Deus ex Candida genetics: overcoming the hurdles for the development of a molecular toolbox in the CTG clade

Nicolas Papon, Vincent Courdavault, Marc Clastre, Andrew J. Simkin, Joël Crèche and Nathalie Giglioli-Guivarc'h

Dominant selectable markers, reporter genes and regulatable systems remain powerful molecular tools for genetic and cell biology studies in fungi. Among Saccharomycotina, it is currently accepted that most species belonging to the genus Candida have adopted a specific codon usage, whereby the CTG codon encodes serine instead of leucine. This group is now widely referred to as the CTG clade. For a long time, this uncommon genetic code has precluded the use of the available Saccharomyces or bacterial markers and reporter systems for genetic studies in Candida species. Over the last 15 years, increasing effort has been made to adapt drug-resistance markers, fluorescent protein variants, luciferase and recombinase genes to favour their expression in species related to the yeast CTG clade. In addition to the growing set of Candida genome sequences, these codon-optimized molecular tools have progressively opened a window for the investigation of the conservation of gene function within Candida species. These technical advances will also facilitate future genetic studies in non-albicans Candida (NAC) species and will help both in elucidating the molecular events underlying pathogenicity and antifungal resistance and in exploring the potential of yeast metabolic engineering.

Introduction

Candida albicans and non-albicans Candida species (NAC) are ascomycetous yeasts which have been heavily studied over the last 20 years due to their clinical importance, their biotechnological interest and their biological control potential.

In humans, some of these fungal agents are responsible for candidiasis and have been described as emerging pathogens characterized by their propensity to develop resistance to antifungal agents during treatment. A modification of candidiasis epidemiology has occurred over the last two decades, in particular with an increase of NAC candidiasis (e.g. Candida dubliniensis, Candida parapsilosis and Candida lusitaniae) and in the variation of the susceptibility of Candida species to antifungal agents. These changes are promoting research programmes aimed at identifying new molecular events supporting pathogenicity and antifungal resistance in these emerging yeasts. The interest in NAC species is notably strengthened by the growing set of available Candida genome sequences provided by the Broad Institute (Butler et al., 2009).

Besides their clinical relevance, various Candida species (e.g. Candida tropicalis, Candida maltosa, Candida famata, Candida rugosa), and also Pichia stipitis, represent powerful biotechnological models for bioconversions or industrial production of value-added metabolites. The efficient ability of certain Candida strains to metabolize plant by-products is being evaluated by a number of research teams in order to produce biocompounds such as antibiotics, vitamins, complex alkanes and biofuel. The trend towards this Candida ‘white biotechnology’ originates from the extraordinary capacities of a small number of Candida species to metabolize C₅ sugars from hemicellulosic waste (Akinterinwa et al., 2008).

In addition, some Candida species have important potential for biological control. Some antagonistic yeasts from the CTG clade (e.g. Candida guilliermondii and Candida oleophila) are used for post-harvest biological control of spoilage fungi during storage of plant-derived products. However, the mechanisms by which these yeasts exert their biocontrol activity have not been fully elucidated (Sundh & Melin, 2011).

All these aspects make Candida species attractive models for exploring new molecular mechanisms relating to pathogenicity, antifungal resistance, metabolite bioproduction and biological control. The development of convenient transformation systems, selectable markers and reporter genes for the genetic manipulation of these yeast species has therefore become indispensable.

Since the late 1990s, it has been accepted that the majority of yeast species belonging to the genus Candida have
adopted a particular codon usage (Kawaguchi et al., 1989; Ohama et al., 1993; Sugita & Nakase, 1999). More recent phylogenetic studies based on whole-genome analysis indicate that the Saccharomyces cerevisiae can be subdivided into two major groups: (i) species whose genomes have undergone a whole-genome duplication (referred to as the WGD clade), including Candida glabrata and yeasts from the genus Saccharomyces, and (ii) species that translate CTG as serine instead of leucine (referred to as the CTG clade), including C. albicans TEF2 sequence under the control of the transcription-regulating region of the C. tropicalis phosphoglycerate kinase (PGK) gene. In 2001, the same team adapted this HPH cassette for gene disruption strategy and validated its usability by creating a homozygous ura3 mutant in C. tropicalis (Hara et al., 2001). During the same time period, Yehuda et al. (2001) reported the direct transformation of the biocontrol yeast C. oleophila by the codon-modified C. tropicalis-adapted HPH cassette. This cassette was notably used to demonstrate the possibility of homologous recombination in this species. Interestingly, in spite of its potential in the CTG clade, the codon-modified HPH cassette (Hara et al., 2000) was apparently forgotten, only to reappear 10 years later (Millerioux et al., 2011; Basso et al., 2010).

Basso et al. (2010) produced a synthetic Escherichia coli hygromycin B resistance gene with optimized C. albicans codons. In this synthetic sequence, the nine CTG codons were substituted by other leucine-encoding codons, and a series of codon modifications were added on the basis of the C. albicans codon usage table. They placed this sequence under the control of the C. albicans TEF2 promoter and the ACT1 terminator and demonstrated that this cassette could be used as dominant selectable marker in a set of C. albicans wild-type isolates. Millerioux et al. (2011) used the codon-modified HPH sequence

Hygromycin B resistance markers

The HPH gene, encoding hygromycin phosphotransferase, which confers resistance to hygromycin B, was the first codon-modified drug-resistance cassette adapted for a Candida species, in 2000 (Fig. 1a; Hara et al., 2000). The authors exchanged the nine CTG codons with CTC codons in the ORF of the bacterial HPH gene using site-directed mutagenesis. To favour its expression in the n-alkane-assimilating yeast C. tropicalis (also known for its clinical relevance), they placed this codon-modified sequence under the control of the transcription-regulating region of the C. tropicalis phosphoglycerate kinase (PGK) gene. In 2001, the same team adapted this HPH cassette for gene disruption strategy and validated its usability by creating a homozygous ura3 mutant in C. tropicalis (Hara et al., 2001). During the same time period, Yehuda et al. (2001) reported the direct transformation of the biocontrol yeast C. oleophila by the codon-modified C. tropicalis-adapted HPH cassette. This cassette was notably used to demonstrate the possibility of homologous recombination in this species. Interestingly, in spite of its potential in the CTG clade, the codon-modified HPH cassette (Hara et al., 2000) was apparently forgotten, only to reappear 10 years later (Millerioux et al., 2011; Basso et al., 2010).

