國立交通大學

分子醫學與生物工程研究所

碩士論文

熱帶念珠菌藉由在 *FCY2* 基因上的無意義突變伴隨異質性缺失的發生獲得抗氟胞嘧啶的能力

Development of flucytosine resistance in *Candida tropicalis* is caused by nonsense mutations in *FCY2*, followed by loss of heterozygosity

研 究 生 ： 陳妍寧

指導教授 ： 楊昀良博士

中 華 民 國 九 十 九 年 七 月

熱帶念珠菌藉由在 *FCY2* 基因上的無意義突變伴隨異質性缺失的發生獲得抗氟胞嘧啶的能力

Development of flucytosine resistance in *Candida tropicalis* is caused by nonsense mutations in *FCY2*, followed by loss of heterozygosity

研究生 ： 陳妍寧　　Student : Yen-Ning Chen

指導教授 ： 楊昀良 博士　　Adviser : Dr. Yun-Liang Yang

國 立 交 通 大 學

分子醫學與生物工程研究所

碩 士 論 文

A Thesis

Submitted to Institute of Molecular Medicine and Bioengineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

in

Molecular Medicine and Bioengineering

July 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

摘要

熱帶念珠菌(*Candida tropicalis*)是一種雙倍體的病原真菌。在念珠菌屬中，熱帶念珠菌極具侵略性且是最常見的非白色念珠菌之一。先前的實驗中發現 31.25%對 flucytosine (5FC)敏感的臨床菌株具有產生抗藥性子代的能力，並進一步證實所觀察到的 5FC 抗藥性與帶有特定一群多型性核苷酸(SNP)的 *FCY2* 基因(會產生運輸 cytosine 的蛋白質)處在同質二倍體的狀態(homozygous state)有關。本研究將針對熱帶念珠菌對 5FC 的抗藥機制做更深入的探討。首先，利用另一個關係較遠的臨床菌株建構 *FCY2* 同質二倍體的突變株，在 5FC 藥物敏感性測試實驗中顯示，只有跟上述帶有同樣一群多型性核苷酸的同質二倍體突變株具有抗此藥物的能力。接著分析對 5FC 極為敏感的臨床菌株在 *FCY2* 基因上的各個多型性核苷酸，發現有三個多型性核苷酸只有在具有抗 5FC 能力的菌株中才存在，經建構這三個多型性核苷酸的點突變株，顯示 G145T 的點突變加上 *FCY2* 基因處在同質二倍體的狀態(145 T/T)與衍生株的抗藥性有關。這個 G145T 的點突變會造成無意義突變(nonsense mutation)使得 *FCY2* 不會合成具有功能的運輸蛋白產物(purine-cytosine permease)。最後，分析全部具有產生抗藥性子代能力的臨床菌株的基因型，發現大部分的親代菌株都帶有 145 T/G 這種基因型。除了分析 *FCY2* 基因與抗 5FC 藥物之間的關係，本研究也針對兩隻臨床菌株和它們的衍生株之間所發生的異質性缺失(loss of heterozygosity; LOH)做探討。藉由基因多型性圖譜觀察到多數的 LOH 涵蓋了大部分染色體的區域，甚至延伸到其端粒，這些 LOH 可能是由於 allelic recombination 或者 break-induced replication 所造成。另外有一個小範圍的 LOH 則可能是由 gene conversion 所產生。總結，在 *FCY2* 基因中產生的無意義突變伴隨大範圍 LOH 的發生是造成本實驗中熱帶念珠菌衍生株產生抗 5FC 的主要原因。

Abstract

Candida tropicalis, a diploid yeast, has been regarded as one of the most invasive and prevalent species of non-albicans *Candida*. Previous work has shown that 31.25% (30/96) of flucytosine (5FC) susceptible clinical isolates could generate drug resistant progeny. It has revealed that there was an association between a set of polymorphic nucleotides in *FCY2* (encoding a cytosine transporter) and the observed drug resistant phenotype. In this study, the molecular mechanisms of 5FC resistance in these *C. tropicalis* have been further investigated. First of all, homozygous mutants carrying the same set of SNPs in *FCY2* under different genetic background displayed similar levels of resistance to 5FC. Second, nucleotide sequences of the *FCY2* alleles in hyper-susceptible strains indicated three SNPs in *FCY2* were unique to those resistant strains. Subsequently, a homozygous SNP at position 145 (T/T) in the *FCY2* gene was proven to be fully accounted for the 5FC resistance observed in the drug resistant derivatives. Since the G145T mutation resulted in a nonsense mutation, it might lead to the loss of purine-cytosine permease. Finally, two types of nonsense mutations (G145T and G201A) were found in *FCY2* of the clinical isolates that had potential to produce 5FC resistant derivatives, where the majority was 145 T/G genotype. Meanwhile, the nature and the extent of loss of heterozygosity (LOH) events have been examined by SNP mapping. The results showed that most of the LOH events covered large chromosomal regions, which even extended to the telomere, suggesting allelic recombination or break-induced replication, whereas one was a localized LOH event, caused by gene conversion. All in all, nonsense mutations in *FCY2* coupled with LOH events are responsible for 5FC resistance in those *C. tropicalis* strains.

誌謝

很幸運在碩班三年學習的道路上能夠有這麼多良師益友相伴，由於你們不吝惜的指導與鼓勵，讓我不但順利取得學位，更獲得許多難能可貴的人生經驗。對於大家的照顧與恩惠，我將銘記在心。

首先，感謝指導教授楊昀良老師給予我在研究上極大的揮灑空間，並適時分享個人的經驗以勉勵學生更加精進，此外還耐心的修改我不純熟的英文論文。同樣地，很感激羅秀容老師不但讓我有機會在國衛院做實驗，更提供我許多研究上的方向與想法，在羅老師身上，我見識到身為一個科學人的熱誠。還要謝謝口試委員藍忠昱老師和梁美智老師給予我許多的寶貴建議，讓我的論文得以更加完備。

接著，感謝所有國衛院的伙伴們，讓我在竹南的日子過得極為充實快樂。文靜又調皮的惠菁學姐；無論身在何地總讓人很有安全感的佳真；率直可愛的小善；很有媽媽的味道的向寧；讓人不想直接誇獎她的小芭樂瑋婷；和我有麥當勞緣的琬立；松鼠、海狸、野雁精神兼具的啟宏；看似溫和實則極具行動力的志兆；名不虛傳的偉得王誌偉；廚藝極佳又很貼心的好鄰居德斌；總是很有朝氣的昱果；說出來的話總讓我嚇一跳的馬靈；有義氣也很有勇氣的奕瑄；笑容可掬的天竺鼠裕傑。

再來要感謝交大實驗室的學長姐以及一群可愛的學弟妹們，讓我在新竹的日子回味無窮，而在竹南時也無後顧之憂。在實驗室和我一同奮鬥到清晨，又帶我去南寮漁港遊玩的淑萍；親切可人又有一副好歌喉的淑貞；熱心服務的敏書；唱歌很會帶動氣氛的旻秀；對於實驗問題總是親切給予指導的志達；口試期間幫了我大忙的好戰友小倩；認真努力的好鄰居馨儀；有著燦爛微笑的阿毛；大家公認的陽光男孩毓駿；對研究極為用心的愛貓學弟重延；有著一股純真氣質的禎憶；總是逗得大家很開心的開心果阿大；積極進取的漂亮學妹幸璇；很有趣的助理阿金吳；乖巧有禮貌的學弟阿賢；化學一把罩也很搞笑的凱薩；很有自己想法的阿白；溫柔有氣質的于鈴學姐；修課的好夥伴音汝。

還要感謝以前工作時的同事們、大學同學以及大學實驗室的同伴們對我的關心以及協助；321 室最親愛的室友芷婷；越陳越香的老朋友們小碗、豬豬、杜杜、小葉、阿福、孟芯，最後，要感謝我最愛的家人，由於你們的支持與鼓勵，讓我得以全心投入做自己想做的事。

Contents

Chapter 3. Results

Contents (Tables)

Contents (Figures)

Appendix

Chapter 1. Introduction

1.1 Clinical significance of *Candida tropicalis*

The prevalence and the severity of fungal infections have significantly increased over the past two decades (Beck-Sague and Jarvis, 1993; Chen *et al.*, 1997; Fridkin and Jarvis, 1996). In particular, *Candida* species, leading to a high mortality rate of patients with candidemia in a range from 35% to 60%, has emerged as the most common cause of opportunistic fungal infections worldwide (Ruan and Hsueh, 2009). Although *Candida albicans* is the major pathogen that causes candidemia, the contributions of other *Candida* species are rising, especially in seriously ill patients. Among the non-albicans *Candida* species, *Candida tropicalis*, *Candida glabrata* and *Candida parapsilosis* are the three most important pathogens responsible for the infections. According to the reports of Taiwan Surveillance of Antimicrobial Resistance of Yeasts (TSARY), *Candida tropicalis* has been identified as the most frequently encountered non-albicans *Candida* species in Taiwan (Yang *et al.*, 2005; Yang *et al.*, 2008). It is notable that the proportion of this pathogen in Taiwan (16-25%) seems to be higher than that in other regions (5-11%), such as North America (Ruan and Hsueh, 2009). In addition, the diploid yeast *C. tropicalis* is more often related to neutropenia and malignancy (Abi-Said *et al.*, 1997; Kontoyiannis *et al.*, 2001; Marr *et al.*, 2000; Wingard, 1995). Previous study has demonstrated that *C. tropicalis* is the second most virulent pathogen comparing to seven other medically important *Candida* species in an immunocompetent animal model (Arendrup *et al.*, 2002). Another research has implied that the gene families likely to involve in pathogenicity and virulence are enriched in *C. tropicalis*, such as genes encoding Als adhesins and Hyr/Iff proteins (Butler *et al.*, 2009). Moreover, a report has indicated the frequent appearance of drug resistance of this yeast. For example, while a 3% incidence of resistance to fluconazole was identified within the *C. albicans*, more than one half of the *C. tropicalis* isolates were resistant to this drug (Law *et al.*, 1996). Because of its high invasiveness, the high incidence of drug resistance and the affinity for causing diseases in immunocompromised patients, *C. tropicalis* has become a clinically significant pathogen (Kothavade *et al.*, 2010).

1.2 Mechanisms of action of antifungal agents

The treatment of fungal infections is a serious issue due to the increase of the immunosuppressed patients and the prevalence of drug resistance. Since fungi and human beings are both eukaryotes, the drug targets are limited to four categories.

(1) Polyene

The polyenes, such as amphotericin B, pimaricin and nystatin, attack ergosterol-containing membranes. As a consequence, pores are formed on the membranes, which cause the leakage of cellular components and, ultimately, cell death (Vanden Bossche *et al.*, 1994). Their usages are compromised by the severe and potentially lethal side effects as a result of the similarity between ergosterol in fungal cell membranes and cholesterol in mammalian cell membranes. Clinically, polyenes are used to treat the systemic fungal infections, and only when the azole treatments are failed (Louie *et al.*, 2001).

(2) Ergosterol biosynthesis inhibitors

This class of antifungals is widely used for the treatment of fungal diseases due to their relatively low cytotoxicity. Thiocarbamates (tolnaftate and tolciclate), allylamines (naftifine and terbinafine) and azoles belong to this class because they all prevent yeast growth by blocking the function of enzymes that are involved in the synthesis of ergosterol from squalene. The first two types of drugs are against squalene epoxidase, encoded by *ERG1*, and thus disturb the conversion of squalene to 2'3-oxidosqualene (Ryder, 1992). The third type of drugs, azoles, includes both imidazoles (ketoconazole and miconazole) and triazoles (fluconazole, itraconazole and voriconazole). They inhibit the function of lanosterol demethylase, a cytochrome p-450 enzyme, encoded by *ERG11*, and lead to the depletion of ergosterol as well as the accumulation of lethal byproducts, such as methylated sterol (Hitchcock, 1991; Kelly *et al.*, 1997). Since azoles are fungistatic drugs, cells become resistant to such type of drugs after long term exposure.

(3) Cell wall synthesis inhibitor

The newest class of antifungals, echinocandins (including caspofungin, micafungin and anidulafungin), interferes with cell wall synthesis by interaction with the β-1,3 glucan synthase (Radding *et al.*, 1998). As the β-1,3 glucan is an essential cell wall component, disruption of its synthesis causes loss of cell wall integrity and severe cell wall stress.

(4) Nucleic acid synthesis inhibitor

Flucytosine (5FC) is one of the oldest antifungal agents, and has an important role in the management of serious candidal sepsis. It was at first synthesized to treat tumors, but its effect on tumor treatment did not meet the expectation. After the antifungal activity of 5FC was proven in animal models, it was used to treat candidosis and cryptococcosis (Vermes *et al.*, 2000). Since 5FC is a fluorinated analogue of cytosine, it is metabolized via the pyrimidine salvage pathway, and converted from a prodrug into a harmful metabolite that inhibits fungal RNA and DNA synthesis (Appendix 1). 5FC is actively transported into cells by the enzyme purine-cytosine permease (PCP). Once inside the fungal cells, it is rapidly converted via 5-fluorouracil (5FU) to 5-fluorouridine monophosphate (FUMP), which is catalyzed by two enzymes, cytosine deaminase and uracil phosphoribosyl transferase (UPRT), respectively. FUMP is in turn phosphorylated by two kinases to 5-fluorouridine triphosphate (FUTP), which disturbs the protein synthesis by incorporation into RNA (Waldorf and Polak, 1983; White *et al.*, 1998). Alternatively, the reduction of FUMP to 5-fluoro-2'-deoxyuridylate (FdUMP) leads to the inhibition of the enzyme thymidylate synthetase, and thus DNA synthesis (Diasio *et al.*, 1978). 5FC has no direct effect on mammalian cells due to their lack of the enzyme cytosine deaminase. However, its monotherapy is limited by the prevalence of primary resistant strains and the frequent development of secondary resistance during treatment. As a result, 5FC has been used only in combination antifungal therapy (Kauffman and Carver, 1997; Sheehan *et al.*, 1999; Viviani, 1995).

1.3 Molecular mechanisms of flucytosine resistance in yeasts

Three possibilities of resistance mechanisms against 5FC have been proposed in

previous studies. First, interference of any protein involved in the pyrimidine salvage pathway could lead to 5FC resistance (Fasoli and Kerridge, 1988; Jund and Lacroute, 1970; Papon *et al.*, 2007) (Appendix 1). For example, a disruption in the *FCY1* gene, encoding cytosine deaminase, as well as a disruption in the *FUR1* gene, encoding UPRT, would inhibit the conversion of 5FC into an antifungal metabolite. Also, the disruption in the *FCY2* gene, encoding PCP, could block the uptake of 5FC into cells. Second, the interference of proteins involved in the pyrimidine de novo biosynthetic pathway, increasing the production of end products, could result in 5FC resistance because of the feedback inhibition on the salvage pathway or the substrate-enzyme competition between UMP and 5FUMP. Different degrees of the feedback inhibition between salvage and de novo pathways have been found in *Saccharomyces cerevisiae*. The presence of the exogenous uracil (an activated precursor) promotes the synthesis of nucleotide through salvage pathway and simultaneously suppresses the de novo pathway (Kern *et al.*, 1991). In contrast, the production of UMP through the de novo pathway might inhibit the salvage pathway. For instance, a disruption in the *URA3* gene, encoding orotidine 5'-phosphate decarboxylase (ODCase), might increase the synthesis of UMP via the de novo synthesis pathway. The overproduction of UMP reduces the metabolism of 5FC by competing with the harmful products (FUMP and FdUTP) for thymidylate synthetase and UMP kinase, or by allosterically inhibiting the UPRT activity (Desnos-Ollivier *et al.*, 2008; Natalini *et al.*, 1979). Third, interference of the regulatory enzymes may also cause the resistant to 5FC due to the loss of feedback inhibition of de novo pyrimidine nucleotide synthesis, enabling the overproduction of UMP (Jund and Lacroute, 1970).

1.4 The association between loss of heterozygosity and antifungal drug resistance

The linkage between genetic alterations and phenotypic changes is more complex in diploid cells than in haploid cells because most phenotypic changes observed in the diploid cells require one more genetic event. The first is to generate a genetic polymorphism, and the second is to replace the original genotype with the new one. The latter event can lead to the expression of the recessive allele or the intensification of the original effects, and is known as loss of heterozygosity (LOH).

This phenomenon has been recently reported for *C. albicans* and its resistance to azoles and micafungin (Coste *et al.*, 2007; Dunkel *et al.*, 2008; Niimi *et al.*, 2010). For example, a disruption in *TAC1* (encoding a transcriptional activator of *CDR* genes) results in constitutive high expression of ATP-binding cassette transporters, which mediates azole susceptibility. Nevertheless, homozygosity for the gain-of-function *TAC1* alleles via LOH is required for a high azole resistance levels in *C. albicans* because the hyperactive alleles are codominant with wild type alleles (Coste *et al.*, 2006). In contrast, the presence of an *MRR1* hyperactive allele (encoding a transcription factor regulating the expression of *MDR1*) leads to the upregulation of another efflux transporter, and thus mediates fluconazole susceptibility. The hyperactive allele in the presence of a wild type allele is enough to confer drug resistance in *C. albicans*; however, higher drug resistance is acquired when the presence of two instead of one such *MRR1* allele is existed (Dunkel *et al.*, 2008). Likewise, a disruption in *GSC1*, a glucan synthase catalytic subunit, alone is enough to generate micafungin resistance, yet two such *GSC1* alleles achieve a higher level of micafungin resistance (Niimi *et al.*, 2010). Previous studies have demonstrated antifungal pressure, similar to UV irradiation, may promote genome alterations, including LOH (Coste *et al.*, 2007; Dunkel *et al.*, 2008; Takagi *et al.*, 2008; Tsang *et al.*, 1999). Analysis of the serial isolates from a patient treated with fluconazole has revealed an intermediate isolate which contains a gain-of-function allele of *MRR1* in the presence of the wild-type allele and displays intermediate drug resistance, while its progeny become homozygous for the hyperactive allele and exhibit increased drug resistance (Dunkel *et al.*, 2008; Franz *et al.*, 1998). The stepwise development of drug resistance provides a competitive advantage for yeasts to adapt for the stress condition. Therefore, the acquisition of mutations in genes involved in drug resistance is, ultimately, coupled with different LOH events, such as gene conversion, allelic recombination/break-induced replication (BIR) or chromosome loss and replication (Andersen *et al.*, 2008; Coste *et al.*, 2007).

