UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Identification and characterization of a mating type-like locus in the "asexual" pathogenic yeast Candida albicans

Permalink
https://escholarship.org/uc/item/38q106tv

Author
Hull, Christina M.

Publication Date
2000

Peer reviewed|Thesis/dissertation

eScholarship.org
Powered by the California Digital Library
University of California

Identification and characterization of a mating type-like locus in the "asexual" pathogenic yeast Candida albicans

by

Christina M. Hull

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

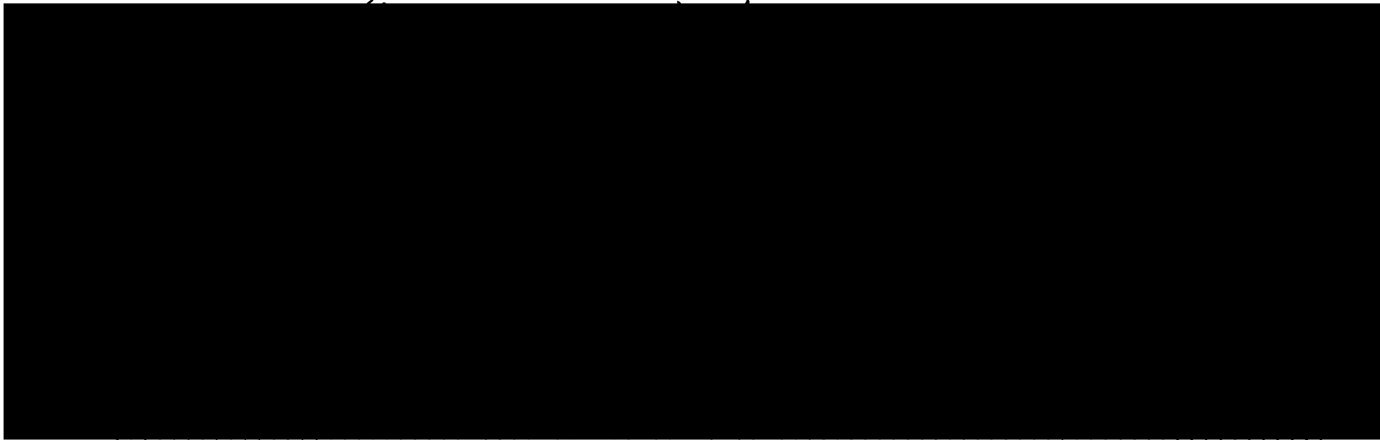
Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO



Date .... University Librarian

Degree Conferred: ....

Copyright 2000
by
Christina M. Hull

I wish to dedicate this thesis to my parents Laura Addie Howell Hull and James Albert Hull. They have inspired me to strive to do my best work, and they have encouraged me through many tears. They have celebrated the good in life with me and helped me to find my way during uncertain times. I have only begun to see the depths of their understanding, and I hope I can learn from their example and guide my children as they have guided theirs. To them I owe a great debt. Their unwavering support of me in my pursuit of happiness is unrivaled, and I am honored to call them my best family and my best friends.

Acknowledgments

Graduate school is about asking interesting questions and then doing whatever it takes to answer them, and I was fortunate to have a team of people behind me helping me take the right path to the answers. First and foremost, this thesis would not have been possible without the support and guidance of Sandy Johnson. Sandy encouraged me to pursue the questions that interested me most. He taught me how to take calculated risks, use the power of logic, and see projects through to completion. His commitment to quality science is unrivaled, and he taught me to do, see, and appreciate the best in science both in terms of experimental rigor and in the culture of science. I cannot thank him enough for this insight. I am also eternally grateful to Sandy for his support of me as an individual. He has supported my choices and helped me to realize my goal of being a scientist and a parent, and his flexibility with my work and education has made this a reality. Science would profit enormously if more principal investigators would see the light and follow his example. We all owe him a great debt.

I am also thankful for a thesis committee with the same commitment to good science. I would have packed my bags and left graduate school a long time ago were it not for the support of Don Ganem who is a super cheerleader with the best ideas and much enthusiasm for the work. His support and intellect were a great boon to my success. The enthusiasm of my thesis committee chair, Ira Herskowitz, also buoyed me through tough times. He has been a super influence with never-ending suggestions, ideas, and clarity that really helped me to see that I was headed in the right direction. I always enjoyed my thesis committee meetings because I knew they would be fun and interesting. This is a wonderful gift from an outstanding committee.

My work at the bench would not have been possible without the lab environment of the Johnson Lab. The collaborative spirit in the lab is excellent, and the individuals involved are exceptional. Thanks again to Sandy for creating a highly functional work environment and filling it with quality people. Burk Braun is a quality person I shared a bay with for over

UCSF LIBRARY

six years, and I now think of him as a member of my family. I thank him for viewing my data with a critical eye, for always making the time to help me, and for knowing the answer to the myriad questions I have asked. Thanks to Rebecca Smith for being such a great labmate, classmate, and friend. The quality of my graduate years was enhanced by her never-ending ability to obtain information, ask the right questions, and make the lab a quality place to work. We were a good team, and I hope that we can work together again someday.

Emily Troemel, Madhu Wahi, Danesh Moazed, and Martha Stark are all dear friends I thank for being a quartet of power and leading me by example down the road of science well done. Thanks to Hien Tran, Andrew Uhl, Ann Cassidy-Stone, David Kadosh, Martha Arnaud, and Nancy Hollingsworth for all the fun times learning about science. Thanks to the new crew Diane Inglis, Sarah Green, Joanne Penko, Annie Tsong, Matt Miller for continuing a tradition of quality and fun.

Special thanks to Ryan Raisner. His work on the *C. albicans* mating experiments was exceptional, and working with him has been a pleasure both professionally and personally. I have no doubt he will go on to do great things.

Thanks also to Sudarsi Desta for making my work in the lab possible with never -ending precision and cheer, and to Sue Adams who pulled me up when I was pushed down and looked after my every need.

To my husband Rob Brazas I can only say the two tiny words, "thank you," and hope that he can understand the depths of their meaning from me. This thesis would not have been possible without him. As a husband, friend, and scientist, he has been a source of support and inspiration that I am grateful for every day. I look forward to continuing to have a home full of love and science with him and our family for a lifetime to come.

Chapters 2 and 3 are reprinted with permission from *Science*.

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

SCHOOL OF MEDICINE

SAN FRANCISCO, CALIFORNIA 94143-0414

Department of Microbiology
and Immunology

UCSF LIBRARY

July 10 , 2000

Science
1200 New York Ave. NW
Washington, D.C. 20005
USA

To whom it may concern,

I would like permission to include in my dissertation a copy of the papers cited below, both of which were published in *Science*. The dissertation will be microfilmed by University Microfilms Incorporated, and they request permission to provide single copies on demand.

Christina M. Hull and Alexander D. Johnson. 1999. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*. *Science* 285:1271-1275.

Christina M. Hull, Ryan M. Raisner, and Alexander D. Johnson. 2000. Evidence for mating of the "asexual" yeast *Candida albicans* in a mammalian host. *Science*. in press.

If possible, please respond by FAX to (415) 476-8201.

Thank you for your consideration.

Sincerely,

Christina M. Hull

Christina M. Hull

Science ● THE GLOBAL WEEKLY OF RESEARCH

Ref.# Hull 00-1390

> *Be sure to quote the above reference number in all correspondence*

DATE: July 13, 2000

TO:

Christina Hull
University of California
Department of Microbiology & Immunology
San Francisco, CA 94143-0414

FROM: Karen Lentz, Permissions Assistant

RE: Your request for permission dated 07/10/00

Regarding your request (see attached copy), we are pleased to grant you non-exclusive, non-transferable permission, **but limited to print and microform formats only**, and provided that you meet the criteria checked below. Such permission is for **one-time** use and therefore does not include permission for future editions, revisions, additional printings, updates, ancillaries, customized forms, any electronic forms, braille editions, translations, or promotional pieces. We must be contacted for permission each time such use is planned. This permission does not apply to figures / artwork that are credited to non-AAAS sources. This permission does not include the right to modify AAAS material.

- ☒ Print the required copyright credit line on the first page that the material appears: "Reprinted (abstracted/excerpted) with permission from [FULL REFERENCE CITATION]. Copyright [YEAR] American Association for the Advancement of Science." Insert the appropriate information in place of the capitalized words.
- ☒ Permission is limited to the number of copies specified in your request or your first printing.
- ☐ Obtain the author's permission. *See original article for author's address.*
- ☐ Remit the following permission fee: $. This letter serves as your invoice. Please make your check payable to "The American Association for the Advancement of Science." A photocopy of this letter must be included with the check. Please remit to SCIENCE, P.O. Box 80144, Baltimore, MD 21280-0144, USA **Please note:** ***Permission is not valid unless payment is received within sixty (60) days of the date on this letter.***
- ☐ Ensure that no other material is attached to the AAAS article copy.
- ☒ AAAS **must** publish the full paper prior to use of any text.
- ☐ Your publication must not be comprised of more than 30% SCIENCE content.

Note: AAAS does not supply photos or artwork. Also, use of the AAAS material must not imply any endorsement by the American Association for the Advancement of Science. **This permission is not valid for the use of the AAAS and/or SCIENCE logos.**

Thank you for writing. If you have any questions please call me at (202) 326-6765 or write to me via FAX at 202 682-0816. For international calls, +1 is the country code for the United States.

Headquarters
1200 New York Avenue, NW, Washington, DC 20005 USA • Telephone: 202 326 7000 • Fax: 202 682 0816

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Identification and Characterization of a Mating Type-Like Locus in the "Asexual" Pathogenic Yeast *Candida albicans*

Christina M. Hull

Abstract

The pathogenic yeast *Candida albicans* can cause infections in humans that are particularly severe in people with compromised immune systems. The properties that allow its colonization have been difficult to study in part because *C. albicans* is a diploid organism with no previously identified sexual cycle. This asexual state in *C. albicans* is curious because closely related yeasts such as *Saccharomyces cerevisiae* have maintained sexual cycles. Additionally curious is the fact that many genes have been identified previously in *C. albicans* that are homologous to *S. cerevisiae* genes encoding components of the mating pheromone response and meiosis pathways. Given that many of these sexual cycle genes are conserved and intact in *C. albicans*, it seemed likely that the regulators of these genes (a1, α1, and α2) may also be conserved and functional. In this work I identified homologs of these master regulators in *C. albicans* that reside in a locus that is similar to the mating type (*MAT*) loci of other fungi and showed that the a1 and α2 proteins function as transcriptional repressors as they do in *S. cerevisiae*. This mating type-like (*MTL*) locus also contains genes never seen before in fungal *MAT* loci: two poly(A) polymerases, two oxysterol binding protein homologs, and two phosphatidylinositol kinase homologs. Subsequent to this discovery, I explored the role of the *MTL* locus in the life cycle of *C. albicans*. Test strains were constructed with deletions in the a1 or α1 and α2 genes to create functional "**a**" and " α" strains which were subsequently shown to mate in an *MTL*-dependent manner in a mouse tail vein model of infection. An analysis of potential haploid-

UCSF LIBRARY

specific target genes confirmed that a1 and $\alpha 2$ are important for the regulation of some sexual cycle homologs, further establishing the relationship between the *MTL* locus and the mating process. The identification of a locus in *C. albicans* that contains genes homologous to the master sexual cycle regulators in *S. cerevisiae* and the demonstration that *C. albicans* mating can occur in a mammalian host have substantially altered the view of *C. albicans* as an asexual organism.

UCSF LIBRARY

Table of Contents:

Chapter 1 - Introduction....................1

Chapter 2 - Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*....................15

Appendices to Chapter 2....................39

Appendix A - Additional evidence for *MTL* locus arrangement....................40

Appendix B - Additional information on **a**1-α2 activity in *C. albicans*....................53

Appendix C - Identification of additional open reading frames in the *C. albicans* *MTL* locus....................59

Appendix D - The *C. albicans MTL* locus is distinct from *MAT* loci in other fungi 63

Appendix E - Possible functions for the *C. albicans MTL* locus....................67

Chapter 3 - Evidence for mating of the "asexual" yeast *Candida albicans* in a mammalian host....................69

Appendix to Chapter 3....................94

Appendix F - Testing for mating in vitro....................95

Chapter 4 - Transcriptional repression of endogenous targets by **a**1-α2 in *Candida albicans*....................96

Chapter 5 - Discussion and Future Directions....................122

Miscellaneous Appendices....................137

Appendix G - Virulence assays for **a**1 deletion and α1+α2 deletion strains in mice....................138

Appendix H - Predicted Splice Junctions for *C. albicans* **a**1....................141

UCSF LIBRARY

List of Tables:

Table 4.1 - Summary of homologs of haploid-specific and **a**-specific genes in *C. albicans* **100**

Table 4.2 - Representatives of genes from computer analysis identified as having **a**1-α2 sites in their promoters....................**104**

Table 4.3 - Summary of results of chromatin IPs and Northern analysis....................**112**

UCSF LIBRARY

List of Figures:

Figure 1.1 - Regulation of the *C. albicans CAG1* gene in *S. cerevisiae*...............................**6**

Figure 1.2 - Functions of the master transcriptional regulators **a**1, α1, and α2 in *S. cerevisiae*...**9**

Figure 1.3 - Configuration of the mating-type (*MAT*) locus in *S. cerevisiae*......................**11**

Figure 2.1 - Features of the *S. cerevisiae* mating-type (*MAT*) locus and the *C. albicans* mating-type-like(*MTL*) locus...**19**

Figure 2.2 - Sequence comparisons between the *C. albicans* and *S. cerevisiae* **a**1, α1 and α2 proteins...**23**

Figure 2.3 - Products of PCR reactions showing deletion of the *C. albicans MTLa1* gene and the absence of genes for **a**1 at other loci...**25**

Figure 2.4 - Identification of an **a**1/α2 transcriptional repression activity in *C. albicans* mediated by the hsg operator and the *MTLa1* gene.......................................**29**

Figure A.1 - DNA matrix comparisons of DNA sequence of *C. albicans MTLa* and *MTLα*..**42**

Figure A.2 - Southern blot showing deletion of the *C. albicans MTLa1* gene...................**44**

Figure A.3 - Products of PCR reactions showing deletion of the *C. albicans MTLα1* gene and the absence of genes for α1 at other loci..**46**

Figure A.4 - Southern blot to show deletion of *C. albicans* α1 gene at *MTL* locus...........**48**

Figure A.5 - Products of PCR reactions showing deletion of the *C. albicans MTLα2* gene and the absence of genes for α2 at other loci..**50**

Figure A.6 - Southern blot to show deletion of *C. albicans* α2 gene at *MTL* locus...........**52**

Figure B.1 - Confirmation that **a**1-mediated repression affects mRNA levels....................**55**

Figure B.2 - **a**1/α2-mediated transcriptional repression in *C. albicans* is dependent on Tup1. ....**57**

Figure D.1 - Mating-type (*MAT*) loci from several fungi..**65**

UCSF LIBRARY

Figure 3.1 - Features of the *S. cerevisiae* mating-type (*MAT*) locus and the *C. albicans* mating type-like(*MTL*) locus..**74**

Figure 3.2 - *MTL* configurations of strains used in test mating experiment......................**78**

Figure 3.3 - FACS and DAPI analysis of prototrophic strains...**81**

Figure 3.4 - Southern analysis showing multiple copies of the *MTL* locus in conjugant strains...**83**

Figure 3.5 - Southern analysis showing multiple copies of the *ADE2* locus in conjugant strains...**86**

Figure 4.1 - Schematic representation of chromatin immunoprecipitation procedure.......**107**

Figure 4.2 - Schematic representation of expected results for an **a**1-α2 regulated gene in a chromatin immunoprecipitation experiment...**109**

Figure 4.3 - Results of PCR analysis of chromatin IPs...**111**

Figure 4.4 - Northern analysis of potential target genes. RNA from wild-type and **a**1 deletion cells was evaluated by Northern blot...**115**

Figure G.1 - Graph displaying a survival curve for mice infected with various strains of *C. albicans*...**140**

Figure H.1 - Predicted and actual splice junctions for *C. albicans* **a**1.............................**143**

UCSF LIBRARY

Chapter 1

Introduction

LSE LIBRARY

Candida albicans is a fungus that lives as a commensal in the bodies of healthy humans. As a member of the normal flora of the digestive tract and vagina, *C. albicans* has no apparent effect on its host under most conditions; however, *C. albicans* has the ability to overgrow its host and cause disease when conditions in the host are altered (*1*). The most clearly defined host condition that leads to *C. albicans* overgrowth is immunosuppression; people with AIDS, chemotherapy patients, and transplant patients are all prime candidates for *C. albicans* overgrowth. This overgrowth occurs primarily on mucosal surfaces in AIDS patients, and these individuals often suffer from severe oral, vaginal, and esophageal infections that are difficult to treat and often recur after treatment (*2*). Chemotherapy and transplant patients are most at risk for *C. albicans* in their bloodstreams. These systemic infections are also difficult to treat and are of particular concern because they often lead to patient mortality. The factors that lead to *C. albicans* overgrowth and infection in these patients are not well understood, and the range of available treatments is not highly effective in immunosuppressed hosts (*3*).

C. albicans infections are also prevalent among apparently healthy individuals. It has been estimated that seventy-five percent of women will experience a vaginal yeast infection in their lifetimes (*4*), and approximately 5% of infants experience oral thrush (*1*) which can be transmitted to their breastfeeding mothers (*5*). Many factors have been blamed for *C. albicans* overgrowth (e.g. hormonal fluctuations, antibiotic use, dietary changes); however, none has been proven to lead to an active infection, and the causes of the overgrowth in these cases are not clear (*6*).

One overriding difficulty in evaluating the factors that lead to *C. albicans* virulence has been the inaccessibility of the *C. albicans* life cycle to genetic study. *C. albicans* is classified as an imperfect (asexual) fungus based on the fact that no discernible haploid state or clear meiotic structures can be identified (*1*). It has been assigned to the catch-all genus for asexual yeasts, *Candida*, and has been characterized as having an anomalous life style ever since. Subsequent to this classification, *C. albicans* was shown to be a diploid

UCSF LIBRARY

organism (*7*) and experiments that identified a substantial number of heterozygous loci in *C. albicans* predicted that existing recessive lethal mutations would render any possible haploids inviable (*7*). This discovery cemented *C. albicans* in the genus of asexual fungi and appeared to eliminate any doubt about its proper classification as asexual. This assessment, however, was difficult to understand in terms of genetic fitness. It was difficult to envision how an organism without a sexual cycle could maintain healthy diversity and overall fitness in the face of deleterious mutations that could not be eliminated from the population through sexual recombination (*8*). In addition, three facts about *C. albicans* were inconsistent with the absence of a sexual cycle. First, population genetics studies suggested that in addition to clonal loci, there were loci in *C. albicans* that were clearly recombining (*9*), although the mechanism for this was completely unknown. Second, related yeasts in the budding yeast family such as *Saccharomyces cerevisiae* and *Kluyveromyces lactis* had retained clear sexual cycles, suggesting that the loss or invisibility of the *C. albicans* sexual cycle occurred relatively recently in evolutionary terms. Third, as improvements in molecular genetics advanced the quest for information about the *C. albicans* life cycle, researchers identified genes in *C. albicans* that were homologous to sexual cycle genes in other organisms including *S. cerevisiae* (e.g. *10*, *11*, *12*, *13*). (Subsequent to the work presented here, scores of these apparent sexual cycle genes have been identified, and in fact, most of the genes involved in mating and meiosis in *S. cerevisiae* have been conserved in *C. albicans*. (*14*))

These three pieces of information hinted that *C. albicans* was not necessarily the asexual organism it was thought to be, and the presence of the apparent mating and meiosis genes suggested that *C. albicans* may indeed possess a cryptic sexual cycle that had so far remained hidden from investigators. Alternatively, the low level recombination in the population studies could have been via some mitotic recombination mechanism, and the conserved sexual cycle gene products could have been co-opted for entirely different purposes. In either case, without mating and meiosis as tools to be used in traditional

UCSF LIBRARY

genetic manipulation, researchers have had a difficult time obtaining information about the features of *C. albicans* that lead to overgrowth and disease. It was clear that learning more about the molecular biology of *C. albicans* would lead to a better understanding of what altered or eliminated its sexual cycle, and hopefully, identify ways to prevent and treat *C. albicans* overgrowth more effectively.

One of the first and most intriguing sexual cycle homologs in *C. albicans* is *CAG1*, which was originally cloned from *C. albicans* using degenerate oligonucleotides in a low stringency polymerase chain reaction (PCR) in an attempt to clone G protein-coupled receptors (*10*). The predicted Cag1 protein is 47% similar to Gpa1, the Gα subunit of the trimeric G protein complex required for pheromone signaling in *S. cerevisiae. CAG1* encodes a protein that complements a *gpa1* deletion in *S. cerevisiae*, and under the control of its own promoter in *S. cerevisiae,* the *CAG1* gene is regulated just like *GPA1*(*10*) (Figure 1.1). That is, this *C. albicans* gene is expressed in *S. cerevisiae* haploid cells and repressed in *S. cerevisiae* **a**/α diploid cells. This conservation of regulation is significant because many genes in *S. cerevisiae* that are necessary for mating are expressed only in haploid cells, are repressed in **a**/α diploids, and are controlled in this manner by the master transcriptional regulators of the *S. cerevisiae* sexual cycle. *CAG1* appears to be regulated in *S. cerevisiae* even though no haploid state has been identified in *C. albicans*. It is unclear why an exclusively diploid organism would contain genes sensitive to differential regulation in haploids and diploids. The conservation of *CAG1* expression in *S. cerevisiae* prompted a search for regulatory sites in its promoter which revealed, in the absence of any other conservation in the promoter region, a binding site for a protein complex required for repression in *S. cerevisiae* **a**/α diploids (*10*). The presence of this gene with regulatory sites intact suggested strongly that homologs of the master regulators of the sexual cycle in *S. cerevisiae* would also be intact and functional in *C. albicans*.

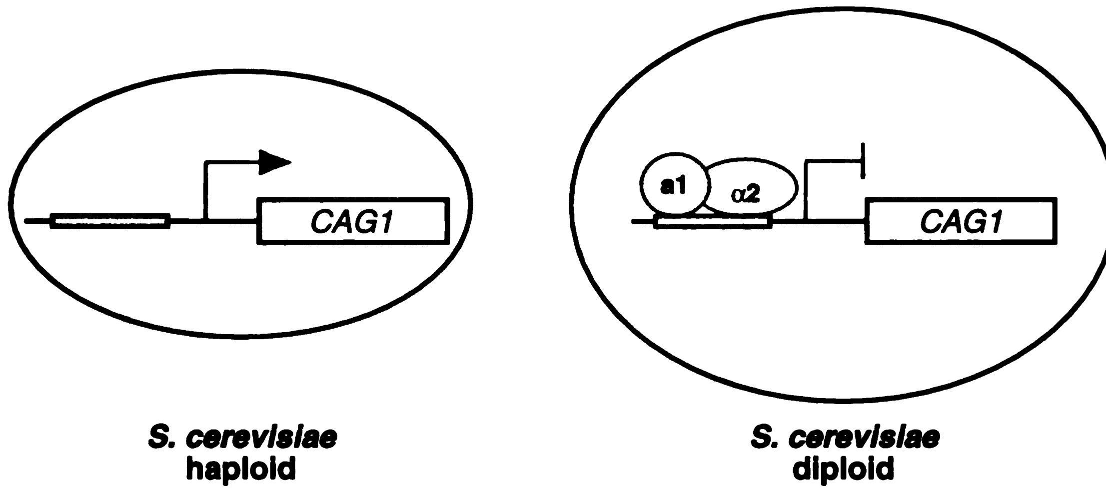
Extensive work on the master transcriptional regulators that control the *S. cerevisiae* sexual cycle has revealed that they are responsible for maintaining the three cell types of *S.*

UCSF LIBRARY

Figure 1.1. Regulation of the *C. albicans CAG1* gene in *S. cerevisiae*. *CAG1* is represented in both haploid and diploid *S. cerevisiae* cells. In the haploid cell, the *CAG1* gene is expressed as indicated by the green arrow. In diploid cells the *CAG1* gene is not expressed as represented by the red bar. This difference in expression in haploid and diploid *S. cerevisiae* cells is a result of the binding in diploid cells of the **a**1-α2 complex to a specific DNA sequence in the *CAG1* promoter. This complex represses a class of genes in *S. cerevisiae* called the haploid-specific genes that is expressed in haploids (in the absence of **a**1-α2) and repressed in diploids (in the presence of **a**1-α2).

UCSF LIBRARY

Figure 1.1



UCSF LIBRARY

cerevisiae (*15*, *16*). Two of the three cell types are the haploid types **a** and α. In the **a** cell, the **a**1 protein is expressed and has no known function in these cells. In the α cell type, the proteins α1 and α2 are expressed and are responsible in conjunction with other proteins for activating the expression of α-specific genes and repressing the expression of **a**-specific genes respectively. This pattern of regulation establishes the α cell type and makes it distinct from the **a** cell type. This expression pattern also makes these cell types competent to mate and form the third cell type in *S. cerevisiae*, the **a**/α cell. The primary signal that mating has occurred is the coming together of the **a**1 and α2 repressor proteins in the same cell to form a heterodimer that binds to DNA and represses the transcription of haploid-specific genes (Figure 1.2). This pattern of repression prevents the **a**/α cell from mating and makes it competent to undergo the processes of meiosis and sporulation.

Part of what allows such efficient regulation of the *S. cerevisiae* sexual cycle is the way in which the master regulators are organized in the genome. The genes are located in a special location termed the mating-type (*MAT*) locus (*15*). This locus is unique because in a diploid cell the **a** and α information is located in the same position in homologous chromosomes creating a region of non homologous DNA sequence between the two chromosomes. This region is created because of the different contribution of information from the two haploid cell types: **a** cells contribute the **a**1 gene in their *MAT* locus, while α cells contribute the α1 and α2 genes in their *MAT* locus. When they are both present in the **a**/α cell, the two regions contain different information on the two otherwise homologous chromosomes (Figure 1.3). This arrangement is common in diverse fungi ranging from budding yeast to mushrooms and represents the sex chromosome equivalent of more complex organisms.

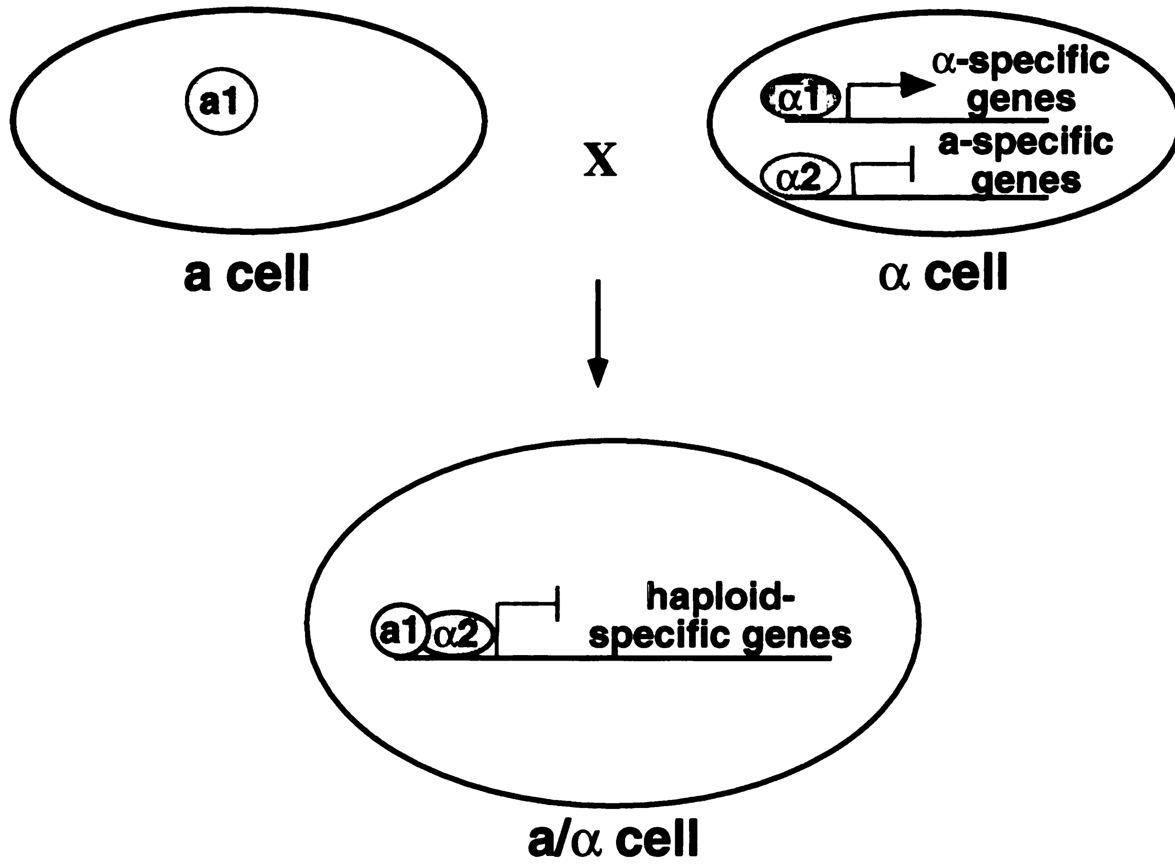
Given the presence of genes in *C. albicans* similar to the sexual cycle genes in *S. cerevisiae,* it was logical to think that the master regulators of those genes may also be present in *C. albicans.* I wanted to clone these regulators from *C. albicans* and understand what roles they could be playing in an organism with no sexual cycle. I began by cloning **a**1

UCSF LIBRARY

Figure 1.2. Functions of the master transcriptional regulators **a**1, α1, and α2 in *S. cerevisiae.* **a**1 is shown in an **a** cell with no function. α1 and α2 are shown in an α cell activating α-specific genes (green arrow) and repressing (red bar) **a**-specific genes respectively. In the **a**/α diploid cell, the proteins **a**1 and α2 are shown forming a heterodimer that binds to DNA and represses (red bar) a class of genes called the haploid-specific genes.

UCSF LIBRARY

Figure 1.2

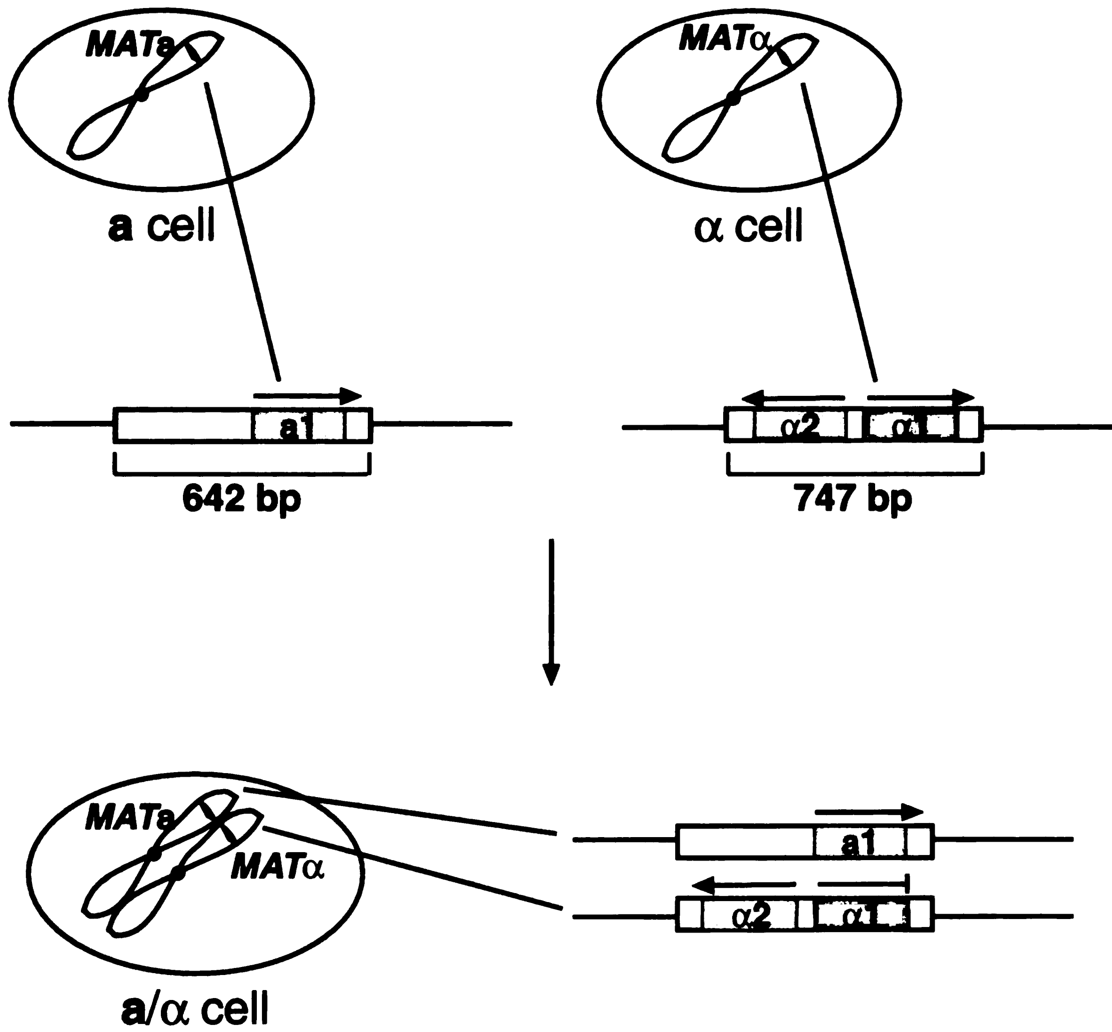


UCSF LIBRARY

Figure 1.3 Configuration of the mating-type (*MAT*) locus in *S. cerevisiae.* In *S. cerevisiae* the *MAT* locus encodes the master transcriptional regulators **a**1, α1, and α2. In **a** cells **a**1 is encoded in the *MAT* locus, and in α cells α1 and α2 are encoded in the *MAT* locus. When the **a** and α cell types mate to form the **a**/α cell, the *MAT* locus information is located on homologous chromosomes, but the information on the two chromosomes is different in the region of the *MAT* locus. The region of heterologous sequence between the chromosomes is 642 base pairs (bp) on the *MATa* segment and 747 bp on the *MAT*α segment. Black arrows indicate the direction of transciption for each gene. **a**1 is repressed by **a**1-α2 in **a**/α cells and is not expressed.