Basso et al. (2010) produced a synthetic Escherichia coli hygromycin B resistance gene with optimized C. albicans codons. In this synthetic sequence, the nine CTG codons were substituted by other leucine-encoding codons, and a series of codon modifications were added on the basis of the C. albicans codon usage table. They placed this sequence under the control of the C. albicans TEF2 promoter and the ACT1 terminator and demonstrated that this cassette could be used as dominant selectable marker in a set of C. albicans wild-type isolates. Millerioux et al. (2011) used the codon-modified HPH sequence

Dominant selection markers

Dominant selection genes are commonly used for genetic transformation of a wide range of organisms. In Candida species, two major groups of selectable genes are now widely employed to carry out genetic modifications: metabolic markers and drug-resistance cassettes.

As early as the 1970s, the development of a procedure for metabolic-marker-mediated yeast transformation usually began by the generation and isolation of an auxotrophic recipient strain. Following the identification of the disrupted metabolic pathway by supplementation with the appropriate amino acid or base, a study of the genetic basis was undertaken by complementation and mapping (Kakar & Magee, 1982). Once identified and isolated, the gene encoding a functional enzyme that complemented the disrupted metabolic pathway was used as dominant marker in the corresponding auxotrophic recipient strain. For example, many C. albicans genetic approaches are based on the BWP17 triple auxotroph (his1 arg4 ura3) in which HIS1, ARG4 and URA3 wild-type genes can be used as selectable markers (Wilson et al., 1999; Samaranayake & Hanes, 2011). These approaches were progressively facilitated by the finding and use of antimetabolites. For example, the introduction of the agent 5-fluoroorotic acid (Boeke et al., 1984) allowed the development of a number of URA3-mediated transformation systems in many fungal species. This compound, toxic for protocytic cells, permits in most cases an easy counter-selection of ura3 or ura5 genotypes from a randomly mutagenized population. In the same way, recent advances have demonstrated the potential of 5-fluoranthranilate for counter-selection of trp1, trp3, trp4 or trp5 genotypes in yeast (Toy n et al., 2000; Cheon et al., 2003, 2009; Lebel et al., 2006; Jones et al., 2008).
(Hara et al., 2000) to transform a series of C. guilliermondii wild-type strains. This codon-modified HPH marker allowed the selection of hygromycin B-resistant transformants, but only when the HPH sequence was placed under the control of the C. guilliermondii PGK promoter and terminator.

To date, the hygromycin-resistance gene is available for efficient transformation of wild-type strains of C. tropicalis, C. oleophila, C. albicans and C. guilliermondii. Candida species appear generally sensitive to hygromycin (used in rich media or buffered minimal media) at concentrations around 400–1000 µg ml⁻¹. Thus, this marker could be potentially adapted for use in various other species of the CTG clade, after appropriate changes to the transcription-regulating sequences.