1.5 Previous works

Overall 97 clinical isolates of *C. tropicalis* were collected from 22 to 24 hospitals in Taiwan during the TSARY programs in 2002 and 2006. According to 5FC

susceptibility testing conducted by the lab fellow Hui-Ching Ko, one was resistant, and 96 remained susceptible (unpublished). Among the susceptible strains, 35 produced progeny within the E-test inhibition ellipses after 24 to 72 h of incubation. Later, 30 isolates were proven to be 5FC resistance by broth microdilution. 5FC resistance in *C. tropicalis* was first investigated in 13 clinical isolate-derivative pairs by sequencing four genes *FCY1*, *FCY2*, *FUR1* and *URA3*, which code for cytosine deaminase, PCP, UPRT, and ODCase, respectively. In six clinical isolate-derivative pairs, a correlation between particular polymorphic nucleotides and resistance to 5FC was found within *FCY2*, as well as LOH events. To examine whether the homozygosity with a set of SNPs at the *FCY2* locus was associated with the 5FC resistance observed in the drug resistant derivatives, the *FCY2* homozygous mutants and the *fcy2* single knockout strains were constructed from one of the clinical isolates, YM020291. Two distinct DNA sequences corresponding to the entire *FCY2* coding region and its flanking sequences were identified by sequencing. For conveniences, $FCY2^R$ was referred to the allele exhibiting the set of SNPs found at the *FCY2* locus carried by the 5FC resistant derivatives, whereas $FCY2^S$ was the allele exhibiting the other genotype. The MICs of 5FC for the clinical isolate YM020291 and its genetic engineered derivatives have revealed that while the $FCY2^R/FCY2^R$ homozygous mutants and the $FCY2^R/fcy2\Delta$ single knockout strains were resistant to 5FC, the $FCY2^R/FCY2^S$ heterozygous clinical isolate, the $FCY2^S/FCY2^S$ homozygous mutants and the $FCY2^S/fcy2\Delta$ single knockout strains were susceptible to 5FC. Sequence analysis showed that seven SNPs were present at the *FCY2* locus, including two at the upstream region, four at the ORF and one at the downstream of this gene. Among them, four polymorphic nucleotides were potentially involved in point mutations linked to 5FC resistance. In the following content, the positions of these polymorphic nucleotides were indicated based on the *FCY2* ORF in yeast MYA-3404 (available on the *C. tropicalis* database), and the number in parentheses represented the SNPs relative to the new coding sequence of this gene, identified in this study. Two polymorphic nucleotides were in the promoter region of *FCY2* gene, including a thymine (T) to cytosine (C) substitution at position -224 (-11) and a guanine (G) to thymine (T) substitution at position -69 (145), which might influence the expression of *FCY2* gene. The third was the substitution of guanine for thymine at nucleotide 273 (486) in the ORF, resulting in the replacement of methionine (ATG) with isoleucine

(ATT) at amino acid position 91 (162) in PCP; the last SNP was a cytosine (C) to thymine (T) substitution at position 1518 (1731) (i.e., the 201 nucleotide downstream of the stop codon of the *FCY2* gene), located in the 3' UTR, which might involve in the translation efficiency and mRNA stability (Chia-Chen Wu, 2009).

1.6 Purpose of this study

The aims of this study were to access the association between the SNPs in *FCY2* and 5FC resistance on molecular levels, and to investigate the nature and molecular basis of the LOH events occurred among two clinical isolates, YM020291 and YM060800, and their derivatives. In the beginning, the effect of the homozygous *FCY2*^{R} allele in another susceptible isolate YM060800 was tested. 5FC resistant phenotype in its *FCY2*^{R}/*FCY2*^{R} homozygous mutants showed that the contribution of the specific *FCY2*^{R} allele to drug resistance was largely comparable in the different strain backgrounds. To further identify which one (or ones) of the polymorphic nucleotides may play a role in the mechanism of 5FC resistance, the DNA sequences of *FCY2* from four highly susceptible clinical isolates were compared to those of *FCY2*^{R} and *FCY2*^{S} alleles. It showed that three SNPs [-69 (145), 273 (486) and 1518 (1731)] were unique to the *FCY2*^{R} allele and might be responsible for the resistance ability in *C. tropicalis*. Introduction of these single-base mutations into the drug susceptible isolate YM020291 revealed that the single mutation G-69T (G145T), resulting in the homozygous SNP (T/T) at position -69 (145), could be solely accounted for the level of 5FC resistance observed in most derivatives of the clinical isolate-derivative pairs. Meanwhile, SNP mapping has shown that four strains were differed by large LOH events (i.e., allelic recombination or break-induced replication), whereas one strain was differed by a localized LOH event (gene conversion).

Chapter 2. Materials and methods

2.1 Materials

2.1.1 Strains

(1) *Escherichia coli* (DH5α)

(2) *Candida* strains: Table 1

2.1.2 Plasmids and primers

Plasmids and primers used in this study are listed in Tables 2 and 3, respectively.

2.1.3 Chemicals and reagents

- **Difco Laboratories**

  Bacto agar (Cat. No. 214040); BHI (Cat. No. 0037-17); LB agar (Cat. No. 244520); LB broth (Cat. No. 244620); Sabouraud dextrose agar (SDA; Cat. No. 291940); YPD broth (Cat. No. 242820); yeast nitrogen base w/o amino acid (Cat. No. 291940)

- **Merck**

  Chloroform (Cat. No. 1.0244511000); Dimethyl sulfoxide (DMSO) (Cat. No. S26740); Sodium dodecyl sulphate (SDS) (Cat. No. 1.01116.0500); Ethanol (Cat. No. K33534874); Ethidium bromide (Cat. No. K27928515); Glucose (Cat. No. K33069537); Isopropanol (Cat. No. K32632434); Sodium acetate (Cat. No. 1.06268.0250); Sodium carbonate ($NaHCO_3$) (Cat. No. A375692); Sodium chloride (NaCl) (Cat. No. K29779304); Sodium hydroxide (NaOH) (Cat. No. B886298); Hydroxymethyl aminomethane hydrogen chloride (Tris-HCl) (Cat. No. 8382T006)

- **Gibco BRL**

  RPMI 1640 pH7.0 (Cat. No. 31800022)

- **Invitrogen**

  1 kb plus DNA ladder (Cat. No. 12308-011); Agarose (Cat. No. 242820)

- **NEB**

  Calf intestinal (CIP; M290); *Hind*III (R0104); *Hinf*I (R0155); *Kpn*I (R0189);

*Sac*I (R0157); *Sac*II (R0157); *Xho*I (R0146)

- **Promega**

  T4 DNA ligase (Cat. No. M-1801)

- **Sigma Chemical Co.**

  Disodium ethylenediaminetetraacetate (EDTA) (Cat. No. D9779); Dithiothreitol (DTT) (Cat. No. D9779); 5-Flucytosine (SI-F7129); Glassbeads (425~600μm) (Cat. No. G9268-500G), Histidine (Cat. No. H8125); Lithium acetate (LiOAC) (Cat. No. L-6883); Lithium chloride (LiCl) (Cat. No. L9650)

- **USB**

  Glycerol (Cat. No. US16374)

- **Qiagen**

  RNeasy Mini Kit (Cat.No. 74106): Buffer RLT, Buffer RW1, Buffer RPE

2.1.4 Buffers

- 10 × Tris-EDTA (10 × TE)

  100 mM Tris-HCl (pH 7.5), 10 mM EDTA (pH 7.5)

- 10 × (1M) LiOAC

  20g LiOAC dissolved in sterile ddH_2O to 200 ml (pH 7.5)

- 50 × TAE [Bio-Rad, Cat. No. 161-0773]

  2 M Tris base, 2 M acetate, 0.05 M EDTA

2.1.5 Medium

- **BHI broth**

  33.7% Calf brain infusion solid, 13.5% Beef heart infusion solids, 27% Proteose peptone, 5.4% Glucose, 13.5% Sodium chloride, 6.7% Disodium phosphate

- **LB (Luria-Bertani) broth**

  1% tryptone, 0.5% yeast extract, 1% NaCl; 1.5% agar for plate cultures

- **YPD medium**

  1% yeast extract, 2% Bacto-peptone, 2% dextrose

- **RPMI 1640 broth (pH 7.0) with 0.165 mol/L MOPS**

  10.4 g RPMI 1640, 34.53 g MOPS (3-[N-morpholino] propanesulfonic acid), 2 g $NaHCO_3$ dissolved in sterile ddH_2O to 1 L (pH 7.0)

2.1.6 Equipments

Peltier Thermal Cycler PTC-200 (MJ Research)
Orbital shaking incubator OSI500 (TKS)
AlphaImager 2000 (Alpha Innotech Corporation)
Vortex-2 genie (Scientific Industry)
Dry bath incubator (Violet Bioscience Inc.)
Brushless microcentrifuge Denville 260D (Scientific Industry)
Power pac 300 (Bio-Rad)
Gene Pulser II Electroporation System (Bio-Rad)
High speed refrigerated centrifuges (Heraes)
Centrifuge (SORVALL RT7)
Mettler AT261 DeltaRange (METTLER TOLEDO)
Mettler GG4002-S (METTLER TOLEDO)
Digitale camaera COOLPIX 900 (Nikon)
Spectrophotometer (Molecular Devices)
pH meter (HANHA instruments)
Hot plates/ Stirrers (CORNING)
Water bath (CHERNG HUEI Co.)
Rotor-Gene™ 3000 (Corbett Research)

2.1.7 Strains and media

Escherichia coli DH5α were used for all plasmid manipulations. They were grown in LB broth supplemented with 100 μg/ml ampicillin (AMP) (for *E.coli* carrying pGEM-T cloning vehicle) or 34 μg/ml chloramphenicol (CAM) (for those carrying pSFS2A), and incubated at 37°C. They were cultivated in YPD medium at 30°C. Solid media were made by adding 2% Bacto agar. Selection of nourseothricin (NAT)-resistance transformants (carrying *SAT1* cassette) was performed on YPD agar plates containing 200 μg/ml NAT. In broth microdilution experiments, *C. albicans* ATCC90028 was used as a reference strain and *Candida krusei* ATCC6258 and *Candida parapsilosis* ATCC22019 were incorporated as quality control strains. All clones were stored as frozen stocks in glycerol at -80°C.

2.2 Methods

2.2.1 DNA methods

Plasmid DNA was purified from *E.coli* based on the instruction of the Plasmid DNA Extraction Mini Sample Kit (Favorgen, Cat. No. FAPDE 001). Genomic DNA of *C. tropicalis* was prepared according to the direction of MasterPure™ Yeast DNA Purification Kit (Epicentre, Cat. No. MPY80200). Restriction digestions (NEB) and DNA ligations (Epicentre, Cat. No. LK6201H) were carried out according to the supplier's recommendations. PCR and enzyme digested products were purified using the FavorPrep GEL/PCR DNA Clean-UP Kit (Cat. No. FAGCK 100) according to the manufacturer's instructions. PCR-generated DNA fragments were cloned into plasmid pGEM-T easy vector (Promega, Cat. No. A1360).

2.2.2 Polymerase chain reaction (PCR)

Typical PCR amplifications from plasmids or genomic DNA were carried out using DreamTaq™ DNA polymerase (Fermentas, Cat. No. EP0701) in a final volume of 25 μl or 50 μl. When PCR was performed in 50 μl volumes, the reaction mixture contained 1.25 U of DreamTaq™ DNA polymerase, 5 μl 10 × DreamTaq™ buffer, 0.2 mM deoxynucleoside triphosphate (dNTP), 2.5 mM $MgCl_2$, 0.5 μM forward and reverse primers, and template DNA (0.01-1 ng for plasmid DNA and 0.1-1 μg for genomic DNA). The PCR cycles comprised an initial denaturation step at 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 50-65°C for 30 s, and 72°C for indicated times (basically, 1 min/kb). After that, a final extension step at 72°C for 5 min. The resulting PCR products were visualized by electrophoresis in 1-2% agarose gels.

Colony PCR:

A colony PCR-based technique was performed using DreamTaq™ DNA polymerase to detect an insert within a plasmid after *E. coli* transformation. Test colonies were inoculated into 12 μl of PCR reaction mix, in which the components and their final concentration remained the same. The PCR conditions were modified as follows: 95°C for 10 min; 30 cycles of 95°C for 1 min, 55°C for 1 min, and 72°C for 2 min; 72°C for 5 min.

Fusion PCR:

Phusion® High-Fidelity DNA Polymerase (NEB, Cat. No. F-530S) was used in the

first two reactions of fusion PCR. PCR was performed in a 50 μl volume containing 1 U of Phusion™ DNA polymerase, 10 μl 5 × Phusion™ HF Buffer, 0.2 mM dNTP, 0.5 μM forward and reverse primers, and 10 ng plasmid DNA. The PCR conditions were as follows: 98°C for 30 s; 30 cycles of 98°C for 10 s, 55°C for 10 s, and 72°C for 15 s; and 72°C for 10 min. In the third step of fusion PCR, Blend Taq- Plus- (Toyobo, Cat. No. TYB-BTQ-201) was used because it produced amplification products that were ready to clone directly into TA cloning vectors. When a final volume was 50 μl, the PCR mixture contained 1.25 U of Blend Taq-Plus-, 5 μl 10 × Buffer for Blend Taq, 0.2 μM dNTP, 0.2 μM forward and reverse primers, and 200 ng template DNA (100 ng each of the two purified PCR products sharing a 27 or 31 bp overlapping sequence). The PCR conditions were as follows: 98°C for 30 s; 30 cycles of 98°C for 10 s, 55°C for 10 s, and 72°C for 15 s; and 72°C for 10 min.

2.2.3 Transformations

The cells of *E. coli* transformation was carried out by following the HIT Non-Heat Shock Transformation Protocol, with some modification. *E.coli* cells HIT-DH5α were used as the competent cells (RBC, Cat. No. RH617). 5 μl of the DNA ligation products were mixed with 50 μl of thawed competent cells by vortex for 1 s. The mixtures were incubated on ice for 20 min, followed by heat shock for 1 min at 42°C, and then placed on ice again for 2 min. The transformants were resuspended in 1ml LB broth and incubated for 1 h at 37°C under agitation (180 rpm). Subsequently, the cells were spread on LB plates containing 100 μg/ml of AMP or 34 μg/ml CAM and incubated at 37°C. Resistant colonies were picked after 1 day of growth and the transformants were streaked on LB plates with 100 μg/ml of AMP or 34 μg/ml CAM for single colonies.

The preparation of yeast competent cells was carried out according to Reuβ *et al*. (2004), and the *C. tropicalis* strains were transformed by electroporation as described previously (Kohler *et al.*, 1997), with slight modifications. Inserts from plasmids LOB381, LOB382, LOB383, LOB384, LOB385 and LOB386 were excised as *Kpn*I–*Sac*I fragments and purified by gel extraction. As soon as 5 μl (~ 2 μg) of the linear DNA fragments was mixed with 50 μl of electrocompetent cells, the electroporation was conducted in a Gene Pulser II Electroporation System (0.2 cm

cuvette, 1.8 kV). The cells were then washed in 1 ml of 1 M sorbitol and resuspended in 1 ml YPD medium. After incubated for 4 h at 30°C under agitation (180 rpm), the yeasts were spread on YPD plates containing 200 μg/ml of NAT and cultured at 30°C. Antibiotic-resistant colonies were picked after 2-3 days of growth and the clones were streaked on YPD plates with 200 μg/ml of NAT.

2.2.4 Constructions of plasmids

The schematic linear map of the plasmids is shown in Figure 5. Two plasmids LOB319 and LOB320, containing the *FCY2*^{S} allele and the *FCY2*^{R} allele, respectively, were designed and created previously (Chia-Chen Wu, 2009). Major components of these plasmids are described as follows:

(1) *FCY2*-A: a 2258 bp *Kpn*I-*Xho*I fragment composed of the entire *FCY2* coding region and its flanking sequences was amplified from the genomic DNA of yeast YM020291, using primer pair HJL1420 and HJL1424.

(2) *FCY2*-B: a 528 bp *Sac*II-*Sac*I fragment including 46 bp 3'end sequence of *FCY2* ORF and its downstream region (from 1531 bp to 1971 bp) as well as 43 bp nucleotide sequence derived from pGEM-T easy vector was amplified from the genomic DNA of yeast YM020291, using a primer pair HJL1422 and HJL1423.

(3) *SAT1* flipper: a cassette which possessed a dominant NAT resistance marker has been designed to contain a few unique restriction sites on the left (*Kpn*I and *Xho*I) and right (*Sac*II and *Sac*I) borders for cloning target sequences by homologous recombination (Reuss *et al.*, 2004).