UCSF LIBRARY

Figure 1.3



UCSF LIBRARY

from a *C. albicans* genomic library using a sequence trace from the Stanford *C. albicans* sequencing project (*17*). In the process, I uncovered an α1 homolog, and as I walked upstream and downstream of **a**1 and α1, I discovered a nearly 9 kilobase region on two chromosomes in *C. albicans* that resembles the mating loci of other fungi. This mating type-like (*MTL*) locus contains homologs of **a**1, α1, and α2 that appear to be arranged on two homologous chromosomes in a manner similar to their arrangement in *S. cerevisiae*. The predicted proteins are conserved along with the relative positions of the genes in the locus, their directions of transcription, and the repression function of **a**1-α2. These similarities make the *MTL* appear to be similar to a typical *MAT* locus; however, unlike other fungi, the *MTL* contains open reading frames for six additional proteins: two poly(A) polymerases, two oxysterol binding protein homologs, and two phosphatidyl inositol kinase homologs. The roles of these additional *MTL* components are unknown, but the presence of the locus with homologs of sexual cycle regulators intact suggests that *C. albicans* may possess the ability to undergo mating and meiosis. I tested the ability of *C. albicans* to mate using specialized *MTL* mutants that were predicted to behave like the **a** and α cells of *S. cerevisiae*. Amazingly, these cells were found to undergo mating when subjected to the conditions found in a mouse host. These experiments confirm that *C. albicans* possesses the inherent ability to mate under the proper environmental conditions and suggests that it may in fact participate in a complete sexual cycle that is dependent on components of the *MTL*. Additional experiments to look at the regulation of potential target genes of **a**1-α2 have identified endogenous genes that are regulated directly by **a**1 and confirmed the conservation of haploid-specific gene regulation and targets. The conservation of **a**1-α2 regulation further connects components of the *MTL* locus to the regulation of a possible sexual cycle. Taken together, the results presented here abolish the old idea of *C. albicans* as an asexual organism and point to a new model in which *C. albicans* can engage in a complete sexual cycle.

UCSF LIBRARY

References and Notes

1. F. C. Odds, *Candida and Candidosis* (Bailliere Tindall, London, ed. 2, 1988).
2. T. C. White, *ASM News* **63**, 427-433 (1997).
3. M. G. Shepherd, R. T. Poulter, P. A. Sullivan, *Annu Rev Microbiol* **39**, 579-614 (1985).
4. J. D. Sobel, *Am Journal Obstet Gynecol.* **1523**, 924-935 (1985).
5. M. J. Heinig, J. Francis, D. Pappagianis, *Journal of Human Lactation* **15**, 281-288 (1999).
6. F. C. Odds, *ASM News* **60**, 313-318 (1994).
7. W. L. Whelan, P. T. Magee, *Journal of Bacteriology* **145**, 896-903 (1981).
8. C. Zeyl and G. Bell, *Nature* **388**, 465-468 (1997).
9. Y. Graser, et al., *Proc Natl Acad Sci U S A* **93**, 12473-7 (1996).
10. C. Sadhu, D. Hoekstra, M. J. McEachern, S. I. Reed, J. B. Hicks, *Molecular and Cellular Biology* **12**, 1977-1985 (1992).
11. A. C. Diener, G. R. Fink, *Genetics* **143**, 769-776 (1996).
12. E. Leberer, et al., *Proc Natl Acad Sci U S A* **93**, 13217-22 (1996).
13. M. Raymond, et al., *Molecular Microbiology* **27**, 587-98 (1998).
14. C. Hull, unpublished observation made possible by the sequence available from the Stanford *Candida albicans* Sequencing project and the Johnson Laboratory *Candida albicans* Genes Database created by B. Braun and curated by B. Braun and A. Uhl.
15. I. Herskowitz, J. Rine, and J. Strathern, in *Molecular and Cellular Biology of the Yeast Saccharomyces* E. W. Jones, J. R. Pringle, J. R. Broach, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992), vol. 2: Gene Expression, pp. 583-656.
16. A. D. Johnson, *Curr Opin Genet Dev* **5**, 552-8 (1995).

UCSF LIBRARY

17. The starting 200 base trace was obtained from the University of Minnesota *Candida albicans* information page at http://alces.med.umn.edu/Candida.html courtesy of the Stanford DNA Sequencing and Technology Center.

UCSF LIBRARY

Chapter 2

Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*

UCSF LIBRARY

This version of Chapter 2 is reprinted here with minor changes and permission from Christina M. Hull and Alexander D. Johnson. 1999. *Science* 285:1271-1275. Copyright 1999 American Association for the Advancement of Science.

Abstract

Candida albicans, the most prevalent fungal pathogen in humans, is thought to lack a sexual cycle. A set of *C. albicans* genes has been identified that corresponds to the master sexual cycle regulators **a**1, α1, and α2 of the *Saccharomyces cerevisiae* mating-type (*MAT*) locus. The *C. albicans* genes are arranged in a way that suggests that these genes are part of a mating-type-like locus that is similar to the mating-type loci of other fungi. In addition to the transcriptional regulators **a**1, α1, and α2, the *C. albicans* mating-type-like locus contains several genes not seen in other fungal *MAT* loci including those encoding proteins similar to poly(A) polymerases, oxysterol binding proteins, and phosphatidylinositol kinases.

UCSF LIBRARY

Text

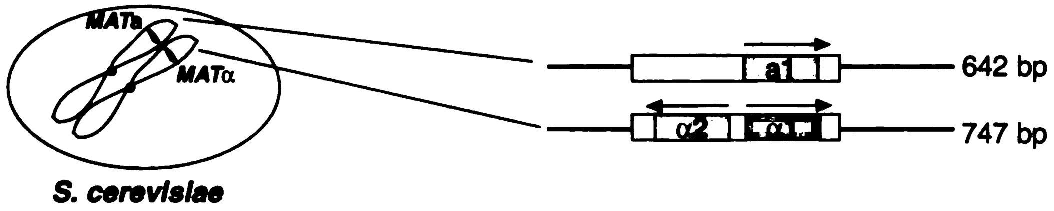
The yeast *Candida albicans* is the most common human fungal pathogen causing most cases of oral and vaginal thrush as well as severe mucosal and systemic infections in immunocompromised individuals (1). A principle difficulty in studying *C. albicans*, compared to other yeasts, is that *C. albicans* has no known sexual cycle and is therefore not amenable to conventional genetic analysis. It is a diploid organism for which no haploid state has been observed, nor has any process resembling meiosis or spore formation been detected. Sexual reproduction in fungi is typically controlled by genes that reside in a genetic locus called a mating-type, or *MAT*, locus. In the budding yeast *Saccharomyces cerevisiae* the genes residing at the *MAT* locus can be either the **a** type or the α type. These genes, which code for transcriptional regulators, specify the three cell types involved in the *S. cerevisiae* sexual cycle (2, 3). A cell containing only the *MATa* locus is an **a** cell; a cell containing only the *MATα* locus is an α cell, and a cell that contains both *MATa* and *MATα* (typically diploid and therefore heterozygous at the *MAT* locus) is an **a**/α cell. *MATa* codes for the homeodomain protein **a**1, which has no known function in **a** cells. *MATα* codes for a homeodomain protein (α2) and an α-domain protein (α1) which cause the repression of **a**-specific genes and the activation of α-specific genes, respectively (Fig. 2.1 panel A). Because the **a**- and α-specific genes encode proteins required for each of the cell types to mate, these changes in gene expression differentiate the **a** cell from the α cell. The product of a successful mating is an **a**/α diploid cell in which **a**1 and α2 are both expressed. The **a**1 and α2 proteins are the key regulators of the **a**/α cell type, and together they bind to a specific DNA sequence to repress the transcription of many target genes (including α1). This **a**1/α2 regulatory activity shuts off the ability of cells to mate and at the same time permits meiosis and sporulation in the presence of the appropriate external nutritional signals.

Figure 2.1. Features of the *S. cerevisiae* mating-type (*MAT*) locus and the *C. albicans* mating-type-like(*MTL*) locus. **A**. The *S. cerevisiae MAT* locus contains open reading frames for three gene regulatory proteins, **a**1, α1, and α2, which are located on homologous chromosomes. The region of heterologous DNA sequence between the two chromosomes is 642 base pairs (bp) for the **a** chromosome and 747 bp for the α chromosome. **B.** The *C. albicans MTL* locus contains open reading frames for nine proteins from four families of proteins: 3 gene regulatory proteins, 2 phosphatidylinositol kinases, 2 oxysterol-binding protein-like proteins, and 2 poly(A) polymerases. The region of DNA sequence that differs between the *MTLa* and *MTLα* segments is 8742 bp for *MTLa* and 8861 bp for *MTLα*. Black arrows indicate the direction of each gene 5' to 3'.

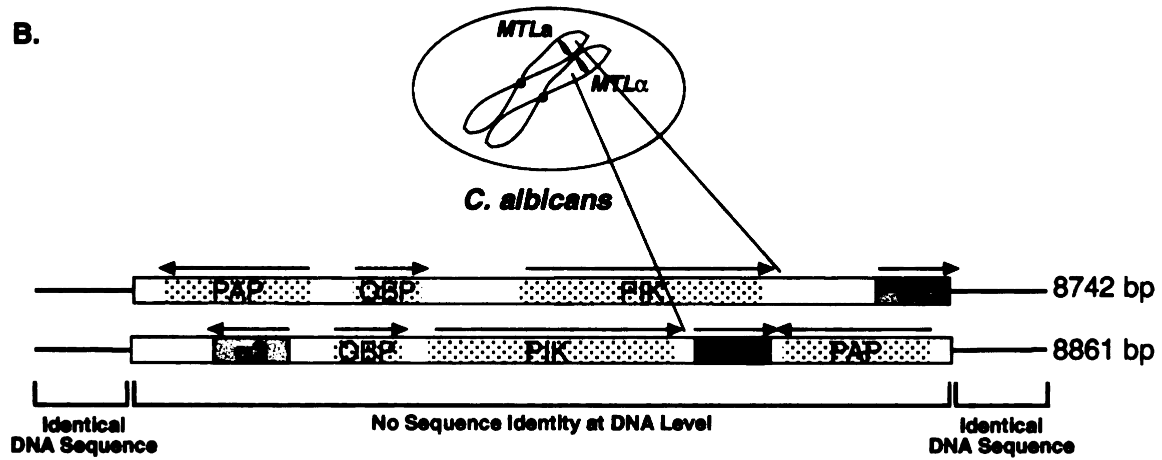
UCSF LIBRARY

Figure 2.1

A.



B.



UCSF LIBRARY

Represented in Figure 2.1 panel B are two genomic fragments from *C. albicans* that contain clusters of genes that bear a striking resemblance to the *MATa* and *MATα* genes of *S. cerevisiae* (denoted mating type-like or *MTLa* and *MTLα* respectively). The *MTLa* gene cluster was obtained by chromosome walking using a lambda library of *C. albicans* genomic fragments. The beginning probe for the walk was based on a sequence trace from the Stanford *C. albicans* Sequencing Project that resembled a portion of the *S. cerevisiae* *MATa1* gene. *MTLα* was obtained by walking downstream of *MTLa* to its flanking DNA sequence and then back into and through *MTLα* (4, 5). *MTLa* contains four open reading frames which encode a gene regulatory protein, a poly(A) polymerase, an oxysterol binding protein-like protein, and a phosphatidylinositol kinase. *MTLα* contains four genes whose products are closely related to those in *MTLa* plus an additional open reading frame coding for another gene regulatory protein. The DNA sequences within *MTLa* and *MTLα* are approximately 48% identical overall; however, the DNA sequences flanking them are greater than 99% identical. We have defined the borders of *MTLa* and *MTLα* as the points at which their DNA sequences become greater than 99% identical.

Although the clusters of genes in the *MTL* locus are much larger than those of the *S. cerevisiae MAT* locus (9 kb vs. 0.7 kb), three features of the *MTL* locus in *C. albicans* are remarkably similar to those of the *MAT* locus in other fungi, especially *S. cerevisiae* (Fig. 2.1). First, three of the proteins coded for by the *C. albicans* locus have predicted amino acid sequences very similar to those of the transcriptional regulators **a**1, α1, and α2 encoded by the *S. cerevisiae MAT* locus. The *C. albicans MTLa* segment codes for a homeodomain protein similar in sequence to the *S. cerevisiae* **a**1 protein (30% identity and 56% similarity over the entire protein and 43% identity and 59% similarity in the homeodomain region). The *C. albicans MTLα* segment codes for an α-domain protein similar to the *S. cerevisiae* α1 protein and for a homeodomain protein similar to the *S. cerevisiae* α2 protein. The predicted *C. albicans* α1 protein is 26% identical (49% similar) to the *S. cerevisiae* α1, and the *C. albicans* α2 is 28% identical (58% similar) to the *S. cerevisiae* α2 protein with

UCSF LIBRARY

particularly strong similarity seen in the homeodomain region (44% identity and 69% similarity) (Fig. 2.2). A second similarity between the *C. albicans MTL* and *S. cerevisiae MAT* loci concerns the overall organization of these three genes. As is true for the *S. cerevisiae* genes, the *C. albicans MTLα1* and *MTLα2* genes are transcribed divergently from one chromosome and the *MTLa1* gene is found on the other chromosome (Fig. 2.1). This feature is particularly striking because gene order and organization are not generally conserved between *C. albicans* and *S. cerevisiae* (6, 7).

A third similarity between *C. albicans* and *S. cerevisiae* is the conserved positions of the introns in the *MATa1* and *MTLa1* genes. In *S. cerevisiae*, the *MATa1* gene is one of only a few genes interrupted by two introns (8). The *C. albicans MTLa1* gene also appears to be interrupted by two introns, one of which is in exactly the same position (in the "recognition" helix of the homeodomain) as that in the *S. cerevisiae MATa1* (Fig. 2.2). The locations of the introns in the *C. albicans MTLa1* gene were initially predicted from the DNA sequence and were verified by observing the sizes of RT-PCR products (9). In *S. cerevisiae*, the *MATα2* gene is free of introns, but the *C. albicans MTLα2* gene contains a single intron located in the same position within the recognition helix of the homeodomain as the C-terminal introns in the *C. albicans MTLa1* and *S. cerevisiae MATa1* genes.

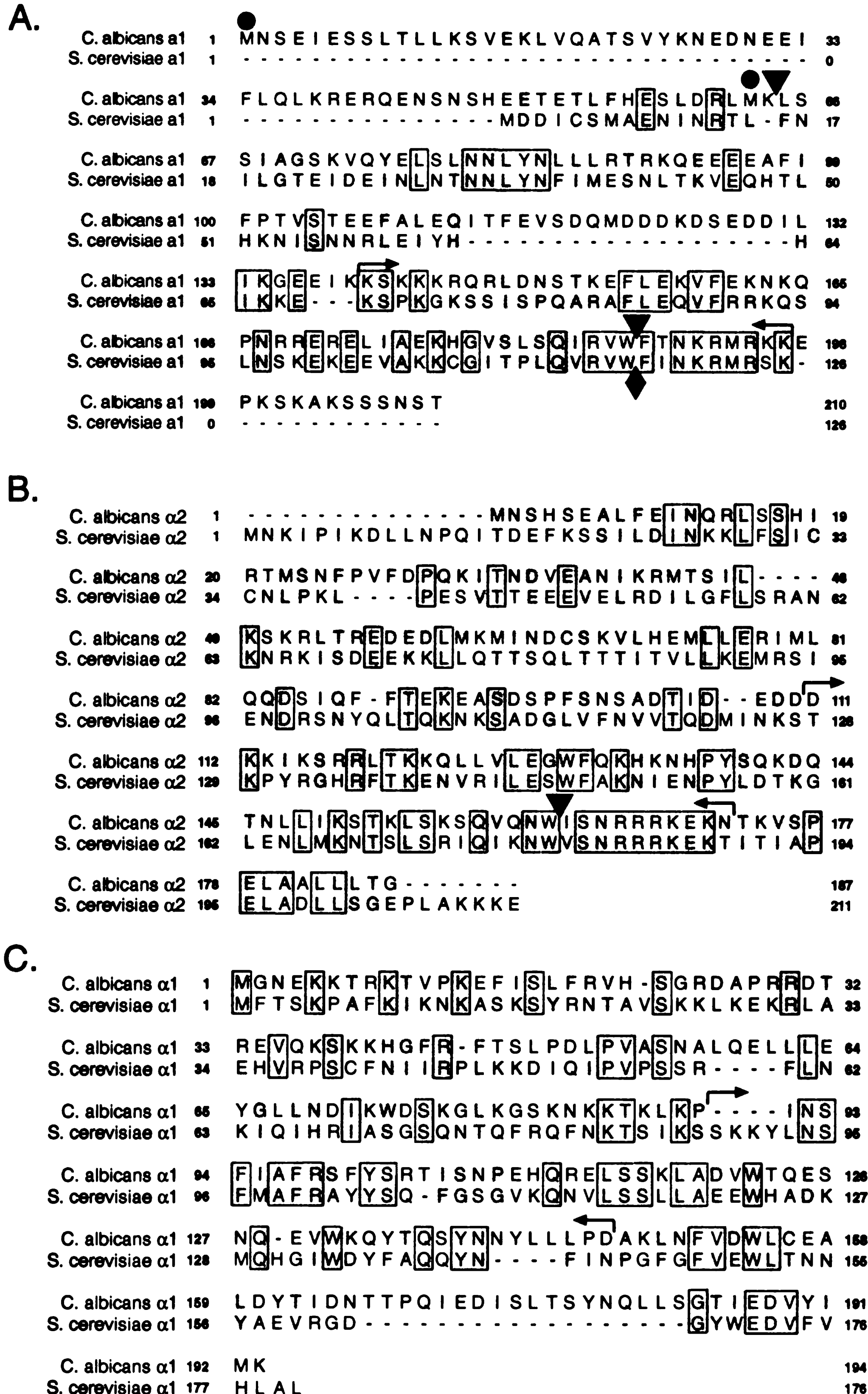
C. albicans is diploid, and several lines of evidence support the idea that *MTLa* and *MTLα* reside at the same position on homologous chromosomes. Two types of *C. albicans MTLa1* deletion mutants were constructed (a complete deletion of the *MTLa1* open reading frame and a deletion of only the homeodomain) by homologous replacement by disrupted genes (10). After a single round of transformation (11), the resultant strains were tested by PCR (12) and Southern analysis (9). The mutant forms of the *MTLa1* gene were readily detected (Fig. 2.3), but the wild-type gene was absent. This result indicates that the *MTLa1* gene is present in only a single (haploid) copy in the *C. albicans* genome, and contrasts with the case for many other genes in *C. albicans*, where two rounds of disruption have been necessary (one for each copy) to destroy a gene. The same approach was utilized to

UCSF LIBRARY

Figure 2.2. Sequence comparisons between the *C. albicans* and *S. cerevisiae* **a**1, α1 and α2 proteins. Alignment performed using the GCG8 program "pileup" (Genetics Computer Group, 575 Science Drive, Madison, WI 53711) and displayed using the program SeqVu 1.1 GES algorithm (Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Sydney, Australia). Boxes indicate sequence identity and shading indicates similar and identical resides. **A.** Protein alignment based on predicted amino sequence for *C. albicans* **a**1. Dots indicate possible start codons for the proteins, and inverted triangles show the positions of introns in the *C. albicans* gene. The first intron is 74 bp in length (*S. cerevisiae* first intron is 51 bp), and the second intron is 256 bp (*S. cerevisiae* second intron is 53 bp). The diamond shows the position of intron 2 in the *S. cerevisiae MATa1* gene. Arrows at *C. albicans* amino acids #140 and #197 delineate the homeodomain. **B.** Protein alignment based on predicted amino acid sequence for *C. albicans* α2. Inverted triangle indicates position of the intron in the *C. albicans* (59 bp) gene, and arrows at *C. albicans* amino acid positions #111 and #172 delineate the homeodomain. **C.** Protein alignment based on predicted amino acid sequence for the *C. albicans* α1 protein. Several α1-like proteins have been identified in ascomycete fungi, and a conserved region designated the "α" domain (29) has emerged. Arrows at *C. albicans* amino acid positions #90 and #146 delineate the α domain.

UCSF LIBRARY

Figure 2.2

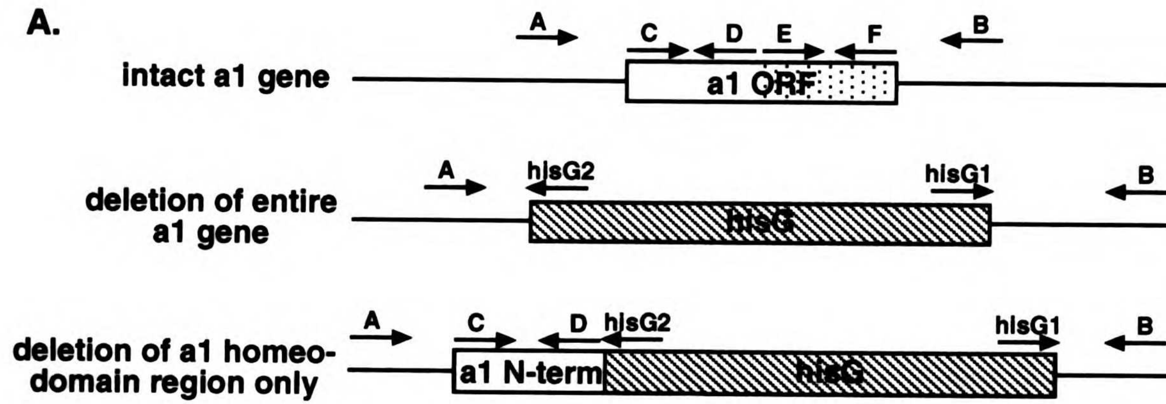


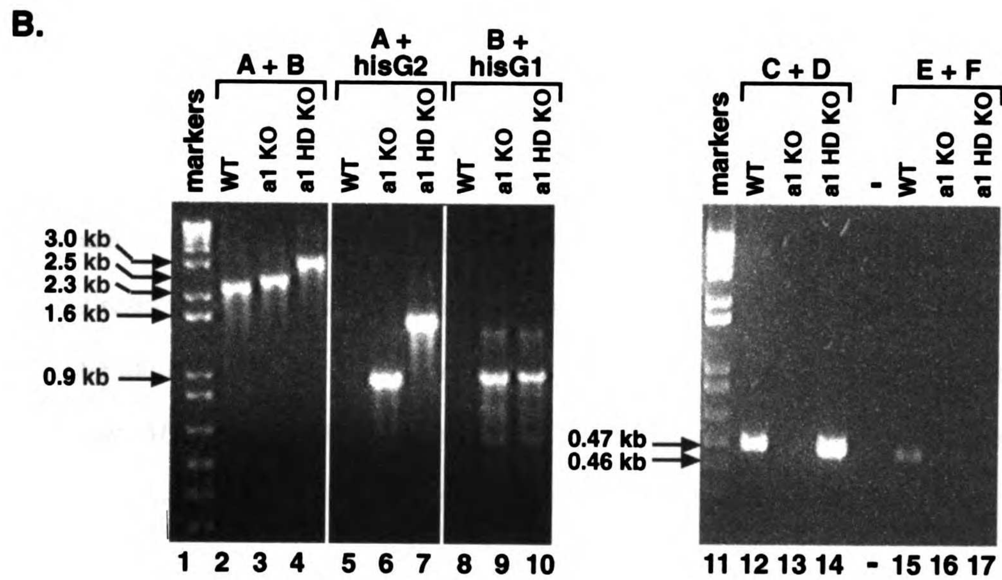
UCSF LIBRARY

Figure 2.3. Products of PCR reactions showing deletion of the *C. albicans MTLa1* gene and the absence of genes for **a**1 at other loci. **A.** Schematic of genomic locus. Labeled arrows indicate locations of primers used in the PCR reactions in panel B. **B.** Agarose gels showing fragments produced in PCR reactions using indicated primer sets on genomic DNA isolated from different strains. Lanes labeled WT contain genomic DNA from wild-type cells as the template; lanes labeled **a**1 KO contain genomic DNA from the complete *MTLa1* deletion strain as template; lanes labeled **a**1 HD KO contain genomic DNA from the *MTLa1* homeodomain deletion strain as template. Predicted sizes of the PCR products are indicated in kilobases (kb) to the left of the panels. All primer sets yielded the appropriately sized products or no product as predicted. Lane 1: DNA marker fragments. Lanes 2-4: Primers A and B on indicated genomic DNA. Lanes 5-7: Primers A and hisG2 on indicated genomic DNA. Lanes 8-10: Primers B and hisG1 on indicated genomic DNA. Lane 11: DNA marker fragments. Lanes 12-14: Primers C and D on indicated genomic DNA. Lanes 15-17: Primers E and F on indicated genomic DNA.

UCSF LIBRARY

Figure 2.3





UCSF LIBRARY

make a complete disruption of the *C. albicans MTLα2* gene. After a single round of transformation, the disrupted allele, but not the naturally occurring *MTLα2* gene, was detected by PCR (9). From these results, we conclude that both *MTLa1* and *MTLα2* exist in only a single copy in the *C. albicans* genome. These results also indicate the absence of silent mating-type "cassettes" in the *C. albicans* genome since PCR primers to the *MTLa1* (Fig. 2.3) and *MTLα2* (9) open reading frames do not detect these genes in the disrupted strains. These data provide strong support for the idea that the *MTL* locus of *C. albicans* is heterozygous, whereas the majority of the *C. albicans* genome is homozygous. Since the two *MTL* loci are each imbedded in nearly identical DNA sequences, the simplest interpretation is that the *MTL* loci reside on homologous chromosomes as they do in *S. cerevisiae*.

One hypothesis for the absence of a sexual cycle in *C. albicans* is that *C. albicans* was originally an **a**/α cell (to use the *S. cerevisiae* nomenclature) and through recombination lost one allele of the mating locus becoming an **a**/**a** or α/α cell unable to undergo meiosis and return to a haploid state (13). The discovery of two *MTL* gene segments in *C. albicans*, one that resembles *MATa* and one that resembles *MATα* appears to rule out this simple idea. The identification of *MTLa* and *MTLα* also suggests that the **a**/α configuration of the *C. albicans MAT*-like locus is a stable one. Since homologous recombination via the sequences flanking the locus could in principle easily result in the loss of either *MTLa* or *MTLα*, it seems likely that some sort of recombinational suppression exists. For example, one or more of the genes in each locus could be essential for cell viability, or the region could be under some type of mechanistic recombinational suppression.

In *S. cerevisiae*, the **a**1 and α2 proteins form a heterodimer which binds to specific DNA sequences (the haploid-specific gene, hsg, operators) and represses transcription of the haploid-specific genes (2, 14, 15). To see whether *C. albicans* has an **a**1/α2 repression activity and whether it is dependent on the genes of the *MTL* locus, we used two different hsg operators: one is the consensus hsg operator from *S. cerevisiae*, and the other is an hsg-

UCSF LIBRARY

like sequence found upstream of the *C. albicans CAG1* gene, which encodes the α subunit of a trimeric G-protein. Although the function of this hsg-like sequence in *C. albicans* is not known, it is known to be recognized by *S. cerevisiae* **a**1/α2 (16). Five *GFP* reporter constructs were analyzed in *C. albicans* cells (17). In each construct, the *GFP* gene was placed under transcriptional control of the *C. albicans ADH1* promoter, and in addition contained: (i) no insert, (ii) three *S. cerevisiae* hsg operator consensus sequences, (iii) three *S. cerevisiae* hsg operators mutated in two nucleotide positions to prevent recognition by **a**1/α2 (18), (iv) three copies of the hsg operator-like sequences found upstream of the *C. albicans CAG1* gene, or (v) three copies of the *C. albicans CAG1* sequence mutated in a way predicted to destroy recognition by **a**1/α2. The operators were inserted in the *ADH1* upstream region 260 base pairs upstream of the *GFP* gene.

C. albicans cells containing the construct that lacks the operator expressed *GFP,* as evidenced by their bright-green fluorescence (Fig. 2.4). The introduction of the *S. cerevisiae* hsg operators into the promoter markedly decreased fluorescence, whereas the mutant *S. cerevisiae* hsg operators showed no significant difference from the control construct that lacks the operators. The presence of the hsg-like sequence from *CAG1* also significantly repressed *GFP* production, and the point-mutations introduced into it relieved this repression.