Nourseothricin resistance markers

During the establishment of a global strategy allowing a large-scale identification of C. albicans essential genes, referred to as GRACE (gene replacement and conditional expression), Roemer et al. (2003) proposed the first codon-modified version of the E. coli SAT-1 gene, encoding streptomycin acetyltransferase, which confers resistance to the antibiotic nourseothricin (Fig. 1a). The unique CTG codon was replaced by a CTT codon, and the modified CaSAT-1 sequence was successfully expressed using the C. albicans ACT1 promoter and PCK1 terminator. Reuss et al. (2004) developed a second version of this codon-optimized SAT-1 by replacing the sole CTG codon by a CTC codon and placing this sequence under the control of the C. albicans ACT1 promoter and the URA3 terminator (referred to as caSAT1). In the same work, they also engineered a flipper variation of this caSAT1 marker (coupled with a codon-modified recombinase system; see below) usable for multiple gene disruption. One year later, Shen et al. (2005) proposed a modified version of the Streptomyces noursei nat1 gene to confer nourseothricin resistance in various Candida species. This synthetic gene, named CaNAT1, was produced by de novo synthesis using the C. albicans codon usage table and was placed under the control of transcription-regulating sequences of the Ashbya gossypii TEF1 gene. These authors showed that the presence of this cassette in the genome of C. albicans does not impair growth or hyphal development of transformed strains. Furthermore, they demonstrated that CaNAT1, driven by the A. gossypii TEF1 gene promoter and terminator, could be used to transform clinical isolates of C. parapsilosis and C. lusitaniae (Shen et al., 2005). Ding & Butler (2007) adapted the caSAT1 flipper cassette (Reuss et al., 2004) to efficiently transform C. parapsilosis clinical isolates. However, they obtained nourseothricin-resistant isolates only after replacing the promoter driving caSAT1 gene expression (C. albicans ACT1 promoter) with the C. albicans ACT1 promoter and PCK1 terminator. Reuss et al. (2004) developed a second version of this codon-optimized SAT-1 by replacing the sole CTG codon by a CTC codon and placing this sequence under the control of the C. albicans ACT1 promoter and the URA3 terminator (referred to as caSAT1). In the same work, they also engineered a flipper variation of this caSAT1 marker (coupled with a codon-modified recombinase system; see below) usable for multiple gene disruption. One year later, Shen et al. (2005) proposed a modified version of the Streptomyces noursei nat1 gene to confer nourseothricin resistance in various Candida species. This synthetic gene, named CaNAT1, was produced by de novo synthesis using the C. albicans codon usage table and was placed under the control of transcription-regulating sequences of the Ashbya gossypii TEF1 gene. These authors showed that the presence of this cassette in the genome of C. albicans does not impair growth or hyphal development of transformed strains. Furthermore, they demonstrated that CaNAT1, driven by the A. gossypii TEF1 gene promoter and terminator, could be used to transform clinical isolates of C. parapsilosis and C. lusitaniae (Shen et al., 2005). Ding & Butler (2007) adapted the caSAT1 flipper cassette (Reuss et al., 2004) to efficiently transform C. parapsilosis clinical isolates. However, they obtained nourseothricin-resistant isolates only after replacing the promoter driving caSAT1 gene expression (C. albicans ACT1 promoter) with the C.
<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene product</th>
<th>Species</th>
<th>No. of CTG codons</th>
<th>Name of adapted version</th>
<th>CTG repl.</th>
<th>Codon opt.</th>
<th>Promoter/terminator</th>
<th>Used in</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPH</td>
<td>Hygromycin B phosphotransferase</td>
<td>Escherichia coli</td>
<td>9</td>
<td>HYG&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Yes</td>
<td>No</td>
<td>CaPGK/CaPGK</td>
<td>C. tropicalis</td>
<td>Hara et al. (2000, 2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CaHygB</td>
<td>Yes</td>
<td>Yes</td>
<td>CaTEF2/CaACT1</td>
<td>C. albicans</td>
<td>Basso et al. (2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPH&lt;sup&gt;+&lt;/sup&gt; (HYG&lt;sup&gt;+&lt;/sup&gt;)</td>
<td>Yes</td>
<td>No</td>
<td>CguiPGK/CguiPGK</td>
<td>C. guillermondii</td>
<td>Millerioux et al. (2011)</td>
</tr>
<tr>
<td>SAT-1</td>
<td>Streptothricin acetyltransferase</td>
<td>Escherichia coli</td>
<td>1</td>
<td>caSAT-1</td>
<td>Yes</td>
<td>No</td>
<td>CaACT1/CaPCK1</td>
<td>C. albicans</td>
<td>Roemer et al. (2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CaSAT1</td>
<td>Yes</td>
<td>No</td>
<td>CaACT1/CaURA3</td>
<td>C. parapsilosis</td>
<td>Reuss et al. (2004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Streptomyces noursei</td>
<td>8</td>
<td>CaNAT1</td>
<td>Yes</td>
<td>Yes</td>
<td>AgTEF1/AgTEF1</td>
<td>C. guillermondii</td>
<td>Gácsér et al. (2007); Nguyen et al. (2009)</td>
</tr>
<tr>
<td>ble</td>
<td>Glyoxalase</td>
<td>Streptomyces hindustanus</td>
<td>5</td>
<td>zeo-n</td>
<td>Yes</td>
<td>No</td>
<td>ScGAL1/ScCYC1</td>
<td>C. rugosa</td>
<td>Tang et al. (2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ScTEF1/ScCYC1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CrLIP3(NS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMH3</td>
<td>Inosine 5’- monophosphate dehydrogenase</td>
<td>Staphylococcus aureus</td>
<td>0</td>
<td>Sa_ble</td>
<td>Yes</td>
<td>No</td>
<td>ScGAP/ScCYC1</td>
<td>C. furinosa</td>
<td>Wang et al. (2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PsTEF1/PsXYL2</td>
<td>P. stipitis</td>
<td>Laplaza et al. (2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Candida albicans</td>
<td>0</td>
<td>IMH3&lt;sup&gt;+&lt;/sup&gt;</td>
<td>No</td>
<td>No</td>
<td>CalIMH3/NS</td>
<td>C. furata</td>
<td>Dmytruk et al. (2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CjTEF1/NS</td>
<td>C. albicans</td>
<td>Köhler et al. (1997)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CalIMH3/CalIMH3</td>
<td></td>
<td>Beckerman et al. (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CalIMH3/CalIMH3</td>
<td>C. tropicalis</td>
<td>Staub et al. (2000b); Wirsching et al. (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CalIMH3/CalIMH3</td>
<td>C. dubliniensis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CalIMH3/CalIMH3</td>
<td>C. parapsilosis</td>
<td>Gácsér et al. (2005); Kosa et al. (2007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CalIMH3/CfIMH3</td>
<td>C. famata</td>
<td>Dmytruk et al. (2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Candida famata</td>
<td>0</td>
<td>Dh_IMH3</td>
<td>No</td>
<td>No</td>
<td>CfJMH3/CfJMH3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Compilation of drug-resistance markers adapted for *Candida* species

No. of CTG codons, number of CTG codons in the original nucleotide sequence; CTG repl., indicates if the CTG codons from the original nucleotide sequence have been replaced; Codon opt., indicates if a series of codon modifications guided by a codon usage table derived from highly expressed genes has been carried out; NS, not specified.
parapsilosis ACT1 promoter (generating the CpsSAT1 version). Millerioux et al. (2011) used the codon-modified sequence cSAT1 (Reuss et al., 2004) to transform C. guilliermondii wild-type strains. The expression of the SAT1 marker was once again only obtained after replacing the cSAT1 transcription-regulating sequences by the C. guilliermondii PGK promoter and terminator. Interestingly, Gács et al. (2007) and Nguyen et al. (2009) achieved the disruption of C. parasplasosis secreted lipase genes and fatty acid synthase genes, respectively, directly using the cSAT1 flipper cassette (Reuss et al., 2004), without changing the regulating sequences.

Growth inhibition of Candida species is observed using a nourseothricin concentration of approximately 100–200 µg ml⁻¹ (in rich media or buffered minimal media). Although quite expensive, nourseothricin and codon-adapted strepto-thricin acetyltransferase genes appear promising to transform a large variety of C. albicans wild-type isolates as well as other Candida species (Shen et al., 2005; Gács et al., 2007; Nguyen et al., 2009). For this reason, cSAT1 has recently been used as a dominant selectable marker to investigate the role of a small number of genes in C. albicans (Walia & Calderone, 2008; Melo et al., 2008; Tsao et al., 2009; Kur et al., 2010).

Phleomycin/zeocin resistance markers

The third codon-modified drug-resistance cassette was originally developed in 2003 (Tang et al., 2003). The authors modified the five CTG codons of the Streptoalloteichus hindustanu ble gene (Sh ble), encoding a glyoxalase, which confers resistance to the bleomycin/phleomycin/zeocin antibiotic group, to transform wild-type strains of C. rugosa (Fig. 1a). In these experiments, the codon-modified version of Sh ble, referred to as zeo-n, was successfully driven by the GAL1 and TEF1 promoters from S. cerevisiae or by the LIP3 promoter from C. rugosa. Three years later, Wang et al. (2006) published a similar approach using a codon-modified sequence of Sh ble, harbouring the five modified CTG codons and driven by the S. cerevisiae GAP gene promoter and CYCI terminator to transform the yeast Pichia farinosa. In the same manner, Laplaza et al. (2006) developed another version of codon-optimized Sh ble by replacing the five CTG codons by TTG, allowing in turn the efficient transformation of strains of Pichia stipitis (formerly Scheffersomyces stipitis, a species belonging to the CTG clade). Interestingly, during the same year, Dmytruk et al. (2006) proposed a different strategy for the development of a zeocin-resistance cassette to transform C. famata wild-type strains. Instead of modifying CTG codons in the well-known Sh ble gene, they derived some benefit from the expanding set of available prokaryote gene sequences by hunting for another bacterial zeocin/pheromycin-resistance gene that does not contain CTG codons. They noticed that the meticillin-resistant Staphylococcus aureus (MRSA) 252 ble gene contains no CTG codons, and demonstrated that this MRSA sequence, under the control of the TEF1 promoter, can be used as dominant selectable marker in C. famata.

