiotic crossing over Annu. Rev. Genet. 4: 295-324.

Hodgman, T. C. 1988. A new superfamily of replicative proteins Science 333: 22-23.

Hoekema, A., R. A. Kastelein, M. Vasser and H. A. deBoer. 1897. Codon replacement in the *PGK1* gene of *Saccharomyces cerevisiae* experimental approach to study the role of biased codon usage in gene expression. Mol. Cell. Biol. 7. 2914-2924.

Holden, G. W., A. Spanos and G. R. C. Banks. 1989. Nucleotide sequence of the *REC1* gene of *Ustilago maydis* Nucleic Acids Res. 17: 10489.

Hollingsworth, N. M. and B. Byers. 1989. *HOP1* a yeast meiotic pairing gene. Genetics 121: 445-462.

Hollingsworth, N. M., L. Goetsch and B. Byers. 1990. The *HOP1* gene encodes a meiosis-specific component of yeast chromosomes. Cell 61: 73-84.

Horowitz, S. and M. A. Gorovsky. 1985. An unusual genetic code in nuclear genes of *Tetrahymena* Proc. Natl. Acad. Sci. USA 82: 2452-2455.

Horowitz, S., J. K. Bowen, G. A. Bannon and M. A. Gorovsky. 1987. Unusual features of transcribed and translated regions of the histone H4 gene family of *Tetrahymena thermophila* Nucleic Acids Res. 15: 141-160.

Hummel, R., P. Norgaard, P. H. Andraesen, K. Skjodt and K. Kristiansen. 1991. A *Tetrahymena* gene encoding a protein with the liver-specific F-antigen Manuscript submitted.

Howard, E. A. and E. H. Blackburn. 1985. Reproducible and variable genomic rearrangements occur in the developing nucleus of the ciliate *Tetrahymena thermophila*. Mol. Cell. Biol. 5: 2035-2050.

Hunt, L. T. and M. O. Dayhoff. 1970. The occurence in proteins of tripeptides ASN-X-SER and ASN-X-THR and of bound carbohydrate. Biochim Biophys Res. Commun. 39: 757-756.

Ikemura, T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms Mol. Biol. Evol. 2. 13-34.

Johnson, R. E., S. T. Henderson, T. D. Petes, S. Prakash, M. Bankmann and L. Prakash. 1992. *Saccharomyces cerevisiae* *RAD5*-encoded DNA repair protein contains DNA helicase and zinc-binding sequence motifs and affects the stability of simple repetitive sequences in the genome. Mol. Cell. Biol. 12: 3807-3818.

Jones, G. H. 1984. The control of chiasmata distribution. In: controlling events in meiosis. C. W. Evans and H. G. Dickinson (eds.) (Cambridge, Society for Experimental Biology). pp. 293-320.

Kao, G., J. C. Shah and M. J. Clancy. 1990. An *RME1*-independent pathway for sporulation control in *Saccharomyces cerevisiae* acts through *IME1* transcript accumulation. Genetics 126: 823-835.

Kassir, Y., D. Granot and G. Simchen. 1988. *IME1*, a positive regulator gene of meiosis in *S cerevisiae* Cell 52: 853-862.

Kataoka, T., D. Broek and M. Wigler. 1985. DNA sequence and characterization of the *S cerevisiae* gene encoding adenylate cyclase. Cell 43: 493-505.

King, P. V. and R. W. Blakesley. 1986. Optimizing DNA ligations for transformation. Focus 8: 1-3.

Klapholz, S. and Esposito, R. E. 1980. Recombination and chromosome segregation during single division meiosis in *spo12-1* and *spo13-1* diploids. Genetics 96: 589-611.

Klapholz, S., C. S. Waddell and R. E. Esposito. 1985. The role of the *SPO11* gene in meiotic recombination in yeast. Genetics 110: 187-216.

Krans, J. A. and R. K. Mortimer. 1991. Nucleotide sequence of the *RAD57* gene in *Saccharomyces cerevisiae*. Gene 105: 139-140.