In this study, two strategies were adopted to introduce a point mutation at one of the three polymorphic sites [-69 (145), 273 (486) and 1518 (1731)] in *FCY2* allele carried by plasmids LOB319 or LOB320. One was performed when the proper restriction sites were present (i.e., the enzyme-digested products included only one target polymorphic nucleotide), and the point mutation was created by exchanging a digest fragment (containing the SNP) between two plasmids LOB319 and LOB320. Alternatively, when no suitable restriction site was found, the point mutation was generated by fusion PCR, in which three separate PCR reactions were conducted. The overlapping regions between two PCR products, separately synthesized from the first two reactions, allowed a third PCR to create the entire *FCY2*-A carrying the expected point mutation. Following are the procedures to construct each plasmid.

(1) LOB381, *FCY2*^{S}-A [G273T (G486T)]-*SAT1-FCY2*^{S}-B: *FCY2* 5' and 3' end separately amplified from LOB319 in two PCR reactions by primer pairs HJL1420 and HJL1743, and HJL1742 and HJL1424. A third PCR was conducted to fuse the two resulting PCR products with primers HJL1420 and HJL1424. *FCY2*^{S}-A [G273T (G486T)] was first cloned into pGEM-T easy vector, and then subcloned into the *Kpn*I-*Xho*I restriction sites of LOB319.

(2) LOB382, *FCY2*^{R}-A [T273G (T486G)]-*SAT1-FCY2*^{R}-B: *FCY2* 5' and 3' end separately amplified from LOB320 by primer pairs HJL1420 and HJL1741, and HJL1740 and HJL1424. A third PCR was conducted with primers HJL1420 and HJL1424. *FCY2*^{R}-A [T273G (T486G)] was first cloned into pGEM-T easy vector, and then subcloned into the *Kpn*I-*Xho*I restriction sites of LOB320.

(3) LOB383, *FCY2*^{S}-A [G-69T (G145T)]-*SAT1-FCY2*^{S}-B: *FCY2* 5' and 3' end separately amplified from LOB319 by primer pairs HJL1420 and HJL2103, and HJL2102 and HJL1424. A third PCR was conducted with primers HJL1420 and HJL1424. *FCY2*^{S}-A [G-69T (G145T)] was first cloned into pGEM-T easy vector, and then subcloned into the *Kpn*I-*Xho*I restriction sites of LOB319.

(4) LOB384, *FCY2*^{R}-A [T-69G (T145G)]-*SAT1-FCY2*^{R}-B: *FCY2* 5' and 3' end separately amplified from LOB320 by primer pairs HJL1420 and HJL2101, and HJL2100 and HJL1424. A third PCR was conducted with primers HJL1420 and HJL1424. *FCY2*^{R}-A [T-69G (T145G)] was first cloned into pGEM-T easy vector, and then subcloned into the *Kpn*I-*Xho*I restriction sites of LOB320.

(5) LOB385, *FCY2*^{S}-A [C1518T (C1731T)]-*SAT1-FCY2*^{S}-B: LOB319 was digested with *Hind*III, and calf intestinal alkaline phosphatase (CIP) was added to prevent self-ligation. After gel extraction, the 2101 bp DNA segment, containing 208 bp 3'end sequence of *FCY2* ORF and its downstream region [including the cytosine at position 1518 (1731)] as well as 1452 bp nucleotide sequence from *SAT1* cassette, was removed from LOB319. The deletion was then replaced by the 2101 bp *Hind*III-*Hind*III DNA fragment of LOB320.

(6) LOB386, *FCY2*^{R}-A [T1518C (T1731C)]-*SAT1-FCY2*^{R}-B: LOB320 was digested with *Hind*III, and CIP was added. After gel extraction, the 2101 bp DNA segment, containing 208 bp 3'end sequence of *FCY2* ORF and its downstream region [including the thymine at position 1518 (1731)] as well as 1452 bp nucleotide sequence from *SAT1* cassette, was removed from LOB320. The deletion was then replaced by the 2101 bp *Hind*III-*Hind*III DNA fragment of LOB319.

2.2.5 Sequence analysis

The *FCY2* ORF was identified by the ORF Finder, available on the National Center for Biotechnology Information website (NCBI; http://www.ncbi.nlm.nih.gov). The homology search was performed with the Basic Local Alignment Search Tool (BLAST) programs, available on the NCBI. The nucleotide and amino acid sequence of *FCY2* and *FCY21* of *C. lucitaniae*, and *FCY2* of *S. cerevisiae* was acquired from the GenBank database (accession nos. AY506866, AY506867 and X51715, respectively); likewise, the nucleotide sequences of *FCY2* of C. tropicalis ODL4-302, ODL3-231 and CBS94 were also acquired from the GenBank database (accession nos. EU327985, EU327984 and EU327983, respectively) (http://www.ncbi.nlm.nih.gov/nuccore); the nucleotide and amino acid sequences of *FCY2* and *FCY21* of *C. albicans* were obtained from the Stanford database (orf19.1357, orf 19.333) (http://www.candidagenome.org/).

2.2.6 Susceptibility testing

Broth microdilution:

Each strain was initially tested for 5FC susceptibility with slight modifications of broth microdilution method recommended by the Clinical and Laboratory Standard Institute (CLSI, 1997). Stock solution of 5FC was prepared by dissolving the drug in sterile distilled water at concentration of 5.12 μg/ml or in DMSO at concentration of l28 mg/ml, and its final concentrations ranged from 0.125 to 64 μg/ml. Strains in the frozen state were directly inoculated to 1 ml BHI, and incubated at 35°C for 24 h. Testing was conducted in 96-well plates with RPMI 1640 buffered to pH 7.0 with 0.165 M MOPS and 2 g/L $NaHCO_3$. The suspensions equivalent to a 0.5 McFarland standard of each strain were diluted to a final cell count of approximately 0.25×10^3 to 1.25×10^3 CFU/ml. After 48 h of incubation at 35°C, the final growth of each strain was measured at 595 nm with Biotrack II plate reader. Photos were taken daily. According to CLSI guidelines, the MIC was defined as the lowest drug concentration that inhibited growth by 50% compared to the drug-free growth control. The classification of the 5FC MIC results as sensitive, intermediate, or resistant was defined as an MIC of ≤4, 8 to 16, ≥32 μg/ml, respectively.

E-test:

The MIC values of some strains were also confirmed by E-test (AB Biodis Solna). Briefly, three colonies were picked and subcultured onto SDA (0.67% bacto-yeast nitrogen base without amino acid, 2% dextrose, 20 mg/l histidine-HCl, 2% agar). After incubated at 35°C for 24 h, the cells were diluted in 0.85% NaCl to achieve 0.5 McFarland turbidity (10^6 to 5×10^6). A sterile swab was soaked in the inoculums suspension and used to streak the entire agar surface of RPMI plate evenly, rotating 90° in three directions. The plates were allowed to dry before E-test strips were applied. Subsequently, they were incubated at 35°C for 48 h. Photos were taken daily. The classification of the 5FC MIC results was followed the CLSI guidelines.

2.2.7 Quantitative analysis of the mRNA level by real-time PCR

The yeasts were grown on YPD agar supplemented with or without 200 μg/ml NAT at 30°C for 2 days. Suspensions were prepared by picking three distinct colonies and grown in YPD broth at 30°C overnight under agitation (180 rpm). The overnight cultures were diluted to an OD_{600} of 0.2 in YPD broth and incubated at 30°C under agitation (180 rpm) until the cultures reached an OD_{600} of 0.5. The broth cultures were individually divided into two portions (50 ml each). While the first portion was the treatment of 0.5 μg/ml of 5FC, the second portion was the addition of sterile water. After incubated at 30°C for 1 h, the cells were pelleted, washed in ice-cold sterile water, and allocated to three vials. After the supernatant were removed, cell pellets were stored at -80°C.

RNA isolation:

Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Cat. No. 74106) associated with the Baseline-ZERO^{TM} DNase set (Epicentre, Cat. No. DB0711K). The procedure was performed following the manufacturer's direction, with some modification. The frozen cells (5×10^7 cells) were loosen by flicking the tube, and resuspended in 600 μl Buffer RLT (containing 1% β-ME). After the addition of the acid-washed glass beads (~ 300 μl), cells were broken using a FastPrep®-24 (5× 30s at 4.5 m/s, with 30 s intervals on ice). The lysates were transferred to a new tube, centrifuged for 2 min at top speed, and transferred to another new tube. An equal volume of 70% ethanol was mixed thoroughly with the homogenized lysate, and the

mixture was transferred to a spin column placed in a 2 ml collection tube, followed by centrifugation for 15 s at top speed, and the flow-through was discarded. The spin column membranes were then washed thrice: first in 700 μl Buffer RW1, followed by centrifugation for 15 s at top speed; second in 500 μl Buffer RPE followed by the centrifugation for 15 s at top speed; third in 500 μl Buffer RPE followed by the centrifugation for 2 min at top speed. The RNeasy spin columns were placed in a new collection tube and centrifuged for 1 min at full speed. Subsequently, they were transferred to a new 1.5 ml tube. 30 μl of RNase-free water was added directly to the spin column membrane and incubated for 10 min, followed by centrifugation for 1 min at top speed. Another 30 μl RNase-free water were added to elute the RNA. Next, 6.5 μl of 10x Baseline-ZERO DNase Reaction Buffer and 2 μl (2 MBU) of Baseline-ZERO DNase were added to the sample. After 1 h of incubation at 37°C, an equal volumn (200 μl) of 4 M LiCl was added to the sample and RNA was precipitated for at least 1 h at -20°C. The RNA precipitate was then centrifuged for at least 30 min at 4°C at top speed, washed with 70% ethanol twice and dissolved in 35 μl of RNase-free water by 5 min incubation at 64°C. The concentrations of RNA were determined with a spectrophotometer at λ_{280} and λ_{260}.

Reverse transcription:

Reverse transcription was performed on 1 μg of total RNA using ImProm-II™ Reverse Transcription System with oligo(dT) according to the instructions of the manufacturer (Promega, Cat. No. A3800). cDNA were used for SYBR green quantitative PCR (Roche, Cat. No. 04 193 850 001) and analyzed by Rotor-Gene 6 software version 6.1 (Info-Zip Pty Ltd.). The cDNA of *FCY2* mRNA normalized with transcripts *ACT1*, using a Rotor-Gene™ 3000 instrument (Corbett Research, Australia). Primers of each gene were designed to amplify approximately 300 bp PCR products (Table 3). The condition for real-time PCR was as follows: denaturation (10 min at 95°C), 40 cycles of repeated amplification (15 s at 95°C and 60 s at 60°C). To eliminate the existence of DNA contamination, 1 μg of total RNA and 1 ng cDNA were separately used as templates and amplified by PCR with the primer pair HJL1937 and HJL1938, of which the sequences were deduced from the intron and exon of the *C. tropicalis EFB1* gene and therefore could distinguish the resulting PCR

product between the genomic DNA (810 bp) and the spliced RNA (588 bp).

2.2.8 Identification of the boundary of loss of heterozygosity (LOH) events

At first, four genes evenly spreading on the chromosome (corresponding to supercontig 2 in the *C. tropicalis* database) (http://www.broadinstitute.org/annotation/genome/candida_group/FeatureSearch.html) were selected. After the sequences of these fragments were analyzed, the state of heterozygosity of each chromosomal region was determined (detail was mentioned in the following contents). Accordingly, the boundary range would be narrowed down to a section where the SNPs have changed from heterozygous state to homozygous state or from homozygous state to heterozygous state. Again, genes evenly spreading in this section were selected and their sequences were analyzed. The process of sequence analysis and boundary localization was repeated until the boundary region was reduced to a range between two genes, and the entire sequence of this region was sequenced to precisely pinpoint the LOH boundary.

The scanning process of SNP markers was carried out by the amplification of specific regions of the chromosome in two clinical isolates YM020291 and YM060800. The PCR products were then sequenced in one direction and the nucleotide sequences were analyzed for polymorphism using the software ChromasPro Version1.21 (Technelysium Pty Ltd.) and Vector NTI software version 10.3.0 (Invitrogen). Heterozygosity was defined by the presence of superimposed peaks in the sequence data. Since the length of PCR products were 800-1600 bp (generally, sequencing data could cover a 400-700 bp region with clarity), sequencing in the opposite direction was performed when no SNP was observed on one end of the fragment. Each primer pair was designed by the Vector NTI software version 10.3.0 (Invitrogen) based on the *C. tropicalis* strain MYA-3404 genome sequence (taxid: 294747) deposited at Broad institute (www.broadinstitute.org/annotation/genome/candida_group/MultiHome.html) (Table 3). The corresponding regions where SNPs were found in the selected locus of the parental strains were further amplified from the genomic DNA of their derivatives using the same primer pairs. Two states were observed at each locus, a heterozygous state or a homozygous state, and the LOH boundaries were defined as the regions flanked by adjacent heterozygous and homozygous SNPs.

Chapter 3. Results

3.1 Phenotypic analysis of the clinical isolates and their homozygous mutants

The 5FC susceptibilities of two clinical isolates YM020291 and YM060800, their homozygous resistant mutants (*FCY2*^{R}/*FCY2*^{R}) and homozygous sensitive mutants (*FCY2*^{S}/*FCY2*^{S}) were determined by broth microdilution method. In previous study, the MICs of 5FC for one of the clinical isolates YM020291 and its genetic engineered derivatives were measured, and revealed that only the *FCY2*^{R}/*FCY2*^{R} homozygous mutants (YLO417 and YLO418) and *FCY2*^{R}/*fcy2*Δ single knockout strains (YLO419 and YLO421) were resistant to 5FC (Chia-Chen Wu, 2009). Here, the association between genotype and 5FC resistance phenotype was confirmed by another isolate, YM060800, with different genetic background, its homozygous resistant mutants YLO440 and YLO441, and homozygous sensitive mutants YLO447 and YLO448. The 5FC MICs of YM020291 and its homozygous mutants were also measured to examine the consistency of their susceptibilities comparing to the previous work. The 5FC susceptibility testing result of the two clinical isolates YM020291 and YM060800 as well as their homozygous mutants at 48 h is given in Figure 1. According to CLSI guidelines, the MIC was defined as the lowest drug concentration that inhibited growth by 50% compared to the drug-free growth control after 48 h of incubation at 35°C. For both clinical isolates YM020291 and YM060800, the MIC was 0.5 μg/ml, and for their homozygous sensitive mutants YLO415, YLO447 and YLO448 (*FCY2*^{S}/*FCY2*^{S}), the MIC was 0.25 μg/ml. In contrast, for the homozygous resistant mutants YLO417, YLO440 and YLO441 (*FCY2*^{R}/ *FCY2*^{R}), the MICs were higher than or equal to 64 μg/ml. The 5FC susceptibility testing was conducted twice in six strains YM020291, YLO417, YLO415, YM060800, YLO441 and YLO447; thus, the MIC values for both clinical isolates and their engineered derivatives from each experiment are summarized in Table 4. The MICs of each strain from the two independent experiments were either identical or within one dilution of each other. Based on the criteria of CLSI (1997), yeast cells with MIC ≤ 4 μg/ml were considered susceptible, while those with MIC ≥ 32 μg/ml were considered resistant.

The results from two different genetic backgrounds showed that the *FCY2*^{R}/*FCY2*^{S} heterozygous and *FCY2*^{S}/*FCY2*^{S} homozygous mutants were susceptible to 5FC, while the *FCY2*^{R}/*FCY2*^{R} homozygous mutants were resistant to 5FC.

3.2 Sequence analysis of the 5' and 3' untranslated regions of the *FCY2* locus in hyper-susceptible strains

Sequence analysis of the *FCY2* locus was performed in four clinical isolates that did not produce 5FC resistant progeny (a.k.a., hyper-susceptible strains) to eliminate the polymorphic nucleotide that was not unique to the *FCY2*^{R} allele. The hyper-susceptible strain was defined as a clinical isolate that could not produce any 5FC resistant derivative in the E-test inhibition ellipse after 72 h of incubation. Four isolates, YM020367, YM020649, YM060302 and YM060559, were chosen because their inhibition zones were clear. PCR products spanning either 5'UTR or 3'UTR as well as the ORFs of the *FCY2* gene were sequenced, and the nucleotide sequences were aligned with those from the *FCY2*^{R} and *FCY2*^{S} alleles in YM020291. One polymorphic nucleotide at position 273 (486), resulting in an amino acid change, has been examined in the hyper-susceptible strains previously. All four strains have a 273 (486) G/G genotype, which is also exhibited in the *FCY2*^{S} allele (Chia-Chen Wu, 2009). The remaining polymorphic sites identified previously were located in the 5' or 3'UTR [at position -224 (-11), -69 (145) or 1518 (1731)]; thus, the SNPs at these positions were examined in the hyper-susceptible strains. The sequencing results are shown in Figures 2, 3 and 4. Alignment of these sequences revealed that the SNP at position -224 (-11) in both alleles of the *FCY2* locus in one strain (YM060559) was thymine, while that in other strains (YM020367, YM020649 and YM060302) was cytosine. In contrast, a -69 (145) G/G genotype and a 1518 (1731) C/C genotype were in all hyper-susceptible strains. These SNPs [at position-69 (145) and 1518 (1731) C/C] were identical to those in the *FCY2*^{S} allele. As a consequence, the possibility that the involvement of the SNP at position -224 (-11) in the drug resistance was ruled out because it was found in the 5FC hyper-susceptible strains not only as cytosine but also as thymine (none unique to the *FCY2*^{R} allele). On the other hand, three polymorphic nucleotides [at position -69 (145), 273(486) and 1518 (1731)] might be involved in the 5FC susceptibility because they exhibited genotypes which were consistent with

those in the *FCY2*^{S} allele.