Although high frequencies of spontaneous resistant mutants can be observed in various Candida isolates (authors’ unpublished observations), Candida species appear generally sensitive to zeocin and phleomycin (used in rich media) at concentrations of 100–400 µg ml⁻¹. Thus, the MRSA CTG codon-less ble gene, controlled by appropriate transcription-regulating sequences, could become a widely used drug-resistance marker in the CTG clade, as is already the case for filamentous fungi (for a review see Weld et al., 2006).

Mycophenolic acid resistance markers

Although it is not strictly a codon-modified version of a resistance gene, a fourth drug-resistance marker currently available for Candida species can be cited to conclude this section. Köhler et al. (1997) demonstrated that plasmid-mediated overexpression of the C. albicans IMH3 gene (encoding inosine 5'-monophosphate dehydrogenase) confers resistance to the drug mycophenolic acid in C. albicans (Fig. 1a). Not long after, this marker was also employed to design a stage-specific virulence and white–opaque switching reporter system in C. albicans (Staib et al., 1999; Strauss et al., 2001). Finally it was adapted as an IMH3 flipper cassette (coupled with an engineered recombinate system; see below) for sequential gene disruption (Wirsching et al., 2000; Morschhäuser et al., 2005). Further studies showed that various mutations leading to amino acid substitutions are required in the IMH3 marker copy to confer mycophenolic acid resistance (Beckerman et al., 2001; Köhler et al., 2005), potentially explaining the highly variable natural sensitivity/resistance toward this drug among Candida species/isolates (Beckerman et al., 2001; Köhler et al., 2005). Nevertheless, this dominant selectable marker has been used over the last decade in a small number of C. albicans genetic studies and has been successfully used to transform strains of C. tropicalis (Beckerman et al., 2001), C. dublindiensis (Staib et al., 2000b, 2001; Wirsching et al., 2001), C. parapsilosis (Gács et al., 2005; Kosa et al., 2007) and C. famata (Dmytruk et al., 2011).

Reporter genes

Fluorescent proteins

Fluorescent proteins (FPs, Fig. 2a) are powerful tools in cell biology since they are often used as transformation reporters (Fig. 2b) as well as for monitoring both protein localization/interaction (Fig. 2c) and promoter activity (Fig. 2d) in a wide range of fungal species. Over the last 15 years, various FP re-engineered sequences with optimized codons and improved brightness have been developed for expression in Candida species (Table 2).

Cormack et al. (1997) published the first version of jellyfish (Aequorea victoria) green fluorescent protein (GFP) with Candida-optimized codons and enhanced brightness. This
A crucial step was then achieved three years later following the development of the yellow fluorescent protein (YFP) and cyan fluorescent protein (CFP) variants (Gerami-Nejad et al., 2001). Using a site-directed mutagenesis approach, the authors modified various amino acid codons of the yEGFP3 gene (Cormack et al., 1997) to produce a YFP (V68L, T203Y substitutions) and a CFP (F64L, Y66W, N146I, M153T, V163A and N164H substitutions). These variants were originally engineered for use in mammalian cells (Heim & Tsien, 1996). In contrast to the strategy detailed previously (Cormack et al., 1997; Morschhäuser et al., 1998), which expressed the GFP variant using autoreplicating plasmids, Gerami-Nejad et al. (2001) used a strategy that allowed GFP, CFP and YFP gene tagging through chromosomal integration of the FP fusion cassettes. This approach allowed the expression of the fusion protein to be visualized from their native promoters but was restricted to C-terminal fusions. Thanks to the generation of strains carrying various pairs of FP-tagged C. albicans proteins, these authors determined that YFP and CFP fusion proteins could be observed sequentially within a single yeast cell. These experiments represented pioneering work utilizing FP dual labelling in the CTG clade, although this strategy appeared unsuitable for N-terminal fusion and could be impaired for genes poorly expressed in the cell. In order to overcome these limitations, Gerami-Nejad et al. (2004) developed additional cassettes allowing N-terminal FP fusions. Furthermore, the presence of the pMET3, pGAL1 or pPCK1 regulatable promoters in this vector set (described in this study) facilitated controllable expression of the fusion proteins at their targeted locus.

Until 2008, the fourth colour from the repertory of FPs, i.e. the monomeric Discosoma sp. red fluorescent protein
Table 2. Compilation of reporter genes adapted for Candida species

