Molecular genetic techniques for gene manipulation in Candida albicans

Qiu-Rong Xu1, Lan Yan2, Quan-Zhen Lv2, Mi Zhou2, Xue Sui3, Yong-Bing Cao2,*, and Yuan-Ying Jiang2,*

1Department of Traditional Chinese Medicine; College of Pharmacy; Fujian University of Traditional Chinese Medicine; Fuzhou, Fujian PR China; 2Center for New Drug Research; School of Pharmacy; Second Military Medical University; Shanghai, PR China; 3School of Life Science and Bio-pharmaceutics; Shenyang Pharmaceutical University; Shenyang, Liaoning PR China

*These authors contributed equally to this work.

Keywords: Candida albicans, gene manipulation, selectable markers, homologous recombination, virulence

Candida albicans is one of the most common fungal pathogen in humans due to its high frequency as an opportunistic and pathogenic fungus causing superficial as well as invasive infections in immunocompromised patients. An understanding of gene function in C. albicans is necessary to study the molecular basis of its pathogenesis, virulence and drug resistance. Several manipulation techniques have been used for investigation of gene function in C. albicans, including gene disruption, controlled gene expression, protein tagging, gene reintegration, and overexpression. In this review, the main cassettes containing selectable markers used for gene manipulation in C. albicans are summarized; the advantages and limitations of these cassettes are discussed concerning the influences on the target gene expression and the virulence of the mutant strains.

Introduction

Candida albicans is one of the most common opportunistic pathogen in humans that causes a number of clinical diseases ranging from superficial infections to life-threatening systemic disease in immunocompromised patients,1 such as those infected with HIV, undergoing cancer chemotherapy and organ transplantation, as well as premature infants.2,3 This pathogenic fungus is becoming the leading cause of nosocomial infections.4 In Kunming, southwest China, the oral yeast colonization rate in AIDS patients (49.5%) was higher than that of healthy people (20.7%).5 C. albicans constituted the most frequent species, accounting for 82.2% of yeast isolates and is the most common lethal species (48.8%) in patients with cancer,5 endotracheal intubation or hypoproteinemia.6 Azole antifungal agents, especially fluconazole (FLC), are used to treat candidiasis. However, the overuse and long-term treatment with FLC have resulted in the emergence of resistance in C. albicans.7-10 It is remaining largely mysterious toward the drug resistance mechanism and the pathways of virulence and pathogenesis in C. albicans.11-14

In recent years, the Candida Genome Database is established to offer the genome sequences of C. albicans for researchers to study freely.15 Several gene manipulation techniques are available to allow powerful genetic approaches to study function of the genes in C. albicans, including gene disruption,16-19 controlled gene expression,16,20,21 protein tagging, gene reintegration and overexpression.16,18,22,23 The principles of these gene manipulation techniques are all based on homologous recombination between the complementary sequences and the genomic sequences in C. albicans. But their applications and the use of selectable markers in these gene manipulation approaches are different. In this review, the main cassettes used for gene manipulation in C. albicans are summarized (Fig. 1; Table 1); the advantages and limitations are discussed.

Cassettes Used in Gene Disruption

The method of targeted gene deletion has enabled the identification of several unique aspects of C. albicans genes that play roles in pathogenesis, virulence, and resistance. Genes that are not essential for survival in C. albicans can be disrupted or replaced by transformation with exogenous DNA. These DNA cassettes used for gene disruption contain the selectable marker, auxotrophic or drug-resistant marker, flanked by sequences homologous to your target gene (YTG).

Gene disruption with nutritional marker for selection

URA blaster cassette

In C. albicans, the URA3 gene encodes the orotidine 5’-monophosphate (OMP) decarboxylase enzyme, catalyzing the conversion of OMP to uridine 5’-monophosphate in the de novo pyrimidine biosynthesis pathway.24 The strain CAI4 (Ura−) (Table 2) with two alleles’ deletion of URA3 is used as the parental strain. The URA blaster cassette carries two direct repeats of the hisG sequences from Salmonella Typhimurium flanking the URA3 gene (Fig. 1.1). After the strain CAI4 being transformed with the URA blaster cassette, the Ura− transformants are selected on uracil-deficient medium. The two direct repeats of the hisG sequences can spontaneously recombine, and then...
Figure 1. Cassettes used for gene manipulation in C. albicans. The figure shows the necessary elements in these cassettes, the light blue color of the elements represents the 5′ and 3′ DNA flanking your target gene (YTG) or the upstream and downstream sequences of YTG. (1) URA blaster cassette. (2) PCR amplifiable marker cassettes from non-C. albicans Candida species. (3) URA flipper cassette. (4) Cre-loxP system. (5) MPAL flipper cassette. (6) SAT1 flipper cassette. (7) "Tet-Off" system. (8) "Tet-On" system. (9) Promoter exchange cassette. (10) Cassettes used for gene reintegration. (11) ACT1P-GFP-ACT1T cassette. (12) The firefly luciferase selectable marker cassette. (13) PCR-mediated gene-tagging cassette. (14) Epitope tagging cassette.
Table 1. Cassettes used for gene manipulation in *C. albicans*