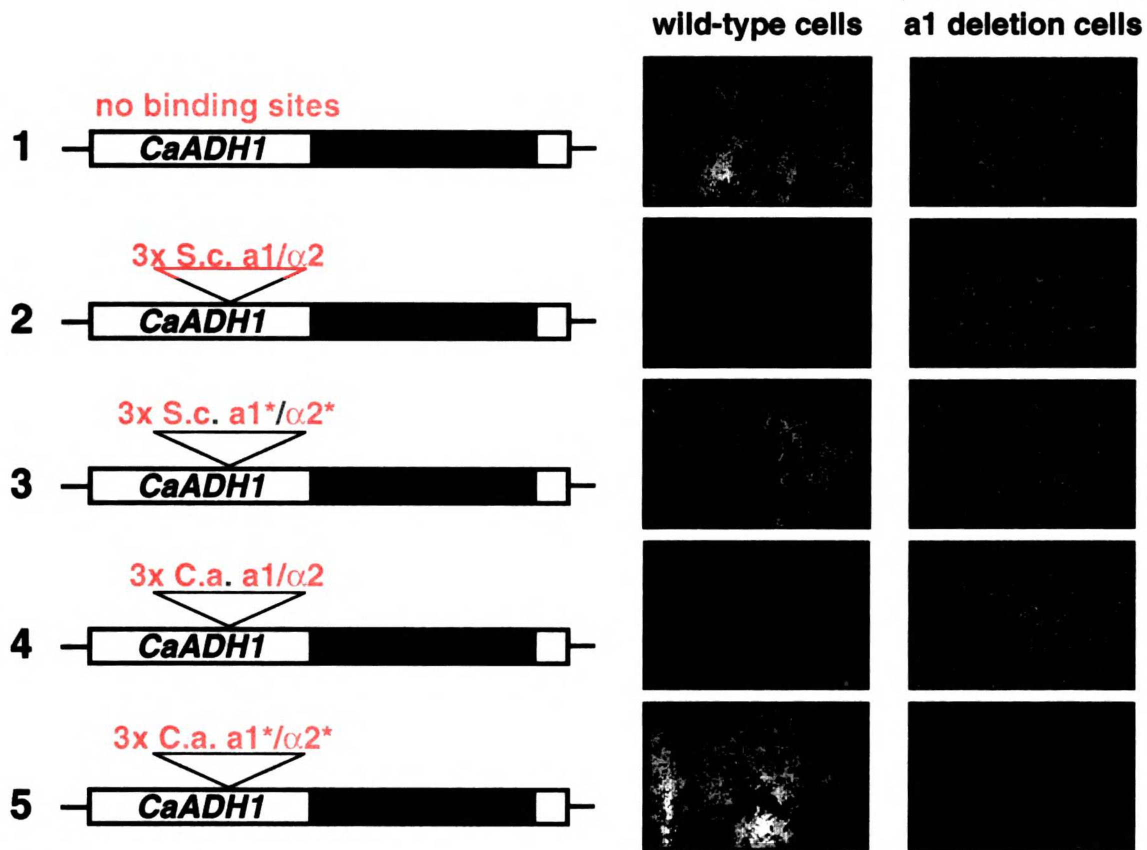
To determine whether the *C. albicans MTL* gene cluster was required for the **a**1/α2-like repression activity, the *GFP* reporters were transformed into *MTLa1* deletion strains and evaluated for fluorescence. In contrast to the wild-type *C. albicans* strains, the *MTLa1* mutant strains showed the same levels of fluorescence for all of the reporter constructs, indicating that the *MTLa1* gene is required for the transcriptional repression activity (Fig. 2.4). Similar behavior was seen for both for the complete deletion of the *MTLa1* gene and for the *MTLa1* homeodomain deletion, consistent with the DNA-binding domain of **a**1 being required for the repression activity (9). Northern analysis also showed that transcription from the reporter constructs containing the functional hsg operators was

UCSF LIBRARY

Figure 2.4. Identification of an **a**1/α2 transcriptional repression activity in *C. albicans* mediated by the hsg operator and the *MTLa1* gene. Transcriptional repression activity determined as a measure of *GFP* fluorescence in a heterologous reporter system in the presence of test DNA binding sites as inserts. In each case the reporter consists of the *C. albicans ADH1* promoter driving *GFP* expression. Reporter 1) no insert, 2) three *S. cerevisiae* hsg operator consensus sequences (S.c. **a**1/α2), 3) three *S. cerevisiae* hsg operators mutated at two nucleotide positions to prevent recognition by **a**1/α2 (S.c. **a**1*/α2*) (18), 4) three copies of the hsg operator-like sequences located upstream of the *C. albicans CAG1* gene (C.a. **a**1/α2), and 5) three copies of the *C. albicans CAG1* sequence mutated in a way predicted to destroy recognition by **a**1/α2 (C.a. **a**1*/α2*). Repression of the reporter is observed only when the reporter contains an intact hsg operator and when the *MTLa1* gene is present.

UCSF LIBRARY

Figure 2.4



UCSF LIBRARY

derepressed in the *MTLa1* deletion mutants compared to the wild-type strain; however, in the absence of **a**1, the functional hsg operators still showed a slight amount of repression when compared to the mutated hsg operators (9). This repression could be due to the effect of the α2 alone, or to a different *C. albicans* activity that has some overlapping function with **a**1/α2. Taken together, these results show that *C. albicans* has an **a**1/α2-like transcriptional repression activity and that this activity is dependent on the *C. albicans MTLa1* gene. We think it likely that *MTLa* and *MTLα* together encode the **a**1/α2 activity. In *S. cerevisiae,* **a**1 and α2 mediate transcriptional repression by bringing the co-repressor Tup1 to DNA, and we know that the **a**1/α2 repressor activity observed in *C. albicans* is at least partially dependent on the *C. albicans* Tup1 protein (9).

In *S. cerevisiae*, genes regulated by the products of the *MAT* locus encode proteins involved in many aspects of the sexual cycle. *C. albicans* contains close relatives of many of these genes including those involved in *S. cerevisiae* mating (e.g. *GPA1* (16), *STE20* (19), *STE6* (20)) and meiosis (e.g. *DMC1* (13)). Although some of the *C. albicans* relatives of the *S. cerevisiae* sexual cycle genes control filamentous growth, the complete functions of most of these genes have not been determined. These observations raise the question of whether any of these *C. albicans* genes are regulated by the *MTL* locus. We have shown that a DNA sequence found upstream of the *C. albicans CAG1* gene (*GPA1* homolog) can bring about *MTL*-dependent transcriptional repression when placed into a test promoter, and we also know that expression of the endogenous *CAG1* gene is indeed regulated by the *MTL* locus (9). Thus, like the analogous situation in *S. cerevisiae*, *C. albicans CAG1* is a natural target of the products of the *MTL* locus. These results have prompted us to reexamine the possibility of a sexual cycle in *C. albicans.* In preliminary experiments, mating in *C. albicans* has not been observed; however, a more extensive analysis is now underway to construct strains with the appropriate configurations of the *MTL* locus and to screen them under numerous environmental conditions, including those conducive to mating in a variety of fungi.

UCSF LIBRARY

In addition to the three transcriptional regulators, **a**1, α1, and α2, six other open reading frames (ORFs) were identified in the *MTL* locus, three in *MTLa* and three in *MTLα* (Fig. 2.1). These additional ORFs are arranged in pairs, one member of which is in *MTLa* and the other in *MTLα*. One pair of the ORFs is similar to the *S. cerevisiae PIK1* gene, a phosphatidylinositol kinase (*PIK*), the second pair of ORFs is similar to the *S. cerevisiae YKR003W*, a member of a class of genes similar to the human oxysterol binding protein (OSBP) gene, referred to here as *OBP*, and the remaining pair of ORFs is similar to *S. cerevisiae PAP1*, a poly(A) polymerase (*PAP*) (21). The proteins encoded by the related ORF pairs are approximately 60% identical to one another. This level of divergence between the members of each gene pair suggests that, although clearly related, they may have subtly different functions from each another. *MAT* loci have been characterized in several fungi, and to date, the proteins encoded by them fall into the categories of sequence-specific DNA binding proteins, pheromones, and pheromone receptors(22, 23, 24, 25, 26). A few other *MAT* genes (*S. pombe Mm*, and *U. maydis LGA2* and *RGA2*) do not closely resemble any known genes, and their functions are as yet unknown. In contrast, six of the nine genes in the *C. albicans MTL* locus code for types of proteins not found in any of the previously described *MAT* loci.

The open reading frames for all nine genes of the *C. albicans MTL* locus appear intact, suggesting that they code for functional proteins. In the case of *C. albicans* **a**1 we know this is the case because it can mediate transcriptional repression from a test promoter bearing an hsg operator. The clustering of all nine genes into the *C. albicans MTL* locus suggests, by analogy with other fungi, that all of the genes are involved in a single biological process. These genes could function to regulate a sexual cycle in *C. albicans* which has remained hidden from investigators, or they could be sexual cycle components derived from an evolutionary ancestor but now be used to regulate another cellular process.

UCSF LIBRARY

References and Notes

1. F. C. Odds, *Candida and Candidosis* (Bailliere Tindall, London, ed. 2, 1988).
2. I. Herskowitz, J. Rine, and J. Strathern, in *Molecular and Cellular Biology of the Yeast Saccharomyces* E. W. Jones, J. R. Pringle, J. R. Broach, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992), vol. 2: Gene Expression, pp. 583-656.
3. A. D. Johnson, *Curr Opin Genet Dev* **5**, 552-8 (1995).
4. Two different sets of overlapping lambda clones were identified using seven ^{32}P labeled PCR products as probes against a lambda zap library of *C. albicans* strain SC5314 genomic fragments (gift from B. Braun). Primer pairs for generating the probe fragments were of the following sequences: **overlapping clone set 1** primer pair 1 5'cgggatcccagatggatgatgataaagatag3' and 5'cgggatcctccagttaaattgtttcatattttg3' primer pair 2: 5'gtctcgaatagaatccaagag3' and 5'gtgcagactgatctgtgtc3' primer pair 3: 5'taggaggtgtaacaccgtatgct3' and 5'ctctgcgatcagcgaccagcaa3' **overlapping clone set 2** primer pair 1: 5'cgggatcccagatggatgatgataaagatag3' and 5'cgggatcctccagttaaattgtttcatattttg3' primer pair 2: 5'cgggatcccaagaactatttccaatccggaa3' and 5'cgggatccaacatcctcaattgtacccg3' primer pair 3: 5'ggagctacaagtatgactgt3' and 5'gcgattctcatctgatcagcta3' primer pair 4: 5'ctctgcgatcagcgaccagcaa3' and 5'cgcagaaacgaagtgcgaga3'. PCR products were labeled with ^{32}P in random prime reactions using Amersham Megaprime DNA Labeling Kit (#RPN1605). Probe hybridizations, plaque lifts, phage purifications and plasmid excisions were carried out according to B. Braun's Plaque lift/Phage screening protocol (http://www.sacs.ucsf.edu/home/JohnsonLab/). For each screening, approximately 10,000 plaques were screened and estimated to represent 2.5 to 5 genome equivalents. Genomic inserts were sequenced in full at the UCSF Biomolecular Resource Center..

UCSF LIBRARY

5. *MTLa* GenBank accession #AF167162 and *MTLα* GenBank accession #AF167163.

6. R. S. Keogh, C. Seoighe, K. H. Wolfe, *Yeast* **14**, 443-57 (1998).

7. H. Chibana, et al., *Genetics* **149**, 1739-52 (1998).

8. A. M. Miller, *Embo J* **3**, 1061-5 (1984).

9. C. M. Hull and A. D. Johnson, unpublished data.

10. Disruption constructs for *MTLa1* and *MTLα2* were constructed using the URA blaster construct pMB-7 (27). The *MTLa1* deletion construct primers 5'cagcacatgcatgcgagctcctttcagatcaagaaacagtttca3' and 5'cagcacatgtcgacggtacccgtcctcagtatagatgc3' yielded a PCR fragment with *Asp*718 and *Sac*I ends, and the primers 5'cagcacatgtcgacggtgccacaataactccactg3' and 5'cagcacatgcatgcgagtgcccgtggtaatacaaa3' yielded a PCR fragment with *Sal*I and *Sph*I ends. The fragments were cut and ligated sequentially into pMB-7. The *Asp*718 to *Sac*I digest was performed as a partial digest on the PCR product due to a *Sac*I site in the fragment. The *MTLa1* homeodomain deletion construct primers 5'cagcacatgcatgcgagctcctttcagatcaagaaacagtttca3' and 5'cagcacatgtcgacggtacccgtcctcagtatagatgc3' yielded a PCR fragment with *Asp*718 and *Sac*I ends, and the primers 5'cagcacatgcatgcgagctcaaagtgtagagaaactagttc3' and 5'cagcacatgtcgacggtacctaactaattattttatttcctcccctttta3' yielded a PCR fragment with *Sal*I and *Sph*I ends. The fragments were cut and ligated sequentially into pMB-7. The *MTLα2* deletion construct primers 5'gaagatctgagctcagtctatcttgatttaggg3' and 5'ggaagatctgtcttgttattgatgtgag3' yielded a PCR fragment with *Bgl*II and *Sac*I ends, and the primers 5'aactgcagcttcgtataggtgtgcacttt3' and 5'aactgcagaagcttgactctttggtcatgccttcc3' yielded a PCR fragment with *Hin*dIII and *Pst*I ends. The fragments were ligated sequentially into pMB-7.

11. Supplemental information on *C. albicans* transformations is available on *Science* Online at www.sciencemag.org/feature/data

UCSF LIBRARY

12. Supplemental information on isolation of *C. albicans* genomic DNA and PCR conditions is available on *Science* Online at www.sciencemag.org/feature/data. Primer sequences used in PCR are as follows: **A** 5'gttacaccacaatcaacacc3', **B** 5'ttacatgttggtgaacctaaag3', **C** 5'cagcacatgcatgcgagctcaaagtgtagagaaactagttc, **D** 5'cagcacatgtcgacggtacctaactaattattttatttcctccccttta, **E** 5'cgggatccgaaacgacaaagactagac **F** 5'cgggatccctaggttgaatttgaacttg, **hisG1** 5'gcgcgtggcgatgcacatggtcag3', **hisG2** 5'gcgcggcggttgagtagctct3'.

13. A. C. Diener, G. R. Fink, *Genetics* **143**, 769-776 (1996).

14. A. D. Johnson, I. Herskowitz, *Cell* **41**, 237-247 (1985).

15. B. J. Andrews, M. S. Donoviel, *Science* **270**, 251-3 (1995).

16. C. Sadhu, D. Hoekstra, M. J. McEachern, S. I. Reed, J. B. Hicks, *Molecular and Cellular Biology* **12**, 1977-1985 (1992).

17. The starting plasmid for *GFP* repression reporters was pY*GFP3* (gift from B. Cormack) (28). pY*GFP3* (pAJ699) was modified by the addition of the *C. albicans URA3* gene into the *Sal*I site of the vector with a PCR fragment with *Sal*I restriction sites on the ends using primers 5'tcgcctcgagtcgacgggcccagtactaataggaattg3' and 5'tctcggagctcgagtcgacgggcccaggaccacctttgattgt3'. The resulting fragment was digested with *Sal*I ligated into the *Sal*I site of pY*GFP3* to create pAJ717. A *URA3* 3' untranslated region was inserted immediately downstream of the *GFP* gene. The *UTR* was generated as a PCR fragment with *Pst*I restriction sites on the ends using primers 5'cgggatccgcgcgctgcagtagttagttaaatgtgaagggggag3' and 5'cgggatccttaattaactgcagaggaccacctttgattgt3'. The resulting fragment was digested with *Pst*I and ligated into the *Pst*I site of pAJ717 to create pAJ724. The *C. albicans ADH1* promoter was inserted immediately upstream of the *GFP* gene. *ADH1-Bgl*II was created in a chimeric PCR reaction to generate an *ADH1* promoter with *Hin*dIII restriction sites on the ends and a *Bgl*II restriction site located between the predicted upstream activating sequences and TATA box. Two starting fragments with

UCSF LIBRARY

overlapping ends were generated with primers 5'cgggatccaagctttaacaaatgaa3', 5'ataagagatctcttgcttgcatgacg3' and 5'cgggatccaagcttaattgtttttgtatttg3', 5'gcaagagatctcttattcagaattttcag3'. These PCR products were mixed in a final PCR reaction to generate the full-length *ADHI-Bgl*II promoter using primers 5'cgggatccaagctttaacaaatgaa3' and 5'cgggatccaagcttaattgtttttgtatttg3'. The resulting *ADH1-Bgl*II promoter fragment was digested with *Hin*dIII and cloned into pAJ724 to generate *CaADH1-GFP* (pAJ868). Reporter plasmids with 3 tandem repressor binding sites were constructed by cloning double-stranded oligonucleotides into the *Bgl*II site of the *CaADH1-GFP* plasmid. In each case two oligos containing three binding sites were kinased and annealed to one another to generate a double-stranded duplex with single-stranded *Bgl*II-compatible overhangs. Single binding sites for each are highlighted in bold text. 1) pAJ888, S. *cerevisiae* haploid-specific gene site 5'gatc**tgatgtaattaattacatga**attgatgtaattaattacatgaattgatgtaattaattacatga3' annealed to 5'gatct**catgtaattaattacatca**attcatgtaattaattacatcaattcatgtaattaattacatca3' 2) pAJ898, *S. cerevisiae* haploid-specific gene site with point mutations that are known to abolish **a**1/α2 binding in *S. cerevisiae* (18). 5'gatc**tgctgtaattaattccatga**attgctgtaattaattccatgaattgctgtaattaattccatga3' annealed to 5'gatc**tcatggaattaattacagca**attcatggaattaattacagcaattcatggaattaattacagca3'. 3) pAJ1081, **a**1/α2 binding sequence from the *C. albicans CAG1* promoter 5'gatc**tgatgtgatttttaacatgg**attgatgtgatttttaacatggattgatgtgatttttaacatgg3' annealed to 5'gatc**ccatgttaaaaatcacatca**atccatgttaaaaatcacatcaatccatgttaaaaatcacatca3'. 4) pAJ933, **a**1/α2 binding sequence from the *C. albicans CAG1* promoter with point mutations analogous to the mutations in the S. cerevisiae binding site that abolish **a**1/α2 binding 5'gatc**tgctgtgatttttaccatgg**attgctgtgatttttaccatggattgctgtgatttttaccatgg3' annealed to 5'gatc**ccatggtaaaaatcacagca**atccatggtaaaaatcacagcaatccatggtaaaaatcacagca3'. Reporter constructs were linearized via a *Bsp*EI site and integrated into the *ADHI*

UCSF LIBRARY

promoter. The pattern of fluorescence was confirmed for 5 transformants for each reporter.

18. Y. Jin, H. Zhong, A. K. Vershon, *Molecular and Cellular Biology* **19**, 585-593 (1999).

19. E. Leberer, et al., *Proc Natl Acad Sci U S A* **93**, 13217-22 (1996).

20. M. Raymond, et al., *Molecular Microbiology* **27**, 587-98 (1998).

21. *PIKa* codes for a predicted protein that is 47% identical to Pik1, and *PIKα* codes for a predicted protein that is also 47% identical to the Pik1 protein. The predicted protein for *OBPa* is 51% identical to the predicted *YKR003W* protein, and the predicted protein for *OBPα* is 60% identical and to the predicted *YKR003W* protein. The predicted protein for *PAPa* is 60% identical to Pap1, and the predicted protein for *PAPα* is 59% identical to Pap1.

22. L. A. Casselton, N. S. Olesnicky, *Microbiology and Molecular Biology Reviews* **62**, 55-70 (1998).

23. B. G. Turgeon, *Annual Reviews in Phytopathology* **36**, 115-137 (1998).

24. M. Kelly, J. Burke, M. Smith, A. Klar, D. Beach, *Embo J* **7**, 1537-47 (1988).

25. F. Banuett, *Trends Genet* **8**, 174-80 (1992).

26. A. Sharon, et al., *Mol Gen Genet* **251**, 60-8 (1996).

27. W. A. Fonzi, M. Y. Irwin, *Genetics* **134**, 717-28 (1993).

28. B. P. Cormack, et al., *Microbiology* **143**, 303-11 (1997).

29. E. Coppin, R. Debuchy, S. Arnaise, M. Picard, *Microbiol Mol Biol Rev* **61**, 411-28 (1997).

30. We thank J. Rine, J. Thorner, I. Herskowitz, D. Ganem and R. Brazas for helpful discussions of the work; I. Herskowitz, R. Smith, D. Inglis, A. Uhl, D. Kadosh, A. Tsong, B. Braun, and R. Brazas for manuscript comments; B. Cormack for use of unpublished reagents. Special thanks to B. Braun for use of his genomic library and technical assistance throughout, to R. Taylor at the UCSF BRC Facility for her

UCSF LIBRARY

sequencing expertise, and to members of the Johnson lab for their support. This
work was supported by NIH grant GM37049 to A.D.J.

Supplemental References and Notes

11. Reporter and deletion constructs were transformed into *C. albicans*. 10 mL overnight yeast peptone dextrose (YPD) cultures grown at 30° C were diluted 0.3 mL into 10 mL fresh YPD and grown at 30°C until OD_{600} reached 0.8-1.0. Cells were pelleted, washed in 1 mL lithium solution (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 0.1 M lithium acetate). Lithium acetate solid was purchased from EM Science, and this brand was found to be critical for success. Cells were resuspended in 0.5 mL lithium solution. 200 µL cells were mixed with 50 µg sonicated salmon sperm DNA and 2.5-5 µg digested plasmid DNA. 1 mL of PEG solution (10 mM Tris-Cl pH 7.5, 1 mM EDTA, 0.1 M lithium acetate, 40% polyethylene glycol 3500) was added. Cells were incubated at 30°C for 3 hours and 42°C for 1 hour. Cells were pelleted, resuspended in 150 µL minimal medium, and plated on selective plates. Transformed deletion strains were counter-selected to eliminate the *URA3* marker gene by growing on yeast synthetic plates containing 0.08% 5-fluoroorotic acid (5-FOA) and 0.025 mg/mL uridine.
12. Genomic DNA was isolated according to Hoffman and Winston (31) with the following exceptions: glass beads were not acid-washed, and cells were agitated on Eppendorf Mixer for 20 minutes at 4°C rather than vortexed, and RNaseA step was omitted. DNA was resuspended in 50 µL dH_2O and 10 µg RNaseA. Genomic DNA from the various disrupted strains was used as template in PCR reactions. 50 µL reactions contained 1µL genomic DNA, 1µM each primer, 200µM each deoxyribonucleotide, 1U Taq DNA polymerase (Boehringer Mannheim), 1X Taq

UCSF LIBRARY

buffer (Boehringer Mannheim). Reactions were carried out in 30 cycles of 94°C for 30 seconds, 58°C for 30 seconds, and 72°C for 2 minutes.

31. C. S. Hoffman, F. Winston, *Gene* **57**, 267-72 (1987)

UCSF LIBRARY

Appendices to Chapter 2

USF LIBRARY

Appendix A

Additional evidence for *MTL* locus arrangement.

One piece of evidence that supports presenting the *C. albicans* genomic fragments containing **a**1, α1, and α2 in a mating type-like locus is that the ends of the fragments contain identical DNA sequence. This fact suggests strongly that the approximately 1.5 kilobase 5' flanks and 3' flanks of the fragments are co-linear in the genome with nearly nine kilobases of intervening sequence between. That intervening sequence contains non-allelic genes (4 on the **a** fragment and 5 on the α fragment) that represent the locus. The sequence relationship between the fragments is represented by dot plot in panel A of Figure A.1.

A second piece of evidence for the (*MTL*) locus arrangement is the confirmation of the prediction that genes on the genomic fragments would be haploid in the *C. albicans* diploid genome. An example of this result was shown for **a**1 by PCR analysis in Chapter 2; however, a further examination was carried out by Southern blot to show the absence of **a**1 after a single round of disruption transformation (Figure A.2). Evidence for haploidy was also confirmed by both Southern Blot and PCR for α1 (Figures A.3 and A.4), and for α2 (Figure A.5 and A.6).

One last piece of information that is made clear from the PCR confirmations of the **a**1, α1, and α2 deletion strains is the fact that no other open reading frames are detectable for any of these genes by PCR or in Southern blots.. They are deleted from the *MTL*, and they appear to be deleted from the genome entirely. This result suggests strongly that there are no silent mating cassettes at any other location in the genome as there are in *S. cerevisiae*. The *MTL* appears to be the only location in the genome for **a**1, α1, and α2 in *C. albicans*.

UCSF LIBRARY

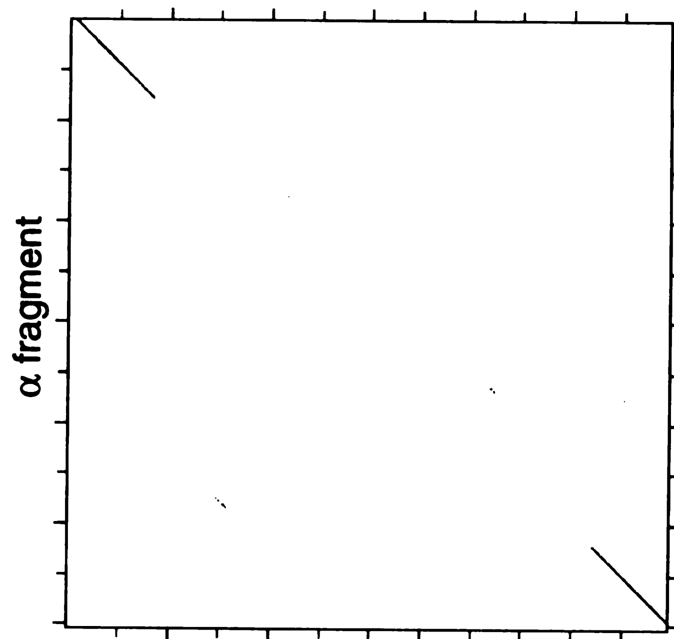
Figure A.1. DNA matrix comparisons of DNA sequence of *C. albicans MTLa* and *MTLα*. (Displayed in DNA Strider 1.2) **A.** Each dot represents 15 nucleotides of identical DNA sequence between the two *MTL* segments. Each hatch mark along the axes represents 1 kilobase of DNA sequence. The plot indicates that while the flanking DNA is nearly identical in sequence (approximately 1.5 kb on each side), the sequences of *MTLa* and *MTLα* differ considerably (over nearly 9 kb). **B.** Each dot represents 7 nucleotides of identical DNA sequence between the two *MTL* segments. The diagonal line in the center of the plot corresponds to the similarity between the *OBPa* and *OBPα* genes and the *PIKa* and *PIKα* genes. The region of similarity in the lower left corresponds to the similarities for the *PAPa* and *PAPα* genes.

UCSF LIBRARY

Figure A.1

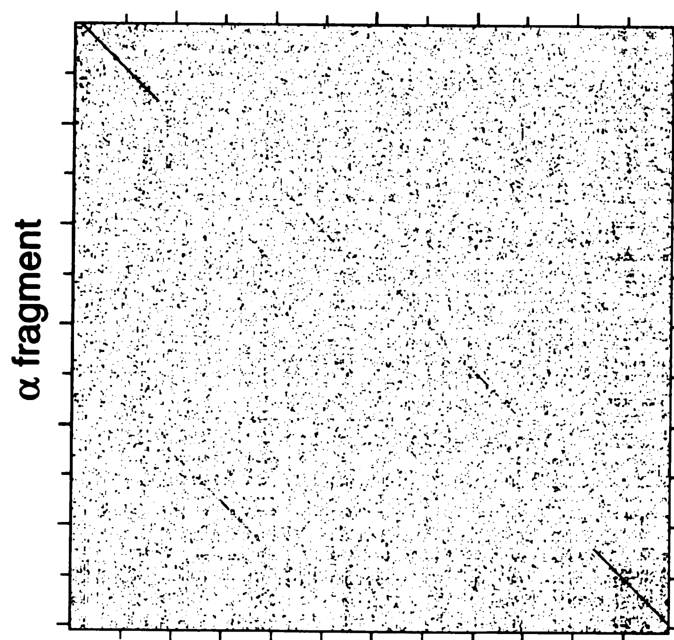
A.

a fragment



B.

a fragment



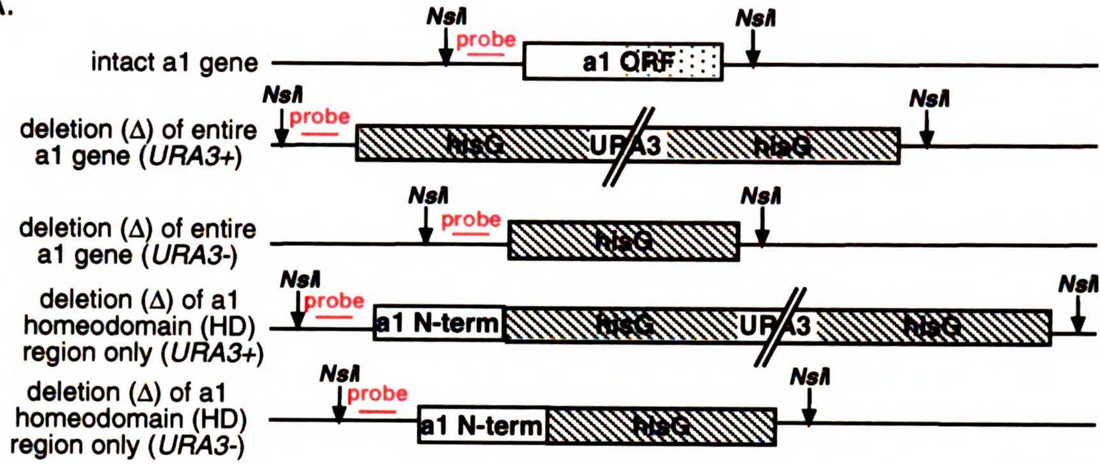
UCSF LIBRARY

Figure A.2. Southern blot showing deletion of the *C. albicans MTLa1* gene. **A.** Schematic of the *MTLa1* gene and various disruption alleles. From top to bottom they are the intact *MTLa1* locus, a full deletion of the *MTLa1* locus prior to counter-selection on 5-FOA, deletion of the locus after counter-selection, a C-terminal deletion prior to counter-selection, a C-terminal deletion after counter-selection. Red bar shows the location of the probe used in the Southern blot in panel B. **B.** Southern blot confirming deletion of *MTLa1* after a single round of transformation. *C. albicans* genomic DNA was cleaved with *Nsi*I prior to electrophoresis. Lane 1: marker lane, Lane 2: band of expected size (2.9 kb) in a wild-type (CAI4) strain, Lane 3: band of expected size (6.0 kb) in single round deletion strain. Note absence of wild-type signal. Lane 4: band of expected size (3.1) kb in counter-selected single round transformant, Lane 5: band of expected size (6.4) kb) in single round homeodomain deletion strain, Lane 6: band of expected size (3.6 kb) in counter-selected single round homeodomain deletion strain.

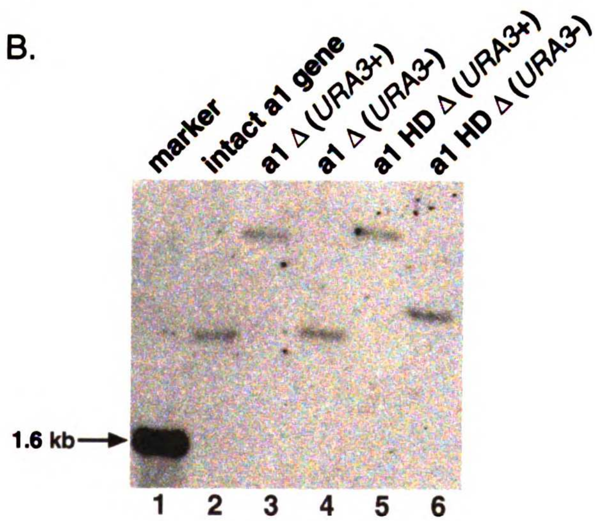
UCSF LIBRARY

Figure A.2

A.



B.



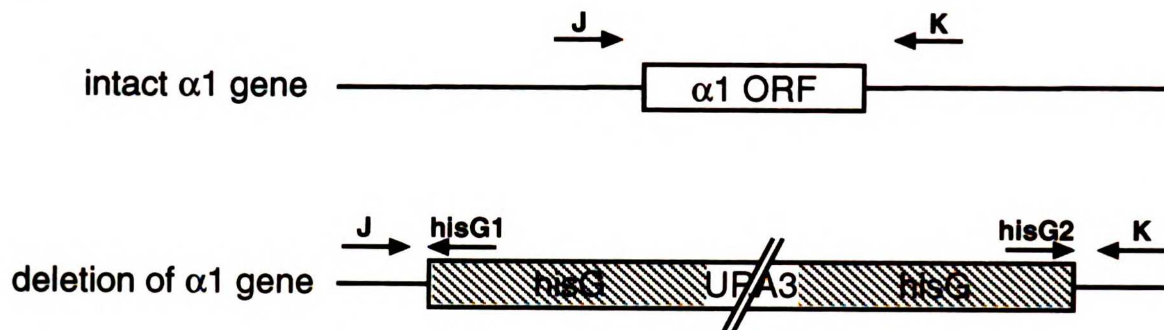
UCSF LIBRARY

Figure A.3. Products of PCR reactions showing deletion of the *C. albicans MTLα1* gene and the absence of genes for α1 at other loci. **A.** Schematic of genomic locus. Labeled arrows indicate locations of primers used in the PCR reactions in panel B. **B.** Agarose gels showing fragments produced in PCR reactions using indicated primer sets on genomic DNA isolated from different strains. Lanes labeled WT contain genomic DNA from wild-type cells as the template; lanes labeled α1 KO #1 and α1 KO #2 contain genomic DNA from two independent α1 deletion strains as template. Predicted sizes of the PCR products are indicated in kilobases (kb) to the left of the panel. All primer sets yielded the appropriately sized products or no product as predicted. Lane 1: DNA marker fragments. Lanes 2-4: Primers J and K on indicated genomic DNA. Lanes 5-7: Primers J and hisG1 on indicated genomic DNA. Lanes 8-10: Primers K and hisG2 on indicated genomic DNA.