No. of CTG codons, number of CTG codons in the original nucleotide sequence; CTG repl., indicates if the CTG codons from the original nucleotide sequence have been replaced; Codon opt., indicates if a series of codon modifications guided by a codon usage table derived from highly expressed genes has been carried out; NS, not specified; NP, native promoter.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene product</th>
<th>Species</th>
<th>No. of CTG codons</th>
<th>Name of adapted version</th>
<th>CTG repl.</th>
<th>Codon opt.</th>
<th>Promoter/terminator Used in</th>
<th>Used in</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
<td>Aequorea victoria</td>
<td>1</td>
<td>yEGFP3</td>
<td>Yes</td>
<td>Yes</td>
<td>CaADH1/CaADH1 CaMAL2/CaMAL2 PstADH2/PaADH2 CpGAL1/CpGAL1 CguiPGK/CguiPGK</td>
<td>C. albicans</td>
<td>Cormack et al. (1997)</td>
</tr>
<tr>
<td>RFP</td>
<td>Red fluorescent protein</td>
<td>Discosoma sp.</td>
<td>5</td>
<td>Venus YFP yEmRFP</td>
<td>No</td>
<td>No</td>
<td>CaADH1/CaADH1 CguiPGK/CguiPGK NP/SsURA3</td>
<td>C. dubliniensis</td>
<td>Staub et al. (2000b)</td>
</tr>
<tr>
<td>LUC</td>
<td>Luciferase</td>
<td>Renilla reniformis</td>
<td>0</td>
<td>DsRFP</td>
<td>Yes</td>
<td>No</td>
<td>GaWH11/GaWH11 CaACT1/CaPGA59 CaMAL2/CaWH11</td>
<td>C. albicans</td>
<td>Srikantha et al. (1996)</td>
</tr>
<tr>
<td>Photinus pyralis</td>
<td>9</td>
<td>FLUC</td>
<td>Yes</td>
<td>No</td>
<td>GaACT1/CaPGA59 CaEN01/CaACT1 CaHWP1/CaACT1</td>
<td>C. albicans</td>
<td>Doyle et al. (2006a, b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gausia princeps</td>
<td>3</td>
<td>gLUC59</td>
<td>Yes</td>
<td>Yes</td>
<td>GaACT1/CaPGA59 CaEF1a/CaPGA59 CaHWP1/CaACT1</td>
<td>C. albicans</td>
<td>Enjalbert et al. (2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAC4</td>
<td>β-Galactosidase</td>
<td>Kluyveromyces lactis</td>
<td>2</td>
<td>LAC4</td>
<td>No</td>
<td>No</td>
<td>CaACT1/CaACT1 CmPGK/NS GpGAL1/CpGAL1</td>
<td>C. albicans</td>
<td>Leuker et al. (1992)</td>
</tr>
<tr>
<td>lacZ</td>
<td>β-Galactosidase</td>
<td>Streptococcus thermophilus</td>
<td>1</td>
<td>lacZ</td>
<td>Yes</td>
<td>No</td>
<td>CaMAL2/CaMAL2</td>
<td>C. albicans</td>
<td>Uhl &amp; Johnson (2001)</td>
</tr>
</tbody>
</table>
(mRFP), had still not been adapted for the expression in the CTG clade. Kepler-Ross et al. (2008) reasoned that the low expression level of the mRFP in C. albicans might be due to the particular codon usage encountered in many Candida species, particularly unread CTG codons and synonymous codons of yeast that are biased toward an A or T at the third position. Ninety-eight per cent of the mRFP codons were shown to contain either a C or a G at the third position. Thus, 218 of 237 codons were modified in the mCherry variant of mRFP (Shaner et al. 2004) to produce a yeast-enhanced mRFP (yEmRFP). Expressed through C. albicans auto-replicating sequence (ARS)-containing plasmids and driven by the C. albicans ADH1 promoter, this enhanced yEmRFP was characterized by high fluorescence detectable at the blastospore and the hyphal stages.

Gerami-Nejad et al. (2009) also reported the construction of a codon-modified mRFP gene, named DsRFP. As with the previously developed yEmRFP, the codon-modified DsRFP signal was detectable in transformants at the colony level. These two later technical advances appeared crucial since they provided a second possibility for FP dual labelling for Candida cell biology by coupling expression of the yEmRFP and the already available yEGFP3 (Cormack et al., 1997; Morschhäuser et al., 1998). Finally, Reijnst et al. (2011) demonstrated yEmRFP/yEGFP3 dual labelling in C. albicans following the co-expression of two eisosome-associated protein fusions. Furthermore, in the same study, the expression of the Venus YFP variant, containing the mutations F46L, F64L, S65G, V68L, Q69M, S72A, Q80R, M153T, V163A, S175G and T203Y compared to the wild-type A. victoria GFP, was carried out for the first time in a Candida species, adding a fifth FP variant usable in the CTG clade. More recently, Milne et al. (2011) developed a series of plasmids allowing the transfer of GFP, YFP, CFP and RFP fusion constructs into C. albicans wild-type strains utilizing the CaNAT1 drug-resistance cassette (Shen et al., 2005) as a selectable marker.

Interestingly, since the development of the first codon-optimized GFP for C. albicans (Cormack et al., 1997; Morschhäuser et al., 1998), FP fusion strategies have only been employed in a few other species belonging to the CTG clade, including C. dubliniensis (Staib et al., 2000b), P. stipitis (Passoth et al., 2003) and C. parapsilosis (Kosa et al., 2007). More recently Courdavault et al. (2011) established a four-colour FP imaging protocol in C. guilliermondii (Fig. 3a) using both mCherry/GFP and YFP/CFP dual labelling (Fig. 3b–d). Additionally, a new GFP variant, CaGFP (including mutations F64L, S65C, V163A and I167T), with improved photo-stability is now available for expression in Candida species (Zhang & Konopka, 2010).

Luciferases

Luciferases are convenient and sensitive reporters for studying gene regulation in many organisms. To date, three types of luciferase are used in bacteria, animals and viruses (Fig. 4a). The firefly (Photinus pyralis) luciferase, which converts the substrate luciferin to oxyluciferin in an ATP-dependent manner, remains the most commonly used luciferase due to various favourable pharmacokinetic parameters. The sea pansy (Renilla reniformis) luciferase and the copepod Gaussia (Gaussia princeps) luciferase, which both produce light emission from coelenterazine in an ATP-independent manner, remain less employed because they usually show limited biodistribution in models and oxidation in serum. Basically, the luminescence of these proteins can be used to analyse cell transformation (Fig. 4b) and promoter regulation/function (Fig. 4c) as previously described for FPs, but they can also be used for in vivo analysis (Fig. 4d). However, the adaptation and use of these reporter systems in Candida species (Fig. 4, Table 2) has been impeded by two main factors: the specific codon usage that modifies the sequence of the expressed heterologous proteins in Candida yeast cells and the restricted permeability of fungal cells, in particular hyphal forms, to the bioluminescent substrates (d’Enfert et al., 2010).