| Cassettes | Parent strains | Selective markers | Application examples | References |
|-----------|----------------|-------------------|---------------------|-----------|
| URA blaster cassette | CAI4 | URA3 | Disrupting gene in auxotrophic strains | 24 |
| URA flipper cassette | CAI4 | URA3 | Disrupting gene in auxotrophic strains | 38 |
| PCR amplifiable URA3 cassette | RM1000, BWP17 | URA3 | Disrupting gene in auxotrophic strains | 19 |
| UAU1 cassette | BWP17 | URA3, ARG4 | Disrupting gene in auxotrophic strains | 43 |
| PCR amplifiable marker cassettes from non-*C. albicans* Candida species | SN87, SN95, SN152 | HIS1, LEU2, ARG4 | Disrupting gene in auxotrophic strains | 45 |
| Cre-loxP system | BWP17 | HIS1, URA3, ARG4 | Disrupting gene in auxotrophic strains | 46 |
| MPAflipper cassette | Any strain | IMH3 | Disrupting gene in any strains | 47 |
| SAT1 flipper cassette | Any strain | SAT1 | Disrupting gene in any strains | 51 |
| “Tet-Off”system | Ca551 | SAT1, URA3 | Repressing the expression of the target gene | 20 |
| “Tet-On” system | CAI4 | SAT1, URA3 | Inducing the normally not expressed gene or inducing the deletion of essential gene | 54 |
| Promoter exchange cassette | BWP17, SN148, SN152 | URA3, HIS1, ARG4, LEU2, SAT1 | Repressing the expression of target gene | 18 and 60 |
| PCR-directed recombination in *S. cerevisiae* | SN152, SN147 | ARG4, SAT1 | Inducing overexpression of target gene | 65 |
| ACT1-GFP-ACT1 cassette | CAI4 | URA3 | To locate the target protein | 76, 110, and 111 |
| The firefly luciferase selectable marker cassette | CAI4 | URA3 | To locate the target protein | 77 and 78 |
| PCR-mediated gene-tagging cassette | CAI4, NGY152 | SAT1 | To locate the target protein | 16 |
| Epitope tagging for antibody detection | BWP17, SN76 | HIS1, URA3, ARG4 | To analyze the protein complex | 23 |

loop out the *URA3* gene and leave behind a short *hisG* sequence in the genome. Cells that have lost *URA3* by homologous recombination can be selected on medium containing 5-fluoroorotic acid (5-FOA), because 5-FOA is toxic to *Ura*+ cells. After another round of transformation and homologous recombination, the cells that are null for the target gene can be generated (Fig. 2).24-25

The *URA* blaster method is a classical and highly effective approach to disrupt the target gene in *C. albicans*. Several parental strains lacking additional nutritional markers such as RM1000 (Ura3− His1−) and BWP17 (Ura3− His1− Arg4−) (Table 2) are constructed by this method.24,25-26 However, the *URA* blaster method has several shortcomings. First, a portion of the 3′ end of the *IRO1* involving in iron utilization was removed during the construction of the parental strain CAI4 (Ura3−), which activated the alternative mechanisms of iron uptake in CAI4. The growth of CAI4 is better than that of the wild-type SC5314 strain in an iron-restricted environment.27-29 Second, one copy of *hisG* sequences left in the chromosome decreases the integration frequency of the disruption cassette at the target locus because of the competition between the endogenous *hisG* sequences and the successive *hisG* cassettes in the second round of transformation.30,31 Third, the flanking direct repeats of this cassette integrated at the *HWP1* locus can negatively influence the expression of the *URA3* in *C. albicans* and affect the virulence phenotype of the *hwp1* null mutants in *C. albicans*.32 Finally, exposure to 5-FOA is potentially mutagenic and could introduce chromosomal rearrangements.32,33 Particularly, the copy number and the changed chromosomal location of the *URA3* gene reduce capability of hyphal morphogenesis, adherence and virulence in *C. albicans*.24,29,33 To ameliorate this problem, it is recommended to place the remaining *URA3* copy at a highly expressed locus, including *ENO1*, *RPS10*, *ARG4*, or at the *URA3* itself native locus.29,34-37 Otherwise, researchers also improve the efficiency of the *URA3* cassette and have applied the PCR amplifiable *URA3* cassette (URA3-dpl200) and the *UAU1* cassette (ura3Δ3′-ARG4-ura3Δ5′) to disrupt the target gene in *C. albicans*.19,26,41,42 The *UAU1* cassette on the Tn7 transposon has been used successfully to disrupt several genes in *C. albicans*.34
PCR amplifiable marker cassettes from non-C. albicans Candida species

These cassettes contain Candida dubliniensis HIS1, Candida maltosa LEU2, or C. dubliniensis ARG4 as selectable markers flanked by two short repeat sequences homologous to the target gene in C. albicans (Fig. 1.2). The auxotrophic parental strains, including SN87 (Leu2− His1−), SN95 (His1− Arg4−), or SN152 (Leu2− His1− Arg4−) (Table 2), are transformed with these cassettes (Fig. 3). These cassettes have been successfully used for large-scale gene deletion studies in C. albicans. The parental strains used for these cassettes display no karyotypic changes and have one copy of URA3 expressed at the native locus, which don’t affect the virulence in C. albicans. So far, there is no evidence showing that the ectopic expression of any of the above nutritional markers would affect the virulence in C. albicans.41,45 However, these cassettes cannot be used to disrupt the genes in the wild-type strains.