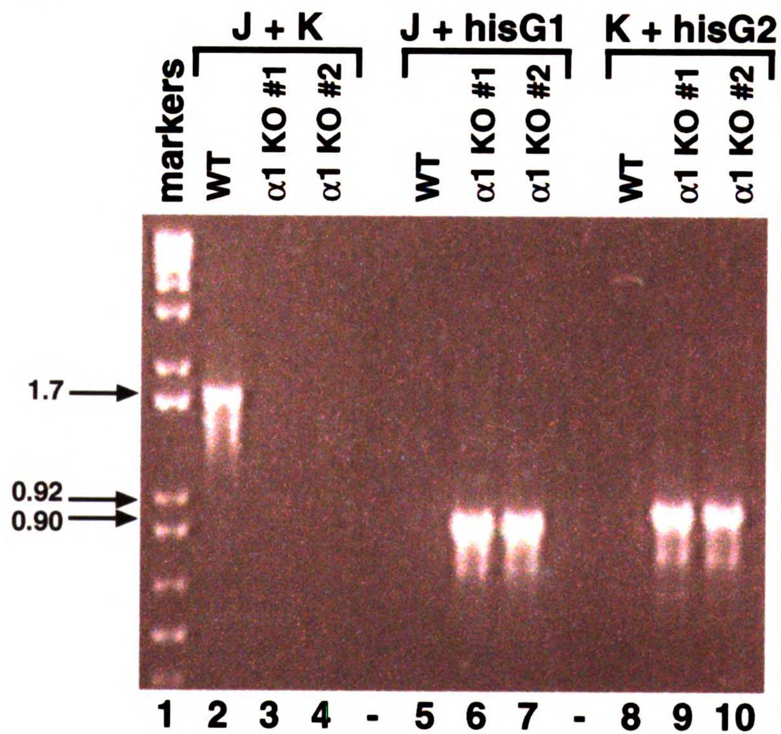
UCSF LIBRARY

Figure A.3

A.



B.



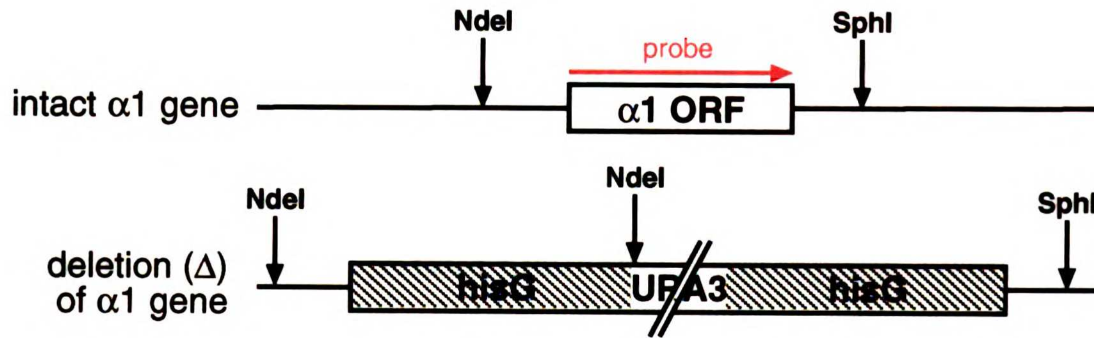
UCSF LIBRARY

Figure A.4. Southern blot to show deletion of *C. albicans* α1 gene at *MTL* locus. **A.** Schematic of genomic locus: intact α1 locus, full deletion of **a**1 locus prior to counter-selection on 5-FOA. Red arrow shows location of probe used in Southern blot. **B.** Southern blot confirming deletion of α1 after single round of transformation. Lane 1: marker lane, Lane 2: band of expected size in a wild-type (CAI4) strain, Lanes 3 and 4: no signal from the α1 open reading frame probe after a single round of transformation in two independent transformants.

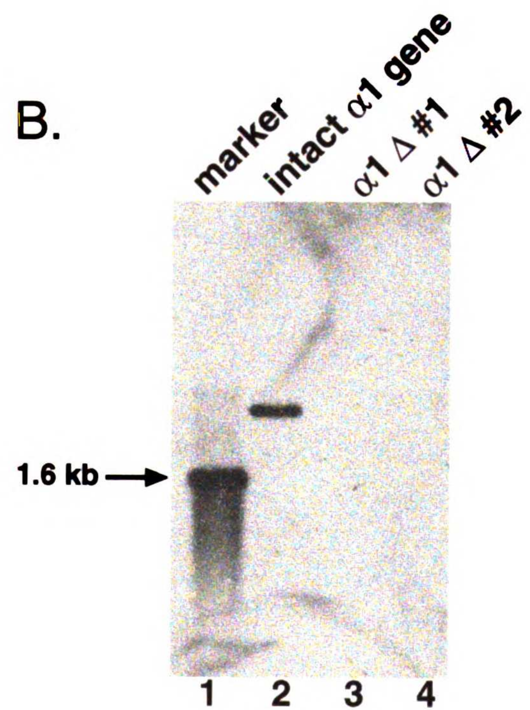
UCSF LIBRARY

Figure A.4

A.



B.

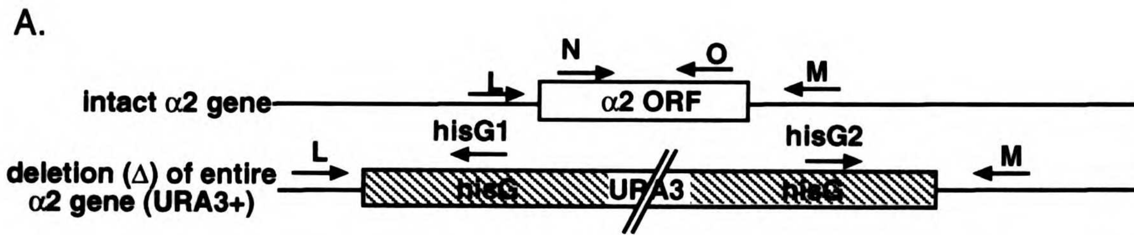


UCSF LIBRARY

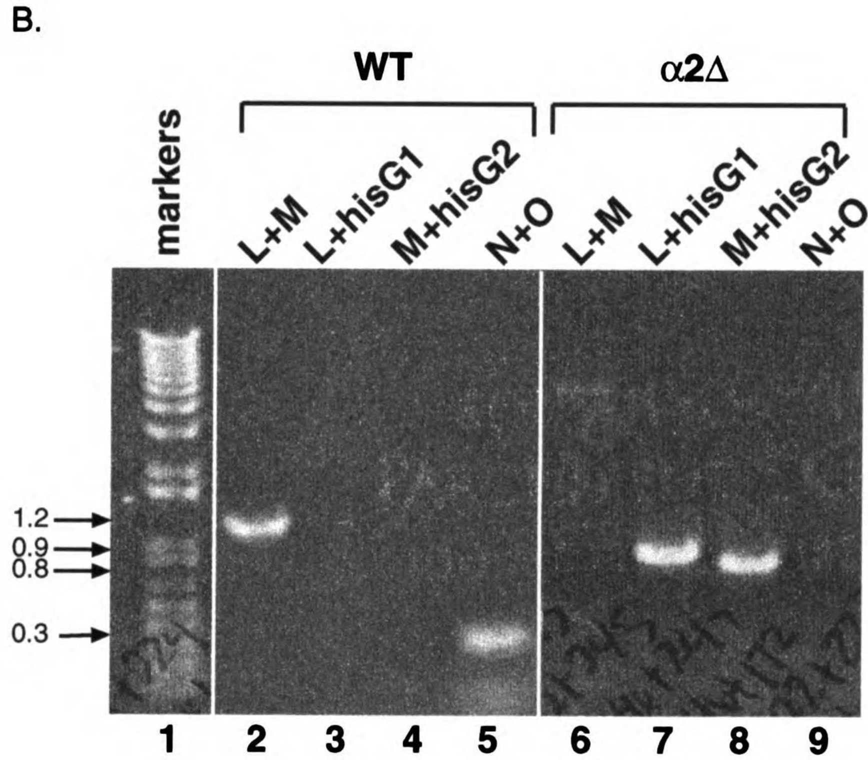
Figure A.5. Products of PCR reactions showing deletion of the *C. albicans MTL*α*2* gene and the absence of genes for $\alpha 2$ at other loci. **A.** Schematic of genomic locus. Labeled arrows indicate locations of primers used in the PCR reactions in panel B. **B.** Agarose gels showing fragments produced in PCR reactions using indicated primer sets on genomic DNA isolated from different strains. Lanes 2-5 labeled WT contain genomic DNA from wild-type cells as the template; lanes 6-9 labeled $\alpha 2\Delta$ contain genomic DNA from an $\alpha 2$ deletion strain prior to counterselection as template. Predicted sizes of the PCR products are indicated in kilobases (kb) to the left of the panel. All primer sets yielded the appropriately sized products or no product as predicted. Lane 1: DNA marker fragments. Lanes 2, 6: Primers L and M on indicated genomic DNA. Lanes 3, 7: Primers L and hisG1 on indicated genomic DNA. Lanes 4, 8: Primers M and hisG2 on indicated genomic DNA. Lanes 5, 9: Primers N and O on indicated genomic DNA.

Figure A.5

A.



B.

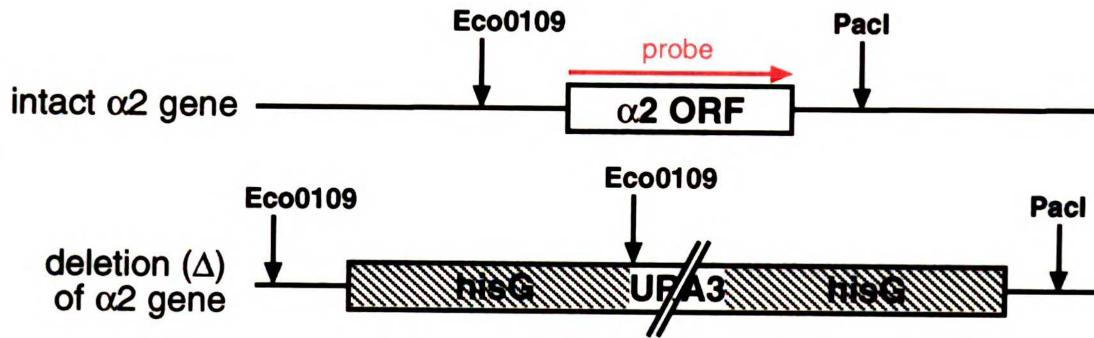


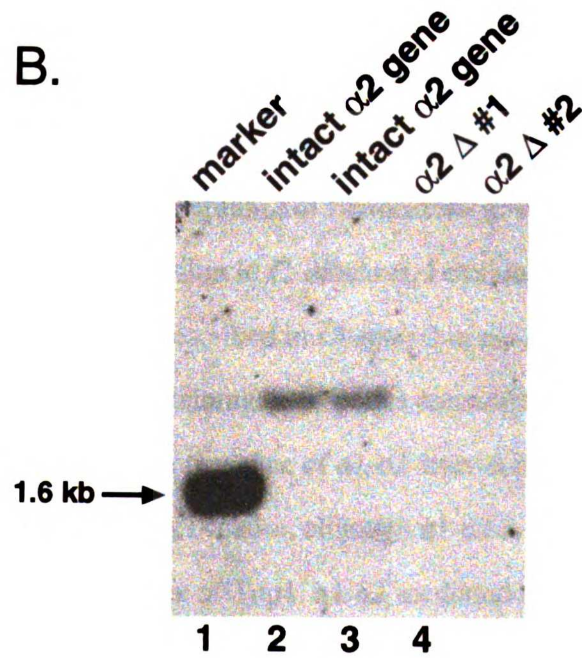
UCSF LIBRARY

Figure A.6. Southern blot to show deletion of *C. albicans* α2 gene at *MTL* locus. **A.** Schematic of genomic locus: intact α2 locus, full deletion of **a**1 locus prior to counter-selection on 5-FOA. Red arrow shows location of probe used in Southern blot. **B.** Southern blot confirming deletion of α2 after single round of transformation. Lane 1: marker lane, Lanes 2 and 3: band of expected size in a wild-type (CAI4) strain, Lanes 4 and 5: no signal from the α2 open reading frame probe after a single round of transformation in two independent transformants.

UCSF LIBRARY

Figure A.6





Appendix B

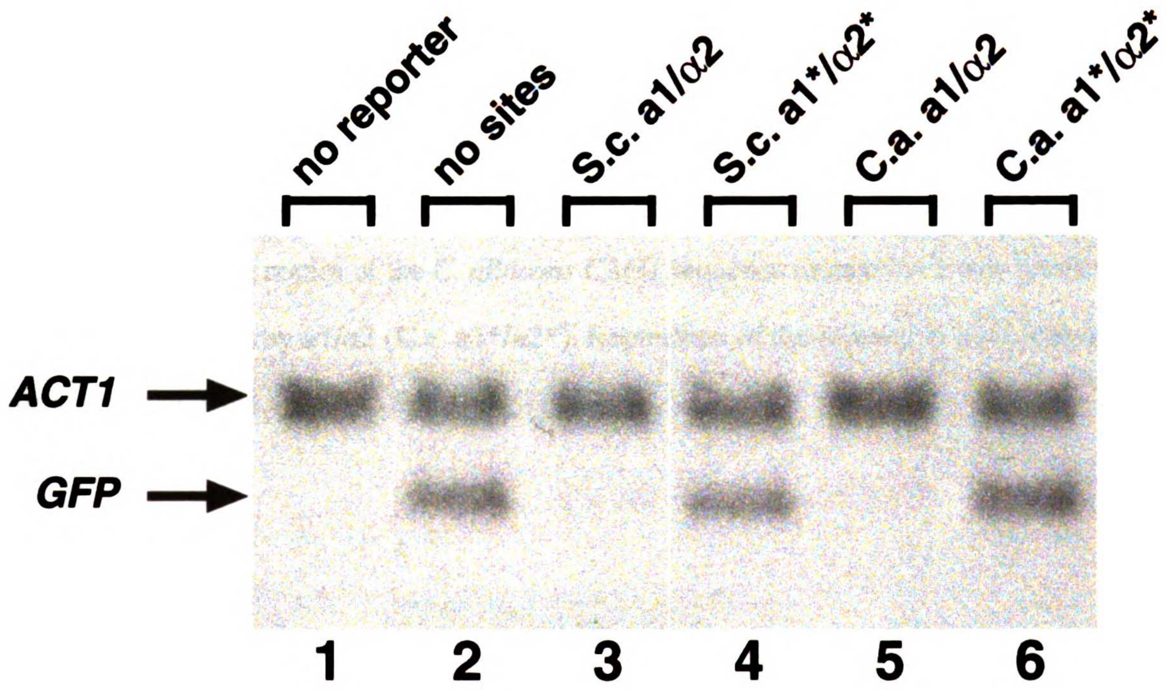
Additional information on a1-α2 activity in *C. albicans.*

As described in Chapter 2, **a**1-α2 activity was detected using heterologous reporter constructs containing **a**1-α2 binding sites upstream of a *C. albicans* modified *GFP* gene. To prove that the apparent repression activity that was being detected was due to regulation at the level of transcription, Northern blots were performed to detect the *GFP* transcripts. Figure B.1 shows the regulation of the *GFP* transcript in response to the presence of **a**1-α2 binding sites in the reporter promoter. This experiment confirms that an activity exists in *C. albicans* that is responsive to **a**1-α2 sites in vivo that represses transcription.

In *S. cerevisiae,* the **a**1-α2 heterodimer binds to a specific DNA site in the promoter of a repressible gene and accomplishes repression by recruiting to the promoter a global repressor complex. This complex is composed of the proteins Ssn6 and Tup1. These proteins have been shown to participate in the repression of scores of genes in *S. cerevisiae,* and Tup1 has been particularly well characterized. It is required for **a**1-α2 repression. Tup1 has also been investigated in *C. albicans. TUP1* was deleted in *C. albicans* by Burk Braun (*1*) and was shown to be a crucial regulator of filamentous growth. In order to better understand **a**1-α2-mediated repression in *C. albicans*, I evaluated repression of the heterologous reporter constructs described in Chapter 2 in *tup1/tup1* deletion cells. As shown in Figure B.2, the cells containing the reporter constructs are constitutively fluorescent. This indicates that the presence of **a**1-α2 sites in the constructs of these strains is not having an effect, presumably because, although **a**1-α2 is binding to the site, no repression can occur in the absence of Tup1. **a**1-α2 mediated repression in *C. albicans* is dependent on Tup1.

Figure B.1. Confirmation that **a**1-mediated repression affects mRNA levels. Actin (*ACT1*) and green fluorescent protein (*GFP*) transcript levels in strains carrying the 5 different reporter constructs were analyzed by Northern blot using *GFP* and *ACT1* sequences as probes. Lane 1 contains RNA isolated from a wild-type strain with no *GFP* reporter construct. Lanes 2-6 contain RNA isolated from strains wild-type strains containing the reporters #1-5 as described in the legend for Figure 2.4. Transcript levels parallel the GFP fluorescence shown in Figure 2.4.

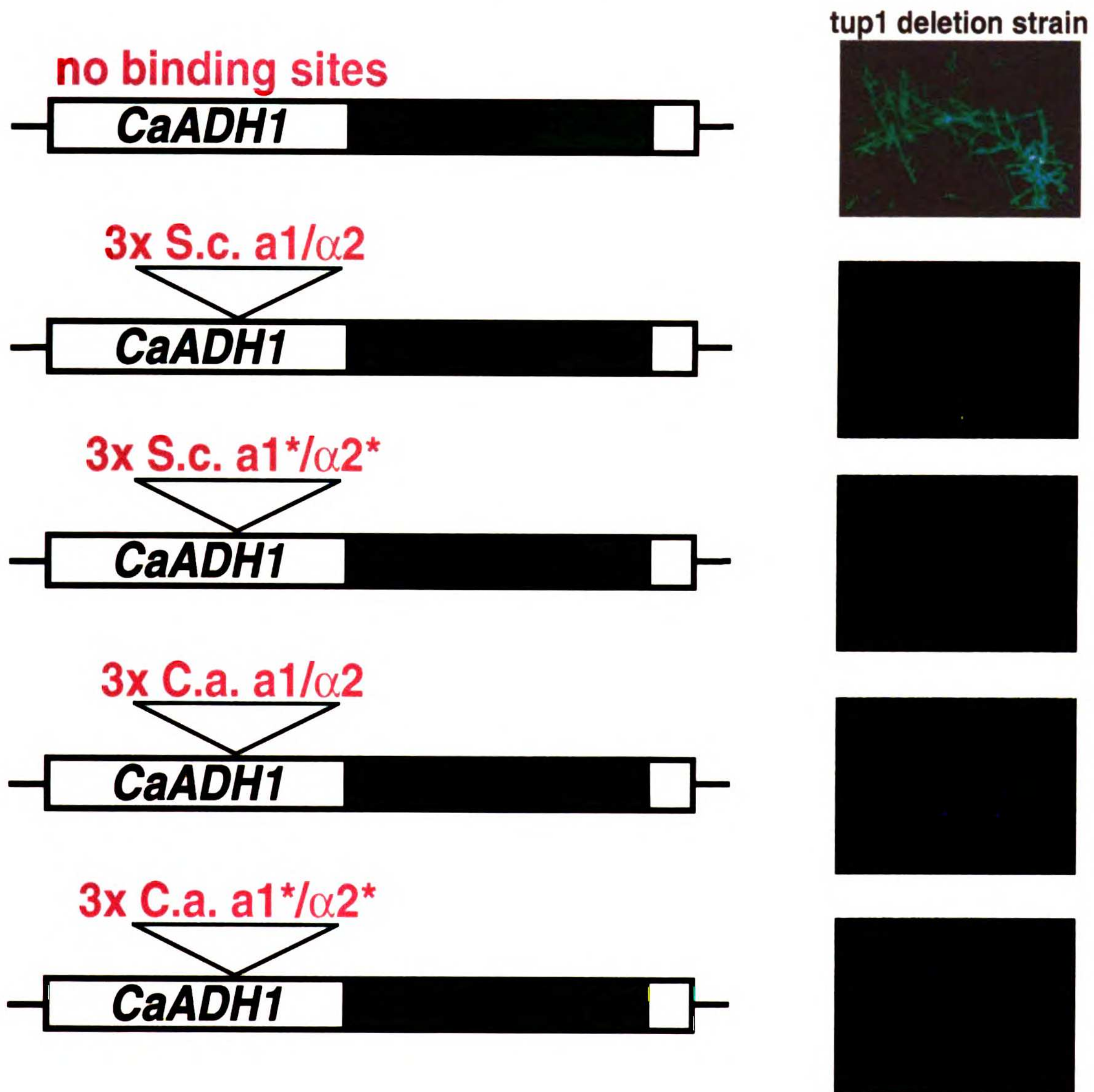
Figure B.1



UCSF LIBRARY

Figure B.2. **a**1/α2-mediated transcriptional repression in *C. albicans* is dependent on Tup1. Transcriptional repression activity determined as a measure of *GFP* fluorescence in a heterologous reporter system in the presence of test DNA binding sites as inserts in a *tup1Δ/tup1Δ* strain. In each case the reporter consists of the *C. albicans ADH1* promoter driving *GFP* expression. Reporter 1) no insert, 2) three *S. cerevisiae* hsg operator consensus sequences (S.c. **a**1/α2), 3) three *S. cerevisiae* hsg operators mutated at two nucleotide positions to prevent recognition by **a**1/α2 (S.c. **a**1*/α2*) (18), 4) three copies of the hsg operator-like sequences located upstream of the *C. albicans CAG1* gene (C.a. **a**1/α2), and 5) three copies of the *C. albicans CAG1* sequence mutated in a way predicted to destroy recognition by **a**1/α2 (C.a. **a**1*/α2*). Repression of the reporter is not observed in the absence of Tup1.

Figure B.2



UCSF LIBRARY

References

1. B. R. Braun and A. D. Johnson. 1997. Control of Filament Formation in *Candida albicans* by the Trancriptional Repressor TUP1. *Science* 277:105-109

UCSF LIBRARY

Appendix C

Identification of additional open reading frames in the *C. albicans MTL* locus

In addition to the three transcriptional regulators, **a**1, α1, and α2, six additional open reading frames (ORFs) were identified, three in *MTLa* and three in *MTLα* (Figure 2.1). These additional ORFs are arranged in pairs, one member of which is in *MTLa* and the other in *MTLα*. The gene pairs have been designated *PIKa/PIKα*, *OBPa/OBPα*, and *PAPa/PAPα* and will be described in more detail below. Although the proteins encoded by the related ORF pairs are approximately 60% identical to one another, they are only weakly similar at the DNA level as shown in a low stringency DNA matrix in panel B of Figure A.1. This level of divergence between the members of each gene pair suggests that, although clearly related, they may have subtly different functions from one another.

One pair of the ORFs is strongly similar to the *S. cerevisiae PIK1* gene, a phosphatidylinositol kinase (*PIK*). The *PIK1*-like gene contained in *MTLa* (designated *PIKa*) codes for a predicted protein that is 47% identical to Pik1, and the *PIK1*-like gene located on *MTLα* (designated *PIKα*) codes for a predicted protein that is also 47% identical to the Pik1 protein. *S. cerevisiae PIK1* gene which encodes a phosphatidylinositol 4-kinase. Although the precise role of the *S. cerevisiae* Pik1 protein is not known, phosphatidylinositol 4-kinases from other organisms function in transducing signals within cells. *PIK1* is an essential gene in *S. cerevisiae*, and conditional alleles suggest that it may be required for cytokinesis (*1*). *PIK1* may also be involved in the pheromone response pathway since overexpression of *PIK1* increases **a** cell sensitivity to α factor (*2*). A third possibility is that *PIK1* may be involved in sensing the sterol composition of different cellular membranes (*3*).

The second pair of ORFs are most similar to the *S. cerevisiae YKR003W*, a member of a class of genes similar to the human oxysterol binding protein (OSBP) gene, referred to

here as *OBP*. The predicted protein for *OBPa* is 51% identical to the predicted *YKR003W* protein, and the predicted protein for *OBPα* is 60% identical and to the predicted *YKR003W* protein. *YKR003W* is similar to the *HES1*, *KES1*, and *OSH1* genes in *S. cerevisiae,* and along with *YHR001W*, appears to complete a five gene family of proteins with strong similarity to the mammalian oxysterol binding protein, OSBP. OSBP may regulate cholesterol biosynthesis in mammalian cells (*4*); however, its biochemical mechanism is not understood in detail. It has been proposed that the *S. cerevisiae* oxysterol binding protein homologs may be involved in regulating synthesis of ergosterol (*5*), the fungal counterpart of cholesterol, or in dispersing sterols to different cellular membranes (*6*).

The remaining pair of ORFs is most similar to *S. cerevisiae PAP1*, a poly(A) polymerase (*PAP*). The predicted protein for *PAPa* is 60% identical to Pap1, and the predicted protein for *PAPα* is 59% identical to Pap1. In *S. cerevisiae*, only one poly(A) polymerase gene is known, and it is essential for cell viability since it is required for the formation of all polyadenylated transcripts (*7*). Although it does not appear to be a selectively regulated process in *S. cerevisiae*, polyadenylation is known in other organisms to be an important step at which the regulation of gene expression can occur (for reviews see *8,9*). Perhaps the *PAP* genes present in the *C. albicans MTL* locus are involved in some form of post transcriptional regulation.

All of the genes in the locus have now been deleted individually, and the transcriptional regulators have been deleted in all possible combinations with each other. Although the regulator deletion strains do not show any gross phenotypes, the total locus deletion strains (*mtlaΔ* and *mtlαΔ*) grow poorly and show altered colony morphologies. The individual *PIK, OBP,* and *PAP* deletion strains have not been rigorously phenotype tested, but they do not show any severe growth defects under YPD growth conditions.

References

1. Garcia-Bustos, J. F., Marini, F., Stevenson, I., Frei, C., and Hall, M. N. (1994). PIK1, an essential phosphatidylinositol 4-kinase associated with the yeast nucleus. Embo J *13*, 2352-61.

2. Flanagan, C. A., Schnieders, E. A., Emerick, A. W., Kunisawa, R., Admon, A., and Thorner, J. (1993). Phosphatidylinositol 4-kinase: gene structure and requirement for yeast cell viability. Science *262*, 1444-8.

3. Thorner, J. personal communication

4. Alphey, L., Jimenez, J., and Glover, D. (1998). A Drosophila homologue of oxysterol binding protein (OSBP)-- implications for the role of OSBP. Biochim Biophys Acta *1395*, 159-64.

5. Jiang, B., Brown, J. L., Sheraton, J., Fortin, N., and Bussey, H. (1994). A new family of yeast genes implicated in ergosterol synthesis is related to the human oxysterol binding protein. Yeast *10*, 341-53.

6. Fang, M., Kearns, B. G., Gedvilaite, A., Kagiwada, S., Kearns, M., Fung, M. K., and Bankaitis, V. A. (1996). Kes1p shares homology with human oxysterol binding protein and participates in a novel regulatory pathway for yeast Golgi-derived transport vesicle biogenesis. Embo Journal *15*, 6447-59.

7. Lingner, J., Kellermann, J., and Keller, W. (1991). Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. Nature *354*, 496-8.

8. Colgan, D. F., and Manley, J. L. (1997). Mechanism and regulation of mRNA polyadenylation. Genes and Development *11*, 2755-66.

9. Seydoux, G. (1996). Mechanisms of translational control in early development. Curr Opin Genet Dev *6*, 555-61.

Appendix D

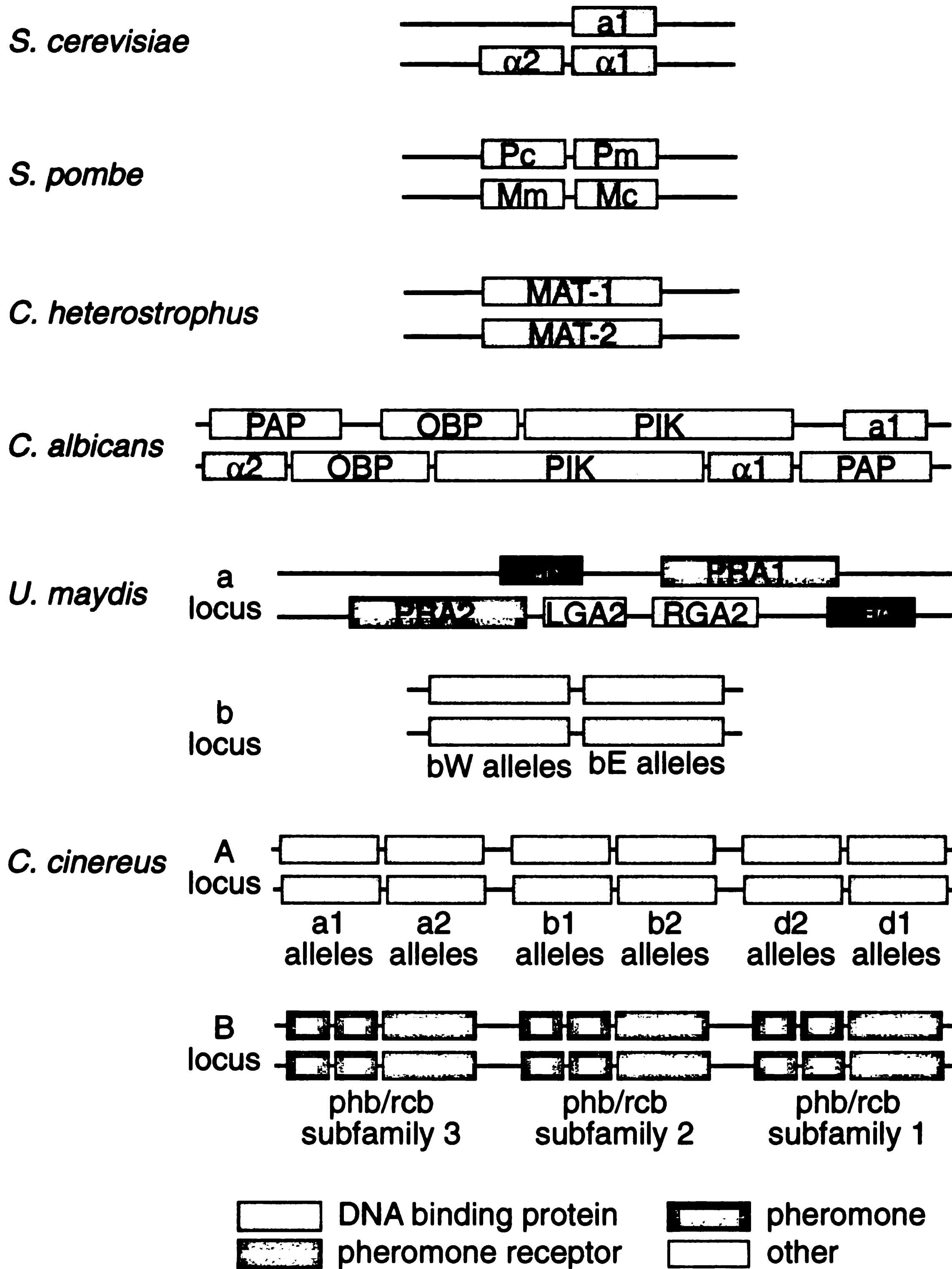
The *C. albicans MTL* locus is distinct from *MAT* loci in other fungi

MAT loci have been identified and sequenced in several fungi, and they typically contain genes of only a few types. In budding and filamentous ascomycetes like *S. cerevisiae*, *S. pombe* (*1*), and *C. heterostrophus* (*2*) the *MAT* locus encodes transcriptional regulators that contain one of three known DNA binding motifs, the homeodomain, α domain, or HMG box (Figure D.1). Typically, these regulators bind to DNA sequences in combination with other DNA binding proteins and regulate the expression of genes involved in mating, meiosis and other aspects of the sexual cycle. In the basidiomycetes such as *U. maydis* and *C. cinereus* (for reviews see *3, 4*), genes involved in mating and meiosis are located at two distinct loci. One of these loci contains genes encoding homeodomain proteins, and the other contains genes encoding pheromones and pheromone receptors. Thus, most of the *MAT* locus proteins described to date fall into the categories of sequence-specific DNA binding proteins, pheromones, and pheromone receptors. A few other *MAT* genes (*S. pombe* Mm, and *U. maydis* lga2 and rga2) do not closely resemble any known genes, and their functions are as yet unknown. In contrast, six of the nine genes in the *C. albicans MTL* locus code for types of proteins not found in any of the previously described *MAT* loci.

Figure D.1. Mating-type (*MAT*) loci from several fungi. The *MAT* loci of the ascomycetes *S. cerevisiae, S. pombe,* and *C. heterostrophus* primarily code for DNA binding proteins (purple) in the homeodomain, α domain, and HMG box classes. The basidiomycetes *U. maydis* and *C. cinereus* each have two genome locations for *MAT* information (the a locus and the b locus), one which codes primarily for homeodomain proteins (purple) and one which codes primarily for pheromones (green) and pheromone receptors (pink). The *C. albicans MTL* locus is exceptional because it codes for three additional classes of proteins (yellow) not seen in other *MAT* loci.