Since the luciferase gene of R. reniformis contains no CTG codon, Srikantha et al. (1996) expressed the luciferase gene in C. albicans and studied its use as a bioluminescent reporter system. Even though it was clearly demonstrated that the activity of this R. reniformis luciferase could be assayed in lysed or intact cells, the presumed low accessibility of the substrate coelenterazine to the cell cytoplasm appeared to be a major drawback for the

---

**Fig. 3.** Application of codon-optimized CFP, GFP, YFP and mCherry genes in a Candida species: C. guilliermondii. (a) Use of FPs as a series of four fluorescent visual selection markers. Cells from GFP-, YFP-, CFP- or mCherry-expressing strains were mixed in distilled water and plated. After 2 days of growth, the freshly formed colonies were observed by epifluorescence microscopy using individual filters. Bar, 1 mm. (b) Subcellular localization of CFP-Skn7p (nuclear, nuc) and YFP-Tub1p (mitotic spindle, tub) fusion proteins in the same cell. (c) Subcellular localization of Abg1p-mCherry (vacuolar, vac) and Ypd1p-GFP (nucleocytosol, cyto) fusion proteins in the same cell. (d) Subcellular localization of YFP-SKL (peroxisomal, perox) and tGcf1p-CFP (mitochondrial, mito) fusion proteins in the same cell. Bars in (b–d), 2 μm.

For further details, see Courdavault et al. (2011).
sensitivity of the system during in vivo experiments. Ten years later, Doyle et al. (2006a) modified the nine CTG codons of the firefly luciferase gene to TTG and attempted its expression in C. albicans, driven by the C. albicans ENO1 gene promoter and ACT1 terminator. In spite of the demonstration that this adapted cassette could be used in C. albicans as a selectable marker, the limiting factor once again appeared to be the substrates availability within the cells due to the low permeability of luciferin through the cell wall, precluding its use for in vivo experiments. However, a later study by Doyle et al. (2006b) reported the first use of a bioluminescent fungal pathogen in mouse infection models, using the previously described P. pyralis luciferase gene expressed by C. albicans cells. In these latter experiments, the luciferin had to be externally supplied to the Candida cells following injection of the substrate into the infected mice. Although this approach was clearly successful in the vaginal candidiasis model, this was not the case for the chronic systemic model of candidiasis, thus limiting the widespread use of such an approach (Doyle et al., 2006b). To address the limitations of the available luciferase reporter systems (in particular luciferase/substrate accessibility), Enjalbert et al. (2009) developed a C. albicans-adapted luciferase reporter system consisting of the expression of a codon-optimized G. princeps luciferase gene fused to the recently characterized glycosylphosphatidylinositol-linked cell wall protein Pga59p (Moreno-Ruiz et al., 2009). This construction circumvents the low cell permeability to luciferase substrates by exposing the fusion proteins to the cell surface, which in turn allows efficient accessibility to external coelenterazine. The authors demonstrated that this luciferase reporter system shows high sensitivity compatible with in vitro gene expression studies and allows the progression of C. albicans infections in live animals to be monitored (except for haematogenously disseminated candidiasis). It is thus possible that this newly adapted luciferase will be useful in future Candida pathobiology studies.

\textbf{\textit{b}}-\textit{Galactosidases}

\textit{E. coli lacZ}, encoding a \textit{b}-galactosidase, has proved to be a highly versatile reporter gene in various models, including bacteria, animal cells and \textit{S. cerevisiae}, and has been used to study many aspects of gene regulation, signal-transduction pathways and other cellular processes. Unfortunately, the coding sequence of the \textit{E. coli lacZ} gene contains 51 CTG codons, which has for a long time prevented the adaptation of this reporter system to Candida species. Leuker et al. (1992) developed the first \textit{b}-galactosidase reporter system in \textit{C. albicans} and \textit{C. tropicalis} by expressing the \textit{Kluyveromyces lactis} LAC4 gene encoding a \textit{b}-galactosidase (Table 2). Placed under the control of the ACT1 promoter, the expression of this gene, which contains two CTG codons, gave rise to a detectable \textit{b}-galactosidase activity that was demonstrated to be proportional to the number of integrated copies of the reporter.
cassette. Since its development, this K. lactis LAC4 reporter system has been used to identify a small number of gene expression studies (Delbrück & Ernst, 1993; Stoldt et al., 1997; Leuker et al., 1997). Although the adaptation of the K. lactis LAC4 gene was also carried out in other Candida species (Masuda et al., 1994; Hará et al., 2001; Kosa et al., 2007), the reliability of this system was called into question by Uhl & Johnson (2001). Using the same conditions, they provided firm evidence that no β-galactosidase activity could be detected using the previously described K. lactis LAC4 reporter system. This notably could be explained by the fact that one of the two leucines encoded by a CTG codon in the K. lactis LAC4 sequence occupies a conserved position requiring a hydrophobic residue in the β-galactosidase family, which could not be converted to serine without a loss of activity. Interestingly, in the same study, Uhl & Johnson (2001) successfully used the lacZ gene from the thermophilic bacterium Streptococcus thermophilus, which contains only one CTG codon in the extreme non-conserved C-terminus of the protein (Table 2). Furthermore, they demonstrated that the wild-type sequence of S. thermophilus lacZ or the codon-optimized version (in which the sole CTG has been replaced by TTA) resulted in equivalent activity in C. albicans. These latter results could explain why the S. thermophilus lacZ gene has been preferentially used in studies of signal-transduction pathways and gene regulation in C. albicans over the last decade (Li et al., 2004; Ramsdale et al., 2008).

Other genetic tools
Recombinases
Since their discovery in the mid-1980s, site-specific recombinases have become convenient molecular tools to achieve targeted genetic modifications in a wide variety of prokaryotic and eukaryotic genomes. To benefit from this strategy, two major recombinases have been engineering over the last 12 years (codon-optimized and with transcription-regulating sequences) to exchange their use in the CTG clade: the FLP/FRT and the Cre-loxP site-specific recombinases (Nunes-Düby et al., 1998; Austin et al., 1981; Sternberg & Hamilton, 1981) (Table 3).

The first recombinase gene specifically engineered for expression in a Candida species was that encoding the FLP recombinase, originally discovered in S. cerevisiae (Vetter et al., 1983). Staib et al. (1999) proposed the first codon-modified version of the S. cerevisiae FLP gene (referred to as caFLP), in which the three CTG codons were replaced by TTG codons. caFLP was driven by the SAP2 (encoding a virulence associated secreted aspartyl proteinase) promoter and ACT1 terminator to monitor, in an infection-stage-specific fashion, the excision of the IMH3 mycophenolic acid resistance gene flanked by direct repeats of the minimal FLP recombination target sequence (FRT). In the same year, the SAP2 promoter-driven caFLP, coupled with the URA3 metabolic marker to produce a C. albicans URA3 flipper cassette, was used to generate homozygous C. albicans mutants disrupted for the CDR4 gene (encoding an ABC transporter) and the MDR1 gene (encoding a membrane transport protein of the major facilitator superfamily), indicating the capacity of this recombinase system to carry out multiple gene disruption in C. albicans (Morschhäuser et al., 1999). Staib et al. (2000a) then developed a second version of the codon-modified FLP gene, referred as caFLP, including the four amino acid exchanges described by Buchholz et al. (1998) in the caFLP sequence (P2S, L33S, Y108N and S294P). These modifications were shown to confer an enhanced activity upon the enzyme.