URA flipper cassette

The URA flipper cassette contains two direct repeats of the minimal FLP recombination target (FRT) sequence flanking the URA3 marker. The site-specific recombinase FLP gene is regulated by secreted aspartyl proteinase (SAP2) promoter (Fig. 1.3). After the insertion of URA3 flipper into the target locus in the parental strain CAI4 (Ura−) (Table 2), the transformants are grown in the yeast carbon base and bovine serum albumin (YCB-BSA; pH 4.0) medium (Table 3) to induce the expression of FLP recombinase which promotes the homologous recombination between the FRT sites to loop out the URA3 gene, making the URA3 marker recyclable for another round of transformation.38

Table 2. The commonly used auxotrophic strains in gene manipulation in C. albicans

| Auxotrophic strains | PhenoType          | Features                                           | References                  |
|---------------------|--------------------|----------------------------------------------------|-----------------------------|
| CAI4                | Ura−               | Decreased virulence; having chromosome 2 trisomy    | 24, 29, 32, 33, 63, 80, 81, and 86 |
| RM1000              | His−, Ura−         | Decreased virulence; lacking the right arm of chromosome 5 | 25, 26, 29, 32, 33, 63, 80, 81, and 86 |
| BWP17               | Arg−, His−, Ura−   | Decreased virulence; lacking the right arm of chromosome 5 | 25, 26, 29, 32, 33, 63, 80, 81, and 86 |
| CaSS1               | His−, Ura−         | Decreased virulence                                | 20, 29, and 33              |
| NGY152              | Ura−               | Full virulence                                     | 16 and 29                   |
| SN76                | Arg−, His−, Ura−   | Decreased virulence; having no chromosome change   | 23, 45, and 87              |
| SN87                | Leu−, His−         | Full virulence                                     | 45                         |
| SN95                | Arg−, His−         | Full virulence; having no chromosome change        | 45 and 87                   |
| SN100               | His−               | Full virulence                                     | 45                         |
| SN148               | Arg−, Leu−, His−, Ura− | Decreased virulence; having no chromosome change | 18, 45, 60, and 87          |
| SN152               | Arg−, Leu−, His−   | Mild virulence defect; having no chromosome change | 18, 45, 60, and 87          |
Compared with the excision of the UR A3 gene by spontaneous recombination between the hisG repeats, the FLP-mediated recombination between the FRT sites to loop out the UR A3 marker is more efficient. The UR A flipper method does not need to use 5-FOA to select the ura3− clones, which would avoid mutagenic potentiality. However, this cassette leaves behind a single FRT sequence in the genome after the excision of UR A3 marker. The multiple FRT sequences left in the genome might recombine induced by the FLP recombinase, resulting in unexpected deletions or chromosome rearrangements. Otherwise, the use of the BSA activates the SAP2 promoter to induce the expression of FLP recombinase in the UR A flipper cassette, whereas the BSA also represses the expression of other genes of SAP family except for SAP2 gene.

**Cre-loxP system**

This system is based on the site-specific recombination between two loxP elements which are catalyzed by the Cre recombinase. This system contains three components, including the disruption cassettes (using of HIS1 and ARG4 as auxotrophic markers flanked by 34 bp loxP elements), the resolving cassette (containing a synthetic, codon-optimized cre gene regulated by MET3 promoter and the reintegration UR A3 marker gene) (Fig. 1.4). After two separate transformations, the two alleles of the target gene in parent strain BWP17 (Ura3− His1− Arg4−) (Table 2) are disrupted by HIS1 and ARG4 markers; then these transformants are selected on medium lacking of histidine during the first transformation or histidine/arginine during the second transformation, but containing methionine and cysteine to repress the expression of the MET3-cre fusion (Table 3).

In the third transformation, the segment of the Cre recombinase and the UR A3 marker flanked by ARG4 sequences is integrated into the ARG4 gene that has disrupted one copy of the target gene in C. albicans. Then, the transformants that are auxotrophic for HIS1, ARG4, and UR A3 were incubated in medium lacking of methionine and cysteine, which can induce the synthetic MET3-cre fusion to produce Cre recombinase and subsequently catalyze the recombination between loxP sites to loop out the HIS1 and the Cre recombinase-UR A3 fragment (Fig. 4).