UCSF LIBRARY

Figure D.1



References

1. Kelly, M., Burke, J., Smith, M., Klar, A., and Beach, D. (1988). Four mating-type genes control sexual differentiation in the fission yeast. Embo J *7*, 1537-47.

2. Turgeon, B. G. (1998). Application of mating type gene technology to problems in fungal biology. Annual Reviews in Phytopathology *36*, 115-137.

3. Banuett, F. (1992). Ustilago maydis, the delightful blight. Trends Genet *8*, 174-80.

4. Casselton, L. A., and Olesnicky, N. S. (1998). Molecular genetics of mating recognition in basidiomycete fungi. Microbiology and Molecular Biology Reviews *62*, 55-70.

Appendix E

Possible functions for the *C. albicans MTL* locus

The open reading frames for all nine genes of the *C. albicans MTL* locus appear intact, suggesting that they code for functional proteins. In the case of *C. albicans* **a**1 we know this is the case because it can mediate transcriptional repression from a test promoter bearing an hsg operator. The clustering of all nine genes into the *C. albicans MTL* locus suggests, by analogy with other fungi, the possibility that all of the genes are involved in a single biological process. These genes could function to regulate a sexual cycle in *C. albicans* which has remained hidden from investigators, or they could be sexual cycle components derived from an evolutionary ancestor but now be used to regulate some other cellular process.

Although too little information exists to discriminate among these and other ideas, some intriguing possibilities exist. According to the sexual cycle model, the *MTLa1*, *MTLα1*, and *MTLα2* genes could regulate the transcription of sexual cycle genes in *C. albicans.* The additional genes in the locus could modify the conditions under which a sexual cycle could occur by sensing sterols in the environment (consistent with the presence of *PIK*- and *OBP*-like genes) or by regulating specific transcripts in response to environmental conditions (consistent with the presence of the *PAP*-like genes). If *C. albicans* indeed lacks a sexual cycle, one possibility is that the *C. albicans MTL* genes could have been co-opted for use in pathogenesis. In particular, the *PIK*- and *OBP*-like genes could be involved in surveying the environment within a mammalian host and eliciting the appropriate response perhaps through the regulators **a**1, α1, and α2. Although the physiological role of the *C. albicans MTL* locus remains to be determined, the identification this locus in *C. albicans* represents the first molecular view of the ways that the *MAT* locus

Chapter 3

Evidence for Mating of the "Asexual" Yeast *Candida albicans* in a Mammalian Host

This version of Chapter 3 is reprinted here with minor changes and permission from Christina M. Hull, Ryan M. Raisner, and Alexander D. Johnson. 2000. *Science* 289: 307-310. Copyright 2000 American Association for the Advancement of Science.

The work in this chapter was carried out by Christina Hull with the technical assistance of SRAI Ryan Raisner. The original deletion strains were constructed by Christina, and Ryan assisted in the construction of the ade2/ade2 derivatives used in the mating experiment. Ryan also carried out half of the mouse matings and all of the Southern blots.

Abstract

Since its classification nearly eighty years ago, the human pathogen *Candida albicans* has been designated as an asexual yeast. We constructed strains of *C. albicans* that are subtly altered at the *MTL* (mating type-like) locus, a cluster of genes that resembles the mating-type loci of other fungi, and found that these derivatives were capable of mating in a mammalian host. *C. albicans* is a diploid organism, but most of the mating products isolated from a mouse host were tetrasomic for the two chromosomes that could be monitored rigorously, and overall exhibited significantly higher than 2n DNA content. These observations demonstrate that *C. albicans* can indeed recombine sexually.

Text

The yeast *Candida albicans* is found as a commensal organism in the digestive tract of mammals. It is also the most common human fungal pathogen, causing both mucosal and systemic infections, particularly in immunocompromised people (*1*). *C. albicans,* a diploid yeast, has been classified as asexual because no direct observation of mating or meiosis has been reported. Population studies of *C. albicans* indicate that some low level genetic exchange may occur (*2*), but there is no conclusive evidence for recent sexual recombination (*3-5*). The apparent absence of sexual reproduction in *C. albicans* is especially intriguing because its relatives in the budding yeast family (e.g., *Saccharomyces cerevisiae* and *Kluyveromyces lactis*) have retained sexual cycles. In fact, sexual reproduction is common throughout the fungal kingdom, ranging from yeasts to mushrooms.

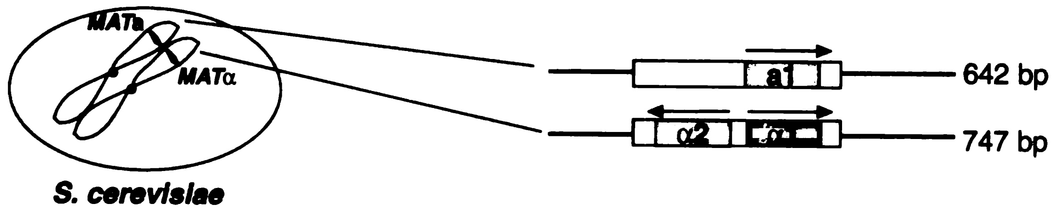
Although mating and meiosis have not been observed for *C. albicans*, homologs of genes that function in both mating (e.g., *GPA1* (*6*), *STE20* (*7*), *STE6* (*8*)) and meiosis (e.g., *DMC1* (*9*)) in *S. cerevisiae* have been identified in *C. albicans.* These genes are intact, code for well-conserved proteins, and several have been shown to complement *S. cerevisiae* mutants. The full range of the functions of these gene products in *C. albicans* has not been determined, but their presence suggests that *C. albicans* may indeed possess a little-used sexual cycle that has so far remained hidden from investigators, or that *C. albicans* may truly lack a sexual cycle, and these conserved gene products have been co-opted for entirely different purposes.

Sexual reproduction in fungi is typically controlled by genes that reside in a specified genetic locus called a mating-type, or *MAT*, locus (Fig. 3.1 panel A). Two features distinguish this locus from other portions of the genome. First, in diploid cells, the *MAT* locus is typically heterozygous; it codes for different (but usually related) genes on each of the two homologous chromosomes. Second, the genes encoded by the *MAT* locus of different fungi generally fall into three specific categories: DNA binding proteins that

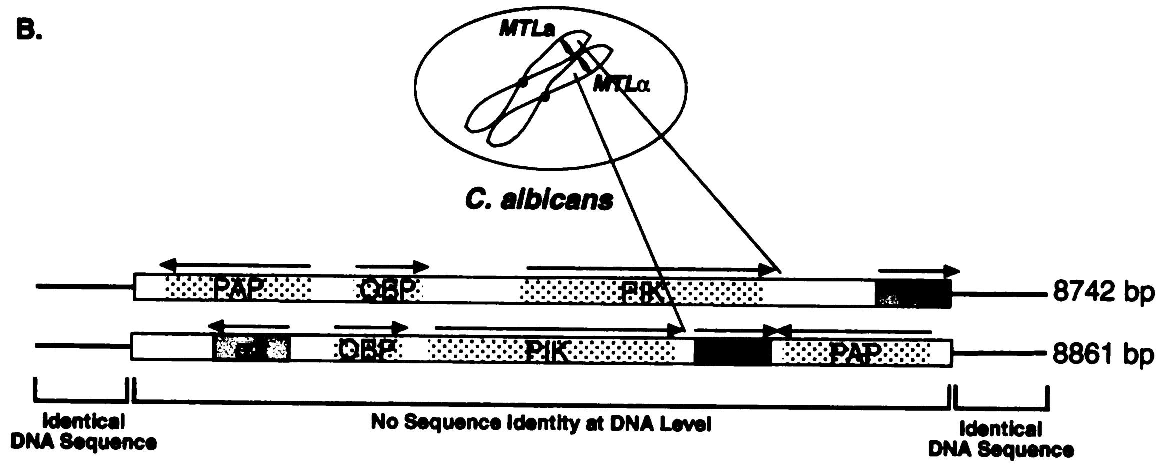
Figure 3.1. Features of the *S. cerevisiae* mating-type (*MAT*) locus and the *C. albicans* mating type-like (*MTL*) locus. Black arrows denote the direction of each gene 5' to 3'. **A.** The *S. cerevisiae MAT* locus contains open reading frames for three gene regulatory proteins, **a**1, α1, and α2, which are located on homologous chromosomes. The region of heterologous DNA sequence between the two chromosomes is 642 base pairs (bp) for the **a** chromosome and 747 bp for the α chromosome. **B**. The *C. albicans MTL* locus contains open reading frames for nine proteins from four families of proteins: three gene regulatory proteins (**a**1, α1, and α2), two phosphatidylinositol kinases, two oxysterol-binding protein-like proteins, and two poly(A) polymerases. The region of DNA sequence that differs between the *MTLa* and *MTLα* segments is 8742 bp for *MTLa* and 8861 bp for *MTLα*.

Figure 3.1

A.



B.



regulate the expression of sexual cycle genes, structural genes that code for mating pheromones, and structural genes that code for mating pheromone receptors (*10-13*).

The genes in the *MAT* locus of the sexually reproducing yeast *S. cerevisiae* have been well characterized (*11, 14*) (Fig. 3.1 panel A). They encode three transcriptional regulatory proteins that, together with other proteins encoded elsewhere in the genome, control the transcription of many target genes. The patterns of expression of these target genes give rise to three distinct types of cells (**a**, α, and **a**/α) that are responsible for the *S. cerevisiae* sexual cycle. The two types of mating cells (**a** and α) are typically haploids and carry different genetic information at the *MAT* locus: **a** cells carry *MATa*, and α cells carry *MAT*α. *MATa* codes for a single regulatory protein (the homeodomain protein **a**1), and *MAT*α codes for two proteins (the homeodomain protein α2 and the alpha domain protein α1). Mating between an **a** cell and an α cell forms the **a**/α cell, which is usually diploid, carries both *MATa* and *MAT*α, and is capable of undergoing meiosis and spore formation.

We recently described a mating-type like (*MTL*) locus in *C. albicans* (Fig. 3.1 panel B) that resembles the mating-type (*MAT*) locus of the sexually reproducing yeast *S. cerevisiae* in two important respects: first, it encodes proteins similar to the three transcriptional regulators encoded by *MAT* (**a**1, α1, and α2), and second, it is heterozygous in the diploid laboratory strain of *C. albicans* we analyzed (SC5314) with an **a**1-like gene carried on one chromosome and α1- and α2-like genes carried on the other. Other similarities in the arrangement of the genes, the positions of introns, and the function of the **a**1 protein (it is a transcriptional repressor in both organisms) further support a close connection between the *MTL* locus of *C. albicans* and the *MAT* locus of *S. cerevisiae* (*15*).

We raised the possibility that, like its counterpart in *S. cerevisiae*, the *MTL* locus could regulate a sexual cycle in *C. albicans*. If the analogy between the *MTL* locus in *C. albicans* and the *MAT* locus of *S. cerevisiae* holds, then the SC5314 laboratory strain of *C. albicans* would be an **a**/α strain and would not be expected to mate. To test the hypothesis that *C. albicans* has the inherent capacity to mate, we genetically altered an SC5314

derivative (CAI4) to create two types of "**a**" strains and two types of "α" strains. For the **a** strains, either the entire *MTLα* locus was deleted from CAI4 (to give an *MTLa/mtlα* strain) or only the α1 and α2 genes were deleted (to give an *MTLa/mtlα1mtlα2* strain). Likewise, α strains were constructed either by deleting the entire *MTLa* locus (*mtla/MTLα*) or by deleting only the **a**1 gene (*mtla1/MTLα*) (*15*, *16*). In addition, *ade2/ade2* (Ade$^-$Ura$^+$) and *ura3/ura3* (Ade$^+$Ura$^-$) derivatives of these strains were constructed to allow successful mating events to be detected by selecting for cells with the ability to grow on media lacking adenine and uracil (Ade$^+$Ura$^+$ prototrophs).

To test for mating, various strains were mixed together as outlined in Fig. 3.2. In addition to mixtures that contained putative mating pairs (an **a** strain and an α strain) we also included pairs that would not be expected to mate; for example, an *MTLa/MTLα ura3/ura3* and an *MTLa/MTLα ade2/ade2* (Test 1, Fig. 3.2). Because mammals are natural hosts for *C. albicans*, we introduced the combinations of strains into mice via tail vein injection (*17*). Pilot experiments had revealed that, following injection, the *ade2/ade2* strains persisted in mice at significantly lower levels than did the *ura3/ura3* strains (*18*). To compensate for this difference we injected three times more *ade2/ade2* cells than *ura3/ura3* cells for each test mating. Twenty-four hours after injection, the mice were sacrificed, the kidneys were removed and homogenized, and the mixture was plated on various media. For the test matings outlined in Fig. 3.2, approximately 10^3 *C. albicans* colonies per kidney formed on rich medium, conditions under which all the starting strains readily grow.

On medium lacking adenine and uracil, we recovered 44 colonies from the **a** + α mix (*MTLa/mtlα1mtlα2* + *mtla1/MTLα*) (Test 2 of Fig. 3.2), one colony from the mixture containing the **a** and α locus deletion strains (*mtlaΔ* + *mtlαΔ*) (Test 3 of Fig. 3.2), and no colonies from the control mix (Test 1 of Fig. 3.2). Additional test matings revealed similar results (Tests 4, 5, and 6 of Figure 3.2). As expected, prototrophs were recovered from Mix 4, and no prototrophs were recovered from mixes 5 and 6. The Mix 4 prototrophs were the mating products of the *mtla1* deletion Ade$^-$ and the *mtlα1mtlα2* deletion Ura$^-$ strains as

Figure 3.2. *MTL* configurations of strains used in test mating experiment. Strains containing different deletions at the *MTL* locus were mixed together and tested for mating in a mouse tail vein model. The *MTLa* and *MTLα* configurations for the different mutants are represented schematically for each of five different injection mixes labeled 1 though 6. Each *MTL* configuration has been labeled "**a**" or "α" to indicate the analogous behavior of a similarly modified diploid in *S. cerevisiae*. Mix 1: intact *MTL* Ade$^-$ + intact *MTL* Ura$^-$. Mix 2: *mtla1* deletion Ade$^-$ + *mtlα1mtlα2* deletion Ura$^-$. Mix 3: *mtlaΔ* Ade$^-$, *mtlαΔ* Ade$^-$ + *mtlαΔ* Ura$^-$, *mtlaΔ* Ade$^-$. Mix 4: intact *MTL* Ade$^-$, *mtla1* deletion Ade$^-$ + intact *MTL* Ura$^-$, *mtlα1mtlα2* deletion Ura$^-$. Mix 5: *mtlaΔ* Ade$^-$, *mtlαΔ* Ade$^-$, *mtla1* deletion Ade$^-$ + intact *MTL* Ura$^-$. Mix 6: intact *MTL* Ade$^-$ + *mtlaΔ* Ura$^-$, *mtlαΔ* Ura$^-$, *mtlα1mtlα2* deletion Ura$^-$.

Figure 3.2

Test Matings

	Ade- Ura+		Ade+ Ura-
1	"a/α"	+	"a/α"
2	"α"	+	"a"
3	"α" "a"	+	"a" "α"
4	"a/α" "α"	+ +	"a/α" "a"
5	"α" + "a" + "α"	+	"a/α"
6	"a/α"	+	"α" + "a" + "a"

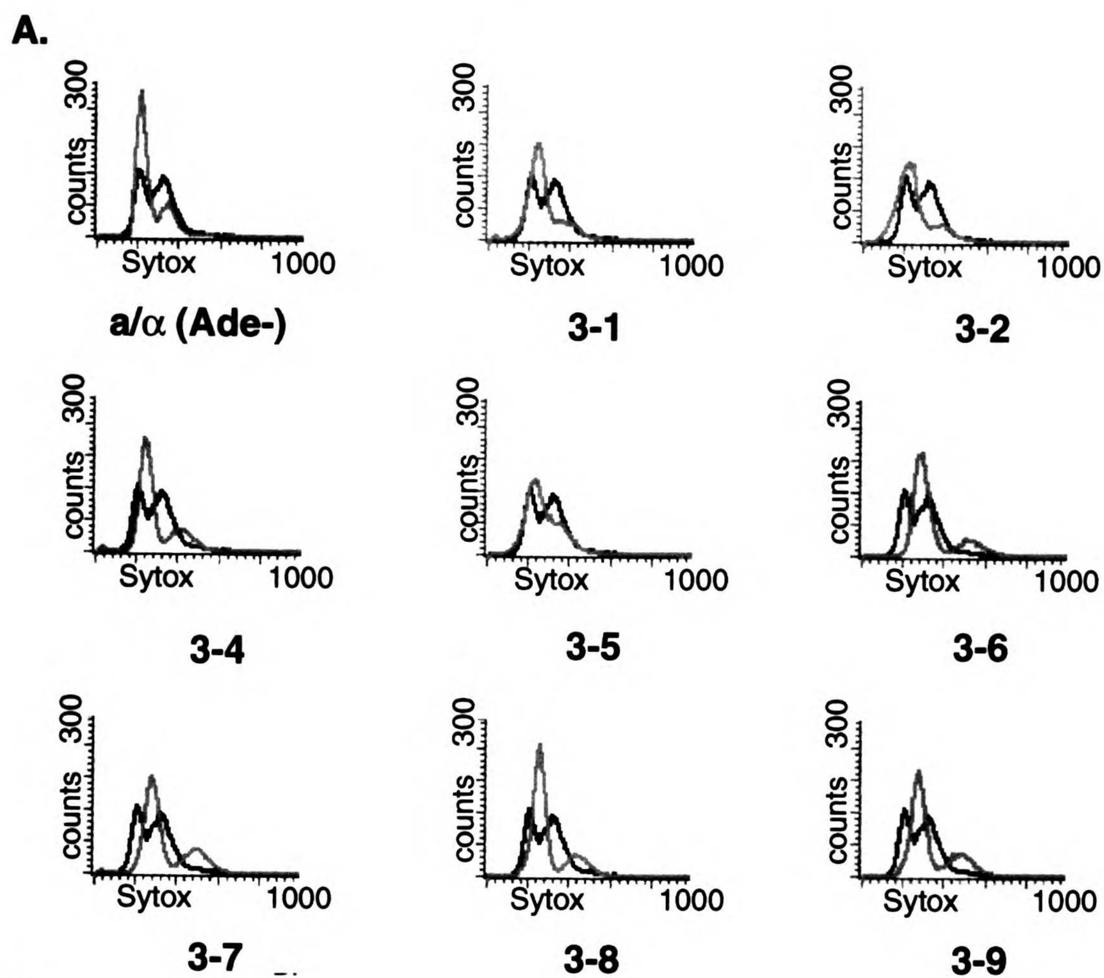
confirmed by Southern blot. Because no prototrophs from this mix resulted from fusion events between strains with intact *MTL* loci (in a mouse in which mating can occur), this result further supports the idea that the mating event is dependent on the **a** and α configurations of the *MTL*. Other control experiments showed that prototrophs did not appear when comparable titers of the starting strains were simply mixed together and incubated on various types of laboratory media and then plated to selective medium (*19*).

After confirming the Ade$^+$Ura$^+$ phenotype of the prototrophs by purifying for single colonies on selective medium, we subjected them to several tests. In a representative sample, treatment with serum elicited rapid germ tube formation, indicating that the prototrophs were indeed *C. albicans* (*18*). FACS analysis (*20*)(Fig. 3.3 panel A) suggested that cells from most of the colonies tested (12 of 15) had significantly increased DNA content compared with that of the starting strains. DAPI staining (*21*)(Fig. 3.3 panel B) showed the presence of a single nucleus in the cells of all of the prototrophs.

We next determined whether the prototrophs carried genetic markers derived from both parents. For the *MTLa/mtlα1mtlα2* x *mtla1/MTLα* cross (Test 2 of Fig. 3.2), we can readily distinguish all four alleles of *MTL*, two from each starting strain (represented in lines 1 through 4 of Fig. 3.4 panel A). By Southern analysis (*22*) we examined sixteen of the prototrophs recovered from this experiment and found that twelve strains carried all four alleles of *MTL* (*MTLa/mtla1/MTLα/mtlα1mtlα2* (Fig. 3.4 panel B, examples shown in lanes 10 through 16)) and four strains carried three alleles of *MTL* (Fig. 3.4 panel B, example shown in lane 9). These four strains represent three of the four possible *MTL* locus combinations (*MTLa/mtla/MTLα, MTLa/mtla/mtlα, MTLa/MTLα/mtlα*). Because the starting strains each carried only one version of *MTLa* and one version *MTLα* (Fig. 3.4 panel B, lanes 7 and 8), we can conclude that all the prototrophs contain genetic information from both of the parent strains, and we will hereafter refer to them as conjugants. We can rule out the possibility that "wild" strains of *Candida* could have confounded this analysis

Figure 3.3. FACS and DAPI analysis of prototrophic strains. Figure shows the results for a random selection of strains; the remaining strains gave similar results. **A.** FACS analysis of prototrophs. Strains were analyzed by fluorescence activated cell scan for DNA fluorescence after treatment. The x-axis of each graph (Sytox) represents a log scale of fluorescence, and the y-axis (Counts) represents a linear scale of cell number. In each case the control strain (CAI4) is in black, and the test strain is in red. Graph a/α (Ade$^-$) shows an overlay of an Ade$^-$ *MTLa/MTLα* strain over the Ura$^-$ *MTLa/MTLα* strain (CAI4). The *mtla1* and *mtlα1α2* starting strains gave similar profiles (*18*). Graphs 3-1, 3-2, 3-4, 3-5, 3-6, 3-7, 3-8, and 3-9 show profiles for eight of eighteen prototrophs, the majority of which (3-2, 3-5, 3-6, 3-7, 3-8, and 3-9) show FACS profiles shifted toward an increase in fluorescence in a manner that is consistent with an increase in DNA content. **B.** DAPI staining to visualize nuclei. Prototrophs were stained with 4,6-Diamidino-2-phenylindole to illuminate their nuclei and determine how many nuclei were visible in the cells. Panels a/α (Ura$^-$), *mtla1*, and *mtlα1mtlα2* show staining of strains prior to injection. Panels 3-4, 3-6, and 3-8 show the staining profile that was observed for all eighteen prototrophs; in every case the cells contain a single nucleus.

Figure 3.3



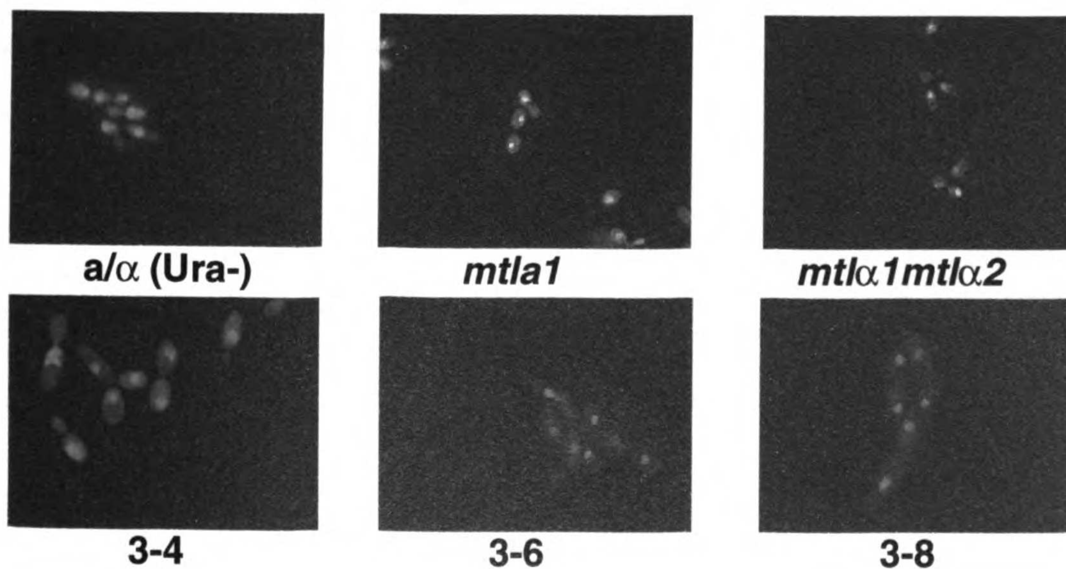
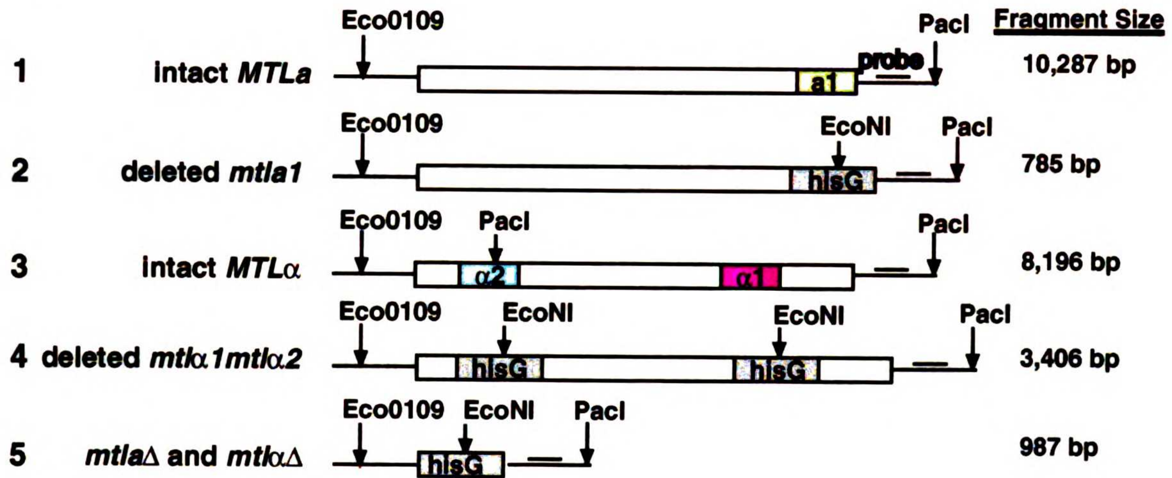


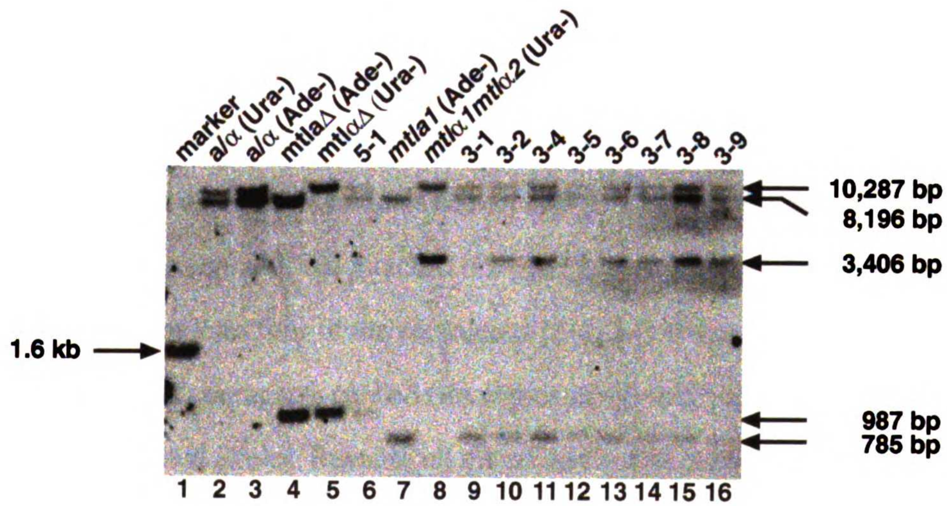
Figure 3.4. Southern analysis showing multiple copies of the *MTL* locus in conjugant strains. **A**. Schematic of *MTL* locus and various disruption alleles. From top to bottom, they are the intact *MTLa* locus, the *MTLa* locus with a complete deletion of the *MTLa1* gene, the intact *MTLα* locus, the *MTLa* locus with a complete deletion of the *MTLα1* and *MTLα2* genes, and a complete deletion of all of the *MTLa* or *MTLα* sequences. The black bar represents the location of the probe used in the Southern blot in panel B. The numbers to the right of the figures represent the sizes of the fragments that will result after digestion with the indicated endonucleases that will be detected by the probe in panel B. **B**. Southern blot confirming the presence of three or more alleles of *MTL* in the conjugant strains. *C. albicans* genomic DNA was isolated from conjugants and subjected to digestion with Eco0109, EcoNI, and PacI prior to electrophoresis. Lane 1: marker lane, Lane 2: *MTLa/MTLα ura3/ura3* auxotroph, Lane 3: *MTLa/MTLα ade2/ade2* auxotroph, Lane 4: *mtlaΔ/MTLα ade2/ade2* auxotroph, Lane 5: *MTLa/mtlαΔ ura3/ura3 2*, Lane 6: prototrophic conjugant from Mating Test 3 of Fig. 2. Lane 7: *mtla1/MTLα ade2/ade2* auxotroph, Lane 8: *MTLa/mtlα1mtlα2 ura3/ura3* auxotroph, Lane 9: prototrophic conjugant from Mating Test 2 of Fig. 2 showing the presence of 3 configurations of *MTL* Lanes 10-16: prototrophic conjugants from Mating Test 2 of Fig. 2 showing the presence of 4 configurations of *MTL*. All strains tested show fragments consistent with the sizes predicted for the various *MTL* configurations shown in panel A.

Figure 3.4

A.



B.



UCSF LIBRARY

because two of the *MTL* alleles in this experiment (one from each starting strain) were recently created for use in this experiment and do not exist outside the laboratory.

The strain recovered from the *MTL* locus deletion experiment (Test 3 of Fig. 3.2) was also analyzed and found to carry *MTLa*, *MTLα*, and at least one *mtl* deletion (represented in lines 1, 3, and 5 of Fig. 3.4 panel A; shown in Fig. 3.4 panel B lane 6). (DNA restriction fragments for the *mtla* deletion and the *mtlα* deletion are identical in size and cannot be distinguished in this experiment). Again, because none of the starting strains in this test mix carried both *MTLa* and *MTLα* (Fig. 3.4 panel B, lanes 4 and 5), we can conclude that the conjugant carried genetic information from each parent. Moreover, although there was a mixture of several different strains injected in this experiment (Test 3 of Fig. 3.2), the conjugant must have formed from one **a** parent and one α parent.

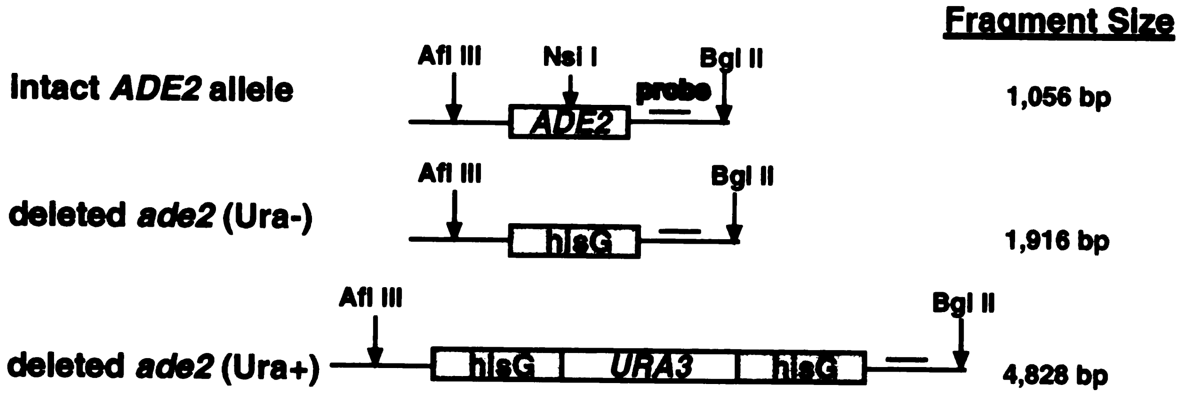
We performed the same type of marker analysis on the conjugants for the *ADE2* locus, which is located on a different chromosome from that of *MTL* (*23*). Of the four *ADE2* alleles that entered the host, three can be unambiguously identified: the two *ade2* disruptions can be distinguished from each other and from the intact *ADE2* alleles, but the two intact *ADE2* alleles cannot be distinguished from each other (Fig. 3.5 panel A). Southern analysis indicated that all of the conjugant strains carried all three of the distinguishable alleles of *ADE2* (examples shown in Fig. 3.5 panel B); moreover, the normalized ratio of the signal of the intact *ADE2* gene to that of either disruptant was approximately 2:1 in at least four of the conjugants (3-4, 3-6, 3-1, and 3-3) (*18*), consistent with the idea that these conjugants carried two copies of the intact *ADE2* gene and one copy each of the two disrupted alleles. Because all the strains in this experiment have the same type of disruption at the *ura3* locus, the analysis of this locus is not informative; the Ura^+ parental strains have *URA3* integrated at the *ade2* locus.