3.3 Construction of plasmids carrying a single mutation at the *FCY2* locus

To examine whether the potential polymorphic nucleotides at position -69 (145), 273 (486) and 1518 (1731) were involved in the function or expression of PCP linked to 5FC resistance, point mutations were generated at one of the polymorphic sites in *FCY2*^{S} or *FCY2*^{R} allele of the clinical isolate YM020291 by the *SAT1* flipping method. The schematic linear map of the plasmids with the locations of each SNP and the relevant restriction sites is shown in Figure 5. The strategies to construct plasmids carrying single mutations at the *FCY2* locus were described in Materials and methods.

For plasmids carrying a single mutation C1518T (C1731T) or T1518C (T1731C) at the downstream of the *FCY2* gene, those with the given inserts were first screened by enzyme digestion to eliminate the existence of self-ligation plasmids. The introduction of the point mutation T1518C (T1731C) into LOB320 would create a *Hinf*I site; on the contrary, the introduction of the point mutation C1518T (C1731T) into LOB319 would lose a *Hinf*I site. Because the plasmids had too many restriction sites for *Hinf*I to produce a distinguishable enzyme digestion pattern, the PCR products were subjected to digestion with *Hinf*I instead of the whole plasmids. Mutations C1518T (C1731T) and T1518C (T1731C) in the *FCY2* gene carried by the plasmids LOB385 and LOB386, respectively, were examined as follows: A 717 bp DNA segment (including the SNP at the 201 nucleotide downstream of the stop codon of the *FCY2* gene) of each plasmid was amplified by primer pair HJL2117 and HJL1424. Digestion of the PCR products with *Hinf*I was expected to yield bands of 64, 233 and 420 bps for LOB385 (Figure 6B, bands a, c, and e), and of 64, 176, 233 and 244 bps for LOB386 (Figure 6B, bands a, b, c, and d). The entire nucleotide sequence of the *FCY2*-A carried by LOB385 or LOB386 was further confirmed by sequencing using primers M13, HJL2124, HJL1208 and HJL2117 to avoid the inclusion of any unexpected mutation. The results of sequence analysis are given in Figure 7.

As to the plasmids LOB381, LOB382, LOB383 and LOB384, the mutations G273T (G486T), T273G (T486G), G-69T (G145T) and T-69G (T145G) were

introduced into either *FCY2*^{S} (LOB319) or *FCY2*^{R} (LOB320) allele by fusion PCR. The resulting PCR products were cloned into pGEM-T easy vector and sequenced using primers T7, HJL1208, HJL1210 and SP6 to avoid the inclusion of any unexpected mutation (Figures 8 and 9). The *Kpn*I-*Xho*I fragments of the plasmids carrying the expected point mutation were then subcloned into LOB319 or LOB320. The introduction of the point mutation T-69G (T145G) into LOB320 would lose an *Nsi*I site, while the introduction of the point mutation G-69T (G145T) into LOB319 would create an *Nsi*I site. Similarly, the introduction of the point mutati T273G (T486G) into LOB320 would create a *Bbs*I site, while the introduction of the point mutation G273T (G486T) into LOB319 would lose a *Bbs*I site. The constructs of the plasmids LOB381, LOB382, LOB383 and LOB384 were examined as follows: The digestion of plasmids LOB383 and LOB384 with *Nsi*I was expected to yield fragment sizes of 1529, 1985 and 6683 bps (Figure 10B, bands a, b and c), and of 1529 and 8668 bps (Figure 10B, bands a and d), respectively. The results showed that the point mutations G-69T (G145T) and T-69G (T145G) in the *FCY2* gene were separately carried by the plasmids LOB383 and LOB384. Likewise, the plasmids LOB381 and LOB382 digested with *Bbs*I were expected to yield fragment sizes of 989, 2115 and 7093 bps (Figure 11B, bands a, c and f), and of 989, 2028, 2115 and 5065 bps (Figure 11B, bands a, b, c and e), respectively. And the results showed the point mutations G273T (G486T) and T273G (T486G) in the *FCY2* gene were separately carried by the plasmids LOB381 and LOB382.

3.4 Construction of G-69T (G145T), T-69G (T145G), G273T (G486T), T273G (T486G), C1518T (C1731T) or T1518C (T1731C) single mutation strain

Each plasmid was digested with *Kpn*I and *Sac*I, and the resulting 2258 bp linear DNA fragment was used to transform strain YM020291 to introduce the single mutation in the selected *FCY2* allele. The integration of the *Kpn*I-*Sac*I fragments of each plasmid occurred at the *FCY2* locus was screened by colony PCR with the primer pair HJL814 and HJL1477, which were located on the *SAT1* cassette and at the downstream of the *FCY2*-B, respectively (Figure 12). Since *C. tropicalis* is a diploid organism, the *Kpn*I-*Sac*I fragments substitute one of the *FCY2* alleles at random.

Therefore, to determine whether the integration was taken place at the specific allele resulting in the identical polymorphic nucleotides present at the same position in both alleles, the single mutation strains were examined as following. To generate distinguishable enzyme-digested products, a DNA fragment containing the point mutation was amplified from the genomic DNA of the transformants. For the constructs YLO454, YLO455, YLO456 and YLO457, a 1846 bp DNA segment [containing the SNP at position 273 (486)] derived from their genomic DNA was amplified by primer pair HJL1207 and HJL1424. After *Bbs*I digestion, the products yielded bands of 1846 bp for YLO454 and YLO455, and of 360 and 1486 bps for YLO456 and YLO457 (Figure 13); for the constructs of YLO466, YLO467, YLO468 and YLO469, a 1106 bp DNA segment [containing the SNP at position -69 (145)] derived from their genomic DNA was amplified by primer pair HJL2104 and HJL1208. Digestion of the PCR products with *Nsi*I was expected to yield bands of 1106 bp for YLO468 and YLO469, and of 388 and 719 bps for YLO466 and YLO467 (Figure 13); for the constructs of YLO470, YLO471, YLO472 and YLO473, a 717 bp DNA segment [including the SNP at position 1518 (1731)] derived from their genomic DNA was amplified by primer pair HJL2117 and HJL1424. Digestion of the PCR products with *Hinf*I was expected to yield bands of 64, 233 and 420 bp for YLO470 and YLO471, and of 64, 176, 233 and 244 bp for YLO472 and YLO473 (Figure 13). Also, nucleotide sequences of each PCR product were examined using primers HJL1207, HJL2104 and HJL2117, corresponding to SNP at 273 (486), -69 (145) and 1518 (1731), respectively (Figure 14).

3.5 Phenotypic analysis of the single mutation strains

The 5FC susceptibilities of the single mutation strains each carrying a point mutation G273T (G486T), T273G (T486G), G-69T (G145T), T-69G (T145G), C1518T (C1731T) or T1518C (T1731C) at the *FCY2* locus were determined by broth microdilution method. The 5FC MICs of the clinical isolate YM020291, its homozygous resistant mutants ($FCY2^R$/ $FCY2^R$) and homozygous sensitive mutants ($FCY2^S$/$FCY2^S$) were also measured as a control to examine the consistency of their susceptibilities comparing to previous results. The 5FC was dissolved in DMSO in the first three experiments, but then the drug was dissolved in sterile distilled water in the

last four experiments; thus, the 5FC susceptibility testing results of each experiment under different conditions were separately given in Tables 4 and 5. The MICs of each strain including clinical isolate YM020291, its homozygous mutants and its single mutation strains were determined after 48 h of incubation at 35°C. According to CLSI guidelines, the MIC was defined as the lowest drug concentration that inhibited growth by 50% compared to the drug-free growth control. The 5FC susceptibility testing was conducted at least three times in 11 strains (i.e., YM020291, YLO417, YLO418, YLO415, YLO454, YLO456, YLO466, YLO467, YLO468, YLO470 and YLO472); thus, the MIC values for these strains from each experiment are summarized in Table 6. The MIC values were listed as follows: (1) For the clinical isolate YM020291 ($FCY2^R/FCY2^S$), its homozygous sensitive mutant YLO415 ($FCY2^S/FCY2^S$), as well as its single mutation strains YLO468 [T-69G (T145G)] and YLO472 [T1518C (T1731C)], the MICs were less or equal to 0.5 μg/ml; (2) For the single mutation strain YLO456 [T273G (T486G)], the MICs were less or equal to 1 μg/ml; (3) For the single mutation strains YLO454 [G273T (G486T)] and YLO470 [C1518T (C1731T)], the MICs were less or equal to 2 μg/ml; (4) For the homozygous resistant mutants YLO417 and YLO418 ($FCY2^R/FCY2^R$), the MICs were higher or equal to 32 μg/ml; (5) For the single mutation strains YLO466 and YLO467 [G-69T (G145T)], the MICs were higher or equal to 64 μg/ml. Susceptibility testing of one reference strain and two quality control strains were used in parallel. All MIC results were within the range given by CLSI except in strain ATCC22019, which gave an increased MIC to 5FC. Based on the criteria of CLSI (1997), yeast cells with MIC $\leq$ 4 μg/ml were considered susceptible, while those with MIC $\geq$ 32 μg/ml were considered resistant. Accordingly, the $FCY2^R/FCY2^S$ heterozygous strains and the $FCY2^S/FCY2^S$ homozygous mutants as well as the G273T (G486T), T273G (T486G), T-69G (T145G), C1518T (C1731T) and T1518C (T1731C) single mutation strains were susceptible to 5FC; on the contrary, the $FCY2^R/FCY2^R$ homozygous mutants and the G-69T (G145T) single mutation strains were resistant to 5FC.

The susceptibility of the clinical isolate YM020291 and its mutant strains to 5FC was also assessed with E-test assay (Figure 15). The MICs were determined after 48 h of exposure to 5FC E-test strip and the values were read at the point where the ellipse intersects the strip. The MIC for homozygous sensitive mutant YLO415 ($FCY2^S/FCY2^S$) was 0.008 μg/ml; for parental strain YM020291 ($FCY2^R/FCY2^S$) and

the G273T (G486T) single mutation strain YLO454 was 0.064 μg/ml; for single mutation strains YLO456 [T273G (T486G)], YLO468 [T-69G (T145G)], YLO470 [C1518T (C1731T)] and YLO472 [T1518C (T1731C)] was 0.023 μg/ml. In contrast, the MIC for homozygous resistant mutants YLO417 and YLO418 ($FCY2^R/FCY2^R$), the G-69T (G145T) single mutation strains YLO466 and YLO467 was higher than 32 μg/ml. Based on the criteria of CLSI (1997), yeast cells with MIC ≤ 4 μg/ml were considered susceptible, while those with MIC ≥ 32 μg/ml were considered resistant. As a result, the $FCY2^R/FCY2^S$ heterozygous strains and $FCY2^S/FCY2^S$ homozygous mutants as well as the G273T (G486T), T273G (T486G), T-69G (T145G), C1518T (C1731T) and T1518C (T1731C) single mutation strains were susceptible to 5FC, whereas the $FCY2^R/FCY2^R$ homozygous mutants and the G-69T (G145T) single mutation strains were resistant to 5FC. Among the 5FC susceptible strains, the MICs for YLO415, YLO456, YLO468, YLO470 and YLO472 were less than that for the parental strain YM020291.

The results of MICs from broth microdilution and E-test indicated that a guanine to thymine substitution at the nucleotide -69 (145) in the $FCY2^S$ genes has contributed to the 5FC resistance in *C. tropicalis*.

3.6 Contribution of the polymorphic nucleotide in the promoter region of the *FCY2* gene

To evaluate whether 5FC resistance observed in the $FCY2^R/FCY2^R$ homozygous strains and the G-69T (G145T) single mutation strains was caused by the differential mRNA levels of *FCY2*, the mRNA levels of this gene were examined in the clinical isolate YM020291, its single mutation strains YLO466 [G-69T (G145T)] and YLO468 [T-69G (T145G)], and its homozygous mutants YLO415 ($FCY2^S/FCY2^S$) and YLO417 ($FCY2^R/FCY2^R$) by real-time quantitative PCR. The effect of the drug on the mRNA levels of *FCY2* gene was also examined. All strains were treated under two conditions (the absence or presence of 0.5 μg/ml of 5FC) at 30°C for an hour. The *FCY2* mRNA levels of the clinical isolate YM020291, its homozygous mutants and its single mutation strains conducted in two independent experiments are shown in Figure 16. For the clinical isolate YM020291, the homozygous resistant strain YLO417, homozygous susceptible strain YLO415, the G-69T (G145T) single

mutation strain YLO466 and the T-69G (T145G) single mutation strain YLO468, the *FCY2* mRNA levels were 1.00±0.00, 0.52±0.02, 1.53±0.07, 0.50±0.11 and 1.70±0.52 folds, respectively, when the strains were not treated with 5FC. For the same strains, the *FCY2* mRNA levels were 1.36±0.05, 0.51±0.01, 2.27±0.12, 0.50±0.20 and 2.70±0.50 fold, respectively, when the yeasts were treated with 0.5 μg/ml of 5FC.The analysis showed that the *FCY2* mRNA existed not only in the wild type strain YM020291 but also in its engineered mutants (YLO415, YLO417, YLO466 and YLO468), and the mRNA levels were in a similar pattern among different treatments (with or without 5FC). That is, a slight increase of the mRNA level in the presence of 0.5 μg/ml of 5FC compared to that in the absence of the drug for all strains except in YLO466. In addition, under both conditions (the treatment with or without 5FC), the highest- and the lowest-levels of *FCY2* mRNA were observed in the homozygous sensitive mutant YLO415 and the homozygous resistant mutant YLO417, respectively, among the clinical isolate YM020291 and its homozygous mutants. Similarly, the highest- and the lowest-levels of *FCY2* mRNA were observed in the single mutation strain YLO468 [T-69G (T145G)] and the single mutation strain YLO466 [G-69T (G145T)], respectively, among the clinical isolate YM020291 and its single mutation strains. The tendency of the expression data was similar between the homozygous mutants and the single mutation strains.

In this analysis, it seemed that the *FCY2* mRNA levels might be influenced by a low concentration (0.5 μg/ml) of 5FC. Besides, the similar mRNA level profiles between the single mutation strains and the homozygous mutant strains implied that the SNP at position -69 (145) in the *FCY2* gene has contributed to the differential mRNA levels observed in this study.

3.7 Nucleotide sequence analysis of the *FCY2* gene

When a BLASTN search was executed to determine if the polymorphic nucleotide at position -69 (145) was located in a crucial regulatory site, it has revealed that the polymorphic site was located in the coding region of the *FCY2* genes. Amino acid alignment of MYA-3404 Fcy2p from *C. tropicalis* database with the *C. albicans* and *S. cerevisiae* purine-cytosine permeases was performed accordingly. The alignment result is given in Figure 17. It has shown that the starting amino acid

methionine of MYA-3404 Fcy2p in the second row was corresponding to the 93rd amino acid of the *S. cerevisiae* Fcy2p in the fifth row and corresponding to the 77th amino acid of the *C. albicans* Fcy21p in the third row. As a consequence, the nucleotide sequences of both *FCY2* alleles in YM020291 were used to examine the *FCY2* coding region by ORF Finder, available on the NCBI website; it showed that the largest frame identified in a 2258 bp DNA segment was 1530 bp long because the 5' end of the newly identified ORF was extended up to 213 bp preceding the former predicted start codon (Figure 18).

Nucleotide analysis of the redefined *FCY2* open reading frame has indicated the new locations of the three selected SNPs as follows: one was shifted from position -69 to position 145; another was from position 273 to position 486; the other was from position 1518 to position 1731 (at the 201 nucleotide downstream of the stop codon of the *FCY2* gene) (Figure 18). Most importantly, the polymorphic nucleotide at position 145 in the $FCY2^{R}$ gene was a guanine-to-thymine substitution, which resulted in a nonsense mutation. There can be no doubt that the truncation at Glu 49 could result in the loss of purine-cytosine permease, which led to the 5FC resistant to *C. tropicalis*. It also revealed a guanine-to-adenine substitution at the nucleotide position 201 in the *FCY2* ORF of the strains MYA-3404, which also led to a nonsense mutation. Clearly, because of the inframe stop codon at Trp 67 in PCP, Met 72 was mistakenly recognized as the start codon in *C. tropicalis* database (Figure 18).

Other than MYA-3404, the nucleotide sequences of the *FCY2* gene in three *C. tropicalis* strains, ODL4-302, ODL3-231 and CBS94, were available on the NCBI database as well. Although they have both 145 G/G and 201 G/G genotypes, their deposit sequences were not completed comparing to the redefined *FCY2* ORF. That is, for strains ODL4-302, ODL3-231 and CBS94, their DNA sequences of *FCY2* separately began at nucleotides 136, 129 and 131, downstream of the new start codon.

3.8 Genotypic characterization of clinical isolates

Nucleotide sequence analysis of the *FCY2* gene in 22 parental strains from 35 clinical isolate-derivative pairs as well as seven previously selected clinical isolate-derivative pairs was performed to determine whether the 145 G/T genotype in the *FCY2* gene was related to the generation of 5FC resistant progeny observed among

the clinical isolates. The analysis of eight other hyper-susceptible clinical isolates was included as reference. Moreover, the nucleotide sequence of *FCY1*, *FCY2*, *FUR1* and *URA3* of the 5FC resistant clinical isolate was examined. The results of 145 and 201 genotypes in the parental strains from 35 clinical isolate-derivative pairs are summarized in Tables 7 and 8, and those in the additional reference strains as well as the 5FC resistant isolate are summarized in Table 9. For 30 flucytosine susceptible clinical isolates that generated 5FC resistant derivatives in their E-test inhibition ellipse, 23 of which carried the heterozygous SNP at position 145 (T/G), yet one of the strains YM020715 carried the heterozygous SNP at position 201 (A/G). On the other hand, for the eight hyper-susceptible strains (YM020367, YM020649, YM060302, YM060509, YM060547, YM060559, YM060647 and YM060828), none of them carried the nonsense mutation at position 145 or position 201. As to the 5FC resistant clinical isolate YM060607, no nucleotide change was detected in the *FCY1* gene, and no noticeable nucleotide change that might associate with the 5FC resistance was identified in the *FCY2* gene. In contrast, nucleotide sequence analysis revealed a heterozygous SNP (C/T) at the position 431 of the *FUR1* gene, where the cytosine-to-thymine substitution resulted in a threonine (ACT) to isoleusine (ATT) change at amino acid 144 in UPRT. It also revealed a heterozygous SNP (A/G) at the position 775 of the *URA3* gene, where the guanine-to-adenine substitution resulted in an alanine (GCC) to threonine (ACC) change at amino acid 259 in ODCase. In conclusion, for clinical isolates that could generate progeny in inhibition ellipse, 76.7% (23/30) carried the 145 T/G genotype in the *FCY2* gene. It implied the 145 T/G genotype exhibited in the *FCY2* gene of the clinical isolates was responsible for the generation of their drug resistant progeny.