The Cre-loxP system doesn’t need to use mutagen 5′-FOA to select the transformants. The synthetic MET3-cre fusion of this system to loop out the loxP-marker-loxP cassettes in C. albicans is efficient and avoids using positive screen method to select for the desired mutants. The knockout strains generated by the

### Table 3. The commonly used regulatable promoters in gene manipulation in C. albicans

| Cassettes                        | Promoters | Inducer          | Gene expression | References        |
|----------------------------------|-----------|------------------|-----------------|-------------------|
| “Tet-On” system                  | TET       | Doxycycline      | Induced         | 55                |
| “Tet-Off” system                 | TET       | Tetracycline     | Repressed       | 20                |
| Marker gene-MET3p                | MET3      | Methionine, cysteine | Repressed     | 18, 59, and 60   |
| Marker gene-MAL2p                | MAL2      | Glucose          | Repressed       | 18, 59, and 60   |
| Marker gene-MAL2p                | MAL2      | Maltose          | Induced         | 18, 59, and 60   |
| G1G1p-GFP cassette              | G1G1      | GlcNAc           | Induced         | 61                |
| URA flipper cassette            | SAP2      | YCB-BSA pH 4.0   | Induced         | 38 and 47–49     |
| MPAα flipper cassette           | MAL2      | Maltose or glucose | Induced       | 51–53             |
| Cre-loxP system                  | MET3      | Methionine, cysteine | Repressed     | 46                |

Figure 3. Strategy of the fusion PCR and heterologous markers used for gene disruption in C. albicans. YTG, your target gene. C. d., Candida dubliniensis; C. m., Candida maltosa.
Cre-loxP system are auxotrophic, unlike those by the positive selection markers. Moreover, the future uses of Cre-loxP system in C. albicans are not limited to gene disruptions; it can be used in the tagging of multiple ORF in situ. However, the Cre-mediated recombination of this system might occur in the absence of explicit MET3-cre fusion induction, which will reduce the stability of cre-containing strains. This method leaves a short loxP sequence in the genome after excision of markers. The multiple loxP sequences left in the genome might also recombine induced by the Cre recombinase, which results in unexpected deletions or chromosome rearrangements.

Gene disruption with drug-resistant marker for selection

Mycophenolic acid-resistant (MPA) flipper cassette

Using the drug-resistant markers, genes of C. albicans can be disrupted in any strains even in clinical isolates. It would be more benefit by using of the drug-resistant marker to study the virulence mechanism of the target genes in C. albicans than using the nutritional markers which might affect the pathways status for the genes of interest.

The inosine monophosphate dehydrogenase (Imh3) directs de novo synthesis of GMP in C. albicans. The mycophenolic acid (MPA) can inhibit the activity of Imh3. The transformants over-expressing IMH3 due to the mutated IMH3 gene are more resistant to MPA. The MPA8 flipper cassette contains the mutated IMH3 gene and two FRT direct repeat sequences flanking the genetically engineered FLP gene regulated by SAP2 promoter.

(Fig. 1.5). After the first transformation, the flanking sequences directly integrate the cassette into the target locus. The transformants are grown in the YCB-BSA medium to activate the SAP2 promoter (Table 3) which promotes the expression of the FLP gene. Then the FLP recombinase mediates site-specific recombination between FRT sites to loop out the MPA8 flipper cassette and makes IMH3 marker recyclable for another round of transformation.

This MPA8 flipper method has been successfully used to study the causal relationships between specific genes and drug resistance in clinical isolates. However, this cassette leaves behind a short FRT sequence in the genome after the excision of the MPA8 flipper cassette. Furthermore, it is time-consuming to screen the MPA8 flipper cassette integrated at the target site of the MPA8 transformants. SAT1 flipper cassette

This cassette consists of a nourseothricin resistance marker C. albicans SAT1 gene and the FLP-mediated recyclable marker system regulated by the MAL2 promoter (Fig. 1.6). After the first transformation, the flanking sequences direct integration of the cassette into the target locus. The resistant transformants are picked on medium containing nourseothricin. Then, the correct transformants are grown in medium containing either maltose or glucose to active the MAL2 promoter. After another round of transformation and selection, the desired homozygous mutants can be generated.

Compared with the MPA8 flipper cassette, the integration of the SAT1 cassette into the correct locus occurs with high specificity. Otherwise, this SAT1 flipper cassette can also be used to reintegrate the intact copy of the target gene into the genome of the null mutant to complement the phenotypes of the mutant.

Cassettes Used in Controlled Gene Expression

Usually the inability to obtain homozygous mutants is considered that the target gene might be essential in C. albicans. To test this, one copy of the target gene can be constructed under a promoter and the other allele is disrupted or replaced. The essentiality of the gene is established by measuring survival of C. albicans after a shift to repressive conditions.