In summary, all the prototrophs we recovered from the "**a**" x "α" crosses contain genetic markers from both parental strains. For the *MTLa/mtlα1mtlα2* x *mtla1/MTLα* cross (Test 2, Fig. 3.2), we have determined that 12 of 16 prototrophs tested contain all four

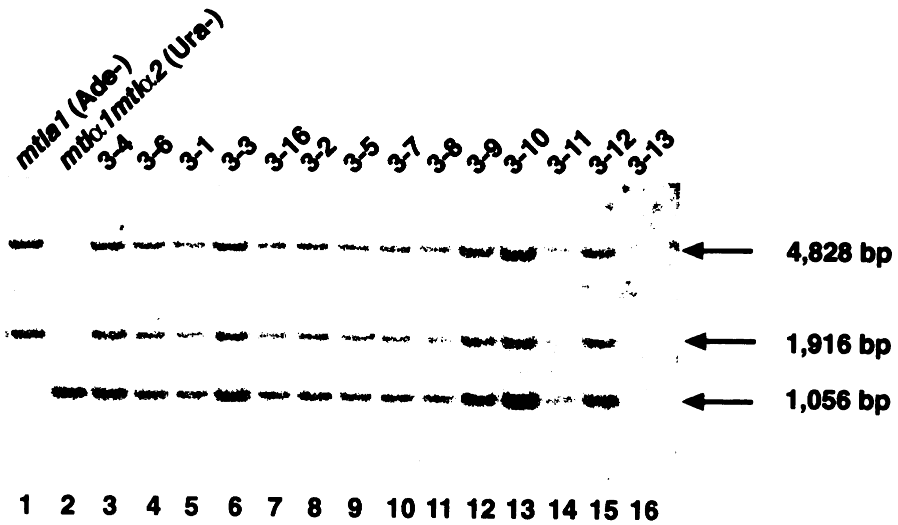
Figure 3.5. Southern analysis showing multiple copies of the *ADE2* locus in conjugant strains. **A**. Schematic of the *ADE2* locus and various disruption alleles. From top to bottom they are the intact *ADE2* locus, the *ADE2* locus deleted and replaced by a single hisG sequence, *ADE2* locus deleted and replaced by a complete hisG-*URA3*-hisG cassette. The black bar represents the location of the probe used in the Southern blot in panel B. The numbers to the right of the figures represent the sizes of the fragments that will result after digestion with the indicated endonucleases that will be detected by the probe in panel B. **B**. Southern blot confirming the presence of at least 3 alleles of *ADE2* in the conjugant strains. *C. albicans* genomic DNA was isolated from conjugants and subjected to digestion with AflIII and BglII prior to electrophoresis. Lane 1: *mtla1/MTLα ade2/ade2* auxotroph, Lane 2: *mtlaΔ/MTLα ura3/ura3* auxotroph, Lanes 3-16: prototrophic conjugants from Mating Test 2 of Fig. 2 showing the presence of at least 3 configurations of *ADE2*. All strains tested show fragments consistent with the sizes predicted for the various *ADE2* configurations shown in panel A.

Figure 3.5

A.



B.



alleles of the *MTL* locus, two derived from each parent. The simplest interpretation of this result is that the two parental strains, which are diploid, mated to form a tetraploid cell. We think it likely that the four prototrophs recovered from this cross that lacked one *MTL* allele arose through chromosome loss or homozygosis of the *MTL* locus following mating, perhaps during the passaging of strains on laboratory media required for their analysis. Although not as definitive as the analysis of the *MTL* locus, the analysis of the *ADE2* locus is consistent with the idea that the prototrophs arose from the formation of a tetraploid strain from the two diploid parents. Since *ADE2* and *MTL* reside on different chromosomes (*23*), we think it likely that most of the prototrophs are tetraploid for at least two of their chromosomes. To date, we have detected mating in five independent crosses; in all cases, mating was observed only between an "**a**" and an "α" strain. Control crosses ("**a**" x "**a**," "α" x "α," and "**a**/α" x various strains) did not produce Ade$^+$Ura$^+$ prototrophs in the same experiments.

Although we cannot rigorously conclude that conjugation occurs only when an "**a**" strain is crossed with an "α" strain, this is the only successful combination we have observed to date, and the results correlate well with this idea. This strong correlation supports the idea that the observed conjugation arose from bonafide mating and not from nonspecific fusion events, as the latter would be expected to form between any two strains, irrespective of the *MTL* configuration. We believe that our demonstration establishes an inherent ability of *C. albicans* to mate and raises the question of why this appears to happen so rarely in nature. Assuming that the *MTLa/MTLα* configuration of *MTL* (as found in SC5314) is the prevalent wild form in *C. albicans*, it is possible that mating requires "**a**" and "α" strains to arise by homozygosis of the *MTL* locus or by chromosome loss. We have found that the "**a**" and "α" derivatives of the SC5314 strain that we used in this work are significantly less virulent than the parent **a**/α strain, as judged in the mouse tail vein model (*18*). These derivatives may be lost quickly from natural populations, and without laboratory intervention, the appropriate pairs of **a** and α cells may arise in the same host

only rarely. We predict that the further development and refinement of this mating reaction will have important consequences for the molecular genetic analysis of *C. albicans*.

Our results raise the possibility that *C. albicans* has a complete sexual cycle, perhaps one in which diploids can mate to form tetraploids, and these tetraploids can undergo meiosis to produce diploids. Experiments are underway to test for meiosis in tetraploid strains of *C. albicans*. It is also possible that following mating, chromosomes could be lost from tetraploids gradually reducing them to the diploid state. The recovery of conjugants that appear trisomic for the *MTL*-containing chromosome provide some support for this notion. Finally, the fact that we readily observed mating in mice but not on a variety of laboratory media suggests the possibility that a signal or condition conducive to mating is provided by the mammalian host.

References and Notes

1. F. C. Odds, *Candida and Candidosis* (Bailliere Tindall, London, ed. 2, 1988).

2. Y. Graser, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12473 (1996).

3. M. Tibayrenc, *Trends Microbiol.* **5**, 253 (1997).

4. R. Vilgalys, Y. Graser, G. Schonion, W. Presber, T. Mitchell, *Trends Microbiol.* **5**, 254 (1997).

5. J. W. Taylor, D. M. Geiser, A. Hurt, V. Koufopanou, *Clin. Micro. Rev.* **12**, 126 (1999).

6. C. Sadhu, D. Hoekstra, M. J. McEachern, S. I. Reed, J. B. Hicks, *Molec. Cell. Biol.* **12**, 1977 (1992).

7. E. Leberer, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13217 (1996).

8. M. Raymond, *et al.*, *Molec. Micro.* **27**, 587 (1998).

9. A. C. Diener, G. R. Fink, *Genetics* **143**, 769 (1996).

10. L. A. Casselton, N. S. Olesnicky, *Microbiol. Molec. Biol. Rev.* **62**, 55 (1998).

11. I. Herskowitz, J. Rine, J. Strathern, in *Molecular and Cellular Biology of the Yeast Saccharomyces*, E. W. Jones, J. R. Pringle, J. R. Broach, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992), vol. 2, pp. 583-656.

12. J. W. Kronstad, C. Staben, *Ann. Rev. Gen.* **31**, 245 (1997).

13. B. G. Turgeon, *Ann. Rev. Phytopathol.* **36**, 115 (1998).

14. A. D. Johnson, *Curr. Opin. Genet. Dev.* **5**, 552 (1995).

15. C. M. Hull, A. D. Johnson, *Science* **285**, 1271 (1999).

16. Disruption constructs for *MTL*, *MTLα1*, and *ADE2* were created using the URA blaster construct pMB-7. The *MTL* locus deletion construct primers 5'tacgagctcggtaccggtcgtccattaggtggtttg3' and 5'ggaagatctgggaggcatgaccaaagagtc3' yielded a PCR fragment with Asp718 and BglII ends, and the primers 5'cgggatcctggaggacagaagaacacag3' and 5'aacatgcatgcaagcttcataatgatactgaggctaag3' yielded a PCR fragment with BamHI and HindIII ends. The fragments were cut and

ligated sequentially into pMB-7. The plasmid was linearized with XmnI for transformation. *MTL* locus deletion strains were confirmed by Southern blot. The *MTLα1* deletion construct primers 5'cggagctcaagcttcacagttgagtatgtggaact3' and 5'cgggtacctgcagatctgttgaagtggtatcttct3' yielded a PCR fragment with HindIII and SphI ends, and the primers 5'cgctgcaggtaccatatacagcaagttgataacatt3' and 5'cgaagcttgagctcgaaattgctcatccttatgtt3' yielded a PCR fragment with Asp718 and SacI ends. The fragments were cut and ligated sequentially into pMB-7. This plasmid was digested with HindIII and SacI to liberate a deletion fragment for transformation. The *ADE2* deletion construct primers 5'tgcacgcaagcttggtaccctcttgggtcacgtgatggt3' and 5'ctaggaagatctctgtaactgcaacatcgtg3' yielded a PCR fragment with Asp718, and BglII ends, and the primers 5'aactgcagggatcctggtgcaggtggtgctgctcat3' and 5'tgcacgcaagctttctagagttagagatgtggtggattgg3' yielded a PCR fragment with PstI and HindIII ends. The fragments were cut and ligated into pMB-7 sequentially. The plasmid was linearized with Xmn I for transformation. Putative *ade2/ade2* transformants were screened by Southern blot and red colony phenotype. *mtlα1mtlα2* strains were constructed by transforming a counter-selected (Ura^-) *mtlα1* strain with an *MTLα2* deletion construct. Ura^-Ade^+ "**a**" and "α" strains were created through counter selection to eliminate the *URA3* marker gene (from the last disruption) by growing on yeast synthetic plates containing 0.08% 5-fluoroorotic acid (5-FOA) and 0.025 mg/mL uridine. Counter-selectants were confirmed by absence of growth on synthetic plates without uracil. Ade^-Ura^+ "**a**" and "α" strains were created by deleting both copies of *ADE2* in the Ura^- strains and allowing the cells to retain *URA3* after the second round of transformation. Because *C. albicans* does not contain silent cassettes of *MTL* (*15*), the various *MTL* deletions would be expected to be sufficient to create strains analogous to **a** and α.

17. Mice used in this experiment were Balb-C. Two mice for each test mating combination were injected with 2.5×10^7 yeast cells into the lateral tail vain. Cells for injection were grown to saturation at 30°C and then back-diluted and grown for 4 hours to mid log-phase. The cells were spun down and resuspended in 3 ml 0.9% saline, and counted using a hemacytometer. To compensate for the *ade2/ade2* strain growth defect in mice, each mouse was injected with a total of 1.875×10^7 cells of the Ade$^-$ strain(s) and 0.625×10^7 cells of the Ura$^-$ strain(s). Mice were sacrificed after 24 hours, and the kidneys were removed, homogenized, and plated in fractions onto standard SD -Ade-Ura plates and incubated at 30°C. A fraction of each sample was also plated to YPD for total colony counts of recovered cells.
18. C. M. Hull, R. M. Raisner, A. D. Johnson, unpublished data.
19. Supplemental information on the in vitro conditions tested is available on *Science* Online at www.sciencemag.org/feature/data/ (see below)
20. Supplemental information on the FACS analysis protocol is available on *Science* Online at www.sciencemag.org/feature/data/ (see below)
21. J. E. Wagstaff, S. Klapholz, R. E. Esposito, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2986 (1982).
22. Supplemental information on the isolation of genomic DNA and Southern blot protocol is available on *Science* Online at www.sciencemag.org/feature/data/ (see below). The Southern blot probe for the *MTL* locus was a 179 bp PCR product of the primers 5'tcccaaaatgaaacaagcaattaaa3' and 5'gccacaataactccactg3'. The probe for the *ADE2* locus was a 476 bp PCR product of the primers 5'aactgcagggatcctggtgcaggtggtgctgctcat3' and 5'tgcacgcaagctttctagagttagagatgtggtggattgg3'.
23. Chromosome information was obtained from http://alces.med.umn.edu/Candida.html. Specific assignments were made by the Stanford *Candida albicans* sequencing project. Sequence data for *C. albicans* was obtained from the Stanford DNA

Sequencing and Technology Center web site at http://www-sequence.stanford.edu/group/candida.

24. The authors wish to acknowledge V. Nguyen and additional members of the J. Li laboratory (UCSF) for assistance with the FACS analysis, A. Uhl and I. Herskowitz for helpful comments on the manuscript, and members of the Johnson Laboratory for continuing support and assistance. Sequencing of *Candida albicans* by the Stanford DNA Sequencing and Technology Center was accomplished with the support of the NIDR and the Burroughs Wellcome Fund. This work was supported by NIH grant GM37049 to A.D.J and a Burroughs Wellcome Merit Award to A.D.J.

Supplemental References and Notes

19. Strains were tested for the ability to mate under several in vitro conditions. Over 1×10^8 cells of the test strains were mixed and allowed to incubate from 2 to 7 days at both 30°C and 37°C on the following solid media before plating to selective media for single colonies: YPD, 10% serum, Spider, corn meal, water agar, sporulation medium (low glucose), SLAD agar, sheep's blood agar, lysed red blood cell agar. Test strains were also mixed and incubated in mouse kidney extracts before transfer to selective media plates. No colonies were recovered from any of the mixes after selection. A series of experiments to test for the production of or response to pheromone was carried out with the test strains and no positive results were obtained.

20. FACS samples were by prepared fixing overnight cultures of cells grown in yeast peptone dextrose (YPD) medium in 70% ethanol overnight at 4°C. The cells were washed and resuspended in 1 ml 50mM Tris-HCl pH 7.5 and sonicated on low power. They were then resuspended in 0.5 mL 2 mg/mL RNaseA solution and

incubated at 37°C for 2 hours, centrifuged and resuspended in 0.2 mL 5 mg/mL pepsin and incubated for 45 minutes at 37°C, washed in 0.5 mL 50 mM Tris-HCl pH 7.5, and sonicated on low power. Samples were stained in a solution of 1μM Sytox Green dye for 1 hour at 4°C and then analyzed by fluorescence activated cell sorter. A total of 20,000 cells were analyzed for each strain.

22. Genomic DNA was isolated according to C. S. Hoffman, F. Winston, *Gene* **57**, 267 (1993) with the following exceptions: glass beads were not acid-washed, and cells were agitated on Eppendorf Mixer for 20 minutes at 4°C rather than vortexed, and RNaseA step was omitted. DNA was resuspended in 50 μL dH_2O and 10 μg RNaseA. In each case 10 μL DNA were digested with excess enzyme (approximately 80 units NsiI) in 100 μL restriction digest reactions overnight. DNA was then precipitated and run on a 1% agarose gel which was then treated for transfer according to *Current Protocols in Molecular Biology*, Ausubel *et. al.*, Eds, (Wiley-Interscience, New York, 1992), vol. 1, pp. 2.9.7-2.9.8. DNA was transferred to Hybond N+ (Amersham Product #RPN303B) and probed and detected according to the Amersham Alk-Phos Direct Kit (#RPN3690).

Appendix to Chapter 3

Appendix F

Testing for mating in vitro.

The effectively **a** and α strains of *C. albicans* were tested for the ability to mate under several in vitro conditions prior to testing in animals. Over 1×10^8 cells of each of the test strains were mixed and allowed to incubate from 2 to 7 days at both 30°C and 37°C on the following solid media before plating to selective media for single colonies: Yeast Peptone Dextrose, 10% serum, Spider, corn meal, water agar, sporulation medium (low glucose), low nitrogen agar, sheep's blood agar, lysed red blood cell agar. Test strains were also mixed and incubated in mouse kidney extracts before transfer to selective media plates. No prototrophic colonies were recovered from any of the mixes after selection. In vitro mating tests were also performed with the *C. albicans* test strains mixed with *S. cerevisiae* strains (CHY222 and CHY223) with no positive results.

A series of experiments to test for the production of or response to pheromone was also carried out with the test strains, and no positive results were obtained. Halo and pheromone challenge tests were carried out according to Fink and Guthrie (Methods in Enzymology Volume 194) with all combinations of the *C. albicans* strains tested against each other and with *S. cerevisiae* strains. Standard *S. cerevisiae* halo tester strains were used for the halo tests (AJY75 and AJY77).

Chapter 4

Transcriptional repression of endogenous targets by a1-α2 in *Candida albicans*

Abstract

Candida albicans is a human fungal pathogen that is responsible for the majority of mucosal and systemic fungal infections in humans. It has traditionally been classified as an asexual organism without the capacity to mate; however, it has recently been shown to contain a mating type-like (*MTL*) locus, whose configuration is important for *C. albicans* mating in a mammalian host. Two of the predicted proteins in the locus are homologous to the *S. cerevisiae* **a**1 and α2 proteins, which are responsible for maintaining the diploid state in *S. cerevisiae* by repressing a specific class of genes (haploid-specific genes). It has been shown previously that the **a**1 and α2 homologs in *C. albicans* are capable of repressing transcription through a specific DNA binding site in a heterologous reporter construct, but their ability to repress biologically relevant targets had not been established. In this work I show that the **a**1-α2 complex of *C. albicans* represses transcription of endogenous genes in *C. albicans* and that these genes are homologous to haploid-specific genes in *S. cerevisiae*. This finding demonstrates the conservation of a complex regulatory circuit between the two yeasts and adds to mounting evidence which suggests that *C. albicans* may engage in a complete sexual cycle.

Introduction

In *S. cerevisiae* the regulators **a**1, α1, and α2 are required to maintain its three cell types. In particular, repression mediated by a complex of the **a**1 and α2 proteins in *S. cerevisiae* is essential for distinguishing the **a**/α diploid from the **a** and α haploids (11, 14). This distinction is important because it prevents diploids from mating and makes them competent to undergo meiosis and sporulation, thus completing the sexual cycle. The establishment of this distinction between haploids and diploids is carried out by a heterodimeric complex composed of **a**1-α2 that binds to DNA and directs the repression of specific set of target genes, the haploid-specific genes (hsg) (9). These genes are expressed in **a** and α haploids but are repressed in diploid cells. In *S. cerevisiae* the hsg are defined by the presence of an **a**1-α2 binding site in their promoters that allows **a**1-α2 to bind and recruit a global repressor complex to turn off the target gene (16). The products of these genes are generally involved in processes in the haploid that are important for mating, and in fact many components of the mating pheromone response pathway are targets of **a**1-α2 repression (11).

The discovery of the mating type-like (*MTL*) locus in *C. albicans* revealed the presence of **a**1 and α2, homologs which were subsequently shown to repress transcription via specific DNA binding sites in heterologous reporter constructs (13). Although this established that the activity was detectable in vivo, it did not test whether the activity was biologically relevant. In order to understand the biological role of **a**1-α2-mediated repression in *C. albicans,* I wanted to identify potential targets of the repressor complex and determine whether or not they were repressed by **a**1-α2 in vivo.

I took two approaches to identify target genes. First, with the assistance of Aaron Straight, an analysis of the available genome sequence of *C. albicans* in April 1999 was carried out (18), which revealed the presence of over 4000 potential **a**1-α2 binding sites in the *C. albicans* genome. This analysis identified several genes with known **a**1-α2 sites and

suggested several others that were possible targets for repression; however, this list did not overlap well with the known *S. cerevisiae* haploid-specific genes and suggested that there could be substantial differences between the two organisms in either the consensus binding site for **a**1-α2 (8) or the pattern of **a**1-α2 gene regulation (or both).

One way to begin to address this question was to take a second approach and test sequence homologs of known **a**1-α2 regulated genes from *S. cerevisiae* for their expression patterns in *C. albicans*. The sequence of the *C. albicans* genome is nearly complete, and sequence comparisons between *S. cerevisiae* and *C. albicans* revealed many homologs of haploid-specific genes (thanks to Andrew Uhl for making many of the comparisons). A list of 14 **a**1-α2 regulated genes from *S. cerevisiae* was compiled (Table 4.1) (23), and several of these genes were tested using chromatin immunoprecipitations (IPs) for the binding of **a**1 to their promoters. The cross-linking of **a**1 to the promoter of a gene was considered evidence that **a**1 was binding to the site in vivo. Subsequent transcript analysis of some of the genes revealed that genes with **a**1 bound to the promoter region were regulated in an a1-α2-specific manner.

Table 4.1. Summary of homologs of haploid-specific and **a**-specific genes in *C. albicans*

haploid-specific genes in *S. cerevisiae*	function in *S. cerevisiae*	homolog in *C. albicans*?
STE3	**a** factor pheromone receptor	*STE3*
GPA1	trimeric G protein subunit Gα	*CAG1*
STE4	trimeric G protein subunit Gß	*STE4*
STE5	MAP kinase cascade scaffold protein	no*
STE18	trimeric G protein subunit Gγ	no
STE12	transcriptional activator	*CPH1*
α1	transcriptional activator of α-specific genes	*MTLα1*
AXL1	**a** factor processing protease	*AXL1*
HO	*MAT* locus endonuclease	no
SST2	involved in desensitization to α factor	*SST2*
RME1	transcriptional repressor of meiosis	*RME1*
FUS3	MAP kinase in mating pathway	*FUS3*
STA1	glucoamylase starch degrading enzyme	*STA1-1, -2, -3, -4, -5*
SGA1	sporulation-specific glucoamylase	*SGA1*

a-specific genes in *S. cerevisiae*	function in *S. cerevisiae*	homolog in *C. albicans*?
BAR1	**a** factor protease	no
MFA1	**a** factor structural gene	no
MFA2	**a** factor structural gene	no
STE2	α factor pheromone receptor	*STE2*
STE6	**a** factor transporter	*HST6*
AGA2	**a**-agglutinin binding	no
ASG7	unknown	no

* "No" in this category indicates that no homolog was apparent after 7 fold sequencing coverage of the *C. albicans* genome.

Materials and Methods

Identification of a1-α2 binding sites in the *C. albicans* genome. **a**1-α2 sites were identified in the Contig-2 compilation of the *C. albicans* genome using an algorithm written and executed by Aaron Straight. The algorithm was based on the consensus binding site for *S. cerevisiae* **a**1-α2, 5' TGATGTAATTAATTACATGA 3' (8), and the search was carried out with the site oriented in both directions (5' to 3' on both strands). The search did not distinguish promoter regions from other regions in the genome.

Chromatin immunoprecipitations on *C. albicans* wild-type and a1 deletion strains. Chromatin IPs were carried out according to Straight et al. (21) with minor changes described here. Chromatin was prepared from wild-type (CHY340) and **a**1 deletion (CHY247) strains. For each of 5 antibodies, 5 μL of antibody was mixed with 500 μL of lysate in the immunoprecipitations. Five different antibodies were used for each strain; data are shown for only two: *S. cerevisiae* α2 antibody bleed #2 from rabbit #2 (shown), *S. cerevisiae* **a**1 antibody bleed #2 from rabbit #2, *C. albicans* **a**1 peptide antibody #1 (83608) (shown), *C. albicans* **a**1 peptide antibody #2 (83607), *C. albicans* **a**1 peptide antibody #3 (83610). PCR reactions were carried out with primer sets to the promoter regions of the potential target genes and control genes. All primer sets were tested on chromatin fragments prior to immunoprecipitation to ensure the presence of equal amounts of chromatin in each reaction and the efficiency of the primers.

Northern analysis on RNA from wild-type and a1 deletion strains. Genes whose promoters appeared to be bound by **a**1 in the chromatin IPs were tested for regulation of their transcripts by Northern blotting in wild-type and **a**1 deletion strains. Total RNA was isolated from the same strains used in the chromatin IPs according to Current Protocols in Molecular Biology (2), Preparation of Yeast RNA by Extraction with Hot Phenol. RNA

samples were quantitated and electrophoresed on a formaldehyde gel, transferred to a nylon membrane (GeneScreen from NEN Research Products #NEF-972) and probed with random-primed ^{32}P PCR generated products that were predicted to hybridize to the open reading frames of the test genes. Gel-purified PCR products were labeled with ^{32}P in random prime reactions using Amersham Megaprime DNA Labeling Kit (#RPN1605). Blots were hybridized in 0.5 M sodium phosphate pH 7.2, 1 mM EDTA pH 8.0, 7% SDS at 65°C overnight and washed with 2X SSC + 0.1% SDS twice for 10 minutes at 65°C and once with 0.5X SSC+0.1% SDS for 10 minutes at 65°C. Washed blots were exposed to Kodak XAR film overnight.

Results

A computer-generated list of potential **a**1-α2 sites in the *C. albicans* genome showed the presence of several known genes in the highest scoring hits. Sites within 500 bp upstream of open reading frames were considered potential regulatory sites, and eight of the higher scoring sites (located upstream of genes) and genes that they may regulate are shown in Table 4.2. The genes fell into three classes. The first class (I) was defined to contain genes with homologs in *S. cerevisiae* that are known **a**1-α2 regulated genes (haploid-specific genes). In this analysis, this class contained a single gene, *CAG1*, which is a homolog of the *S. cerevisiae* haploid-specific gene *GPA1*, the Gα subunit of the trimeric G-protein complex involved in *S. cerevisiae* mating (4). *CAG1* was already known to contain an hsg regulatory site in its promoter (19). The second class (II) is made up of genes that are sequence homologs of genes involved in the *S. cerevisiae* sexual cycle that are not **a**1-α2 regulated in *S. cerevisiae* but contain **a**1-α2 sites in *C. albicans*. This class contains two genes, *HST6* and *STE2*. In *S. cerevisiae,* the *HST6* homolog is *STE6,* a transporter of **a**-factor (17), and *STE2* is the receptor expressed in **a** cells that is responsible for sensing α factor (3). Both of these genes in *S. cerevisiae* are regulated in a cell-type specific manner (both are **a**-specific genes (10, 22)) but neither one is regulated by **a**1-α2 in *S. cerevisiae*. Like *CAG1*, *HST6* had previously been found to contain an **a**1-α2 site in its promoter (19). The third class (III) contains genes that were identified as having **a**1-α2 binding sites in *C. albicans* but do not have these sites in *S. cerevisiae* and are not known to be important in the sexual cycle or other related process. These genes were not included in subsequent analyses. Surprisingly, this sequence analysis identified only one gene whose homolog is known to be regulated by **a**1-α2 in *S. cerevisiae* (class I). If the regulation of haploid-specific genes is the same, or at least very similar, between *S. cerevisiae* and *C. albicans*, many genes in the first class would have been identified. This disparity between the presence of the **a**1-α2

Table 4.2 Representatives of genes from computer analysis identified as having **a**1-α2 sites in their promoters

binding site sequence	potentially regulated gene	class
TGATGTAAAAATTAACATGG	***CAG1***	I
AGATGTACTNTGTAACATGA	***HST6***	II
GTATGTTAAAAATCACATCT	***STE2***	II
TGATGTTGATGTTTACTTGA	*MSN5*	III
TGATGTAATTTATGACATGC	*YDR132C*	III
TCATGTCTACATCTGCATCA	*GTR1*	III
TCATGGTAAATATTACATCA	*RRP45*	III
AAATGTAAATAATTTCATCA	*GAS1*	III

consensus sequence

TGATGTAATTAATTACATGA

regulatory site and the presence of the apparent haploid-specific gene homologs suggested that **a**1-α2 regulation may be working to control very different genes in *C. albicans* or that the binding site for **a**1-α2 has diverged significantly from the *S. cerevisiae* consensus sequence. A change from the consensus, particularly in the spacing of the **a**1 and α2 half sites would likely result in the exclusion of regulated genes from this analysis (8). In order to determine whether homologs of **a**1-α2 regulated genes from *S. cerevisiae* were indeed regulated in *C. albicans*, I took a complementary approach and used chromatin immunoprecipitations (Figure 4.1) to identify homologs whose promoters were bound directly by the **a**1 protein.

Chromatin immunoprecipitations were performed on two strains of *C. albicans* (wild-type and **a**1 deletion) with two antibodies each (*S. cerevisiae* anti-α2, *C. albicans* anti-**a**1). After cross-linked chromatin lysates were subjected to immunoprecipitation, the cross links were reversed, and recovered target sequences were subjected to polymerase chain reaction (PCR) with primers to the promoters of eleven potential target genes and two control genes not expected to be bound by **a**1. As schematized in Figure 4.2, a positive signal represents the presence of fragments in the PCR reaction that were bound by **a**1. Reactions that result in a positive signal in the wild-type strain and no signal for the **a**1 deletion strain represent fragments that are bound directly by **a**1. In the presence of **a**1, the fragments are precipitated by the antibody; in the absence of **a**1, the fragments are not precipitated by antibody. (Genes that are bound by **a**1-α2 would be expected to give the same pattern with the α2 antibody.) Figure 4.3 shows the results of PCR for each gene. As summarized in Table 4.3, all of the potential target genes tested appear to have **a**1 bound to their promoters, while the control gene *ADE2* does not appear to be subject to **a**1 binding. Surprisingly, the *ACT1* control appeared to have **a**1 bound to its promoter.

Thus far, three of the genes that gave positive results in the chromatin precipitations have been analyzed by Northern blot. RNA from the same strains used in the chromatin IPs was prepared and subjected to Northern analysis using probes to the open reading frames

Figure 4.1. Schematic representation of chromatin immunoprecipitation procedure. Cultures of test strains are grown to late log phase and then treated with formaldehyde to cross-link proteins to DNA and to each other. The cells are then lysed and the resulting extract is sonicated to shear the DNA to an average size of approximately 500 bp. Chromatin lysates are then immunoprecipitated with antibodies to the proteins of interest. After immunoprecipitation, the cross links are reversed by heating, the DNA is purified, and fragments of interest are detected by polymerase chain reaction.

Figure 4.1

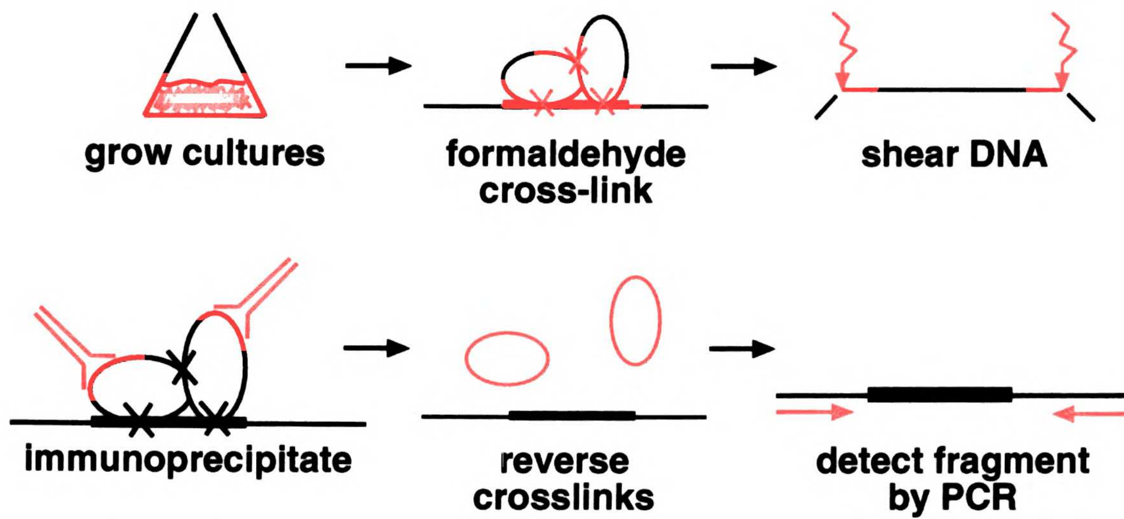


Figure 4.2. Schematic representation of expected results for an **a**1-α2 regulated gene in a chromatin immunoprecipitation experiment. Lanes 1 and 2 show PCR products derived from a wild-type strain in which the primers are to an **a**1-α2 regulated gene. In this case, the products are detected because the antibodies in the lower left panel precipitated the proteins **a**1 and α2 and the DNA to which they were crosslinked. Lanes 3 and 4 show no PCR products because in an **a**1 deletion strain, there is no protein bound to the DNA (and no crosslinking to the DNA), so the antibodies do not precipitate any DNA fragments as indicated in the lower right panel.