3.9 Characterization of loss of heterozygosity events in clinical isolate-derivative pairs

Two clinical isolates (YM020291 and YM060800), and their derived strains (YM0202091-1, YM020291-2, YM060800-1 and YM060800-2) were included in this study. According to previous work, the clinical isolates and their derivatives were differed by at least one LOH event at the *FCY2* locus located on the chromosome (corresponding to supercontig 2 on the *C. tropicalis* database). Therefore, SNP

mapping were conducted to determine the extent of LOH events on the chromosome containing *FCY2* gene that differentiated the strains from each analyzed pair. The strategy to search for the boundary regions was described in Materials and methods. In brief, regions containing SNPs in the parental strain YM020291 or YM060800 were identified by sequencing, where the heterozygous SNP was determined based on the presence of two coincident peaks. Those SNP markers were subsequently used to indicate the heterozygous or homozygous state in the corresponding area in their derivatives. Since the LOH was first observed in the *FCY2* gene, the border was approached from the ORF toward both ends of the chromosome.

The SNP map of the chromosomes of two clinical isolates and their derivatives is shown in Figure 19. Location of the selected loci and the amount of identified SNPs as well as the status of each locus (heterozygosity or homozygosity) are detailed in Figure 20. Among 89 regions of the chromosome analyzed in the clinical isolates (YM020291 and YM060800), 64 were polymorphic (i.e., at least one heterozygous SNP present in one of the clinical isolates), and total 233 SNP markers were identified. Other regions were excluded because of their poor sequence quality or the lack of polymorphic nucleotide. Comparison of the chromosome of the clinical isolate YM020291 and its derivatives revealed the following results: first, of the 34 polymorphic regions identified in YM020291, 23 consecutive regions located on the left end side were homozygous in YM020291-1, while the remaining 11 consecutive regions located on the right end side were heterozygous. Second, of 35 polymorphic regions identified in YM020291, 28 consecutive regions located in the left end side were homozygous in YM020291-2, while seven consecutive polymorphic regions located on the right end side were heterozygous. As a result, derivatives YM020291-1 and YM020291-2 separately differed from the isolate YM020291 by an LOH that encompassed approximately 2.19 and 2.25 Mb of the chromosome.

For clinical isolate YM060800, 21 consecutive regions on the left end side of the chromosome preceding the gene CTRG_01804, which located at the position 1565480 to 1572091, were devoid of the SNP markers, when compared to the relevant polymorphic sites found in YM020291. The nucleotide sequence close to the telomere of the chromosome (i.e., a 3080 bp DNA fragment corresponding to nucleotides from position 20 bp to 3100 bp) were analyzed in clinical isolates (YM020291 and YM060800) and both derivatives of YM020291 (YM020291-1 and YM020291-2). Six SNP markers recognized in YM020291 have confirmed that homozygosity was

retained at the end of the chromosome in the strains YM020291-1, YM020291-2 and YM060800 (Figure 21). Accordingly, comparison of the chromosome of the clinical isolate YM060800 and its derivatives began at the region approximately 1.56 Mb from the telomere and revealed the results as follows: first, of the 32 polymorphic regions identified in YM060800, 31 consecutive polymorphic regions located in the left end side were homozygous in YM060800-1, while one region (including 43 SNP markers) located on the right end side were heterozygous. Second, of 30 polymorphic regions identified in YM060800, two heterozygous segments (12 consecutive regions were on the left, while 9 were on the right) were flanking a homozygous segment (including 9 consecutive regions) in YM060800-2. As a result, YM060800-1 differed from its parental strain YM060800 by an LOH that encompassed approximately 0.7 Mb of the chromosome. As to YM060800-2, the LOH event was restricted to an extent smaller than 181 kb. In conclusion, four of the five LOH events (including one observed in the clinical isolate YM060800, which might be a result of LOH events) covered a large region from 2.19 to 2.23 Mb in the chromosome containing *FCY2* gene, while one was remained heterozygous of the chromosome except a 181 kb internal homozygous segment.

3.10 Characterization of the loss of heterozygosity boundaries

Six LOH borders were determined and further pinpointed to a region flanked by the adjacent homozygous and heterozygous SNPs. They included four at the right end side of the chromosome downstream of the *FCY2* gene, and two preceding the *FCY2* gene. Since these chromosome fragments probably underwent a recombination event that resulted in the observed LOH, a DNA segment overlapping the LOH boundary region was sequenced to investigate features that may be associated with the recombination events, such as the location of the boundary (at intragenic or intergenic region) and the frequency of SNP around that area. The boundary region was referred to the area flanked by adjacent heterozygous and homozygous SNPs. The SNP maps and the nucleotide sequences of the six LOH boundaries are depicted in Figures 22, 23, 24, 25, 26, 27 and 28. For the strain YM020291-1, its LOH boundary (within position 2194213 and 2194621 of supercontig 2) fell in the 5' UTR of CTRG_02080, covering a 409 bp segment. Eight homozygous SNPs were found clustered in the 5'

end of this gene as well as its promoter region within 510 bp long (Figure 22). For the second strain YM020291-2, the homozygous SNP at position 2253701 was located in the promoter region of the gene CTRG_02104, whereas the heterozygous SNP at position 2255494 was in the coding sequence of the gene CTRG_02105. Since no SNP maker was present within a 1794 bp segment, this LOH boundary (within position 2253701 and 2255494) was across the intergenic region containing the gene CTRG_02104 (Figure 23). As to the third strain YM060800-1, its LOH boundary (within position 2300598 and 2300901) was located in the gene CTRG_02123, covering a 304 bp segment. 47 SNPs were identified in a 2500 bp segment of this gene (Figure 24). For the fourth strain YM060800-2, two LOH boundaries were either located in an intergenic region (between position 2028007 and 2028323; 317 bp) preceding the *FCY2* gene, or between two genes CTRG_02086 and CTRG_02087 on the right end side of the chromosome downstream of the *FCY2* gene (within position 2208445 and 2209266; 822 bp). Three SNP clusters were observed around the left border. More specifically, eight SNPs were in the gene CTRG_02009 within a 559 bp segment and five SNPs located within 680 bp region adjacent the LOH boundary, followed by 10 SNPs located in another gene CTRG_02010 within a 239 bp segment (the data does not show on the map). On the contrary, one cluster was observed on the right border, which was in the CTRG_02087 with five SNPs within a 471 bp segment (between position 2209266 and 2209736 of supercontig 2) (Figures 25 and 26). For the clinical isolate YM060800, the homozygous SNP at position 1561562 was located in the gene CTRG_01802, whereas the heterozygous SNP at position 1564848 was in the promoter region of the gene CTRG_01803. Since no SNP maker was present within a 3287 bp segment, this LOH boundary (within position 1561562 and 1564848 of supercontig 2) was across the intergenic region containing one gene CTRG_01803 (Figure 27). Furthermore, a BLASTN search has performed to search for recombination hot spots of the six LOH boundaries. However, no detectable hot spots were found (Figure 29). Taken together, for the six LOH boundaries reported in this study, two were located in intergenic promoter-containing intervals (for the border of YM020291-1 and for the left border of YM060800-2) and one existed in a gene with abundant SNPs (the border for YM060800-1). Three were narrowed down to a region between two genes including at least one of these intervals (i.e., 821 bp for the left border of YM060800-2, 1794 bp for YM020291-2 and 3287 bp for YM060800).

Chapter 4. Discussion

4.1 The redefinition of *FCY2* open reading frame in *C. tropicalis*

Multiple sequence comparison revealed a new 5' end to the *FCY2* ORF. The amino acid sequences of both *C. tropicalis* Fcy2p (one was available in the database and the other was deduced from the $FCY2^{S}$ allele in YM020291) were aligned with *C. lucitaniae* and *S. cerevisiae* PCPs. These yeasts were chosen because their *FCY2* coding sequence was available and the length of each ORF was comparable to their mRNA length (Chapeland-Leclerc *et al.*, 2005; Schmidt *et al.*, 1984; Weber *et al.*, 1990). The result suggested that there was very likely a new 5' end to the *FCY2* ORF, which is 213 nt upstream of the previous one presented in MYA-3404 from the *C. tropicalis* database.

Another evidence consistent with the existence of a new 5' end was given by the comparison of the sequence context around both start codons in *C. tropicalis FCY2* gene. According to the study in *S. cerevisiae*, a consensus sequence of 5'-(A/U)A(A/C)A(A/C)A(A/C)AUGUC(U/C)-3' (the position of the start codon is underlined), was found around transcriptional start codons (Hamilton *et al.*, 1987). Kozak and Liu *et al.* have reported when two AUG codons were present in the 5' proximal region, the translation of mRNA most frequently initiated at the AUG codon that located in the most optimal context (i.e., the sequence most resembling the consensus). In their works, the initiation of translation would rather choose a downstream AUG in an optimal context than an upstream located AUG in a poor context (Kozak, 1984; Liu *et al.*, 1984). The sequences flanking both start codons of *FCY2* were therefore evaluated accordingly. The first ATG codon of the newly defined open reading frame was located in a sequence context similar to that found in most *S. cerevisiae* mRNA (Hamilton *et al.*, 1987), such as the presence of a conserved adenine residue at positions -1 and -3 as well as a cytosine at position +5 and a thymine at position +6 in the yeast consensus initiation site; also, the complete absence of guanine residues in the seven bases preceding the start codon and the usage of the stabilizing amino acid Ala at the second position (Figure 18). In contrast, these features were not observed in the ATG codon identified by the *C. tropicalis* database. For these reasons, it is most likely that the methionine triplet boxed in red in Figure 18 is the authentic initiation

codon. The resulting ORF corresponding to the *FCY2* gene of *C. tropicalis* could encode a deduced protein, Fcy2p, of 509 amino acid residues.

4.2 Contribution of the *FCY2* gene to flucytosine susceptibility in *C. tropicalis*

4.2.1 LOH at the *FCY2* locus was associated with 5FC resistance

The 5FC susceptibility of the clinical isolate YM020291 and its genetic modified counterparts was examined in previous work. The MIC for isolate YM020291 and its $FCY2^S/FCY2^S$ homozygous mutants YLO415 and YLO416 was 0.5 μg/ml, and the MIC for its $FCY2^R/FCY2^R$ homozygous mutants YLO417 and YLO418 was higher than 64 μg/ml (Chia-Chen Wu, 2009). In this study, another clinical isolates YM060800 was chosen to create the *FCY2* homozygous mutation strains for the purpose to confirm the results. Both strains carried the least amount of SNPs at the *FCY2* locus, which made it easier to examine the factors that would influence the 5FC resistance to *C. tropicalis*. Isolates YM020291 and YM060800 were separately collected from middle and southern areas of Taiwan during the TSARY programs in 2002 and 2006. Also, they had different MLST diploid sequence types (DST 155 and DST 200, respectively) (Chou *et al.*, 2007; Li *et al.*, 2009). Thus, they were considered having different genetic backgrounds, and factors that contributed to the drug-resistance phenotype could be evaluated under different genomic backgrounds. Here, 5FC susceptibilities were analyzed among both clinical isolates as well as their engineered derivatives to determine whether the homozygosity at the *FCY2* locus responsible for the 5FC susceptibility in *C. tropicalis*. In agreement with previous result, the wild type strains YM020291 and YM060800 as well as their $FCY2^S/FCY2^S$ homozygous mutants were susceptible to 5FC, while their $FCY2^R/FCY2^R$ homozygous mutants were resistant to this drug. All in all, it has demonstrated unambiguously that loss of both $FCY2^S$ alleles or the existence of the *FCY2* homozygous resistant alleles was the reason for 5FC resistance in *C. tropicalis*, suggesting that the $FCY2^S$ allele appeared to be dominant and indicating the linkage between 5FC resistant phenotype and the occurrence of loss of heterozygosity.

4.2.2 The nonsense mutation in the *FCY2* gene, accompanied by LOH,

contributing to 5FC resistance in *C. tropicalis*

The effect of each polymorphic site on the 5FC susceptibility to *C. tropicalis* was further determined by evaluating the MICs of each single mutation strain. Although two kinds of solvents (DMSO and water) were used to dissolve the antifungal agent 5FC in the broth microdilution experiments, the results from each independent experiment has indicated that the usage of DMSO or water as solvents did not have obvious impact on the 5FC susceptibility to each strain. In fact, the strains that were resistant to 5FC in DMSO according to the MICs were still resistant to 5FC in water, and vice versa. Among the clinical isolate YM020291 and its genetically engineered strains, the results from E-test and broth microdilution have shown that only G145T single mutation strains and $FCY2^{R}/FCY2^{R}$ homozygous mutants were resistant to 5FC. Since they both exhibited a 145 T/T genotype, it implied the contribution of the SNP at position 145 in the *FCY2* gene, resulting in a nonsense mutation, to 5FC susceptibility.

Additional evidence on the link between nonsense mutations in the *FCY2* gene, coupled with LOH, and the 5FC resistance in *C. tropicalis* was demonstrated in 13 clinical isolate-derivative pairs based on the sequence analysis and their 5FC susceptibility testing results. Total 35 clinical isolates were divided into five groups according to the existence of nonsense mutations and the MICs for their derivatives. The 13 previous selected clinical isolate-derivative pairs were separated into groups A, B and C, while the remaining pairs were in groups D and E (Tables 7 and 8). For groups A to C, their entire coding sequence of *FCY2* gene as well as the flanking regions was confirmed by sequencing, which is different from groups D and E where only the 5' end of the coding sequence was checked. In group B, the MICs for the derivatives were considered as resistance, but no noticeable nucleotide change that were known to mediate 5FC resistance was observed at the *FCY2* locus as well as other three loci *FCY1*, *FUR1* and *URA3*, which were analyzed previously. It indicated that there might be other mechanisms contributed to the 5FC resistance in *C. tropicalis*. In group C, the derivatives were determined to be susceptible by broth microdilution. It might be caused by picking the colonies growing at the margin of the inhibition ellipse, which was, in fact, out of the inhibition zone. For strains that were classified in groups B and C, no LOH at the *FCY2* locus was detected in the derivatives. In contrast, the linkage between genotype and phenotype was discovered among the rest seven clinical isolate-derivative pairs, which were classified as group

A. In this group, while the parental strains were susceptible to 5FC, their progeny were resistant to the drug, except YM060088 clinical isolate-derivative pair (one of the two progeny exhibited intermediate susceptibility). This intermediate phenotype might be contributed by other mechanisms involving in the 5FC toxicity. Sequence analysis has shown that the *FCY2* locus was converted from the heterozygous state in parental strains to the homozygous state in their progeny. Most importantly, in six clinical isolate-derivative pairs, the SNP at position 145 was converted from the T/G genotype to the T/T genotype, while in the YM020715 clinical isolate-derivative pair, the SNP at position 201 was converted from the A/G genotype to A/A genotype. No matter which mutation (G145T or A201T) occurred in the *FCY2* gene, they would result in a nonsense mutation. In conclusion, nonsense mutations in one of the *FCY2* alleles, followed by the LOH, were responsible for the 5FC resistance observed in *C. tropicalis.*

4.2.3 Isolates exhibiting 145 T/G genotype might represent a subgroup associated with the generation of 5FC resistant derivatives

The data from multilocus sequence typing (MLST) and pulse-field gel electrophoresis (PFGE) revealed the distribution of 145 T/G and 201 A/G genotype among various subgroups of the clinical isolates, which is shown in Table 10 (Chou *et al.*, 2007). Total 97 clinical strains were divided into five groups according to their genotype and 5FC susceptibility phenotype. In brief, strains produced resistant progeny were put into groups 1, 2 and 3; in turn, each group represented the clinical isolates carrying heterozygous SNP (201 A/G), another heterozygous SNP (145 T/G), and homozygous SNPs (145 G/G and 201 G/G). Hyper-susceptible strains were classified as group 4 and the only 5FC resistant clinical isolate from the TSARY program was classified as group 5. In previous research, the clinical isolates were divided into six PFGE major types (I-VI). The strain YM020715 in group 1 belonged to type V, whereas the strains (YM020693, YM020274 and YM020291) in group 2 belonged to type III. It has suggested that strains between groups 1 and 2 were epidemiologically unrelated. In addition, while type II and IV were closely associated with the fluconazole susceptibility claimed by Chou *et al*. (2007), type III might relate to the 145 T/G phenotype, indirectly associated with the 5FC resistance. Similarly, three major clonal clusters (1-3) were identified in previous work (Chou *et al.*, 2007). Interestingly, clonal cluster 1 was corresponding to part of the group 2 because eight

strains with known DST number were included in this clonal cluster (4 in DST 134, 2 in DST 90, 1 in DST 153 and 1 in DST 155). Also, clonal cluster 1 was related to one major clonal cluster (clade 2) of isolates from the UK and the USA (Chou *et al.*, 2007; Tavanti *et al.*, 2005), which implied the prevalence of the strains exhibiting 145 T/G genotype in the *FCY2* gene. In addition, several PFGE subtypes were associated with the isolates in this cluster, indicated the existence of more than one clone possessed the heterozygous SNP at position 145 (T/G). One strain in group 3 belonged to clonal cluster 3 (DST 139). The reasons for the generation of 5FC resistant progeny in this group were supposed to be involved in other mechanisms. Two biggest populations of the clinical isolates were in DST 140 and DST 98, which belonged to clonal cluster 2. They were composed of the clinical isolates that did not generate resistant progeny, and sequence analysis of five strains that belonged to DST 140 (YM060173, YM060509 and YM060828) or DST 98 (YM060547 and YM060647) has shown the homozygous SNPs (145 G/G and 201 G/G). Chen *et al*. declared the existence of another subgroup representing 5FC resistant clinical isolates, where the majority was DST 164. However, no clinical isolates with known DST in this study belonged to the subgroup (Chen *et al.*, 2009). In conclusion, heterozygous SNP at position 145 (T/G) in *FCY2* gene may be a common genotype associated with the generation of 5FC resistant progeny in *C. tropicalis*.