Michel et al. (2002) showed that placing caFLP under the control of the SAP2 inducible promoter (more precisely the C. albicans URA3 flipper cassette) allowed the development of a strategy for the identification of essential genes in C. albicans. As a continuation of this pioneering work demonstrating the potential of the codon-modified FLP gene, two studies proposed an alternative inducible promoter, MAL2 (replacing the SAP2 promoter), to drive the recombinase gene (Sánchez-Martínez & Pérez-Martín, 2002; Reuss et al., 2004). In these approaches, FLP expression could be induced by growth of the transformants in maltose-containing medium. Sánchez-Martínez & Pérez-Martín (2002) first developed a strategy permitting the integration/excision of foreign DNA using a MAL2 promoter-driven caFLP. In addition, Reuss et al. (2004) developed the well-known caSAT1 flipper cassette consisting of the MAL2 promoter-driven caFLP coupled with the caSAT1 marker and flanked by direct repeats of the FRT sequence. After the selection of nourseothricin-resistant transformants (insertion of the caSAT1 flipper sequence), FLP-mediated excision of the caFLP-caSAT1 cassette from the genome could be achieved by culturing the transformants in maltose-containing medium. More recently, Park & Morschhäuser (2005) developed a tetracycline-inducible version of the caFLP system (see next section). The caSAT1 flipper cassette of Reuss et al. (2004) was adapted by Ding & Butler (2007) to efficiently transform C. parapsilosis clinical isolates. The authors were unable to use the caSAT1 flipper construct directly. To obtain maltose-induced excision of the cassette, they therefore replaced the promoter driving expression of the caFLP gene by the C. parapsilosis MAL2 promoter (CpFLP). This CpSAT1 flipper cassette was then used to sequentially generate a ura3/ura3 auxotrophic strain and to disrupt the BCR1 gene (involved in biofilm formation). In contrast to Ding & Butler (2007), Gäsper et al. (2007) and Nguyen et al. (2009) successfully achieved sequential disruptions with FLP-mediated excision of the cassette in C. parapsilosis, using the caSAT1 flipper cassette directly (without changing the regulatory sequences).

The second type of recombinase, Cre, was adapted for use in the CTG clade in 2005. Dennison et al. (2005) produced the first version of a synthetic codon-modified cre gene in which all the 18 CTG codons were replaced by TTG and driven by the inducible C. albicans MET3 promoter. To
Table 3. Compilation of recombinases and other regulatable systems adapted for *Candida* species

No. of CTG codons, number of CTG codons in the original nucleotide sequence; CTG repl., indicates if the CTG codons from the original nucleotide sequence have been replaced; Codon opt., indicates if a series of codon modifications guided by a codon usage table derived from highly expressed genes has been carried out; NS, not specified.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene product</th>
<th>Species</th>
<th>No. of CTG codons</th>
<th>Name of adapted version</th>
<th>CTG repl.</th>
<th>Codon opt.</th>
<th>Promoter/terminator</th>
<th>Used in</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLP</td>
<td>Recombinase</td>
<td>Saccharomyces cerevisiae</td>
<td>3</td>
<td>caFLP</td>
<td>Yes</td>
<td>No</td>
<td>CaSAP2/CaACT1</td>
<td><em>C. albicans</em></td>
<td>Staib <em>et al.</em> (1999); Morschhäuser <em>et al.</em> (1999)</td>
</tr>
<tr>
<td>cre</td>
<td>Recombinase</td>
<td>P1 bacteriophage</td>
<td>18</td>
<td>CpFLP</td>
<td>Yes</td>
<td>No</td>
<td>CaSAP2/CaACT1</td>
<td><em>C. albicans</em></td>
<td>Staib <em>et al.</em> (2000a); Reuss <em>et al.</em> (2004)</td>
</tr>
<tr>
<td>tTA</td>
<td>Tetracycline-dependent transactivator</td>
<td>Escherichia coli</td>
<td>1 (3)</td>
<td>tTA‡</td>
<td>Yes</td>
<td>No</td>
<td>CaENO1/CaWH11</td>
<td><em>C. albicans</em></td>
<td>Nakayama <em>et al.</em> (2000)</td>
</tr>
<tr>
<td>tetR</td>
<td>Tetracycline-dependent transactivator</td>
<td>Escherichia coli</td>
<td>1 (1)</td>
<td>carT‡</td>
<td>Yes</td>
<td>No</td>
<td>CaADH1/CaACT1</td>
<td><em>C. albicans</em></td>
<td>Park &amp; Morschhäuser (2005)</td>
</tr>
</tbody>
</table>

*This version includes the four amino acid exchanges P2S, L33S, Y108N and S294P described by Buchholz *et al.* (1998).

†tTA consists of the codon-modified *E. coli* tetR gene fused to the codon-modified nucleotide sequence encoding the *S. cerevisiae* Hap4 activation domain. The tetR gene contains one CTG codon and the nucleotide sequence encoding the ScHap4 activation domain contains three CTG codons.

‡carT is composed of the reverse codon-modified *E. coli* tetR gene (carTetR) fused to the codon-modified nucleotide sequence encoding the *S. cerevisiae* Gal4 activation domain. The carTetR gene contains one CTG codon and the nucleotide sequence encoding the ScGal4 activation domain contains one CTG codon. The carTetR gene product contains the five amino acid exchanges (S12G, E19G, A56P, D149E and H179R) described by Urlinger *et al.* (2000).
demonstrate its potential in C. albicans, they carried out the sequential disruption of the ADE2 and MET15 genes. One year later, Laplaza et al. (2006) engineered a second version of a codon-modified cre gene by changing the 18 CTG codons to various Candida frequently used codons. Since their development, the codon-modified FLP and Cre systems have been rarely employed in Candida species. Nevertheless, because FLP- or Cre-mediated multiple gene disruption strategies do not always require the handling of toxic compounds, unlike the widely used URA3 blaster systems (utilizing the harmful 5-fluoroorotate), their use may well become more widespread in future.