Figure 4.2

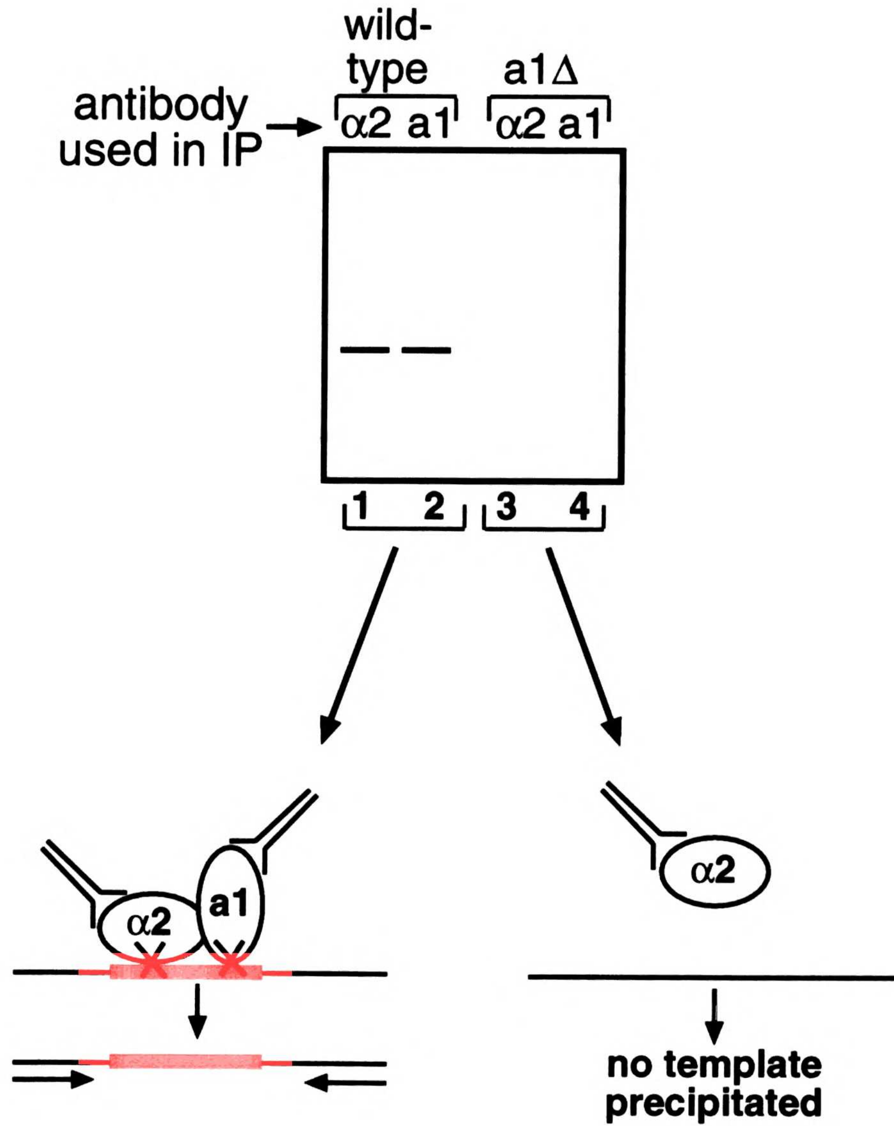


Figure 4.3. Results of PCR analysis of chromatin IPs. Each panel shows the results of PCR analysis of precipitated DNAs. Fragments were electrophoresed on agarose gels. Each is labeled under the panel with the name of the gene whose promoter region was tested by PCR. Each test is 4 lanes: *S. cerevisiae* α2 antibody and *C. albicans* **a**1 antibody each used to precipitate chromatin from wild-type and **a**1 deletion strains. Lanes 1, 10, 15, 28, 29, 42, and 51 are marker lanes. Lanes 2-5 show the results for the control gene *ADE2*. Lanes 6-9 show the PCR signal with the *ADE2* primers on chromatin prior to immunoprecipitation. Lanes 11-14 show the results for *ACT1*. Lanes 16-19 show the results for *CAG1*. Lanes 20-23 show the results for *RME1*. Lanes 24-27 show the results for *STE3*. Lanes 30-33 show the results for *STA1-1*. Lanes 34-37 show the results for *STA1-2*. Lanes 38-41 show the results for *CPH1*. Lanes 43-46 show the results for *HST6*. Lanes 47-50 show the results for *STE2*. Lanes 52-55 show the results for *MTLα1*. Lanes 56-59 show the results for *ALS1-2*. Lanes 60-63 show the results for *FUS3*. Although the wild-type signals for *ALS1-2* and *FUS3* are difficult to discern in this figure, they are visible in the original photo.

Figure 4.3

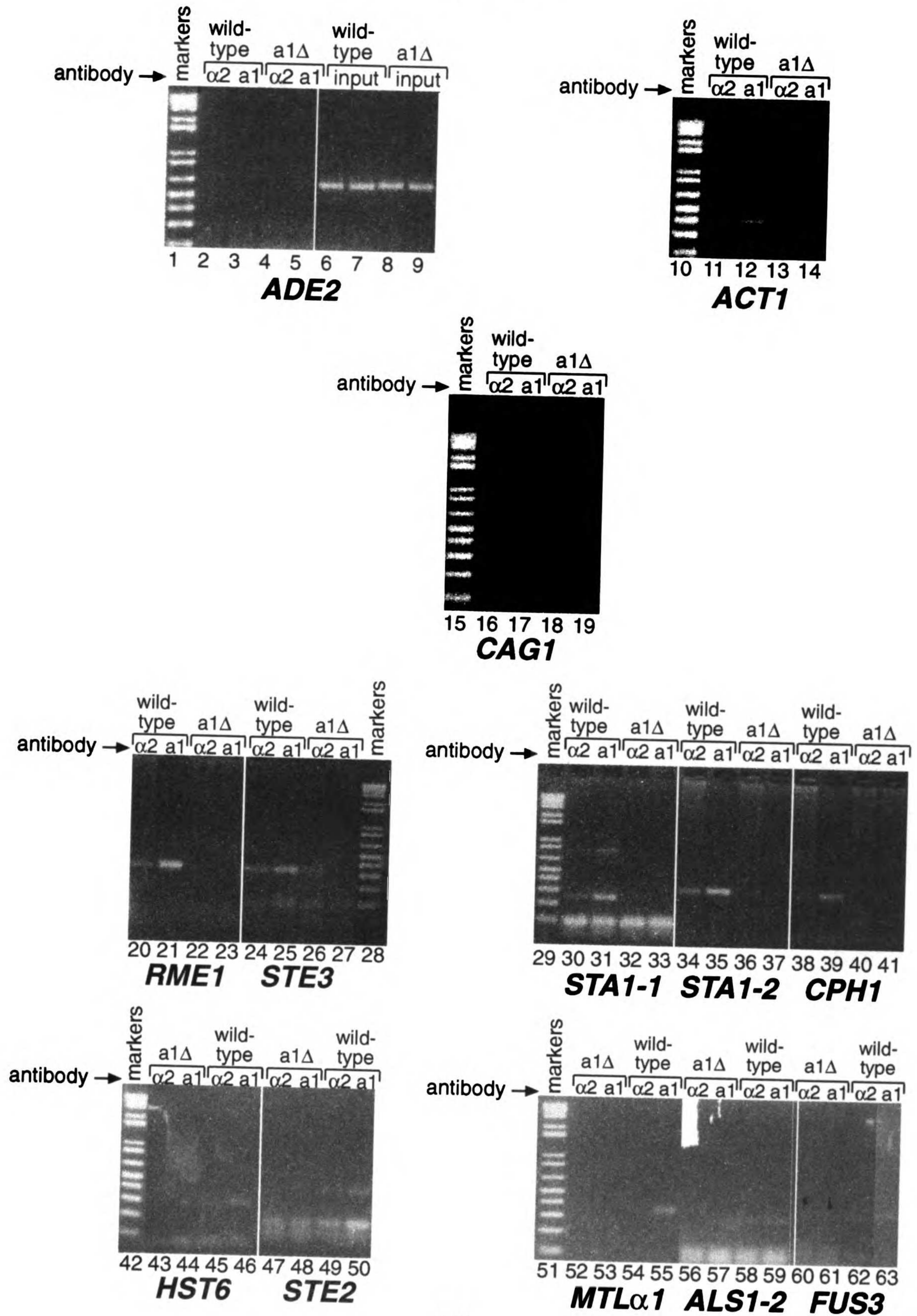


Table 4.3 Summary of results of chromatin IPs and Northern analysis

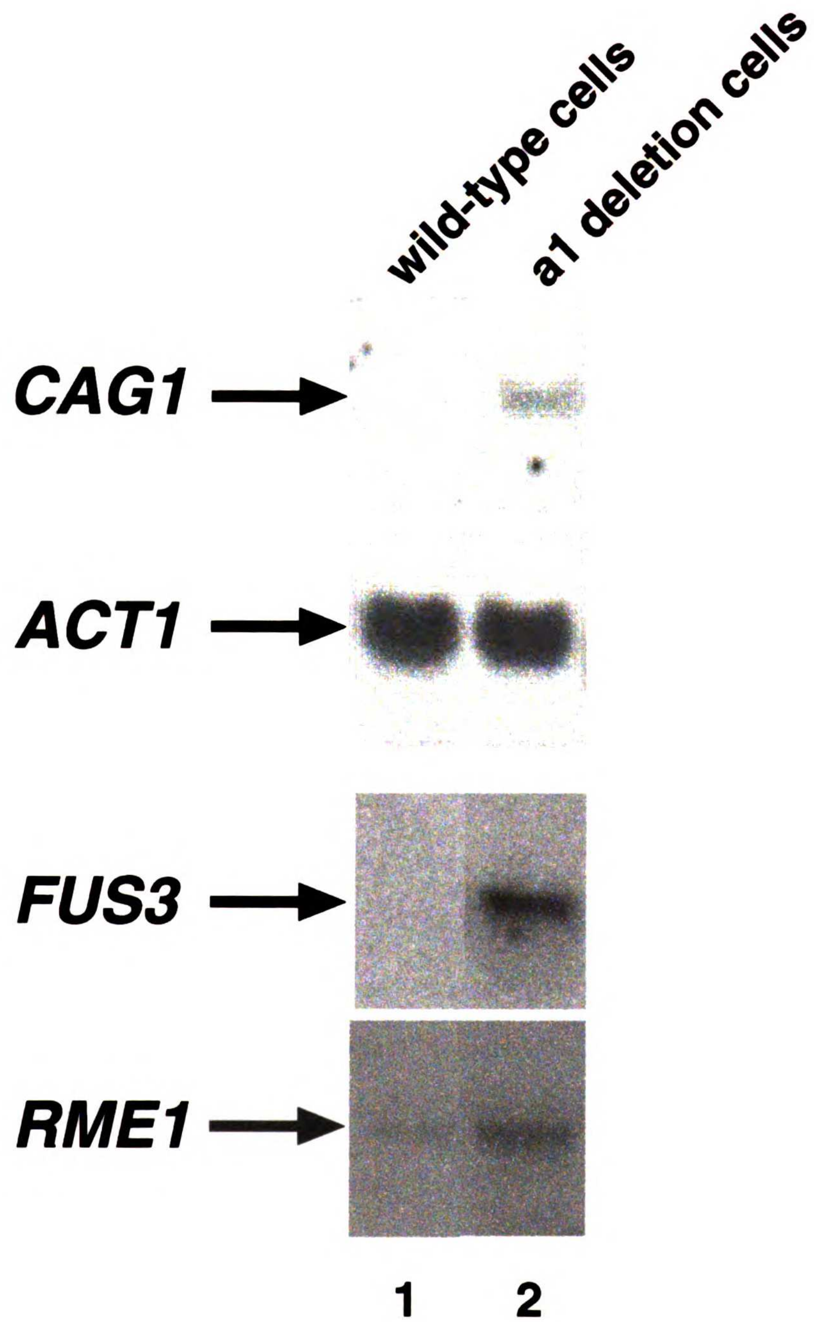
genes tested	bound by **a**1?	regulated?	identified in computer analysis?
STE3	yes	not tested	no
CAG1	yes	yes	yes
STE4	not tested	not tested	no
CPH1	yes	no tested	no
MTLα1	yes	not tested*	no
AXL1	not tested	not tested	no
SST2	not tested	not tested	no
RME1	yes	yes	no
FUS3	yes	yes	no
STA1-1	yes	not tested	no
STA1-2	yes	not tested	no
SGA1	not tested	not tested	no
ALS1-2	yes	not tested	no
STE2	yes	not tested	yes
HST6	yes	not tested	yes
ACT1	yes	no	no
ADE2	no	not tested	no

*Although *MTLα1* was not evaluated by Northern blot, reverse-transcription PCR reactions have consistently shown partial regulation of this transcript by **a**1-α2 (12) consistent with *α1* regulation in *S. cerevisiae.*

of the genes of interest. Figure 4.4 shows the regulation of *CAG1*, *RME1*, and *FUS3*. Even though *ACT1* appeared to be bound by **a**1 in the chromatin IPs, it does not appear to be subject to transcriptional regulation by **a**1. The absence of any regulation of *ADE2* by **a**1 has not been confirmed by Northern blot. The results for *ACT1* indicate that it is possible to obtain positive results by chromatin IP that suggest **a**1 binding and not see evidence for regulation by Northern blot. Further control genes will have to be tested to ensure that this is not a global problem that will impair the specificity of the experiment. Work by others (Hien Tran, personal communication) has resulted in this same pattern for *ACT1* with other antibodies. Perhaps there is something unusual about the *ACT1* promoter that leads to false positive signals by chromatin IP. Based on the *ADE2* result, it appears that the chromatin IPs are likely to be specific. Testing additional control genes and conducting further Northern analysis on target genes and control genes will help to confirm the results presented here.

Figure 4.4. Northern analysis of potential target genes. RNA from wild-type and **a**1 deletion cells was evaluated by Northern blot. Lane 1 contains RNA from wild-type cells. Lane 2 contains RNA from an **a**1 deletion strain. Transcripts of potentially regulated genes were detected with ^{32}P labeled probes predicted to hybridize to the open reading frames of the genes. The transcripts correspond to the gene named to the left of the panels. Note: *FUS3* and *RME1* were not on the same blot as *CAG1* and *ACT1*. Also, the derepression of *FUS3* in **a**1 deletion cells is enhanced by growth in Lee's media, which is the comparison shown here (although the derepression clearly occurs under YPD growth conditions as well).

Figure 4.4



Discussion

In this chapter I present preliminary data to suggest that **a**1-α2 mediated repression occurs on endogenous targets in *C. albicans* and that three of the targets genes tested (*CAG1, FUS3,* and *RME1*) are repressed by **a**1-α2 and have **a**1 bound to their promoters. This work also shows that the promoters of several additional *C. albicans* genes appear to be bound directly by **a**1, but the regulation of these genes by **a**1-α2 has not yet been tested. Two of these genes, *HST6* and *STE2,* are additionally interesting because they have clear **a**1-α2 sites in their promoters even though they are **a**-specific genes in *S. cerevisiae* (10, 22) (genes regulated by a complex composed of Mcm1 and α2 through a specific DNA binding site that is different than the **a**1-α2 binding site) (15). Most of the **a**-specific genes do not have apparent sequence homologs in *C. albicans*, and the two that do (*HST6* and *STE2*) have **a**1-α2 repressor sites in their promoters. The data presented here indicate that **a**1 binds directly to the promoters of these two genes. This apparent difference in regulation suggests that cell types could be defined differently in *C. albicans* than they are in *S. cerevisiae.* Defining the **a**/α cell through **a**1-α2-mediated repression, however, looks so far to be very similar between the two organisms.

Two of the regulated genes, *CAG1* and *FUS3,* are clearly regulated at the level of transcription and suggest that **a**1-α2 has clear repression targets in vivo. These genes are the first targets of the repression activity that have been identified in *C. albicans* and mimic the regulation of their *S. cerevisiae* homologs well. In *S. cerevisiae,* the homologs of *CAG1* and *FUS3* are both components of the pheromone response pathway (4, 7) required for mating in **a** and α haploids. In diploid *S. cerevisiae,* both *GPA1* and *FUS3* are repressed by **a**1-α2 (1, 6). This repression is essential because among other things, it shuts down the pheromone response pathway and indicates to cells that they are diploid. It appears from the conservation of this regulation for *CAG1* and *C. albicans FUS3* that the repression of these genes is also important for the *C. albicans* diploid cell. *RME1* is also subject to **a**1-α2

regulation, but the repression appears to be only partial because some amount of transcript is detectable in wild-type cells. This is the same pattern observed for *RME1* in *S. cerevisiae* (5). In each of these cases the pattern of repression is the same between *C. albicans* and *S. cerevisiae*, but the sites for repression in the promoters of *FUS3* and *RME1* must clearly be quite different in *C. albicans*. The absence of obvious **a**1-α2 binding sites in the promoters of these genes suggests that repression by this complex can proceed via a binding site that is quite different from the *S. cerevisiae* site. Clearly, **a**1-α2 repression can occur via the *S. cerevisiae* site as seen for *CAG1*, but the *FUS3* and *RME1* results suggest that it can occur through a different sequence as well.

Because these experiments were carried out in the **a**1 deletion strain only, one might propose that the repression could be very different in *C. albicans* and depend only on **a**1. That is, in the absence of α2, **a**1 could bind to DNA and direct repression on its own. Two pieces of evidence argue against this possibility. First, the chromatin IPs were carried out with an *S. cerevisiae* α2 antibody that appears to cross react with a protein in *C. albicans*. The most likely candidate for this cross-reactivity is α2. If, in fact, the antibody recognizes α2, then the binding of α2 in *C. albicans* is dependent on the presence of **a**1 because no target promoters immunoprecipitate with the α2 antibody in the **a**1 deletion strain. Because it appears that both **a**1 and α2 are bound to the test promoters, it seems unlikely that only **a**1 would be responsible for the repression activity. The second piece of evidence against **a**1 binding on its own is the clear dependence of the *S. cerevisiae* **a**1 on the tail of α2 for ordering the **a**1 DNA binding domain and allowing binding (20). **a**1 in *C. albicans* would have to behave quite differently, and the high level of identity in the DNA binding domain of **a**1 between the two organisms argues that they may behave in a similar manner (13). On the other hand, one cannot rule out that the *S. cerevisiae* α2 antibody could be cross reacting with the *C. albicans* **a**1 proteins and giving the appearance of α2 binding. Chromatin immunoprecipitation experiments with chromatin from an α2 deletion strain should resolve any ambiguity because they will discriminate among the different binding patterns. For

example, if **a**1 mediates repression on its own, then the α2 deletion strain should look just like the wild-type strain with respect to **a**1 binding. Accordingly, if the *S. cerevisiae* α2 antibody is crossreacting with **a**1 in *C. albicans*, then the binding should be insensitive to changes in the α2 deletion strain. These remote possibilities aside, it appears based on the data presented here that **a**1-α2-mediated repression occurs on genes in *C. albicans* that resemble haploid-specific genes in *S. cerevisiae* and that this repression can occur through a DNA binding site that is different than the site in *S. cerevisiae.*

The promoter sequences upstream of the various test genes in these experiments have been evaluated in an effort to identify potential **a**1-α2 binding sites, but this analysis has not revealed clear candidates for DNA binding by **a**1-α2. The clearest way to identify the binding sites would be to test fragments of the promoters of interest with in vitro DNA binding assays. Using promoter fragments from the different promoters as probes for binding in gel retardation assays with purified **a**1 and α2 proteins from *C. albicans* would be the most direct way to identify the sites to which they bind. Once a clear consensus of the site is developed, then this sequence could be used in a computer analysis of the *C. albicans* genome again to attempt to identify additional target genes which could then be tested by chromatin IP and Northern for direct binding and regulation by **a**1-α2.

Understanding the kinds of genes regulated by the **a**1-α2 complex in *C. albicans* will give a better understanding of the processes regulated by these proteins and perhaps shed additional light on why a diploid organism such as *C. albicans* would retain the ability to carry out haploid-specific gene regulation. A better understanding of the conservation of molecular regulation between *C. albicans* and *S. cerevisiae* would also provide information about the way in which complex regulatory circuits are maintained and evolve between organisms over time. Perhaps these changes could lend insight into what makes the pathogenic *C. albicans* behave differently than its non pathogenic relatives such as *S. cerevisiae.*

References

1. **Ammerer, G.** 1994. Sex, stress and integrity: the importance of MAP kinases in yeast. Curr Opin Genet Dev **4:**90-5.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1992. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
3. **Blumer, K. J., J. E. Reneke, and J. Thorner.** 1988. The *STE2* gene product is the ligand-binding component of the alpha-factor receptor of *Saccharomyces cerevisiae*. J Biol Chem **263:**10836-42.
4. **Clark, K. L., D. Dignard, D. Y. Thomas, and M. Whiteway.** 1993. Interactions among the subunits of the G protein involved in *Saccharomyces cerevisiae* mating. Mol Cell Biol **13:**1-8.
5. **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 **a**1-alpha2. Genes Dev **5:**1982-9.
6. **Dietzel, C., and J. Kurjan.** 1987. The yeast *SCG1* gene: a G alpha-like protein implicated in the **a**- and alpha-factor response pathway. Cell **50:**1001-10.
7. **Elion, E. A., B. Satterberg, and J. E. Kranz.** 1993. *FUS3* phosphorylates multiple components of the mating signal transduction cascade: evidence for *STE12* and FAR1. Mol Biol Cell **4:**495-510.
8. **Goutte, C., and A. D. Johnson.** 1994. Recognition of a DNA operator by a dimer composed of two different homeodomain proteins. EMBO Journal **13:**1434-1442.
9. **Goutte, C., and C. Johnson.** 1993. Yeast **a**1 and $\alpha2$ homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein. Journal of Molecular Biology **233:**359-371.

10. **Hartig, A., J. Holly, G. Saari, and V. L. MacKay.** 1986. Multiple regulation of *STE2*, a mating-type-specific gene of *Saccharomyces cerevisiae*. Mol Cell Biol **6:**2106-14.

11. **Herskowitz, I., J. Rine, and J. Strathern.** 1992. Mating-type determination and mating-type interconversion in *Saccaromyces cerevisiae*, p. 583-656. *In* E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), Molecular and Cellular Biology of the Yeast *Saccharomyces*, vol. 2: Gene Expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

12. **Hull, C.** unpublished data. .

13. **Hull, C. M., and A. D. Johnson.** 1999. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*. Science **285:**1271-5.

14. **Johnson, A. D.** 1995. Molecular mechanisms of cell-type determination in budding yeast. Curr Opin Genet Dev **5:**552-8.

15. **Keleher, C., S. Passmore, and A. D. Johnson.** 1989. Yeast repressor α2 binds cooperatively with the yeast protein Mcm1. Molecular and Cellular Biology **9.**

16. **Keleher, C., M. J. Redd, J. Schultz, M. Carlson, and A. D. Johnson.** 1992. SSN6/TUP1 is a general repressor of transcription in yeast. Cell **68:**709-719.

17. **Michaelis, S.** 1993. *STE6*, the yeast **a**-factor transporter. Semin Cell Biol **4:**17-27.

18. **Sequence** data for *C. albicans* Contig 2- was obtained from the Stanford DNA Sequencing and Technology Center web site at http://www-sequence.stanford.edu/group/candida.

19. **Sadhu, C., D. Hoekstra, M. J. McEachern, S. I. Reed, and J. B. Hicks.** 1992. A G-protein subunit from asexual *Candida albicans* functions in the mating signal transduction pathway of *Saccharomyces cerevisiae* and is regulated by the **a**1-α2 repressor. Molecular and Cellular Biology **12:**1977-1985.

20. **Stark, M. R., D. Escher, and A. D. Johnson.** 1999. A trans-acting peptide activates the yeast **a**1 repressor by raising its DNA-binding affinity. Embo J **18:**1621-9.

21. **Straight, A. F., W. Shou, G. J. Dowd, C. W. Turck, R. J. Deshaies, A. D. Johnson, and D. Moazed.** 1999. Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. Cell **97:**245-56.

22. **Wilson, K. L., and I. Herskowitz.** 1986. Sequences upstream of the *STE6* gene required for its expression and regulation by the mating type locus in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A **83:**2536-40.

23. Much of the data and information used in compiling the tables for this Chapter was obtained from the Yeast Protein Database from Proteome, Inc. and from the Johnson Laboratory *Candida albicans* gene database developed by B. Braun and A. Uhl.

24. Special thanks to A. Straight for assistance with the computer analysis, R. Raisner for technical assistance, A. Uhl for sequence comparisons, and the Stanford *Candida albicans* Sequencing Project for their sequence of the *C. albicans* genome. Sequencing of *Candida albicans* by the Stanford DNA Sequencing and Technology Center was accomplished with the support of the NIDR and the Burroughs Wellcome Fund.

Chapter 5

Discussion and Future Directions

Summary

The work presented in this thesis describes the discovery of a mating type-like locus (*MTL*) in *C. albicans* and shows that components of the locus are capable of functioning in vivo on endogenous targets to mediate transcriptional repression. It also shows that the *MTL* locus can function to control a previously unidentified mating reaction. These discoveries have radically changed our view of the *C. albicans* life cycle.

Candida albicans is not asexual

Until this work, *C. albicans* was described as asexual, and the changes that distinguished it from its sexual relatives were not understood. Models have been proposed to explain the absence of a sexual cycle in *C. albicans*, but the work I have presented eliminates many of these models and sets a new paradigm for understanding its life cycle. In particular, because the locus I discovered has **a** information on one chromosome and α information on the other (it is **a**/α), it eliminates the possibility that *C. albicans* cannot mate because it has undergone a mitotic recombination event at its mating type-like locus rendering it either **a/a** or α/α (effectively eliminating one mating type) (1). The strain background that I studied is derived from the common laboratory strain SC5314, and although other more wild isolates of *C. albicans* have subsequently been shown to be **a/a** or α/α (P.T. Magee, unpublished results), the existence of all three possible genotypes (**a**/α, **a/a**, and α/α) rules out the single mating type hypothesis as a way to explain the absence of mating and meiosis in this organism. Furthermore, the idea that *C. albicans* is asexual was wholly eliminated when I showed that derivatives of *C. albicans* expected to behave like the **a** and α cell types of *S. cerevisiae* can mate in a mouse host, resulting in the first demonstration of *C. albicans* mating. Given this mating result, **a/a** and α/α strains would be expected to behave like the **a** and α mating types of *S. cerevisiae* and mate in the same way that the artificially derived "**a**"

and "α" strains do. Perhaps one explanation for the apparent low level recombination detected in *C. albicans* isolates in population genetics studies (3) is the infrequent meeting of **a**/**a** and α/α strains in a single host that leads to mating and some as yet uncharacterized form of recombination and meiosis in a diploid-tetraploid life cycle. There are many reasons why these events may not have been detected previously, including infrequent mitotic recombination events leading to **a**/**a** and α/α strains, poor persistence of these strains in the host (our **a** and α strains are not as virulent as **a**/α; see Appendix G), and thus few opportunities for the strains to come into contact with one another and mate. Although the work presented here abolishes the idea that *C. albicans* cannot mate, it is clearly only the beginning in understanding how and when the mating process occurs in *C. albicans*.

C. albicans mating in vivo and in vitro

Experiments to test mating between *C. albicans* **a** and α strains revealed that *C. albicans* has the ability to mate under the proper conditions. Although I am confident that the prototrophs we have recovered and analyzed are the product of a mating event, we have little information about what factors govern the mating process. In these experiments, mating between marked **a** and α strains of *C. albicans* occurs only in mice, occurs sporadically, and appears to be a low frequency event. I have attempted to achieve mating between these strains under many in vitro conditions without success. Work by the Magee group at the University of Minnesota has resulted in mating between strains in vitro (although this is not a highly efficient process either) (10). Based on our discovery of the *MTL* locus, they created **a** and α strains by inducing specific chromosome loss rather than making precise gene deletions, and their strains are not the SC5314 background. Our strains do not mate under their conditions, and it will be interesting to determine what the differences are between the strains. Perhaps there are components on the *MTL* chromosome that suppress mating in vitro that are lost in their strains and thus allow mating under in vitro conditions.

Mating strains in our SC5314 background created by chromosome loss will need to be tested for the ability to mate in vitro. Another possibility is that there is a difference in the strain backgrounds that causes a difference in response to mating signals. It is certainly clear from the work in *S. cerevisiae* that background differences can have striking phenotypic results. This is particularly noticeable with respect to spore formation; *S. cerevisiae* geneticists know well that some backgrounds sporulate exceedingly well, while other barely sporulate at all. Perhaps the same is true for mating in *C. albicans* and the responses to mating signals are altered in different strain backgrounds.

Without mating in vitro in the standard lab background SC5314, it will be very difficult to test mutants for their ability to mate and learn more about the mating process in *C. albicans*. Other strains have not been developed for use with molecular biology tools (gene deletion strategies, expression plasmids, reporter constructs, etc.), so a decision will have to be made about whether to pursue SC5314 to mating in vitro or to develop the tools in other backgrounds to carry out molecular genetic studies. It would be ideal if the condition or factor necessary for SC5314 mating in vitro could be identified and supplied. One way to attempt to identify a possible cellular factor is to attempt crosses between different backgrounds. Crosses need to be attempted between our strains and the Magee strains. In the event that mating occurs in vitro between the two backgrounds, the arduous process of creating a library from the mating strain and looking for complementation of mating in the non-maters could be undertaken. Another possible strategy is to use *S. cerevisiae* as a mating partner. Perhaps the factor necessary for mating can be supplied by *S. cerevisiae*, and the presence of this factor could provide another tool for identifying what the requirements are for SC5314 mating in vitro. Efficient mating in vitro would be a first step toward developing a genetic system for studying *C. albicans*.

Meiosis in *C. albicans*?

Another important step in developing a genetic system for *C. albicans* would be to identify a meiotic phase of its life cycle. In *S. cerevisiae*, the process of meiosis is understood quite well, and the genes that control it have been characterized (8). Just like the mating genes, most of the meiosis genes of *S. cerevisiae* have homologs in *C. albicans*(5). The presence of these genes strongly suggests that meiosis can occur. To date, however, no clear process resembling meiosis has been identified in *C. albicans*. No meiotic structures such as spores or fruiting bodies have been identified, and the only possible meiotic structures that have been seen in *C. albicans*, chlamydospores, are not haploid (12). Because no clear functions have been previously identified for chlamydospores, they have been classified as apparent dead-end developmental structures in *C. albicans* (12). In addition, attempts to recover haploids by selecting for recessive markers after exposure to different environmental conditions have failed as no clearly haploid *C. albicans* cells have ever been identified (16). It seems that *C. albicans* either has no meiosis cycle at all, or it is an unusual or low frequency process that has yet to be identified. Maybe the process of meiosis occurs very rarely and only a few haploids survive. It is also possible that looking for haploid cells is not the right approach. Perhaps the *C. albicans* life cycle is altered such that it is a naturally diploid-tetraploid organism. This pattern would eliminate the haploid cell and circumvent the predicted problems with recessive lethal mutations that would lead to haploid cell inviability. It is also a formal possibility that after mating, no meiosis occurs, and the strains are returned to a diploid state through chromosome loss. There is obviously a great deal of work to be done to distinguish among all of the possibilities for meiosis, and a logical place to start is with the conjugants recovered from the mating experiments in mice. These strains are good candidates to test for meiosis because they are nearly tetraploid (circumventing predicted problems with inviable haploids), and they are marked in such a way that the loss of markers (via meiosis or chromosome loss) can be monitored relatively easily. It is not

clear what the appropriate signals for meiosis would be in *C. albicans,* and our preliminary experiments to induce meiosis in vitro using starvation and other conditions have failed to produce any visually distinguishable meiotic structures. Given that the mouse was a successful medium for mating and that mammals are the most common environment for *C. albicans*, it is logical that the mouse might also be suitable for meiosis and that the processes are both taking place in the mammalian host. In order to test this possibility conjugants are being passed through mice and monitored for marker phenotypes that suggest the absence of two of the copies of a given marker. Colonies that lose at least two copies of a given gene could be evaluated for additional genome changes that would suggest that they are the products of a meiotic event rather than strains that are simply subject to genomic instability. Comparing the karyotypes of such strains with the starting conjugants and regular diploids may reveal differences in the numbers of each chromosome. These strains, if they are truly diploid, would suggest that meiosis is occurring in mice and would also suggest that at some time in the mouse, meiosis-specific genes are being expressed. It will be very difficult to prove meiosis based on chromosome number because the strains could return to diploidy from a higher number through simple chromosome loss. However, the expression of meiosis-specific genes would be a strong indicator of a meiotic process, rather than simple chromosome loss. Experiments detecting the expression of meiosis-specific genes during passage through the mouse would be a strong indicator of meiosis that could then direct more specific efforts to identify the signals that allow meiosis to occur in vivo and possibly in vitro. The use of FLP recombinase in conjunction with a selectable marker has been shown to be effective for detecting the expression of a given gene in vivo in *C. albicans* even if the expression event is transient (15). This system could be used with the promoter sequences of a meiosis-specific gene homolog *DLH1*. The expression of *DLH1* has never been detected in *C. albicans* (1), so this gene is a good candidate for testing for expression during an infection. Perhaps *DLH1* is being expressed in the mouse in a specific time and/or place and the expression of this gene will shed some light on the

possibility for meiosis of *C. albicans*. Many different genes could potentially be tested for expression, and once this expression profile has been determined, specific efforts to determine the details of when, where, and why expression is occurring can be undertaken. Genes expressed during mouse infection could be involved in a meiotic process for *C. albicans*, but they might also be important for the infection process itself. Understanding the functions of these genes may shed light on both the *C. albicans* sexual cycle and the process of infection. It is known in *Cryptococcus neoformans* that virulence is associated with only one of the cell types (18), suggesting that the sexual cycle and virulence are linked. If such a connection exists in *C. albicans*, it will be interesting to see what the molecular relationships are between mating, meiosis, and pathogenesis.