4.3 Other possibilities that might involve in flucytosine resistance in *C. tropicalis*

Although this study has confirmed the expression of *FCY2* mRNA and the attribution of this gene to 5FC resistance in *C. tropicalis*, it could not completely explain the 5FC susceptibility in all derivatives. For example, one of the derivatives of the isolate YM060088 carrying 145 T/T genotype exhibited an intermediate phenotype, and some derivatives were resistant to 5FC without disrupting any protein that was known to involve in 5FC resistance mechanisms. A BLASTP search has indicated the existence of another homolog gene CTRG_00460. The amino acid sequence of this gene shared an identity of 43% with the *C. tropicals* Fcy2p, deduced from the $FCY2^{S}$ allele. Therefore, it is possibly that this *FCY2*-like gene contributed to the 5FC toxicity as well. It was similar to that found in *C. lusitaniae* and *C. albicans*,

where two *FCY2*-like genes were identified so far (Chapeland-Leclerc *et al.*, 2005; Hope *et al.*, 2004). In *C. albicans*, there was no evidence that either one or both genes (i.e., *FCY21* and/or *FCY22*) took part in encoding a functional PCP. However, in *C. lusitaniae,* the expression of mRNA was detected only in one of the genes, *FCY2*. The other gene was considered to be a pseudogene (Chapeland-Leclerc *et al.*, 2005). According to multiple sequence alignment analysis, *C. tropicalis* Fcy2p shared 65% amino acid sequence identity with *C. lucitaniae* Fcy2p and shared 43% identity with *C. lucitaniae* Fcy21p, suggesting that Fcy2p in *C. tropicalis* might correspond to the functional PCP. The research also demonstrated disruption of *FCY2* gene alone could confer the 5FC resistance in *C. lusitaniae* (Chapeland-Leclerc *et al*., 2005). However, another report has implied that Fcy2p is not solely responsible for 5FC uptake in *S. cerevisiae*. Four other *FCY2* homologues, *FCY21*, *FCY22*, *TPN1* and *FUR4*, were also contributed to 5FC toxicity, although in a less effective way (Paluszynski *et al.*, 2006).

All in all, the studies of *S. cerevisiae* and *C. lusitaniae* have indicated that there was only one functional PCP in each yeast species, although the existence of more than one *FCY2*-like pseudogene and could contribute to the 5FC toxicity. Accordingly, it would be necessary to study the expression of the *FCY2*-like genes in *C. tropicalis* to confirm that apart from *FCY2* gene whether other *FCY2*-like genes such as its paralogous gene CTRG_00460 may contribute to 5FC toxicity.

4.4 The cause of differential mRNA levels of *FCY2* in yeasts carrying a nonsense mutation

The mRNA levels of the *FCY2* gene in the G145T single mutation strains and the $FCY2^R/FCY2^R$ homozygous mutants were lower than those in the clinical isolates, homozygous susceptible strains and the T145G single mutation strains. It indicated the involvement of the nonsense mutation in the decrease of the *FCY2* mRNA level. Since the polymorphic nucleotide was located in the open reading frame instead of the promoter region, it is surprising to detect the influence of this SNP on the mRNA level. In fact, this phenomenon might be associated with a mechanism called nonsense-mediated mRNA decay, which leads to the initiation of mRNA decay (Jacobson and Peltz, 1996). The event that nonsense mutation in a gene reduced the abundance of the mRNA transcribed from that gene happened in both prokaryotes and eukaryotes

(Maquat, 1995; Peltz *et al.*, 1994). The position where nonsense mutation occurred was important for the triggering of the rapid decay of mRNA. Previous researches have shown that nonsense mutation located within the first two-thirds to three-quarters of the coding region would facilitate the decay rate of the encoded transcripts up to 20 folds (Hagan *et al.*, 1995; Yun and Sherman, 1995). In this study, the observed nonsense codon was located at amino acid position 49 in the deduced Fcy2p (509 a.a.), so it was within the range where a nonsense codon may increase the degradation rate of the *FCY2* transcript. Supposedly, one of the functions of nonsense-mediated mRNA decay is to minimize the possibility of any harmful products derived from the transcripts with a nonsense mutation. This mRNA surveillance function raised the question why the nonsense mutation maintained in a significant number of the clinical isolates (24.7%), as the cells worked so hard to protect themselves from the unnecessary deleterious transcripts. Previous researches have implied the activation of the salvage pathway, which could cause the accumulation of high concentration of pyrimidine nucleotides in cells, might be toxic for the cells (Blondel *et al.*, 2004; Seron *et al.*, 1999). Therefore, the existence of this nonsense mutation in *FCY2* gene, resulting in the inactivation of PCP, could give the mutation cells a selective advantage over other cells that did not possess this mutation.

4.5 Characterization of the loss of heterozygosity events

In this study, while one clinical isolate YM020291 remained heterozygous in all the polymorphic regions, the other isolate YM060800 showed consecutive homozygous regions on the left end side of the chromosome, extending up to 1.56 Mb. This large region of homozygosity also reported in another *C. tropicalis* strain MYA-3404, as well as other *Candida* species, such as *C. albicans*, *L. elongisporus* (Butler *et al.*, 2009). Because of the large chromosomal regions devoid of informative SNP markers, the identification of LOH events on the left end side of the chromosome was limited in the derivatives of YM060800. Apart from the absence of the useful SNP markers, sequence quality could also cause a problem in the SNP mapping process. For instance, the existence of mononucleotide repeats typically disturbed the analysis of nucleotides that were after the 3′ end of the repeat, which led to unclear or weak signals. These poor sequence data sometimes resulted in a false discovery of

SNP markers, and therefore produced a misjudgment of the polymorphic state. As a result, the polymorphic states of regions surrounding the LOH boundaries require additional sequencing check, such as using different primers to examine the corresponding regions or choosing adjacent areas to reevaluate the polymorphic states. The sequencing profile also revealed a region, sometimes followed mononucleotide repeats, exhibited abundant of superimposed peaks which extended for a long distance. For example, nucleotide sequence downstream of the gene CTRG_02086 in the clinical isolate YM060800 (from position 2208744 to 2208817, designated according to their relative location in the supercontig 2) has shown the extended superimposition pattern, whereas the corresponding region in its derivative YM060800-2 did not exhibit this pattern. The nucleotide sequence was compared between the clinical isolate YM060800 and its derivative YM060800-2. It has indicated that the cause of this sequencing profile was a deletion of adenine at position 2208754 of supercontig 2, and therefore generated a deletion polymorphism. Accordingly, while the polymorphic site at position 2208754 was heterozygous in the parental strain, it was homozygous in its derivative. In this case, it implied a smaller LOH boundary of 512 bp long (Figure 30).

SNP mapping has shown that four strains (YM020291-1, YM020291-2, YM060800-1 and YM060800) exhibited an extended homozygosity to the end of the supercontig 2. Although genome assemblies of *C. tropicalis* have not been mapped to chromosome yet, the distance from the 5' end of the supercontig 2 to the telomeric repeat was estimated to be approximately 4,930 bases (Butler *et al.*, 2009), suggesting the end of the supercontig 2 was likely compatible to the end of the chromosome in *C. tropicalis*. As a result, the LOH events that encompassed a large distal region (from 0.7 to 2.25 Mb) in four strains might be due to break-induced replication (BIR) or allelic recombination, whereas the small LOH region observed in the strain YM060800-2 might be caused by gene conversion (Cullen *et al.*, 2007). The majority of LOH events reported in this study were attributed to BIR/allelic recombination; however, the number of strains analyzed here could not draw any conclusion to the frequency of different LOH events regarding to the generation of 5FC resistance to *C. tropicalis in vitro.*

Six LOH boundaries were further investigated to identify features that might be involved in recombination. According to Baudat and Nicolas (1997), regions tending to introduce a double strand break (DSB), the initiation event of homologous

recombination, mostly located in intergenic promoter containing intervals (either one promoter/one terminator combination or two divergent promoters) (Baudat and Nicolas, 1997). In this study, two LOH boundaries (YM020291-1 and the left border in YM060800-2) fell in these intervals while three boundaries, covering a larger area, included such regions. However, it was surprising to observe one LOH boundary was falling within a coding region (CTRG_02123). In conclusion, of the six LOH boundaries, five were likely to be located in regions tending to initiate homologous recombination. However, except for the location of the boundaries, no strong evidence supports other features relating to the recombination was associated with the boundaries acquired in this study.

Chapter 5. Future work

5.1 Evaluation of the mechanism of 5FC resistance in *C. tropicalis*

The mechanisms of 5FC resistance of the clinical isolate YM060607 and the derivatives of seven clinical isolates (YM020438, YM020671, YM060075, YM060097, YM060210, YM060369 and YM060616) require further investigation. Based on the nucleotide sequence analysis, the 5FC resistant clinical isolate, YM060607, has a heterozygous SNP (C/T) at the position 431 of the *FUR1* gene, and a heterozygous SNP (A/G) at the position 775 of the *URA3* gene. Both single nucleotide variations (C431T in *FUR1* and G775A in *URA3*) lead to single amino acid changes (T144I in UPRT and A259T in ODCase), which might influence the functions of proteins involved in the 5FC resistance mechanisms. The contribution of these nucleotide changes to the 5FC resistance can be further investigated by constructing the point mutation strains. Genes, other than *URA3*, encoding proteins that are involved in the de novo pathway may play a role in 5FC resistance. Therefore, these genes can be examined if there are correlations between particular polymorphic nucleotides in the coding region and the resistance to 5FC. Also, alternations on gene expression, which may be due to changes in the 5' or 3'UTR region, or the interference of regulators of genes associated with 5FC resistance, could result in 5FC resistance. Accordingly, the mRNA level of these genes should be investigated. In addition, as mentioned in the discussion 4.3, the *FCY2*-like gene, CTRG_00460, might involve in 5FC toxicity and, therefore, the disruption of this gene is likely to influence the 5FC resistance in *C. tropicalis*. However, the contribution of this homolog gene to the 5FC resistance needs to be examined.

5.2 Examination of the features of the boundaries that might be related to recombination event.

At least three known factors might be associated with recombination and therefore relate to the LOH boundaries (i.e., the GC content, the SNP frequency and the distance to tRNA). Previous research has implied that high GC contents at local

scales might stimulate recombination rate (Marsolier-Kergoat and Yeramian, 2009). Nevertheless, a gene is often characterized by having a higher GC content relative to the background GC content for the entire genome and the length of a gene is directly proportional to a higher GC content (Pozzoli *et al.*, 2008). Therefore, the GC content of the boundaries with a comparable length scale could compare to either the background GC content or the GC content of the coding regions. Furthermore, the frequency of SNPs also plays a role in the recombination, which is considerable higher in hot spots relative to the genomic average (Brandstrom *et al.*, 2008). SNP frequency of one of the LOH boundaries, located in a coding region, is higher (one SNP per 67 bases) than the average rate of polymorphism in *C. tropicalis* (one SNP per 576 bases) (Pozzoli *et al.*, 2008); however, whether the SNP frequency around this boundary is actually higher than other regions of supercontig 2 needs to be determined by additional static analyses. Apart from the location of the LOH boundary and the SNP frequency, several researches have proposed the association of tRNA with gross chromosomal rearrangements (Diogo *et al.*, 2009; Dunham *et al.*, 2002; Dunn *et al.*, 2005). Here, two tRNA genes are found near the boundary of YM060800 (the closest distance of which was 3 kb), but their associations with the boundary should be determined by additional static analyses. Last but not the least, the sequence of each boundary can be compared and analyzed to search for hot spots of the recombination.

Reference

Abi-Said, D., Anaissie, E., Uzun, O., Raad, I., Pinzcowski, H., and Vartivarian, S. (1997) The epidemiology of hematogenous candidiasis caused by different Candida species. *Clin Infect Dis* **24**: 1122-1128.

Andersen, M.P., Nelson, Z.W., Hetrick, E.D., and Gottschling, D.E. (2008) A genetic screen for increased loss of heterozygosity in Saccharomyces cerevisiae. *Genetics* **179**: 1179-1195.

Arendrup, M., Horn, T., and Frimodt-Moller, N. (2002) In vivo pathogenicity of eight medically relevant Candida species in an animal model. *Infection* **30**: 286-291.

Baudat, F., and Nicolas, A. (1997) Clustering of meiotic double-strand breaks on yeast chromosome III. *Proc Natl Acad Sci U S A* **94**: 5213-5218.

Beck-Sague, C., and Jarvis, W.R. (1993) Secular trends in the epidemiology of nosocomial fungal infections in the United States, 1980-1990. National Nosocomial Infections Surveillance System. *J Infect Dis* **167**: 1247-1251.

Blondel, M.O., Morvan, J., Dupre, S., Urban-Grimal, D., Haguenauer-Tsapis, R., and Volland, C. (2004) Direct sorting of the yeast uracil permease to the endosomal system is controlled by uracil binding and Rsp5p-dependent ubiquitylation. *Mol Biol Cell* **15**: 883-895.

Brandstrom, M., Bagshaw, A.T., Gemmell, N.J., and Ellegren, H. (2008) The relationship between microsatellite polymorphism and recombination hot spots in the human genome. *Mol Biol Evol* **25**: 2579-2587.

Butler, G., Rasmussen, M.D., Lin, M.F., Santos, M.A., Sakthikumar, S., Munro, C.A., Rheinbay, E., Grabherr, M., Forche, A., Reedy, J.L., Agrafioti, I., Arnaud, M.B., Bates, S., Brown, A.J., Brunke, S., Costanzo, M.C., Fitzpatrick, D.A., de Groot, P.W., Harris, D., Hoyer, L.L., Hube, B., Klis, F.M., Kodira, C., Lennard, N., Logue, M.E., Martin, R., Neiman, A.M., Nikolaou, E., Quail, M.A., Quinn, J., Santos, M.C., Schmitzberger, F.F., Sherlock, G., Shah, P., Silverstein, K.A., Skrzypek, M.S., Soll, D., Staggs, R., Stansfield, I., Stumpf, M.P., Sudbery, P.E., Srikantha, T., Zeng, Q., Berman, J., Berriman, M., Heitman, J., Gow, N.A., Lorenz, M.C., Birren, B.W., Kellis, M., and Cuomo, C.A. (2009) Evolution of pathogenicity and sexual reproduction in eight Candida genomes. *Nature* **459**: 657-662.

Chapeland-Leclerc, F., Bouchoux, J., Goumar, A., Chastin, C., Villard, J., and Noel, T. (2005) Inactivation of the FCY2 gene encoding purine-cytosine permease promotes cross-resistance to flucytosine and fluconazole in Candida lusitaniae. *Antimicrob Agents Chemother* **49**: 3101-3108.

Chen, K.W., Chen, Y.C., Lin, Y.H., Chou, H.H., and Li, S.Y. (2009) The molecular

epidemiology of serial Candida tropicalis isolates from ICU patients as revealed by multilocus sequence typing and pulsed-field gel electrophoresis. *Infect Genet Evol* **9**: 912-920.

Chen, Y.C., Chang, S.C., Sun, C.C., Yang, L.S., Hsieh, W.C., and Luh, K.T. (1997) Secular trends in the epidemiology of nosocomial fungal infections at a teaching hospital in Taiwan, 1981 to 1993. *Infect Control Hosp Epidemiol* **18**: 369-375.

Chou, H.H., Lo, H.J., Chen, K.W., Liao, M.H., and Li, S.Y. (2007) Multilocus sequence typing of Candida tropicalis shows clonal cluster enriched in isolates with resistance or trailing growth of fluconazole. *Diagn Microbiol Infect Dis* **58**: 427-433.

Coste, A., Turner, V., Ischer, F., Morschhauser, J., Forche, A., Selmecki, A., Berman, J., Bille, J., and Sanglard, D. (2006) A mutation in Tac1p, a transcription factor regulating CDR1 and CDR2, is coupled with loss of heterozygosity at chromosome 5 to mediate antifungal resistance in Candida albicans. *Genetics* **172**: 2139-2156.

Coste, A., Selmecki, A., Forche, A., Diogo, D., Bougnoux, M.E., d'Enfert, C., Berman, J., and Sanglard, D. (2007) Genotypic evolution of azole resistance mechanisms in sequential Candida albicans isolates. *Eukaryot Cell* **6**: 1889-1904.

Cullen, J.K., Hussey, S.P., Walker, C., Prudden, J., Wee, B.Y., Dave, A., Findlay, J.S., Savory, A.P., and Humphrey, T.C. (2007) Break-induced loss of heterozygosity in fission yeast: dual roles for homologous recombination in promoting translocations and preventing de novo telomere addition. *Mol Cell Biol* **27**: 7745-7757.