Tetracycline-regulatable systems

The TetR/tetO tetracycline-regulatable system has emerged as a powerful system to control gene expression in prokaryotic and eukaryotic cells. Originally, the archetypal E. coli TetR/tetO system was composed of two modules: the TetR transactivator and the tetO operator from the promoter of tetracycline-resistance genes. In the absence of tetracycline, TetR binds the tetO sequence and represses the expression of fused tetracycline-resistance genes. When added, tetracycline binds the TetR repressor and provokes its dissociation from the tetO operator, inducing the expression of tetracycline-resistance genes (Gossen et al., 1993) (Table 3).

The first tetracycline-regulatable system usable in Candida species was adapted by Nakayama et al. (2000) (Fig. 5a). This system, referred to as Tet-off, consists of the codon-modified E. coli TetR fused to the activation domain of the Hap4 transcriptional activator (tTA) and the TR promoter, including a yeast minimal promoter element coupled with a tetracycline operator sequence (tetO). The four CTG codons of the gene encoding the fusion transactivator tTA were replaced by other leucine codons. In this regulatable system, in the absence of tetracycline, the tTA transactivator binds the tetO sequence, allowing the active expression of the target gene. The addition of tetracycline inhibits the binding of tTA to tetO, thus turning off the target gene expression (Fig. 5a). This Tet-off system has become a versatile tool to investigate the function of essential genes (which cannot be deleted from the genome; e.g. Yang et al., 2006) as well as virulence-associated genes (Saville et al., 2003). In this way, Roemer et al. (2003) used this latter development to design the C. albicans GRACE strategy.

Park & Morschhäuser (2005) proposed a second tetracycline-regulatable system referred to as the Tet-on system, consisting of a reverse tetracycline-controlled transactivator (cartTA) and the Ptet promoter (Fig. 5b). The codon-modified cartTA gene was composed of the cartTetR nucleotide sequence fused to the sequence encoding the S. cerevisiae Gal4 activation domain (displaying a CTG to TTG codon replacement). The cartTetR sequence was generated following codon modifications in the TetR gene introducing the five amino acid exchanges (S12G, E19G, A56P, D149E and H179R) described by Urlinger et al. (2000) to invert the effect of tetracycline on the expressed transactivator cartTA. The Ptet promoter includes a minimal promoter element from the C. albicans OP4 gene fused to the tetO tetracycline operator sequence. In this latter strategy, the expression of a target gene can be accurately positively regulated by the addition of tetracycline to engineered cells (Fig. 5b). Since its construction, this regulatable system has facilitated a number of investigations into the role of specific genes on the behaviour of C. albicans (Bernardo et al., 2008; Saville et al., 2006; Staib et al., 2008; Carlisle et al., 2009; Becker et al., 2010).

Concluding remarks

Candida species are attractive models for exploring new molecular mechanisms relating to pathogenicity, antifungal resistance, metabolite bioproduction and biological control. As a consequence, and predominantly since the 1990s, the establishment of convenient transformation systems as well as the development of selectable markers and reporter genes has become indispensable to carry out genetic manipulations in these yeast species. As underlined in this paper, the establishment of such molecular tools was for a long time impeded due to the codon usage in the CTG clade. However, as soon as PCR-mediated mutagenesis techniques became available at the end of the 1990s, the Candida researcher community (particularly C. albicans researchers) progressively adapted Saccharomyces genetic tools for use in Candida species. Recently, this tendency was once again illustrated by the adaptation of the Saccharomyces two-hybrid system for use in C. albicans (Stynen et al., 2010). After 15 years of research, the Candida clade toolbox is now considered to be reasonably well developed. Some research groups currently propose global strategies employing a wide range of dominant markers, reporter genes and other protein tags for use in particular Candida species, as is the case for the PCR-amplified functional analysis (PFA) cassette series for C. albicans (Gola et al., 2003; Schaub et al., 2006; Reijnst et al., 2011; Milne et al., 2011).

As highlighted in the present review, three major genetic modifications have to be carried out for the adaptation of selectable markers, reporter genes or regulatable systems for their expression in the CTG clade. First, the substitution of all CTG codons of the expressed sequence by other leucine-encoding codons appears indispensable in all studies, with the exception of CTG codons located at the extreme 3′ end of the sequence and encoding a non-conserved leucine residue (Uhl & Johnson, 2001). Secondly, in the case of a complete synthetic sequence, a series of codon modifications guided by a codon usage table derived from highly expressed genes in the target species can act positively on the expression of the sequence (Cormack et al., 1997; Shen et al., 2005). For example, the nucleotides G or C should be omitted at the third position
of each codon for expression in C. albicans (Keppler-Ross et al., 2008). Finally, the suitability of Candida transcription-regulating 5' and 3' regions used to drive the expression of a heterologous codon-optimized sequence in various species of the CTG clade remains controversial. As an example, although the C. albicans ACT1 promoter was shown by different teams to be able to drive the expression of various codon-modified sequences in a series of Candida species (Gácser et al., 2007; Nguyen et al., 2009), this same promoter proved ineffective for the expression of heterologous genes in the hands of other researchers (Ding & Butler, 2007; Millerioux et al., 2011).

In conclusion, it is possible that the range of Candida-adapted molecular tools will increase thanks to two major aspects. First, the progressive reduction in the price of selection drugs (hygromycin B, nourseothricin, zeocin...), substrates (luciferin, coelenterazine...), and automated gene synthesis will foster the development of genetic tools that will become commonplace in research laboratories for a number of Candida species. Furthermore, the growing set of sequenced Candida genomes (Butler et al., 2009) and the availability of databases of metabolic pathways (Rossignon et al., 2008; Skrzypek et al., 2010) will facilitate the isolation of homologous constitutive or regulatable promoters and terminators for use in specific Candida species. In a wider perspective, the increasing number of sequenced genomes (specifically prokaryotic genomes) will help in identifying species/isolates harbouring interesting CTG codon-less coding sequences.

Acknowledgements

We acknowledge the Broad Institute Fungal Genome Initiative for making complete genome sequence of various Candida species available. We thank 'Le STUDIUM' (Centre region, France) for the financial support of A. J. Simkin.

References