Characterizing the unexpected genes in the *MTL* locus

The genes in the *MTL* locus are prime candidates for study to see if there are relationships between the sexual cycle and infection. These genes are still relatively uncharacterized, and a comprehensive analysis of their functions will be an important part of understanding the sexual cycle. The *MTL* locus genes fall into two categories: three transcriptional regulator genes (**a**1, α1, and α2) which will be discussed in more detail below and six somewhat unexpected genes (*PAPa, PAPα, OBPa, OBPα, PIKa, and PIKα*). Thus far we know little about these unanticipated genes in the locus. We know that they are not found in the *MAT* loci of other fungi, but beyond that, we do not have much additional information about these genes in *C. albicans.* We can, however, make some guesses about two of the genes based on what we know about *S. cerevisiae.* It appears that the *PAP* genes may be essential genes in *C. albicans* because there are no other poly(A) polymerases in the genome, and this is an essential function in most organisms (9). We know that we can delete one or the other (*PAPa* or *PAPα*) from the genome without any gross defects, so it will be interesting to see if these polymerases have completely overlapping functions or if their primary functions

have been altered. For example, one of them could be the housekeeping poly(A) polymerase while the other could participate in a regulated polyadenylation step of some sort (14). The *OBP* and *PIK* genes are more difficult to speculate about because their functions in *S. cerevisiae* are not clear. It is fun to propose, however, that products of the *PIK* and *OBP* genes could be involved in a sensor mechanism that detects sterols in the environment and transmits signals to transcriptional regulators to activate the relevant response pathways. These gene products could participate in detecting the conditions under which mating or meiosis should occur, or perhaps they could elicit the appropriate response to a host environment and lead to invasion and establishment of infection. Although this speculation is entertaining, the realities of *PAP*, *OBP*, and *PIK* functions will become clear when the deletion mutants are thoroughly phenotype tested. To date, all of the single gene deletions of the *PAP*, *OBP*, and *PIK* genes have been made, and strategies have been worked out for testing which of the genes are essential and which can be deleted from both the **a** and α chromosomes.

There are three classes of phenotypic tests that will help elucidate the roles of these genes: First, there are gross phenotypic tests of the strains to look at ability to grow and respond to a wide variety of in vitro conditions. It will be particularly interesting to see if there are any changes in the *OBP* deletion strains in response to sterols in the growth medium or in sensitivity to azoles (17). Such responses would suggest that the *OBP*s might be involved in sterol sensing or ergosterol biosynthesis. Other in vitro tests such as these may give clues about what pathways to pursue in future studies. The second class of phenotypic tests involves testing the mutant strains in vivo in virulence studies. In conjunction with other testing, seeing effects on virulence may help narrow down what roles the locus genes are playing in mating and/or virulence. Finally, in the third kind of phenotypic tests, the mutant strains can be tested and evaluated using microarrays to give a molecular phenotype. By comparing the expression profiles of genes between deletion and wild-type strains and among mutant strains, one should be able to determine whether the

components of the locus affect the same or different pathways and obtain information about which pathways are being affected. This analysis should be an efficient way to determine whether the components of the locus are primarily involved in the same processes (as they are in other fungi) or whether they participate in different pathways in *C. albicans.* If they are all involved in the same process, it will be informative to see what the process is and how such different genes have come together to carry it out. If they are involved in different processes, it will be interesting to derive how the locus evolved to contain such different genes and to see how the divergence of the **a** and α copies of the genes has affected regulation of the processes. Taken together, the results of the phenotype tests should generate a fairly clear picture of what the *OBP, PAP,* and *PIK* genes are doing in the *MTL.*

Further characterization of the transcriptional regulators in the *MTL* locus

The initial characterization of **a**1, α1, and α2 has shown that **a**1 and α2 function as transcriptional repressors, that **a**1 and α1+α2 deletion strains can mate, that these same deletion strains have decreased virulence in mice, and that none of the deletion strains constructed (all combinations of the three genes) has any obvious phenotype in vitro. The best path to understanding more about these genes will be to analyze the **a**1, α1, and α2 deletion strains by microarray and obtain a molecular phenotype for each of these strains. This analysis will be highly useful for investigating the regulatory roles of the transcriptional regulators in the locus. By comparing microarray results for all of the locus genes, it may help show what overlap, if any, the transcriptional regulators have with the other genes in the locus. Perhaps this process could define common pathways for the locus components. This analysis will also reveal whether the transcriptional regulators are controlling the same kinds genes in *C. albicans* that they are controlling in *S. cerevisiae.* If they regulate the same kinds of genes, then it is likely that they are responsible for controlling cell type in *C. albicans*, which will provide even stronger evidence for the

presence of a complete sexual cycle. If different classes of genes are regulated, then one can look at the possibilities that the sexual cycle is similar but carried out differently at the molecular level (same process; different genes) or that the regulatory genes control processes different than they do in *S. cerevisiae* (same genes; different processes). The mating experiments using strains with specific regulator deletions already suggest that these proteins control cell type (apparent mating between **a** and α strains only), so it will be interesting to see if the array analysis sheds any additional light on how they might be carrying out this control

There are still a lot of unanswered questions about **a**1 and α2 in *C. albicans* and many interesting leads on differences in **a**1-α2 regulation between *C. albicans* and *S. cerevisiae*. The first and most obvious difference lies in the differences in the sequences of the proteins. Although the predicted proteins from *C. albicans* are clear homologs of the *S. cerevisiae* proteins, the conservation of sequence is in the DNA binding domains of the proteins. There is almost no significant identity in the amino-terminal regions of **a**1 and α2 between *C. albicans* and *S. cerevisiae*. These regions are the predicted interaction domains between the two proteins (4), so it is possible these proteins interact in a fundamentally different way in *C. albicans* that would alter their binding specificity or change how they interact with other proteins. Kelly Komachi identified a region in the amino terminus of α2 that may be important for interacting with the global repressor Tup1 (7). This region is absent from the *C. albicans* α2, but repression still appears to be Tup1-dependent in *C. albicans* as shown in Appendix B. Perhaps the differences between the sequences could be useful in identifying the determinants of specific interactions between the repression complex and its DNA binding partners.

Other differences between the genes sequences may also be useful. One area of interest is the intron architecture of the genes. In *S. cerevisiae* **a**1 there are two introns that are spliced making it one of only a few genes with multiple introns in *S. cerevisiae* (11). In the *C. albicans* **a**1 there are also two introns, one in the same position as in *S. cerevisiae*,

and one in an unrecognizable part of the sequence. Interestingly, the 3' splicing signal in the first intron does not appear to conform to consensus predictions (13). As shown in Appendix H by RT-PCR, splicing using the less common of two 3' splice signals results in a larger protein than was initially predicted based on sequence analysis alone. Also interesting is the presence of an intron in the *C. albicans* α2 sequence (in the same position in the homeodomain region as in **a**1) that does not exist in the *S. cerevisiae* α2. The confirmation of the splice junctions for these proteins should be carried out using cDNA clones, which have yet to be recovered for **a**1 and α2.

It will also be intriguing to see if **a**1 and α2 from *C. albicans* have the same DNA binding properties in vitro as the *S. cerevisiae* proteins. **a**1-α2 binding in *S. cerevisiae* has been studied in great detail (6), and the binding site and requirements for complex binding have been determined (2). It will be interesting to see if the *C. albicans* **a**1 and α2 bind in the same way or whether the differences in the protein sequences radically alter how they interact with DNA. In vitro experiments with purified proteins could be carried out, but not until cDNAs are recovered for both of the *C. albicans* **a**1 and α2 genes. If the proteins were purified, they could be used in bandshift experiments to determine what sequences comprise the **a**1-α2 binding site in *C. albicans*. Although the *S. cerevisiae* consensus site works for repression in vivo in *C. albicans*, it is not clear that this is the optimal site or that it resembles other sites in the genome. The differences in amino terminal sequences between the *C. albicans* and *S. cerevisiae* proteins could lead to changes in the spacing of the **a**1 and α2 half sites in the *C. albicans* binding site. Differences between binding sequences between *C. albicans* and *S. cerevisiae* may explain why targets of repression were identified in the absence of obvious **a**1-α2 binding site. In vitro studies to test specific binding should resolve this apparent conflict.

Conclusion

I began the work presented in this thesis with a desire to answer a single question: Why do haploid-specific gene regulatory sites exist in an organism that is exclusively diploid? In the process of attempting to answer that question, I cloned homologs of the *S. cerevisiae* master sexual cycle regulators **a**1, α1, and α2 and found that they are located in a locus in *C. albicans* that resembles the mating loci of other fungi, especially *S. cerevisiae*. The locus is much larger than the *S. cerevisiae* locus and contains open reading frames for six additional proteins that have never been seen before in mating loci. Having identified the locus, I was able to make deletions of **a**1 and of α1, and α2 to create strains that appear to mate like the **a** and α cells of *S. cerevisiae* under the conditions provided by a mouse host. Furthermore, experiments to explore what genes *C. albicans* **a**1 and α2 might be repressing in vivo have revealed some interesting differences between **a**1-α2 mediated repression in *C. albicans* and *S. cerevisiae.* Although I did not completely answer the question I set forth originally, along the way, I learned a great deal about the *C. albicans* life cycle and resolved the great mystery that had been at the core of *C. albicans* biology: Why is *C. albicans* asexual? The answer is that *C. albicans* is not asexual. It contains master sexual cycle regulators in a mating-type locus, and those regulators control cell type in a manner that is likely to be similar to *S. cerevisiae.* Although the big question about mating has been answered, there are many more questions that are now apparent. Does mating occur in nature and how is the process regulated? Does *C. albicans* undergo meiosis? What are the signals that control mating (and possibly meiosis)? Is there a connection between a sexual cycle and pathogenesis? What processes does the locus control? What are the targets of **a**1-α2-mediated repression and how does is this repression carried out? How are cell types defined? With the cloning of the *C. albicans MTL* locus and recovery of cells from mating experiments, the tools are in hand to answer all of these questions. With a more complete understanding of the *C. albicans* sexual cycle, perhaps it can be developed into a genetic system that can be used to

References

1. **Diener, A. C., and G. R. Fink.** 1996. *DLH1* is a functional *Candida albicans* homologue of the meiosis-specific gene *DMC1*. Genetics. **143:**769-776.
2. **Goutte, C., and A. D. Johnson.** 1994. Recognition of a DNA operator by a dimer composed of two different homeodomain proteins. EMBO Journal. **13:**1434-1442.
3. **Graser, Y., M. Volovsek, J. Arrington, G. Schonian, W. Presber, T. G. Mitchell, and R. Vilgalys.** 1996. Molecular markers reveal that population structure of the human pathogen Candida albicans exhibits both clonality and recombination. Proc Natl Acad Sci U S A. **93**(22)**:**12473-7.
4. **Ho, C. Y., J. G. Adamson, R. S. Hodges, and M. Smith.** 1994. Heterodimerization of the yeast MATa1 and MATalpha 2 proteins is mediated by two leucine zipper-like coiled-coil motifs. Embo J. **13**(6)**:**1403-13.
5. **Hull, C. M.** unpublished observations. .
6. **Johnson, A. D.** 1992. A combinatorial regulatory circuit in budding yeast, p. 975-1006. *In* S. L. McKnight and K. R. Yamamoto (ed.), Transcriptional Regulation. Cold Spring Harbor Press.
7. **Komachi, K., M. J. Redd, and A. D. Johnson.** 1994. The WD repreats of Tup1 interact with the homeodomain protein alpha2. Genes and Development. **8:**2857-2867.
8. **Kupiec, M., B. Byers, R. Esposito, and A. Mitchell.** 1997. Meiosis and Sporulation in *Saccharomyces cerevisiae*, p. 889-1036. *In* E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), The Molecular and Cellular Biology of the Yeast *Saccharomyces*, vol. 3: Cell Cycle and Cell Biology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
9. **Lingner, J., J. Kellermann, and W. Keller.** 1991. Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. Nature. **354**(6353)**:**496-8.

10. **Magee, B. B. M. a. P. T.** 2000. Induction of mating in *Candida albicans* by construction of *MTLa* and *MTLalpha* strains. Science. **in press**.

11. **Miller, A. M.** 1984. The yeast *MATa1* gene contains two introns. Embo J. **3**(5)**:**1061-5.

12. **Odds, F. C.** 1988. *Candida* and *Candidosis,* 2 ed. Bailliere Tindall, London.

13. **Rymond, B., M. Rosbash.** 1992. Yeast Pre-mRNA Splicing, p. 143-192. *In* E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), The Molecular and Cellular Biology of the Yeast *Saccharomyces*, vol. 2: Gene Expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

14. **Seydoux, G.** 1996. Mechanisms of translational control in early development. Curr Opin Genet Dev. **6**(5)**:**555-61.

15. **Staib, P., M. Kretschmar, T. Nichterlein, G. Kohler, S. Michel, H. Hof, J. Hacker, and J. Morschhauser.** 1999. Host-induced, stage-specific virulence gene activation in *Candida albicans* during infection. Mol Microbiol. **32**(3)**:**533-46.

16. **Summers, E.** personal communication. .

17. **White, T. C.** 1997. Antifungal Drug Resistance in *Candida albicans*. ASM News. **63:**427-433.

18. **Wickes, B. L., M. E. Mayorga, U. Edman, and J. C. Edman.** 1996. Dimorphism and haploid fruiting in *Cryptococcus neoformans*: association with the alpha-mating type. Proc Natl Acad Sci U S A. **93**(14)**:**7327-31.

Miscellaneous Appendices

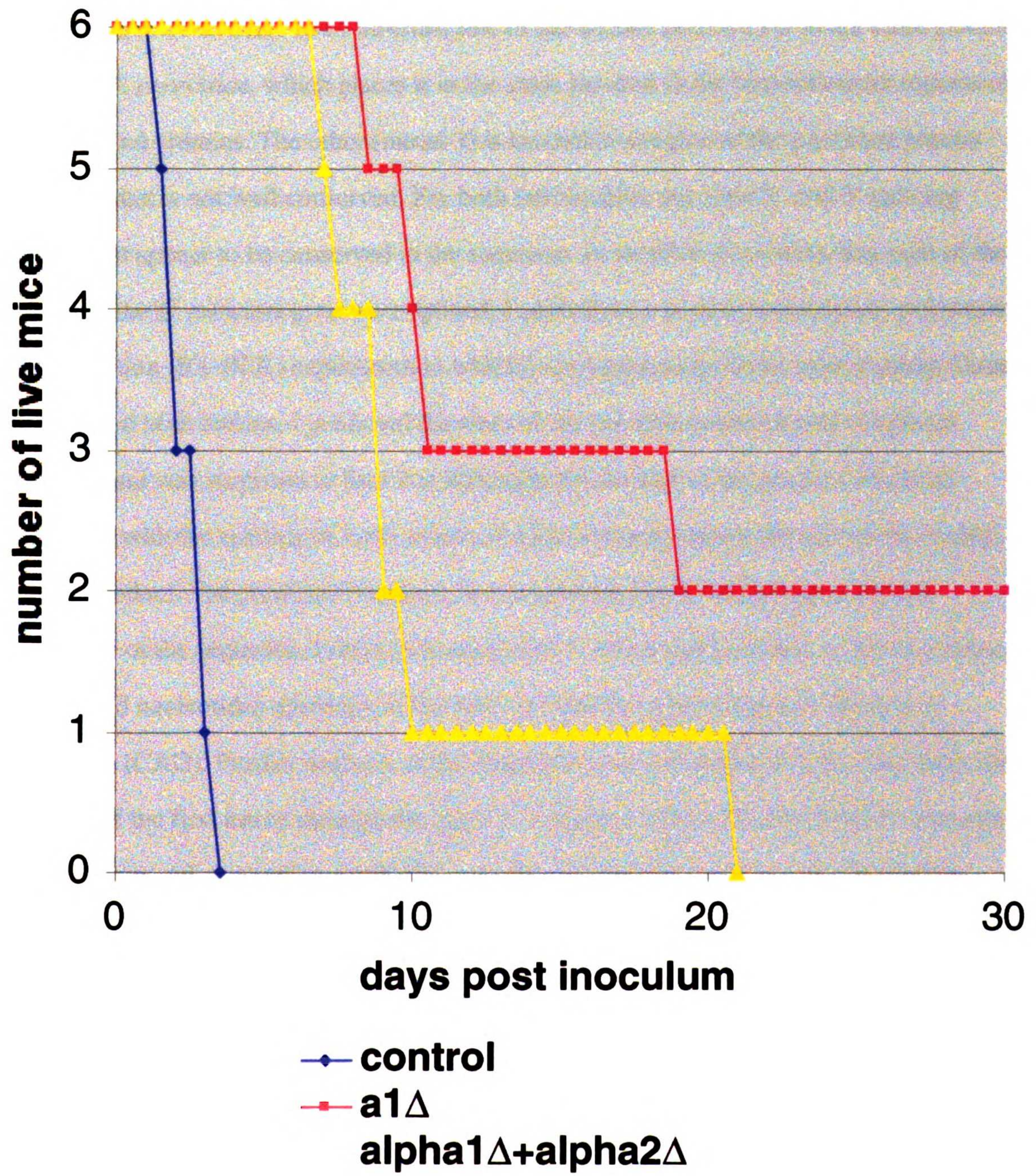
Appendix G

Virulence assays for a1 deletion and α1+α2 deletion strains in mice.

In order to understand better what role the *MTL* locus might be playing in *C. albicans,* we tested two of the transcriptional regulator deletion strains in a systemic model of infection in mice to look at virulence of the different strains. In this experiment we tested the **a**1 deletion strain and the α1+α2 deletion strain and compared them with a wild-type strain. Healthy female Balb-C mice were injected with 10^6 *C. albicans* cells and monitored for rates of deterioration. The graph in figure G shows that the **a**1 and α1+α2 strains had a strong effect on strain virulence with the **a**1 deletion taking three times as long to lead to mouse morbidity in the mice that succumbed at all and the α1+α2 strains taking substantially longer to cause disease as well. All of the mice in this experiment were dissected and analyzed histologically, and in the cases of the mice infected with deletion strains, significantly fewer cells were found in the organs of the animals. The conclusion from this experiment is that mice injected with *C. albicans* cells that were deleted for **a**1 or α1+α2 were substantially less affected than mice injected with wild-type *C. albicans.* Although this result does not necessarily indicate that the transcriptional regulators are directly involved in virulence, it does indicate that they are in some way important for infecting mammalian hosts. These data also helped us to predict how the deletion strains would persist in animals in the mating experiments carried out in Chapter 2.

Figure G.1. Graph displaying a survival curve for mice infected with various strains of *C. albicans*. The x axis represent the number of days post inoculum, and the y axis represents the number of live mice. The blue line shows the number of mice alive over time for the wild-type control strain. The yellow line shows the number of mice alive over time for the α1+α2 deletion strain. The red line show the number of mice alive over time for the **a**1 deletion strain. All remaining mice were sacrificed at 30 days.

Figure G.1



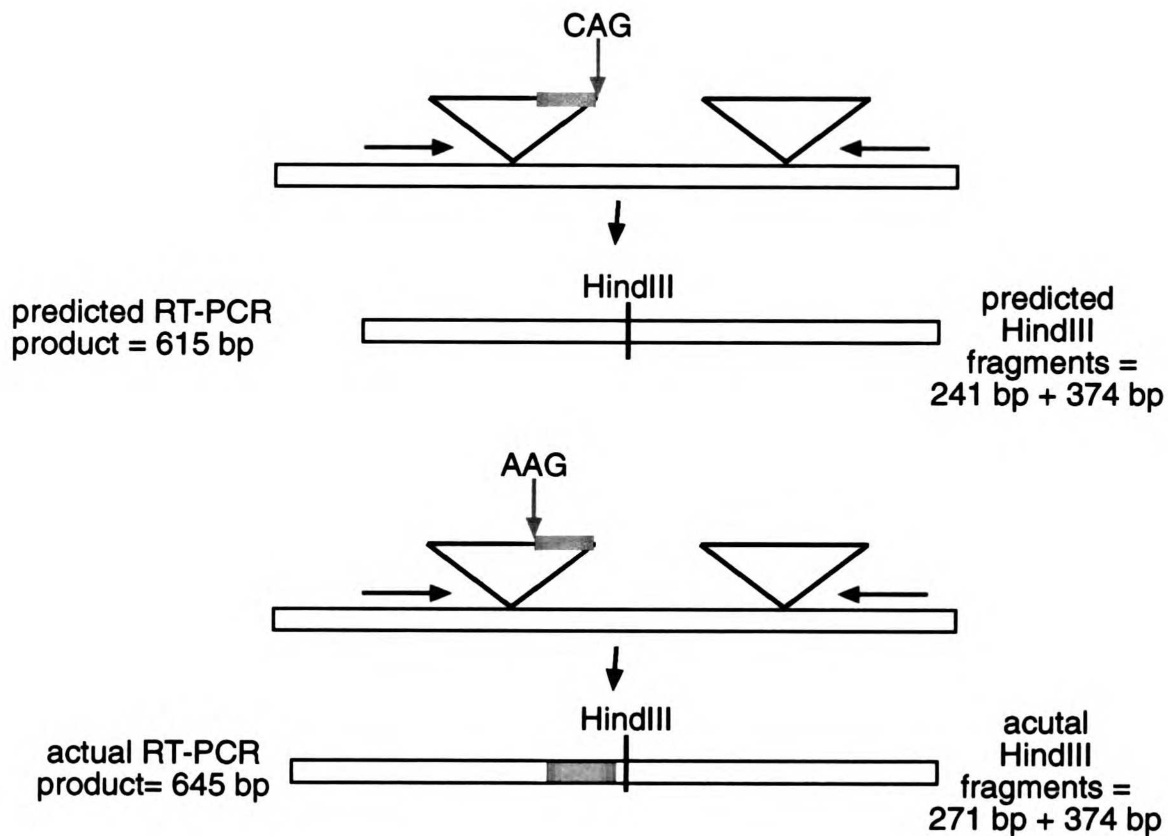
Appendix H

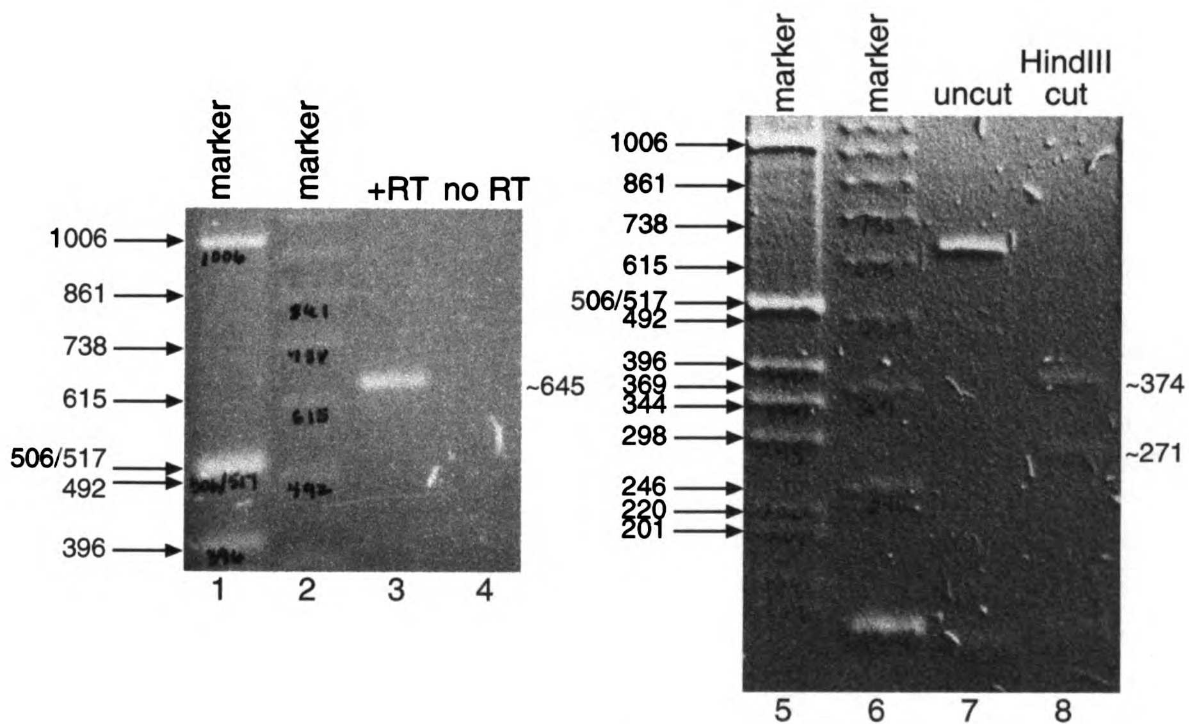
Predicted Splice Junctions for *C. albicans* a1

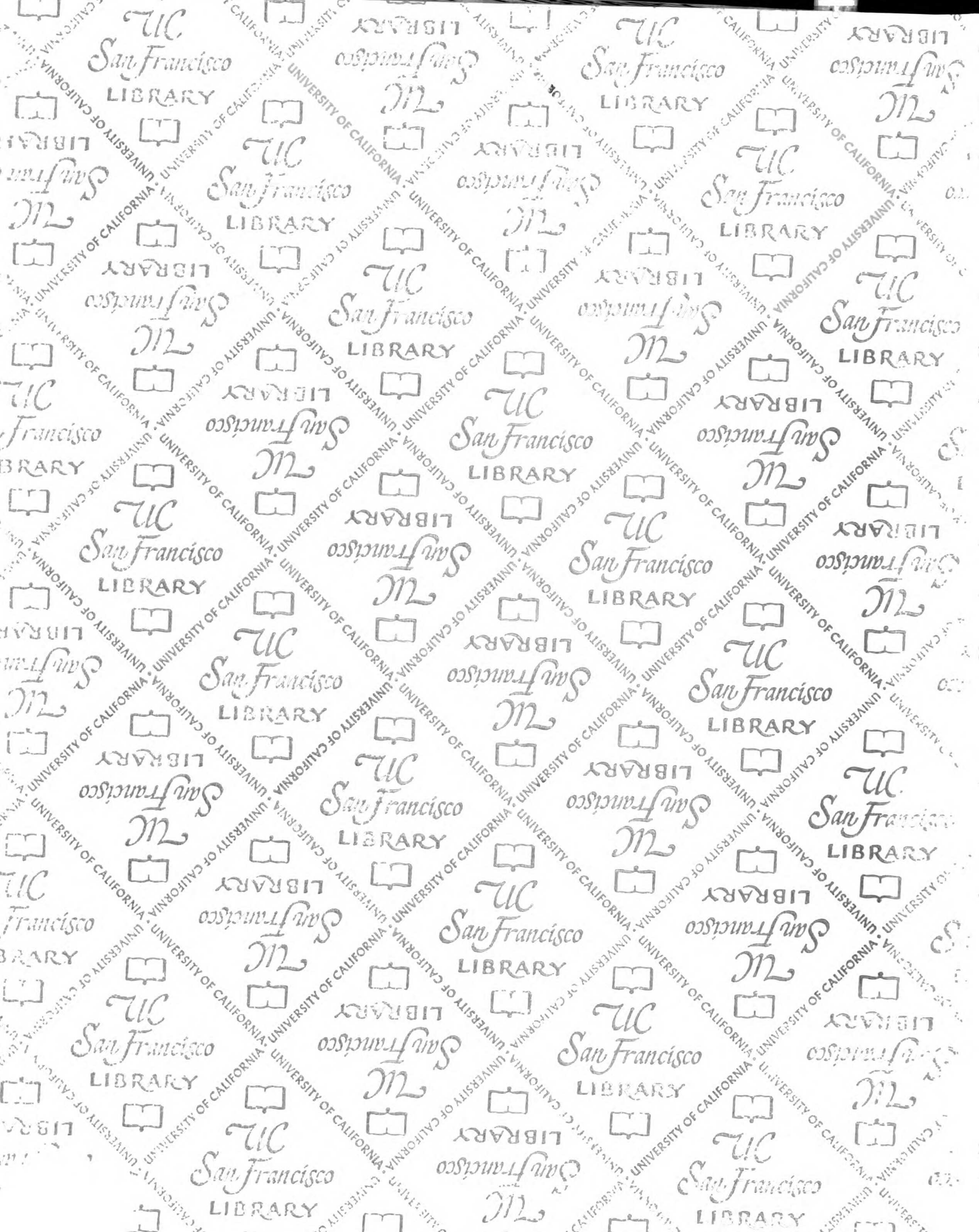
One interesting feature of the *C. albicans* **a**1 gene is that it contains two introns like its homolog in *S. cerevisiae*. In *C. albicans* one of the introns (intron 2) is in the same position as it is in *S. cerevisiae,* which places it in the same location in the homeodomain regions of the predicted proteins. The other (intron 1) is located in a region of the predicted protein sequence that is not well conserved. For both introns there are clear 5' and 3' splicing signals that appear to be conserved in the sequence. In an effort to confirm that both of the introns in the *C. albicans* gene were spliced, I carried out a reverse transcription-polymerase chain reaction (RT-PCR) experiment in which I used primers to the **a**1 open reading frame that flanked both introns. I predicted the sizes of the possible combinations of spliced products and was surprised to find that althought the the size of the product was most consistent with the splicing of both introns, it did not exactly match the size of the doubly spliced product. The predicted fragment was larger than predicted, and upon further inspection of the sequence, I identified a different 3' splice signal (AAG) in the first intron residing 30 nucleotides upstream of the more commonly used consensus signal in *S. cerevisiae* (CAG). Further analysis of the fragments revealed fragment sizes consistent with splicing of the first intron through the AAG 3' sequence rather than the CAG 3' sequence. This different splice junction resulted in a predicted protein 10 amino acids larger than protein I had intially predicted. The sequence of a cDNA clone of the *C. albicans* **a**1 would clarify any ambiguities in splice junctions.

Figure H.1. Predicted and actual splice junctions for *C. albicans* **a**1 . The top portion of the figure shows a schematic representation of the predicted and actual products of RT-PCR experiments. The predicted product was based on the splicing of both introns using the most commonly used consensus sequences for splicing in *S. cerevisiae*. The actual products appear to result from splicing using a less common 3' splice junction in the first intron. Green arrows represent the PCR primers used. The inverted triangles represent the introns. The red boxes indicate the area that is spliced out with the intron in the predicted sequence but appears to remain in the sequence in the actual splicing event. The blue arrows represent the two different 3' splice junctions for the first intron. The junctions were confirmed by analysis of recovered PCR fragments shown in the lower portion of the figure. All marker sized are given in base pairs (bp). Lanes 1 and 2 are marker lanes. Lane 3 shows the fragment resulting from the RT-PCR running at a molecular weight greater than the predicted 615 bp. Lane 4 is a negative control showing no product in the absence of reverse transcriptase. Lanes 5 and 6 are marker lanes. Lane 7 shows the actual RT-PCR product again prior to digestion with HindIII. Lane 8 shows the fragments that result from digestion with HindIII. All RT-PCR reactions were carried out using Ready-to-Go RT-PCR beads from Amersham-Pharmacia Biotech.

Figure H.1







For reference

Not to be taken from the room.

7064788

3 1378 00706 4788