“Tet-Off” system

The “Tet-Off” system combines gene replacement and conditional expression (GRACE) to assess gene essentiality. This
system contains a chimeric trans-activator protein and a tetracycline responsive promoter system. In the first step, the precise gene replacement of one allele of the target gene in the parental strain CαSS1 (Ura3− His3−) (Table 2) is made by a cassette containing the HIS3 selectable marker, these transformants were selected on YNB medium to obtain the HIS3 prototrophs which can be used to additional transformation. In the second step, the controllable expression of the remaining allele is made by a Tet regulatable promoter system using a codon optimized SAT1 selectable maker which is driven by the ACT1 promoter (Figs. 1.7 and 6). The desired transformants can be obtained by selecting on medium containing nourseothricin. The transactivator protein binding to the Tet promoter results in constitutive expression of the target gene regulated by the Tet promoter system in the absence of tetracycline. In contrast, in the presence of tetracycline, the association between the transactivator protein and the Tet responsive promoter is disrupted, which leads to repressing of the expression of the target gene (Table 3).20

In the GRACE strains, the transactivator and the URA3 marker gene are integrated into the LEU2 locus. Therefore, GRACE strains can be forced to lose the transactivator and the URA3 marker gene on 5-FOA-containing medium. Thus strains that are unable to survive in the medium containing 5-FOA may be identified as those carrying essential gene.20 Furthermore, the strains obtained by this method have been used for high throughput drug screening. However, this method is less suitable for inducing the expression of specific target genes under conditions where they are normally not expressed, because this system is not always feasible to efficiently remove the tetracycline from cells grown in the presence of the drug environment.20,54

“Tet-On” system

The “Tet-On” system contains different elements and can be applied to induce or repress genes in C. albicans. It permits to induce expression of genes in conditions where they are normally not expressed in the wild-type genome. In such case, the Tet-On system contains reverse tetracycline-controlled transactivator (rTA) regulated by ADH1 or OP4 promoter, the URA3 or SAT1 gene, the target gene regulated by the rTA-dependent promoter, and the flanking ADH1 sequences (Fig. 1.8A). The cassette would integrate into one copy of the ADH1 in a single transformation step. In the presence of the doxycycline, the binding of the rTA to the Tet responsive promoter leads to the expression of the target gene.

Tet-inducible gene expression system is also useful to induce the deletion of the essential genes and to create conditional lethal C. albicans mutants. In such case, the “Tet-On” system contains rTA regulated by ADH1 promoter, the URA3 gene, and the FLP recombinase gene regulated by the rTA-dependent promoter (Fig. 1.8C). Before using this inducible gene deletion system, the parental strain CAI4 (Ura3−) (Table 2) is transformed with the URA flipper cassette (Fig. 1.3) to delete one allele of target gene. Then one copy of target gene is reintegrated into the ACT1 gene locus by a cassette that contains the target gene and the MPA8 marker flanked by FRT sites (Fig. 1.8B).25 This is followed by the
disruption of the remaining allele of target gene by the UR A flipper cassette. Subsequently, the “Tet-On” cassette is transformed to integrate into ADH1 locus (Fig. 7). Providing doxycycline for the cells, the binding of the rtTA to the Tet responsive promoter results in the expression of the recombinase FLP gene, which induces the homologous recombination between the FRT sites to excise the target gene along with the MPA5 marker (Table 3). The mutants losing the target gene can be identified by the gain of MPA susceptibility. On the contrary, cells remain MPA resistance in the absence of doxycycline, demonstrating that the FLP recombinase is not expressed and that no excision occurs.

This method does not depend on nutrient changes in the medium but simply supplies an inducible or repressible substance which does not affect metabolism. In contrast to the “Tet-Off” system, the “Tet-On” system also allows an inducer to express the genes of interest in conditions. However, the “Tet-On” system requires higher concentration of doxycycline to induce gene expression.54 Due to the doxycycline is cytotoxic,36 its concentration should be controlled. And dimethyl sulfoxide can be used to enhance doxycycline-dependent protein expression in Tet-On cells.37 Furthermore, the chelation between the doxycycline and iron interferes with iron homeostasis and thus reduces resistance to FLC in C. albicans.38 Otherwise, this system requires more transformation steps to disrupt two copies of the target gene than the “Tet-Off” system does.54

Promoter exchange cassette

The regulatable MAL2 or MET3 promoter following one marker gene (Fig. 1.9) has been successfully used to regulate the expression of target gene in C. albicans heterozygous strains. Marker genes include ARG4, CaHIS1, CdHIS1, CmLEU2, SAT1, and URA3. The strains BWP17 (Ura3· His3· Arg4·), SN148 (Ura· His· Leu· Arg·), and SN152 (His· Leu· Arg·) (Table 2) are usually used as parental strains. The promoter exchange cassette carries its 5’ end flanking sequences homologous to the non-coding upstream sequences of the target gene and its 3’ end flanking sequences homologous to the starting sequences of the target ORF. After the transformation, the cassette integrates into the genome of heterozygous strain, where the MAL2 or MET3 promoter is right before the target gene. The methionine and cysteine would shut off the MET3 promoter and repress the target gene (Table 3). Since the MAL2 promoter is shut down by glucose and activated by maltose, glucose-containing medium is used to repress the target gene (Table 3). These promoter exchange modules provide rapid way to study the gene function in C. albicans. Furthermore, this method does not recycle marker gene and is very convenient.58,59,60 In addition, the PGLd/5-GFP cassette can be induced by N-acetylglucosamine (GlcNAc) (Table 3) to express the GFP.61