Kyte, J. and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol Biol. 157: 105-132.

Landschulz, W. H., P. F. Johnson and S. L. McKnight. 1988 The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins Science 240: 1759-1764

Leem, S.-H. and H. Ogawa. 1992. The *MRE1* gene encodes a novel protein kinase homologue required for meiotic recombination in *Saccharomyces cerevisiae* Nucleic Acids Res 20 449-457

Lipman, D. J. and W. R. Pearson. 1985 Rapid and sensitive protein similarity searches Science 227: 1435-1441

Malavasic, M. J. and R. T. Elder. 1990. Complementary transcripts from two genes necessary for normal meiosis in the yeast *Saccharomyces cerevisiae* Mol. Cell. Biol. 10: 2809-2819

Malone, R. E. and R. E. Esposito. 1980 The *RAD52* gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast Proc. Natl. Acad Sci. USA 77: 503-507

Malone, R. E. and R. E. Esposito. 1981. Recombinationless meiosis in *Saccharomyces cerevisiae* Mol. Cell. Biol. 1. 891-901.

Malone, R. E. 1983. Multiple mutant analysis of recombination in yeast. Mol. Gen. Genet. 189: 405-412.

Malone, R. E. 1990. Dual regulation of meiosis in yeast Cell 61: 375-378

Malone, R. E., S. Bullard, M. Hermiston, R. Rieger, M. Cool and A. Galbraith. 1991. Isolation of mutants defective in early steps of meiotic recombination in the yeast *Saccharomyces cerevisiae* Genetics 128: 79-88.

Martindale, D. W., C. D. Allis and P. J. Bruns. 1982. Conjugation in *Tetrahymena thermophila*: a temporal analysis of cytological stages. Exp. Cell. Res. 140: 227-236.

Martindale, D. W. and P. J. Bruns. 1983. Cloning of abundant mRNA species present during conjugation of *Tetrahymena thermophila* identification of mRNA species present exclusively during meiosis. Mol. Cell. Biol. 3: 1857-1865.

Martindale, D. W., C. D. Allis and P. J. Bruns. 1985. RNA and protein synthesis during meiotic prophase in *Tetrahymena thermophila* J. Protozool 32: 644-649.

Martindale, D. W., H. M. Martindale and P. J. Bruns. 1986. *Tetrahymena* conjugation-induced genes: structure and organization in macro- and micronuclei. Nucl. Acids Res. 14: 1341-1353.
Martindale, D. W. 1989 Codon usage in *Tetrahymena* and other ciliates. J. Protozool. 36: 29-34.

Martindale, D. W. 1990. A conjugation-specific gene *(cnjC)* from *Tetrahymena* encodes a protein homologous to yeast RNA polymerase subunits (RPB3, RPC40) and similar to a portion of the prokaryotic RNA polymerase α subunit (rpoA). Nucleic Acids Res. 18: 2953-2959.

Matsumoto, K., I. Uno, Y. Oshima and T. Ishikawa. 1982. Isolation and characterization of yeast mutants deficient in adenylate cyclase and cAMP-dependent protein kinase. Proc. Natl. Acad. Sci. USA 79: 2355-2359.

Matsumoto, K. 1983. Initiation of meiosis in yeast mutants defective in adenylate cyclase and cyclic AMP-dependent protein kinase. Cell 32: 417-423.

Matsuura, A., M. Treinin, H. Mitsuzawa, Y. Kassir, I. Uno and G. Simchen. 1990. The adenylate cyclase/protein kinase cascade regulates entry into meiosis in *Saccharomyces cerevisiae* through the gene *IME1* EMBO J. 9: 3225-3232.

Mayo, K. A. and E. Orias. 1981. Further evidence for lack of gene expression in the *Tetrahymena* micronucleus. Genetics 98: 747-762.

McLeod, M., M. Stein and D. Beach. 1987. The product of the mei^{3+} gene, expressed under control of the mating-type locus, induces meiosis and sporulation in fission yeast. EMBO J. 6: 729-736.