Desnos-Ollivier, M., Bretagne, S., Bernede, C., Robert, V., Raoux, D., Chachaty, E., Forget, E., Lacroix, C., and Dromer, F. (2008) Clonal population of flucytosine-resistant Candida tropicalis from blood cultures, Paris, France. *Emerg Infect Dis* **14**: 557-565.

Diasio, R.B., Bennett, J.E., and Myers, C.E. (1978) Mode of action of 5-fluorocytosine. *Biochem Pharmacol* **27**: 703-707.

Diogo, D., Bouchier, C., d'Enfert, C., and Bougnoux, M.E. (2009) Loss of heterozygosity in commensal isolates of the asexual diploid yeast Candida albicans. *Fungal Genet Biol* **46**: 159-168.

Dunham, M.J., Badrane, H., Ferea, T., Adams, J., Brown, P.O., Rosenzweig, F., and Botstein, D. (2002) Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **99**: 16144 -16149.

Dunkel, N., Blass, J., Rogers, P.D., and Morschhauser, J. (2008) Mutations in the multi-drug resistance regulator MRR1, followed by loss of heterozygosity, are the

main cause of MDR1 overexpression in fluconazole-resistant Candida albicans strains. *Mol Microbiol* **69**: 827-840.

Dunn, B., Levine, R.P., and Sherlock, G. (2005) Microarray karyotyping of commercial wine yeast strains reveals shared, as well as unique, genomic signatures. *BMC Genomics* **6**: 53.

Fasoli, M., and Kerridge, D. (1988) Isolation and characterization of fluoropyrimidine-resistant mutants in two Candida species. *Ann N Y Acad Sci* **544**: 260-263.

Franz, R., Kelly, S.L., Lamb, D.C., Kelly, D.E., Ruhnke, M., and Morschhauser, J. (1998) Multiple molecular mechanisms contribute to a stepwise development of fluconazole resistance in clinical Candida albicans strains. *Antimicrob Agents Chemother* **42**: 3065-3072.

Fridkin, S.K., and Jarvis, W.R. (1996) Epidemiology of nosocomial fungal infections. *Clin Microbiol Rev* **9**: 499-511.

Hagan, K.W., Ruiz-Echevarria, M.J., Quan, Y., and Peltz, S.W. (1995) Characterization of cis-acting sequences and decay intermediates involved in nonsense-mediated mRNA turnover. *Mol Cell Biol* **15**: 809-823.

Hamilton, R., Watanabe, C.K., and de Boer, H.A. (1987) Compilation and comparison of the sequence context around the AUG startcodons in Saccharomyces cerevisiae mRNAs. *Nucleic Acids Res* **15**: 3581-3593.

Hitchcock, C.A. (1991) Cytochrome P-450-dependent 14 alpha-sterol demethylase of Candida albicans and its interaction with azole antifungals. *Biochem Soc Trans* **19**: 782-787.

Hope, W.W., Tabernero, L., Denning, D.W., and Anderson, M.J. (2004) Molecular mechanisms of primary resistance to flucytosine in Candida albicans. *Antimicrob Agents Chemother* **48**: 4377-4386.

Jacobson, A., and Peltz, S.W. (1996) Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annu Rev Biochem* **65**: 693-739.

Jund, R., and Lacroute, F. (1970) Genetic and physiological aspects of resistance to 5-fluoropyrimidines in Saccharomyces cerevisiae. *J Bacteriol* **102**: 607-615.

Kauffman, C.A., and Carver, P.L. (1997) Antifungal agents in the 1990s. Current status and future developments. *Drugs* **53**: 539-549.

Kelly, S.L., Lamb, D.C., Kelly, D.E., Manning, N.J., Loeffler, J., Hebart, H., Schumacher, U., and Einsele, H. (1997) Resistance to fluconazole and cross-resistance to amphotericin B in Candida albicans from AIDS patients caused by defective sterol delta5,6-desaturation. *FEBS Lett* **400**: 80-82.

Kern, L., de Montigny, J., Lacroute, F., and Jund, R. (1991) Regulation of the pyrimidine salvage pathway by the FUR1 gene product of Saccharomyces

cerevisiae. *Curr Genet* **19**: 333-337.

Kohler, G.A., White, T.C., and Agabian, N. (1997) Overexpression of a cloned IMP dehydrogenase gene of Candida albicans confers resistance to the specific inhibitor mycophenolic acid. *J Bacteriol* **179**: 2331-2338.

Kontoyiannis, D.P., Vaziri, I., Hanna, H.A., Boktour, M., Thornby, J., Hachem, R., Bodey, G.P., and Raad, II (2001) Risk Factors for Candida tropicalis fungemia in patients with cancer. *Clin Infect Dis* **33**: 1676-1681.

Kothavade, R.J., Kura, M.M., Arvind, V.G., and Panthaki, M.H. (2010) Candida tropicalis: its competitive prevalence, pathogenicity versus host defense and increasing resistance to fluconazole. *J Med Microbiol*.

Kozak, M. (1984) Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. *Nucleic Acids Res* **12**: 3873-3893.

Law, D., Moore, C.B., Joseph, L.A., Keaney, M.G., and Denning, D.W. (1996) High incidence of antifungal drug resistance in Candida tropicalis. *Int J Antimicrob Agents* **7**: 241-245.

Li, S.Y., Yang, Y.L., Lin, Y.H., Ko, H.C., Wang, A.H., Chen, K.W., Wang, C.W., Chi, H., and Lo, H.J. (2009) Two closely related fluconazole-resistant Candida tropicalis clones circulating in Taiwan from 1999 to 2006. *Microb Drug Resist* **15**: 205-210.

Liu, C.C., Simonsen, C.C., and Levinson, A.D. (1984) Initiation of translation at internal AUG codons in mammalian cells. *Nature* **309**: 82-85.

Louie, A., Kaw, P., Banerjee, P., Liu, W., Chen, G., and Miller, M.H. (2001) Impact of the order of initiation of fluconazole and amphotericin B in sequential or combination therapy on killing of Candida albicans in vitro and in a rabbit model of endocarditis and pyelonephritis. *Antimicrob Agents Chemother* **45**: 485-494.

Maquat, L.E. (1995) When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *Rna* **1**: 453-465.

Marr, K.A., Seidel, K., White, T.C., and Bowden, R.A. (2000) Candidemia in allogeneic blood and marrow transplant recipients: evolution of risk factors after the adoption of prophylactic fluconazole. *J Infect Dis* **181**: 309-316.

Marsolier-Kergoat, M.C., and Yeramian, E. (2009) GC content and recombination: reassessing the causal effects for the Saccharomyces cerevisiae genome. *Genetics* **183**: 31-38.

Natalini, P., Ruggieri, S., Santarelli, I., Vita, A., and Magni, G. (1979) Baker's yeast UMP:pyrophosphate phosphoribosyltransferase. Purification, enzymatic and kinetic properties. *J Biol Chem* **254**: 1558-1563.

Niimi, K., Monk, B.C., Hirai, A., Hatakenaka, K., Umeyama, T., Lamping, E., Maki,

K., Tanabe, K., Kamimura, T., Ikeda, F., Uehara, Y., Kano, R., Hasegawa, A., Cannon, R.D., and Niimi, M. (2010) Clinically significant micafungin resistance in Candida albicans involves modification of a glucan synthase catalytic subunit GSC1 (FKS1) allele followed by loss of heterozygosity. *J Antimicrob Chemother* **65**: 842-852.

Paluszynski, J.P., Klassen, R., Rohe, M., and Meinhardt, F. (2006) Various cytosine/adenine permease homologues are involved in the toxicity of 5-fluorocytosine in Saccharomyces cerevisiae. *Yeast* **23**: 707-715.

Papon, N., Noel, T., Florent, M., Gibot-Leclerc, S., Jean, D., Chastin, C., Villard, J., and Chapeland-Leclerc, F. (2007) Molecular mechanism of flucytosine resistance in Candida lusitaniae: contribution of the FCY2, FCY1, and FUR1 genes to 5-fluorouracil and fluconazole cross-resistance. *Antimicrob Agents Chemother* **51**: 369-371.

Peltz, S.W., He, F., Welch, E., and Jacobson, A. (1994) Nonsense-mediated mRNA decay in yeast. *Prog Nucleic Acid Res Mol Biol* **47**: 271-298.

Pozzoli, U., Menozzi, G., Fumagalli, M., Cereda, M., Comi, G.P., Cagliani, R., Bresolin, N., and Sironi, M. (2008) Both selective and neutral processes drive GC content evolution in the human genome. *BMC Evol Biol* **8**: 99.

Radding, J.A., Heidler, S.A., and Turner, W.W. (1998) Photoaffinity analog of the semisynthetic echinocandin LY303366: identification of echinocandin targets in Candida albicans. *Antimicrob Agents Chemother* **42**: 1187-1194.

Reuss, O., Vik, A., Kolter, R., and Morschhauser, J. (2004) The SAT1 flipper, an optimized tool for gene disruption in Candida albicans. *Gene* **341**: 119-127.

Ruan, S.Y., and Hsueh, P.R. (2009) Invasive candidiasis: an overview from Taiwan. *J Formos Med Assoc* **108**: 443-451.

Ryder, N.S. (1992) Terbinafine: mode of action and properties of the squalene epoxidase inhibition. *Br J Dermatol* **126 Suppl 39**: 2-7.

Schmidt, R., Manolson, M.F., and Chevallier, M.R. (1984) Photoaffinity labeling and characterization of the cloned purine-cytosine transport system in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **81**: 6276-6280.

Seron, K., Blondel, M.O., Haguenauer-Tsapis, R., and Volland, C. (1999) Uracil-induced down-regulation of the yeast uracil permease. *J Bacteriol* **181**: 1793-1800.

Sheehan, D.J., Hitchcock, C.A., and Sibley, C.M. (1999) Current and emerging azole antifungal agents. *Clin Microbiol Rev* **12**: 40-79.

Takagi, Y., Akada, R., Kumagai, H., Yamamoto, K., and Tamaki, H. (2008) Loss of heterozygosity is induced in Candida albicans by ultraviolet irradiation. *Appl Microbiol Biotechnol* **77**: 1073-1082.

Tavanti, A., Davidson, A.D., Johnson, E.M., Maiden, M.C., Shaw, D.J., Gow, N.A., and Odds, F.C. (2005) Multilocus sequence typing for differentiation of strains of Candida tropicalis. *J Clin Microbiol* **43**: 5593-5600.

Tsang, P.W., Cao, B., Siu, P.Y., and Wang, J. (1999) Loss of heterozygosity, by mitotic gene conversion and crossing over, causes strain-specific adenine mutants in constitutive diploid Candida albicans. *Microbiology* **145 (Pt 7)**: 1623-1629.

Vanden Bossche, H., Marichal, P., and Odds, F.C. (1994) Molecular mechanisms of drug resistance in fungi. *Trends Microbiol* **2**: 393-400.

Vermes, A., Guchelaar, H.J., and Dankert, J. (2000) Flucytosine: a review of its pharmacology, clinical indications, pharmacokinetics, toxicity and drug interactions. *J Antimicrob Chemother* **46**: 171-179.

Viviani, M.A. (1995) Flucytosine--what is its future? *J Antimicrob Chemother* **35**: 241-244.

Waldorf, A.R., and Polak, A. (1983) Mechanisms of action of 5-fluorocytosine. *Antimicrob Agents Chemother* **23**: 79-85.

Weber, E., Rodriguez, C., Chevallier, M.R., and Jund, R. (1990) The purine-cytosine permease gene of Saccharomyces cerevisiae: primary structure and deduced protein sequence of the FCY2 gene product. *Mol Microbiol* **4**: 585-596.

White, T.C., Marr, K.A., and Bowden, R.A. (1998) Clinical, cellular, and molecular factors that contribute to antifungal drug resistance. *Clin Microbiol Rev* **11**: 382-402.

Wingard, J.R. (1995) Importance of Candida species other than C. albicans as pathogens in oncology patients. *Clin Infect Dis* **20**: 115-125.

Yang, Y.L., Li, S.Y., Cheng, H.H., and Lo, H.J. (2005) Susceptibilities to amphotericin B and fluconazole of Candida species in TSARY 2002. *Diagn Microbiol Infect Dis* **51**: 179-183.

Yang, Y.L., Wang, A.H., Wang, C.W., Cheng, W.T., Li, S.Y., and Lo, H.J. (2008) Susceptibilities to amphotericin B and fluconazole of Candida species in Taiwan Surveillance of Antimicrobial Resistance of Yeasts 2006. *Diagn Microbiol Infect Dis* **61**: 175-180.

Yun, D.F., and Sherman, F. (1995) Initiation of translation can occur only in a restricted region of the CYC1 mRNA of Saccharomyces cerevisiae. *Mol Cell Biol* **15**: 1021-1033.

吳佳真 (2009) 熱帶念珠菌失去異質合子性導致 5-Flucytosine 抗藥性及鑑別 *CaNDT80* 活化區重要胺基酸之研究 Loss of heterozygosity contributes to the resistance to 5-flucytosine in *Candida tropicalis* and the study of identifying the crucial amino acids in the activation domain of *CaNDT80*

Table 1. Yeast strains used in this study

Strain	Parental strain	*FCY2* genotype[a]	Source
YM020291		wild-type, flucytosine susceptible	TSARY[b]
YM020291-1	YM020291	*FCY2*^{R}/ *FCY2*^{R}	Hui-Ching Ko
YM020291-2	YM020291	*FCY2*^{R}/ *FCY2*^{R}	Hui-Ching Ko
YLO417	YM020291	*FCY2*^{R}/ *FCY2*^{R} :: *SAT1-FLIP*	Chia-Chen Wu
YLO418	YM020291	*FCY2*^{R}/ *FCY2*^{R} :: *SAT1-FLIP*	Chia-Chen Wu
YLO415	YM020291	*FCY2*^{S}/ *FCY2*^{S} :: *SAT1-FLIP*	Chia-Chen Wu
YLO454	YM020291	*FCY2*^{R}/ *fcy2*^{S}(G486T) :: *SAT1-FLIP*	this work
YLO455	YM020291	*FCY2*^{R}/ *fcy2*^{S}(G486T) :: *SAT1-FLIP*	this work
YLO456	YM020291	*FCY2*^{S}/ *fcy2*^{R}(T486G) :: *SAT1-FLIP*	this work
YLO457	YM020291	*FCY2*^{S}/ *fcy2*^{R}(T486G) :: *SAT1-FLIP*	this work
YLO466	YM020291	*FCY2*^{R}/ *fcy2*^{S}(G145T) :: *SAT1-FLIP*	this work
YLO467	YM020291	*FCY2*^{R}/ *fcy2*^{S}(G145T) :: *SAT1-FLIP*	this work
YLO468	YM020291	*FCY2*^{S}/ *fcy2*^{R}(T145G) :: *SAT1-FLIP*	this work
YLO469	YM020291	*FCY2*^{S}/ *fcy2*^{R}(T145G) :: *SAT1-FLIP*	this work
YLO470	YM020291	*FCY2*^{R}/ *fcy2*^{S}(C1731T) :: *SAT1-FLIP*	this work
YLO471	YM020291	*FCY2*^{R}/ *fcy2*^{S}(C1731T) :: *SAT1-FLIP*	this work
YLO472	YM020291	*FCY2*^{S}/ *fcy2*^{R}(T1731C) :: *SAT1-FLIP*	this work
YLO473	YM020291	*FCY2*^{S}/ *fcy2*^{R}(T1731C) :: *SAT1-FLIP*	this work
YM060800		wild-type, flucytosine susceptible	TSARY[b]
YM060800-1	YM060800	*FCY2*^{R}/ *FCY2*^{R}	Hui-Ching Ko
YM060800-2	YM060800	*FCY2*^{R}/ *FCY2*^{R}	Hui-Ching Ko
YLO440	YM060800	*FCY2*^{R}/ *FCY2*^{R} :: *SAT1-FLIP*	Chia-Chen Wu
YLO441	YM060800	*FCY2*^{R}/ *FCY2*^{R} :: *SAT1-FLIP*	Chia-Chen Wu
YLO447	YM060800	*FCY2*^{S}/ *FCY2*^{S} :: *SAT1-FLIP*	Chia-Chen Wu
YLO448	YM060800	*FCY2*^{S}/ *FCY2*^{S} :: *SAT1-FLIP*	Chia-Chen Wu

[a] *FCY2*^{R}, 5-FC resistant allele; *FCY2*^{S}, 5-FC susceptible allele; *SAT1-FLIP*, the *SAT1* flipper cassette

[b] TSARY, Taiwan Surveillance of Antimicrobial Resistance of Yeasts

Table 2. Plasmids used in this study

Lab No.	Backbone	description	source
LOB319	pSFS2A	pSFS2A-derived plasmid containing an insertion of *FCY2*^{S} allele and its flanking sequences, which was amplified from YM020291.	Chia-Chen Wu
LOB320	pSFS2A	pSFS2A-derived plasmid containing an insertion of *FCY2*^{R} allele and its flanking sequences, which was amplified from YM020291.	Chia-Chen Wu
LOB381	LOB319	Mutation of LOB319, introducing a point mutation G273T (G486T) into *FCY2*^{S} allele.	this work
LOB382	LOB320	Mutation of LOB320, introducing a point mutation T273G (T486G) into *FCY2*^{R} allele.	this work
LOB383	LOB319	Mutation of LOB319, introducing a point mutation G-69T (G145T) into *FCY2*^{S} allele.	this work
LOB384	LOB320	Mutation of LOB320, introducing a point mutation T-69G (T145G) into *FCY2*^{R} allele.	this work
LOB385	LOB319	Insertion of the *Hind*III-*Hind*III fragment from LOB320	this work
LOB386	LOB320	Insertion of the *Hind*III-*Hind*III fragment from LOB319	this work

Table 3. Primers used in this study.