**Cassettes Used for Gene Reintegration and Overexpression**

In the process of constructing gene deletion mutants in C. albicans, the ectopic expression of selectable marker can mislead phenotypes of the null mutants.35,62,63 It is necessary to reintegrate one copy of the intact target gene into the null mutant to restore the wild-type phenotypes. Several integrating vectors such as Clp10, Clp20, and Clp30 have been used efficiently for generating the complemented strains.46 These plasmids carrying the HIS1, URA3, and ARG4 markers are usually used for reintegrating the markers or target genes into the RP10 locus of the auxotrophic strains.34 Otherwise, the SAT1 flipper cassette containing the complete ORF as well as upstream and downstream flanking sequences of the target gene can also be applied to reintegrate one allele of target gene into the native locus of the homozygous mutants (Fig. 1.10A).51 Furthermore, the technique of the PCR-directed recombination in S. cerevisiae can be used to construct the integrating cassette which reintegrates one copy of the intact target gene into the LEU2 locus of the null mutants in C. albicans (Table 1). The homologous recombination of the PCR fragments and the linearized pRS316 vector in S. cerevisiae can be constructed as the integrating cassette containing (S’ to 3’) a Pmel restriction site, the upstream sequences of LEU2 ORF in C. albicans, the ARG4 gene of C. dubliniensis (used as a selectable marker), the intact target gene ORF plus its promoter sequences, the downstream sequences of the LEU2 ORF in C. albicans, and another Pmel site (Fig. 1.10B).64-66 Meanwhile, the specific gene overexpression strains created by using of the highly active promoter to replace the endogenous promoter of the target gene is important as well for studying the gene function...
in *C. albicans*. The highly active promoters such as NAD-linked glyceraldehyde-3-phosphate dehydrogenase (*TDH3*) promoter, PEP carboxykinase (*PCK1*) promoter, and hexosaminidase (*HEX1*) promoter are often fused with the ORF of target gene to promote the expression of the target gene in *C. albicans*.65,67,68

**Table 4.** The commonly used tagging in *C. albicans*

| Epitope tagging for antibody detection | Fluorescent markers |
|---------------------------------------|---------------------|
| Tagging | Maker genes | **Tagging** | **Maker genes** |
| HA | URA3, HIS1, ARG4 | GFP | URA3, CaHIS1, CmHIS1, ARG4, CmLEU2, SAT1 |
| MYC | URA3, HIS1, ARG4 | YFP | SAT1, URA3 |
| TAP | URA3, HIS1, ARG4 | CFP | SAT1, URA3 |
| GST | URA3 | RFP | SAT1, URA3 |
| His9 | URA3 | luc | URA3 |
| V5–6 × His | URA3, SAT1 | |

**Cassettes Used in Protein Tagging**

The molecular techniques of epitope tags are usually applied to protein purification, localization and detection.23,69-72 The cassettes used for tagging proteins with yellow, cyan, and green
fluorescent proteins (YFP, CFP, and GFP) at C-terminus has been applied to analyze the expression and localization of target protein within living cells in *C. albicans*.18,73 Tagging at the N-terminus usually replace the native promoter with a regulatable or constitutive promoter cassette.74 The small epitopes poly-His, V5, GST, HA, MYC, and FLAG are applied to research the assemble process of target protein complexes in *C. albicans* (Table 4).18,23,75

**ACT1P-GFP-ACT1T cassette**

This cassette contains GFP gene regulated by ACT1 promoter and URA3 used for selectable marker (Fig. 1.11). The flanking sequences of the cassette insert into the target locus through homologous recombination. Then the GFP gene of the integrative transformants exhibits a homogeneous, constitutive fluorescent phenotype, which can be applied to study the gene expression and evaluate the gene activation during pathogen–host cell interactions. The major advantage of this cassette is that gene induction can be monitored in a single cell of living organisms.76

**The firefly luciferase selectable marker cassette**

This cassette contains the firefly luciferase gene regulated by ENO1 promoter and the URA3 gene as a selectable marker gene (Fig. 1.12). The cassette inserts into the 5′ end of one allele of the target gene in *C. albicans* through homologous recombination. Then the expression of the firefly luciferase increases the light emission of the transformed colonies. The major feature of this cassette is that it can be used as a sensitive reporter to analyze gene function both in laboratory and clinical isolates of *C. albicans*. Furthermore, the expression of luciferase might be used as a selection strategy to introduce other genes into *C. albicans* cells to investigate the effect of these genes in the cells. Otherwise, the strains that constructed by this cassette can be used in animal infection models to monitor the location of the infection by the virtue of their bioluminescence.77,78

**PCR-mediated gene-tagging cassette**

The PCR-mediated gene-tagging cassette is based on combing the RFP, CFP, YFP, and GFP variants with the nourseothricin resistance marker CaNAT1 regulated by the heterologous TEF1 promoter and terminus (Fig. 1.13). The parental strain CAI4 (Ura*) or NGY152 (Ura*) (Table 2) is transformed with this cassette which inserts into the 3′ end of the target gene directly. Then the strains grow in selective medium to select for nourseothricin-resistant transformants (Fig. 8).16 The major feature of this cassette is to be applied to study protein functional in *C. albicans*. Furthermore, this cassette requires the using of lower concentrations of nourseothricin. However, similar to the use of the SAT1 marker, the expression of the CaNAT1 marker takes time.16