Menees, T. M. and G. S. Roeder. 1989. *MEI4*, a yeast gene required for meiotic recombination. Genetics 123: 675-682.

Menees, T. M., P. B. Ross-Macdonald and G. S. Roeder. 1992. *MEI4*, a meiosis-specific yeast gene required for chromosome synapsis. Mol. Cell. Biol. 12: 1340-1351.

Mitchell, A. P. and I. Herskowitz. 1986. Activation of meiosis and sporulation by repression of the *RME1* product in yeast. Nature 319: 738-742.

Mitchell, A. P., S. E. Driscoll and H. E. Smith. 1990. Positive control of sporulation-specific genes by the *IME1* and *IME2* products in *Saccharomyces cerevisiae* Mol. Cell. Biol. 10: 2104-2110.

Murray, J. M., A. M. Carr, A. R. Lehman and F. Z. Watts. 1991. Cloning and characterization of the *rad9* DNA repair gene from *Saccharomyces cerevisiae*. Nucleic Acids Res. 19: 3525-3531.

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

Neigeborn, L. and A. P. Mitchell. 1991. The yeast *MCK1* gene encodes a protein kinase homolog that activates early meiotic gene expression. Genes and Development 5: 533-548.

Neves, A. M., P. Guerreiro, L. Miquerol and C. Rodrigues-Pausada. 1991. Molecular cloning and expression of a *Tetrahymena pyriformis* ubiquitin fusion gene coding for a 53-amino-acid extension protein. Mol. Gen. Genet. 230: 186-192.

Nicolas, A., D. Treco, N. P. Schultes and J. W. Szostak. 1989. An initiation site for meiotic gene conversion in the yeast Saccharomyces cerevisiae. Nature 338: 35-39.

Nomoto, M., N. Imai, H. Saiga, T. Matsui and T. Mita. 1987. Characterization of two types of histone *H2B* genes from macronuclei of *Tetrahymena thermophila* Nucleic Acids Res. 15: 5681-5697.

Ochman, H., A. S. Gerber and D. L. Hartl. 1988. Genetic applications of an inverse polymerase chain reaction. Genetics 120: 621-623.

Orias, E. 1986. Ciliate conjugation. In: J. G. Gall (ed.) Molecular Biology of the ciliated protozoa. Academic Press, New York, New York. pp. 45-84.

Padmore, R., L. Cao and N. Kleckner. 1991. Temporal comparison of recombination and synaptonemal complex formation during meiosis in *S. cerevisiae* Cell 66: 1239-1256.

Park, E., S. N. Guzder, M. H. M. Koken, I. Jasper-Dekker, G. Weeda, J. H. J. Hoeijmakers, S. Prakash and L. C. Prakash. 1992. *RAD25 (SSL2)*, the yeast homolog of the human xeroderma pigmentosum group B DNA repair gene, is essential for viability. Proc. Natl. Acad Sci. USA 89: 11416-11420.

Passmore, S., G. T. Maine, R. Elble, C. Christ and B. Tye. 1988. *Saccharomyces cerevisiae* protein involved in plasmid maintenance is necessary for mating of *MAT*α cells. J. Mol. Biol. 204: 593-606.

Pearson, W. R. and D. J. Lipman. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci USA 85: 2444-2448.

Preer, Jr., J. R. and L. B. Preer. 1979. The size of macronuclear DNA and its relationship to models for maintaining genetic balance. J. Protozool. 26: 14-18.

Raikov, I. B. 1969. The macronucleus of ciliates. In: T. T. Chen (ed.) Research in Protozoology. Pergamon Press, Oxford, New York, 3: 1-13.

Ray, Jr., C. 1956. Meiosis and nuclear behavior in *Tetrahymena pyriformis*. J. Protozool. 3: 86-96.

Resnick, M. A. and P. Martin. 1976 The repair of double-strand breaks in the nuclear DNA of *Saccharomyces cerevisiae* and its genetic control. Mol. Gen. Genet. 143: 119-129.

Resnick, M. A. 1987. Investigating the genetic control of biochemical events in meiotic recombination. In: Meiosis, P. B. Moens (ed.) New York, Academic Press, pp. 119-129.

Roberts, B. 1989. Nuclear location signal-mediated protein transport. Biochim. Biophys. Acta 1008: 263-280.

Robinson, L. C., J. B. Gibbs, M. S. Marshall, I. S. Sigal and K. Tatchell. 1987. *CDC25*: a component of the *RAS*-adenylate cyclase pathway in *Saccharomyces cerevisiae*. Science 235: 1218-1221.

Robinson, L. C. and K. Tatchell. 1991. *TSF1*: a suppressor of *cdc25* mutations in *Saccharomyces cerevisiae*. Mol. Gen. Genet. 230: 241-250.

Rockmill, B. and G. S. Roeder. 1988. *RED1*, a yeast gene required for the segregation of chromosomes during the reductional division of meiosis. Proc. Natl. Acad. Sci. USA 85: 6057-6061.

Rockmill, B. and G. S. Roeder. 1990. Meiosis in asynaptic yeast. Genetics 126: 563-574.

Rogers, M. B. and K. M. Karrer. 1989. Cloning of *Tetrahymena* genomic sequences whose message abundance is increased during conjugation. Devel. Biol 131: 261-268.

Rosendahl, G., P. H. Andraesen and K. Kristiansen. 1991. Structure and evolution of the *Tetrahymena thermophila* gene encoding ribosomal protein L21. Gene 98: 161-167.

Rothstein, R. 1983. One-step gene disruption in yeast. Meth. Enzymol. 101: 670-676.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

Sambrook, J., E. F. Fritsch and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring harbor, New York.

Sander, M., K. Lowenhaupt, W. S. Lane and A. Rich. 1991. Cloning and characterization of Rrp1, the gene encoding *Drosophila* strand transferase: carboxy-terminal homology to DNA repair endo/exonucleases. Nucleic Acids Res. 19: 4523-4529.

Sanger, F., S. Nicklen and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Schulman, I. G., T. Wang, M. Wu, J. Bowen, R. G. Cook, M. A. Gorovsky and C. D. Allis. 1991. Macronuclei and micronuclei in *Tetrahymena thermophila* contain high-mobility-group-like chromosomal proteins containing a highly conserved eleven-amino-acid putative DNA-binding sequence. Mol. Cell. Biol. 11: 166-174.

Scott, M. P., J. W. Tamkun and G. W. Hartzell III. 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta 989: 25-48.

Shah, J. C. and M. J. Clancy. 1992. *IME4*, a gene that mediates *MAT* and nutritional control of meiosis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 12: 1078-1086.

Shero, J. H. and P. Hieter. 1991. A suppressor of a centromere DNA mutation encodes a putative protein kinase (*MCK1*). Genes and Devel. 5: 549-560.

Shinohara, A., H. Ogawa and T. Ogawa. 1992. Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. Cell 69: 457-470.

Smith, H. E. and A. P. Mitchell. 1989. A transcriptional cascade governs entry into meiosis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 9: 2142-2152.

Smith, H. E., S. S. Y. Su, L. Neigeborn, S. E. Driscoll and A. P. Mitchell. 1990. Role of *IME1* expression in regulation of meiosis in *Saccharomyces cerevisiae* Mol. Cell. Biol. 10: 6103-6113.

Stargell, L. A., K. M. Karrer and M. A. Gorovsky. 1990. Transcriptional regulation of gene expression in *Tetrahymena thermophila*. Nucleic Acids Res. 18: 6637-6639.

Strich, R., M. R. Slater and R. E. Esposito. 1989. Identification of negative regulatory genes that govern the expression of early meiotic genes in yeast. Proc. Natl Acad. Sci. USA 86: 10018-10022.

Studier, F. W., A. H. Rosenberg, J. J. Dunn and J. W. Dubendorff. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Meth. Enzymol. 185: 60-89.