**Epitope tagging cassette**

The construct for in vivo protein tagging in *C. albicans* combines the selectable marker genes such as URA3, HIS1, and ARG4 with the protein tagging of TAP, HA, and MYC (Fig. 1.14) by PCR-mediated homologous recombination.23 The insertion of epitope tags enables detection of target gene by immunodetection. The use of the tagged proteins is to perform chromatin immunoprecipitation coupled with microarray analysis (ChIP-ChIP) or tandem affinity purification (TAP). The MYC and HA epitope tags can be used for co-immunoprecipitation (co-IP) or immunoprecipitation (IP) experiments to validate formation of protein complex, while the TAP tag is often applied to obtain high purified protein samples for mass spectrometry analysis of protein complexes. In such case, the strain BW2517 (Ura3· His1· Arg4·) or SN76 (Ura· His1· Arg·) (Table 2) can be used as parental strain. After the transformation, the tag is integrated into 3′ terminus of one allele of the target gene. This PCR tagging cassettes allow the rapid biochemical analysis of tagged proteins and have been successfully used to analyze the genome-wide location in *C. albicans*.23,79

**Conclusion**

Nowadays, several cassettes have been used to manipulate the *C. albicans* genome to study the function of the genes. The URA blaster cassette is the first used method. However, the expression of selectable auxotrophic marker URA3 affects the virulence,
adhesion, and morphogenesis in C. albicans. The ura3 null mutant strains such as CAI4, RM1000, and BWP17 are deficient in virulence.34,29,32,33,63,80,81 The ectopic expression of URA3 results in 30% of published virulence-deficient mutants having misattributed to the deletion of genes.29 To overcome these problems, researchers proposed other solutions which are the reason that generated many other gene manipulation cassettes, such as the PCR amplifiable markers from non- C. albicans Candida species. This method has reintegrated the URA3 gene into its native locus, so the constructed strains by this method have full virulence such as SN95 (His- Arg*), SN87 (His- Leu*), SN100 (His*), and SN152 (His- Leu- Arg*) (Table 2).45 Among the use of URA3 as a selectable marker for gene deletion, the UAU1 cassette is the prominent method, because it is useful to create C. albicans null mutants with a single transformation step without leaving behind any chromosomal rearrangements. Nevertheless, the use of ARG4, LEU2, and HIS1 as selectable markers for deleting the gene of interest in C. albicans is the best way among avoid using URA3 as nutritional selectable marker, because this method needs two transformation steps to create C. albicans null mutants and doesn’t need additional steps to excise the markers. So far, there is no evidence showing that the ectopic expression of the remaining nutritional markers LEU2, HIS1, or ARG4 affects the virulence of C. albicans. However, the auxotrophic markers in laboratory strains can still be restrictive because of the target virulence of C. albicans resistant markers, such as LEU2 and doesn’t need additional steps to excise the markers. So far, needs two transformation steps to create null mutants using is the best way among avoid C. albicans ARG4 ing behind any chromosomal rearrangements. Nevertheless, the URA3 is the prominent method, because it is useful to create C. albicans URA3 null mutants (Table 2).95 To overcome these prob- results in 30% of published virulence-deficient mutants having somal aneuploidies and changes in chromosomal copy numbers often arise as a response to stress.85,87,90-93 Therefore, C. albicans mutants should be tested for aneuploidy before being used in further studies.

In general, the perfect cassette of gene manipulation should be that the expression of selectable markers do not affect the virulence of strains, the constructed parent strains have no karyotypic changes, as well as the use of reagents do not generate the potential mutant.

To modify and improve the available methods is the first way to conceive the perfect method. Gene expression at an ectopic site is lower relative to that at the native locus, which could lead to virulence defect and growth defects.94-96 However, Gerami-Nejad et al.97 have identified a neutral intergenic region NEUT5L, which facilitates the integration and expression of ectopic genes. They constructed a series of integrated shuttle vectors containing 550 bp sequences homologous to NEUT5L and three selectable markers (NAT1, the recyclable URA3-dpl200 or URA3). Ectopic genes integrated at NEUT5L by these vectors do not influence growth rates and allow native-locus expression levels. Thus the NEUT5L is an ideal genomic locus for the integration of exogenous DNA in C. albicans. Otherwise, Lai et al.98 have modified the vector pTET25 into the pTET25M, so that the URA3 gene flanked by dpl200 can be used repetitively. The pTET25M vectors not only allow ectopic expression of target proteins in a “Tet-On” system with either a C-terminal 6× His epitope or N-terminal or a C-terminal GFP tag, but also possess a Ura-blaster cassette to allow reintroducing a URA3 marker.