Sonnenborn, T. M. 1975. *Tetrahymena pyriformis* In: Handbook of genetics. R. C. King (ed.) Vol II, Plenum, New York, pp. 433-467.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

Sugai, T. and K. Hiwatashi. 1974. Cytological and autoradiographic studies of the micronucleus at meiotic prophase in *Tetrahymena pyriformis*. J. Protozool. 21: 542-548.

Sugino, A., J. Nitiss. and M. A. Resnick. 1988. ATP-independent DNA strand transfer catalysed by protein(s) from meiotic cells of the yeast *Saccharomyces cerevisiae* Proc. Natl. Acad. Sci. USA 85: 3683-3687.

Suhr-Jessen, P. B. 1984. Stage-specific changes in protein synthesis during conjugation in *Tetrahymena thermophila* Exp. Cell Res. 151: 374-383.

Sun, H., D. Treco, N. P. Schultes and J. Szostak. 1989. Double-strand breaks at an initiation site for meiotic gene conversion. Nature 338: 87-90.

Sun, H., D. Treco and J. W. Szostak. 1991. Extensive 3' overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the *ARG4* recombination initiation site. Cell 64: 1155-1161.

Surosky, R. T. and R. E. Esposito. 1992. Early meiotic transcripts are highly unstable in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 12: 3948-3958.

Swanson, C. P., T. Merz and W. J. Young. 1981. In: cytogenetics: the chromosome in division, inheritance and evolution, 2nd ed. Prentice-Hall, Inc., Englewood Cliffs, New York, pp 200-282

Takahagi, M., H. Iwasaki, A. Nakata, and H. Shinagawa. 1991. Maolecular and genetic analysis of *Escherichia coli ruvC* gene, which encodes a holliday junction-specific endonuclease. J. Bacteriol. 173: 5747-5753.

Tanaka, K., K. Matsumoto and A. Toh-e. 1989. *IRA1*, an inhibitory regulator of the *RAS*-cyclic AMP pathway in *Saccharomyces cerevisiae* Mol. Cell. Biol. 9: 757-768.

Tanaka, K., M. Nakafuku, F. Tamanoi, Y. Kaziro, K. Matsumoto and A. Toh-e. 1990a. *IRA2*, a second gene of *Saccharomyces cerevisiae* that encodes a protein with a domain homologous to mammalian *ras* GTPase-activating protein. Mol. Cell. Biol 10: 4303-4313.

Tanaka, K., M. Kakafuku, T. Satoh, M. S. Marshall, J. B. Gibbs, K. Matsumoto, Y. Kaziro and A. Toh-e. 1990b. *S cerevisiae* genes *IRA1* and *IRA2* encode proteins that may be functionally eqivalent to mammalian *ras* GTPase activating protein. Cell 60. 803-807.

Tanaka, K., B. K. Lin, D. R. Wood and F. Tamanoi. 1991. *IRA2*, an upstream negative regulator of *RAS* in yeast, is a RAS GTPase activating protein. Proc. Natl. Acad. Sci. USA 88: 468-472.

Taylor, F. M. and D. W. Martindale. 1993. Retroviral-type zinc fingers and glycine-rich repeats in a protein encoded by *cnjB*, a *Tetrahymena* gene active during meiosis. Manuscript submitted.

Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77: 5201-5205.

Thompson, E. A. and G. S. Roeder. 1989. Expression and DNA sequence of *RED1*, a gene required for meiosis I chromosome segregation in yeast. Mol. Gen. Genet. 218: 293-301.

Tishkoff, D. X., A. W. Johnson and R. D. Kolonder. 1991. Molecular and genetic analysis of the gene encoding the *Saccharomyces cerevisiae* strand exchange protein Sep1. Mol. Cell. Biol. 11: 2593-2608.

Toda, T., I. Uno, T. Ishikawa, S. Powers T. Kataoka, D. Broek, S. Cameron, J. Broach, K. Matsumoto and M. Wigler. 1985. In yeast, *RAS* proteins are controlling elements of adenylate cyclase. Cell 40: 27-36.

Toda, T., S. Cameron, P. Sass, M. Zoller, J. D. Scott, B. McMullen, M. Hurwitz, E. G. Krebs and M. Wigler. 1987a. Cloning and characterization of *BCY1*, a locus encoding a regulatory subunit of the cAMP dependent protein kinase in yeast. Mol. Cell. Biol. 7: 1371-1377.

Toda, T., S. Cameron. P. Sass, M. Zoller and M. Wigler. 1987b. Three different genes in *S. cerevisiae* encode the catalytic subunits of the cAMP-dependent protein kinase. Cell 50: 277-287.

Tondravi. M. M., R. L. Willis, H. D. Love, Jr. and G. A. Bannon. 1990. Molecular characterization of *SerH3*, a *Tetrahymena thermophila* gene encoding a temperature-regulated surface antigen. Mol. Cell. Biol. 10: 6091-6096.

Vershon, A. K., N. M. Hollingsworth and A. D. Johnson. 1992. Meiotic induction of the yeast *HOP1* gene is controlled by positive and negative regulatory sites. Mol. Cell. Biol. 12: 3706-3714.

Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gray. 1982. Distantly related sequences in the α- and ß- subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and common nucleotide binding fold. EMBO J. 1: 945-951

Wang, H.-T., S. Frackman, J. Kowalisyn, R. E. Esposito and R. Elder. 1987. Developmental regulation of *SPO13*, a gene required for separation of homologous chromosomes at meiosis I. Mol. Cell. Biol. 7: 1425-1435.

Watanabe, Y., Y. Iino, K. Furuhata, C. Shimoda and M. Yamamoto. 1988. The *S pombe mei2* gene encoding a crucial molecule for commitment to meiosis is under regulation of cAMP. EMBO J. 7: 761-767.

von Wettstein, D., S. W. Rasmussen and P. B. Holm. 1984. The synaptonemal complex in genetic segregation. Annu. Rev. Genet. 18: 331-413.

White, E. M., C. D. Allis, D. S. Goldfarb, A. Srivastra, J. W. Weir and M. A. Gorovsky. 1989. Nucleus-specific and temporally restricted localization of proteins in *Tetrahymena* macronuclei and micronuclei. J. Cell. Biol. 109: 1983-1992

Whyte, W., L. H. Koepp, J. Lamb, J. C. Crowley and D. B. Kaback. 1990. Molecular cloning of chromosome I DNA from *Saccharomyces cerevisiae*: isolation, characterization and regulation of the *SPO7* sporulation gene. Gene 95: 65-72.

Wolfe, J., B. Hunter and W. S. Adair. 1976. A cytological study of micronuclear elongation during conjugation in *Tetrahymena* Chromosoma 55: 289-308.

Yamauchi, M. 1991. The sequence flanking translational initiation site in protozoa. Nucleic Acids Res. 19: 2715-2770.

Yao, M.-C. and M. A. Gorovsky. 1974. Comparison of the DNA sequences of *Tetrahymena* macro- and micronuclei. Chromosoma 48: 1-18.

Yao, M.-C. 1986. Amplification of ribosomal RNA genes. In: J. G. Gall (ed.) Molecular Biology of ciliated protozoa. Academic Press, New York, New York. pp. 179-201.

Yao, M.-C. 1989. Site-specific chromosome breakage and DNA deletion in ciliates: In: D. E. Berg and M. M. Howe (eds.) Mobile DNA. American Society for Microbiologists, Washington, DC. pp. 715-734.

Yao, M.-C. and Yao, C.-H. 1991. Transformation of *Tetrahymena* to cycloheximide resistance with a ribosomal protein gene through sequence replacement. Proc. Natl. Acad. Sci. USA 88: 9493-9497.

Yoshida, M., H. Kawaguchi, Y. Sakata, K. Kominami, M. Hirano, H. Shima, R. Akada and I. Yamashita. 1990. Initiation of meiosis and sporulation in *Saccharomyces cerevisiae* reqires a novel protein kinase homologue. Mol. Gen. Genet. 221: 176-186.