The second way is to optimize and apply the gene manipulation methods used in other fungal species. For example, RNA interference techniques are used to disrupt the target genes in other fungal species such as Aspergillus fumigatus, Cryptococcus neoformans, and Aspergillus nidulans. Nowadays, this method is also proving to be successful to disrupt the gene of interest in C. albicans.99 Moreover, Vieira et al.100 have constructed three C. albicans integrative vectors such as Clp10, Clp20, and Clp30 which are ligated to 2 μ S. cerevisiae sequences, so that they are able to in vivo maintain and clone by gap repair within S. cerevisiae. These vectors are especially useful for the integration of genes into C. albicans genome that cannot be reproduced in Escherichia coli because of their toxic effects. Furthermore, Su-Kim et al.101 have adopted double-joint PCR with NAT-split markers to construct the gene disruption cassettes in C. neoformans, which generates higher targeted-integration frequency. The NAT-split marker is referred as the two portions of DNAs of the NAT dominated selectable marker containing 200bp overlapping truncations in between, named the 5′-NAT-split marker and the 3′-NAT-split marker. During the double-joint PCR, the 5′-flanking region of target gene joints with the 5′-NAT-split marker and the 3′-flanking region of target gene joints with the 3′-NAT-split marker. Then, the two double-joint PCR fragments
are combined and are introduced into the strains. To obtain the desired deletion mutants in \textit{C. neoformans}, researchers only have to screen a small number of transformants.

More recently, the transcription activator-like effector nucleases (TALENs) which contain a modular DNA-binding domain and a FokI endonuclease monomer, are fusion proteins and work in pairs. When two TALENs bind to their target DNA sequences, the FokI monomers will dimerize and introduce a double-strand DNA break within the specific binding site. Then the disrupted DNA can either be repaired by non-homologous end-joining (NHEJ) or homologous recombination, which leads to deletion/insertion mutations, specific mutation sites or specific sequence additions. This technology have been successfully used for gene manipulation in zebrafish, \textit{Drosophila}, \textit{Arabidopsis}, mice, human cell lines, and so on.\textsuperscript{102-105} Otherwise, in the clustered regularly interspersed short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system, the targeting CRISPR RNA (crRNA) fused with the trans-activating crRNA can generate double stranded DNA breaks (DSBs) in the target site. Then the DSBs can be repaired either through NHEJ or through homologous recombination, which would result in random deletions or insertions. This system has also been used for genome editing in human cell lines, mouse, zebrafish, and \textit{Caenorhabditis elegans}, and so on.\textsuperscript{106-109}

In summary, there are many other novel gene manipulation strategies that should be considered for future investigations in \textit{C. albicans}. We should make full use of our resources to improve the method used in gene manipulation in \textit{C. albicans}.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

L. Yan is supported by Natural Science Foundation of China (31000079). J. Jiang is supported by China National 973 Program (2013CB531602) and Natural Science Foundation of China (81330083). Y. Cao is supported by Natural Science Foundation of China (81173100). The authors would like to apologize to all researchers whose important works were unable to be cited because of space limitations.
29. Brand A, MacCallum DM, Brown AJ, Gow NA. Ecopic expression of URA3 can influence the virulence phenotypes and protocine of Candida albicans but can be overcome by targeted reintegra-

31. tion of URA3 at the RPS10 locus. Eukaryot Cell 2009; 8:490-9; PMID:19502833; http://dx.doi.

33. 10.1128/EC.4.3.561-569.2004

35. tion of Candida albicans and other pathogenic

37. Candida albicans confers resistance to the specific inhibitor

39. 2006; 11004171; http://dx.doi.org/10.1128/EC.

41. mycophenolic acid. J Bacteriol 1997; 179:2331-8;

43. 2001; 2:15664973; http://dx.doi.org/10.1128/

45. Strains and strategies for large-scale gene disruption in the diploid human fungal pathogen Candida albicans. Eukaryot Cell 2005; 4:298-309; PMID:1570792; http://dx.doi.org/10.1128/EC.4.2.298-309.2005

47. expression and gene deletion in Candida albicans. Yeast 2000; 16:3856-65; PMID:10844673; http://dx.doi.org/10.1006/yea.2000.18899

53. Flanking direct repeats of hisG alter URA3 marker

55. the role of the MDR1 gene in fluconazole-resistant, clinical Candida albicans isolates. Mol Microbiol 2000; 36:856-65; PMID:10844673; http://dx.doi.org/10.1128/EC.3.4.900-909.2004

57. URA3 targeted to the ENO1 locus. Infect Immun 2000; 68:5953-9; PMID:10992507; http://dx.doi.org/10.1128/EC.3.4.900-909.2004

59. Reuss O, Vik A, Kolter R, Morschhäuser J. The

71. Krawchuk MD, Wahls WP. High-efficiency gene

73. Sasse C, Morschhäuser J. Gene deletion in Candida albicans wild-type strains using the SAT1 flipper, an optimized tool for gene disruption in Candida albicans. Gene 2004; 341:19-27; PMID:15474295; http://dx.doi.org/10.1006/jfin.

75. Jen S, Guo W, Kohler JR. CaNATI, a heterolo-

77. of doxycycline with fluconazole against Candida

79. Goshorn AK, Mitchell AP, Contestabile DM, Fetherston RW, Backen AC, Broadbent ID, Fetherston RW, Backen AC, Broadbent ID, Fetherston RW, Backen AC, Broadbent ID, Fetherston RW, Backen AC, Broadbent ID, Fetherston RW.