Name	Sequences (5' to 3')	Position
HJL814	CTCAACATGGAACGATCTAGC	*CaSAT1*: 465 bp to FRT
HJL1207	TGCCCATAAATTAAATGCAGAA	*CtFCY2*: +126~+147
HJL1208	GGAAGCAACAAACCCAAAAA	*CtFCY2*: +863~+844
HJL1210	GTGAAAACGAGCCAATCCAT	*CtFCY2*: +1550~+1530
HJL1420	ggtaccTCAACTCAACCCCAAAGT [b]	*CtFCY2*: -287~-270
HJL1422	CGGATTCAATGTAGCCAG	*CtFCY2*: +1485~+1502
HJL1423	GTCATTCCATGTCGTGGT	*CtFCY2*: +1971~+1954
HJL1424	ctcgagGTCATTCCATGTCGTGGT [c]	*CtFCY2*: +1971~+1954
HJL1477	CTGTTGCTCCAGGTGAATCA	*CtFCY2*: +2065~+2084 [a]
HJL1740	GTGCAGTTAATATCAT**G**TCTTCCGCTC [d]	*CtFCY2*: +470~+496
HJL1741	GAGCGGAAGA**C**ATGATATTAACTGCAC [d]	*CtFCY2*: +496~+470
HJL1742	GTGCAGTTAATATCAT**T**TCTTCCGCTC [d]	*CtFCY2*: +470~+496
HJL1743	GAGCGGAAGA**A**ATGATATTAACTGCAC [d]	*CtFCY2*: +496~+470
HJL2100	CATAAATTAAATGCA**G**AAACTAAAGGTATTG [e]	*CtFCY2*: +130~+160
HJL2101	CAATACCTTTAGTTT**C**TGCATTTAATTTATG [e]	*CtFCY2*: +160~+130
HJL2102	CATAAATTAAATGCA**T**AAACTAAAGGTATTG [e]	*CtFCY2*: +130~+160
HJL2103	CAATACCTTTAGTTT**A**TGCATTTAATTTATG [e]	*CtFCY2*: +160~+130
HJL2104	CTTCTCCTTAACTACCTTTTCCTCC	*CtFCY2*: -243~-219
HJL2117	TCATCTCATTTCATTTGGTACAAAGG	*CtFCY2*: +1255~+1280
HJL2124	CAGACTTTGTTGGTGGTAGAAC	*CtFCY2*: +692~+713
HJL1937	TCCGACTTCTCCAAAGTTGA	*CtEFB1*: +10~+29
HJL1938	TTGGACGTGGTCTTCATCTTC	*CtEFB1*: +819~+799
HJL2299	GTAATGCTGCTACTTTAGCC	*CtFCY2*: +1139~+1158
HJL2300	TTCACCGATTTGTCTACCA	*CtFCY2*: +1437~+1419
HJL2301	ATCACCATTGGTAACGAAAG	*CtACT1*:+637~+656
HJL2302	TGGAAAGTAGACAAGGAAGC	*CtACT1*:+953~+934

(Continued)

[a] The position of this primer is at the downstream of the stop codon of the *FCY2* gene.

[b] Lower case shows restriction sites for *Kpn*I.

[c] Lower case shows restriction sites for *Xho*I.

[d] Primers that were used to introduce single mutation at position 486 by fusion PCR. The single mutation is indicated by the underlined bold letters.

[e] Primers that were used to introduce single mutation at position 145 by fusion PCR. The single mutation is indicated by the underlined bold letters.

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL1494	AAGAAACAGAATACGATGAGTGATG	CTRG_01076: +58~+82
HJL1495	AGGTTGGTACTGGGTAACATGA	CTRG_01076: +1071~+1050
HJL1496	ATGCCAAAGAAAACTCAGGG	CTRG_01434: +1~+20
HJL1497	GGGTACATTTAGAAAGACTAGCCG	CTRG_01434: +1063~+1040
HJL1498	TGTCATCATCTTCATCTAATCAAGC	CTRG_01816: +2~+26
HJL1499	CCATCTCCACCAAGATAACCAA	CTRG_01816: +1052~+1031
HJL1500	ATTGATTAGCGACCCAACTGC	CTRG_02123: +111~+131
HJL1501	CCCATTGTTCCAAACCCTTT	CTRG_02123: +1054~+1035
HJL1621	TCGTTACTGGAGCAACAGGATTC	CTRG_01886: +23~+45
HJL1622	TAAGGTAAGGAATCGCCAAGG	CTRG_01886: +1037~+1017
HJL1623	TTTCATCTACTGCTGCTCCTGG	CTRG_01926: +3968~+3989
HJL1737	TGCTGGAGATGATGAGGATGAAG	CTRG_01926: +4974~+4952
HJL1627	AACACCTGAAGATACCAAACTGTC	CTRG_02012: +12~+35
HJL1628	AGTGGAAGTAGCAGTAGTAGCAGTG	CTRG_02012: +1530~+1506
HJL1629	AGGGATTCAGCAAGGTGACTTTG	CTRG_02080: +192~+214
HJL1630	TCTTTCTCCTCATCTTTCTGTTCG	CTRG_02080: +1145~+1122
HJL1631	TTGGGCTGATGATGCTGTTG	CTRG_02099: +609~+628
HJL1632	TTTCACCCGCTCTCCATCTC	CTRG_02099: +1468~+1449
HJL1717	AAACGCCATTAAGAGATTCAGTCC	CTRG_01843: +39~+62
HJL1718	TTCAATGCTAGTGACAGTGTTGGTG	CTRG_01843: +1077~+1053
HJL1719	TGCCGTAAAGATATTGGGATTCAG	CTRG_01858: +48~+71
HJL1720	TGGGTGTGACTTATGGTTGATGAC	CTRG_01858: +1080~+1057
HJL1721	AAGGTTACAGGATGCCGTGGAG	CTRG_02083: +96~+117
HJL1722	TGAGATCCCAGACGCAAATCAG	CTRG_02083: +1238~+1217
HJL1723	TGTATGCCGATAAGACAGTACCTGG	CTRG_02089: +2~+26
HJL1724	ATCCAAACAATCAGATAAGCCACC	CTRG_02089: +1227~+1204
HJL1725	CCATTCAAGGCAGACATAGATGA	CTRG_02092: +3001~+3023
HJL1726	AATCTCCTGCCTTGTAGAAATGG	CTRG_02092: +3799~+3777
HJL1727	ATGACAACGACAGTGCCATTG	CTRG_02095: +1~+21
HJL1728	TTACCAGCAGCACCTTCATTAAC	CTRG_02095: +938~+916
HJL1729	TTCGCACACACACACACTCTCTC	CTRG_02103: -459~-437

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL1730	GATAAGCCACCAGTTGAAGAAGATG	CTRG_02103: +884~+860
HJL1758	TGTTGTTGTTGTTTTCTTGCTC	CTRG_02108: -44~-23
HJL1732	TCGGCATAGCGTTGGCATAC	CTRG_02108: +1200~+1181
HJL1832	TATCGTTGGTCTAGTGTTGTGAATC	CTRG_02109: -274~-250
HJL1833	ATCCATATTCTCCATTCCTTCATC	CTRG_02109: +834~+811
HJL1733	AAAGAAGAGAGGTTTGTTTGGTTCC	CTRG_02114: -231~-207
HJL1734	ACTTGATGGTGGAGTAATCGGC	CTRG_02114: +821~+800
HJL1735	AAAGAATGTGATACGAGCCAAGC	CTRG_02117: -411~-389
HJL1736	AGGGCTCTAGTCAACTCAACATCAC	CTRG_02117: +786~+762
HJL1759	CATCAACAACAACAAGAACATCG	CTRG_01823: +97~+119
HJL1760	ACCTGGGTGTTGAATGTACTGTC	CTRG_01823: +965~+943
HJL1769	CAAAGTTGTGGTGAAGACCTCATC	CTRG_01837: +192~+215
HJL1770	ACTGTGTCTCATCATCATTGGCAC	CTRG_01837: +991~+968
HJL1824	GCTTCTTGGATATGAACACTAACC	CTRG_01944: +11~+34
HJL1825	AAATAGTTGGGTGTTGACCTTG	CTRG_01944: +964~+943
HJL1826	TCCTTCCTACATACATTGCCTCG	CTRG_01991: -403~-381
HJL1827	TGGAGCAGGAACAGGAGCAG	CTRG_01991: +735~+716
HJL1828	AGCATCATTACAGCCTACCACTC	CTRG_02034: +327~+349
HJL1829	AGAAGAATCAGATGTCACCTCAGC	CTRG_02034: +1641~+1618
HJL1830	CTAAGATGTTGTTCAAGAACCGTG	CTRG_02082: +1175~+1198
HJL1831	GACACTCTTCAACTATGTGCTTACC	CTRG_02082: +2236~+2260
HJL1834	GTCACACAGCCAAGAGATAATAGTC	CTRG_02119: +3~+27
HJL1835	ATGACCTGCCTGAAGTAATATCAC	CTRG_02119: +927~+904
HJL1836	AAATGGGATGAATGGGTAGG	CTRG_01820: +223~+242
HJL1837	CAGTTGTAGATACACCATCCATTG	CTRG_01820: +901~+878
HJL2113	CATACGAGGCTAGAACGATTGG	CTRG_01817: -374~-353
HJL2114	TGGAAATATCCCGAATCTCTACAAG	CTRG_01817: +963~+939
HJL2115	GGTCAACAAGATGCCATTCTG	CTRG_01993: +352~+372
HJL2116	AACCTTCAGGAGTTTCACCAAC	CTRG_01993: +1345~+1324
HJL2129	ATACCTCTTTGTAGAGTGTGAAACG	CTRG_02120: -472~-448
HJL2130	TTCCATCATACTGACCCTTTG	CTRG_02120: +767~+787

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL2131	GTATGGCTGCTTGTTTGGTAAC	CTRG_02122: -2~+20
HJL2132	GGCAACCAACCTCTGAACAT	CTRG_02122: 712~731
HJL2179	GAAAGAGAAAGATGGCTCGTTCATC	CTRG_01999: +14~+38
HJL2180	CATTTGTCCTCATTGTGCTGAATG	CTRG_01999: +927~+950
HJL2183	AAGATCCCCTGTTTGTTCCCTCAAC	CTRG_02008: -40~-16
HJL2184	CAATGAACTTCTTTCCCACCAGC	CTRG_02008: +1075~+1097
HJL2189	ACCAGACATCCCGATTTATAGTTC	CTRG_02080: +161~+138
HJL2190	AGCAACACTGTGATGTAAGTTGTTC	CTRG_02080: -977~-953
HJL2191	CACTGTCAAACAGCATTTCCAG	CTRG_02081: +1224~+1203
HJL2192	GAGATCCAAGTTGTCCATCATTC	CTRG_02081: +39~+61
HJL2193	TGTCTTTCACCAAGGGAGAAGTTTG	CTRG_02081: +233~+209
HJL2194	GAATGTCAAAGCTGTTGCCAAGA	CTRG_02082: +2886~+2908
HJL2195	AAGGGATCTTTGCGTTGCTTC	CTRG_02082: +3192~+3172
HJL2196	GAAGTTGTTACCAGCAGGTGATAG	CTRG_02082: +2079~+2102
HJL2197	CGAAAGATTAGCCACTCAATATG	CTRG_02009: +1113~+1135
HJL2198	TGCTGAAACTCCTCCAAGTC	CTRG_02009: +2211~+2192
HJL2199	TGAGTGAGGGAGTATGTTACCAAG	CTRG_02010: -447~-424
HJL2200	CCTGAAGGATGATGTCTTGTTG	CTRG_02010: +1265~+1244
HJL2201	GATGGTGATGGTGAAGGTGATA	CTRG_02011: -431~-424
HJL2202	ATTCTTCGGAAATAACGGTAGC	CTRG_02011: +529~+508
HJL2203	TAATGACGATCCCACTTATTGGAAG	CTRG_01138: -2~+23
HJL2204	TCACCTGAAACCAATGAACAACAC	CTRG_01138: +1090~+1067
HJL2205	CATTCGTCCTTCAAGTTGGTTTG	CTRG_01190: +102~+124
HJL2206	TTCTTATGAGGTTTCTCACTTCCCA	CTRG_01190: +698~+674
HJL2207	ATACCCACAGGACTGTTCCATTG	CTRG_01239: +540~+562
HJL2208	CGCACCTGGATCTTTGTAGTCAG	CTRG_01239: +1635~+1613
HJL2209	TGGTTTCAGTTACTTTGGGAAC	CTRG_01337: -46~-25
HJL2210	CAGCGTCAGTATATCCCATTATG	CTRG_01337: +886~+864
HJL2211	CCGATTTCACTATTGATTCACATGG	CTRG_01490: +140~+164
HJL2212	CCCAAAACAGCTCTTTCACTAACAA	CTRG_01490: +1163~+1139
HJL2213	ACCACAATCATCCAGAAAAGGGAAG	CTRG_01532: +39~+63

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL2214	GCAACTGGAACTCTTGATTCAACCT	CTRG_01532: +1187~+1163
HJL2215	GTCGCAAGAAACGATGTTGAACATG	CTRG_01598: +463~+487
HJL2216	CCTTCTTTTGACCCAATCTATGTGC	CTRG_01598: +1555~+1531
HJL2217	TCCTAGAAGTTTGAAAAGCGAATCC	CTRG_01630: +21~+45
HJL2218	TGAATCTGCTCTTCTTGGCGATATT	CTRG_01630: +1074~+1050
HJL2219	TCTGCTAGAACCTCCTTGAAGAACG	CTRG_01722: +7~+31
HJL2220	CCAGCATCTTTCAAAGCCTTCTTA	CTRG_01722: +1055~+1032
HJL2221	ACCACCCTTGAAAAGAGTTCTATC	CTRG_02010: +39~+62
HJL2222	AAACAACCATCGCAAGACAAGG	CTRG_02084: +333~+354
HJL2223	GCCTTCACTTCCTTAATACCCAATT	CTRG_02084: +1388~+1364
HJL2224	TTTTGACGATTGAGAGTAAGTGC	CTRG_02086: -301~-279
HJL2225	AATCGTCATCATCATCATCGTC	CTRG_02086: +752~+732
HJL2226	CTCGTCTTCATCTTCATCATCG	CTRG_02010: +1265~+1244
HJL2229	AACCGATCAACATTTGGATTACTCC	CTRG_01092: +61~+85
HJL2230	GTTATAGCACCGTCATTGGTCAAAT	CTRG_01092: +1244~+1220
HJL2231	TAACAGCCATTTGGGACGAAGC	CTRG_01100: -446~-425
HJL2232	AAGTGAAGGTATTCGGTGGTGTTC	CTRG_01100: +806~+783
HJL2233	AACTGAAATACGCAAGGGTTCTGG	CTRG_01112: -74~-52
HJL2234	TTGAAACTGTGGTACTGGTTGTGG	CTRG_01112: +1225~+1202
HJL2235	TGCTTGGATTATGGGTTATTACAGG	CTRG_01120: +412~+436
HJL2236	TCCCTGTGTCAATATCGCCTTG	CTRG_01120: +1708~+1687
HJL2237	TGGCTGACCCATTGGCATC	CTRG_01128: +302~+320
HJL2238	GTCTCAACTGGGTACAACGCAAG	CTRG_01128: +1540~+1518
HJL2239	ACCACTACCACAACCACCACAAC	CTRG_02104: -36~-14
HJL2240	GACAAACAAACAAACAAGGCAGC	CTRG_02104: +1264~+1242
HJL2241	CATCATTTCAACAACAACAACCACC	CTRG_02105: -79~-55
HJL2242	AAATTCATAGATGCGAGGGAGGTAG	CTRG_02105: +1264~+1242
HJL2243	TTTCATCCAGTTCAGTTCCTCATTC	CTRG_02106: -70~-46
HJL2244	GGGTTCAGGGAATTTTGTAGCAG	CTRG_02106: +1362~+1340
HJL2247	ATCCAAAGCGTTCAAGACTTC	CTRG_02120: +396~+376
HJL2248	GCTGGTGGTATGGTTGAACAAATAC	CTRG_02087: +43~+67

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL2249	GCACCCAACACATATCCATTTTCAT	CTRG_02087: +1182~+1158
HJL2250	TGGTGTACTTGAATCTTCTGCCATC	CTRG_02088: -497~-473
HJL2251	TTCAAAGTGTCAGGAAATGCTGG	CTRG_02088: +731~+709
HJL2252	TGATTGATCTTGCATTCGACC	CTRG_02009: +611~+591
HJL2253	TATCCTGCTGCTTCAGTACAAAG	CTRG_02009: -703~-681
HJL2254	GCTTCATCATTTCCAATCATTG	CTRG_02009: -323~-344
HJL2255	GTTCAACAATCATCAAGTGATCC	CTRG_02009: -1684~-1662
HJL2256	CACGTCTGATTCTTACACAAGTTAG	CTRG_02010: -1973~-1949
HJL2257	ATCCAGAATAATAGGTACACCAAAG	CTRG_02010: -792~-817
HJL2258	GAATGATGGACAACTTGGATCTC	CTRG_02081: +39~+61
HJL2259	ACAAAACATTCACTCCCACAACAAC	CTRG_01080: -41~-17
HJL2260	CGTGGATGAGGAGATCAACATTGG	CTRG_01080: +1407~+1384
HJL2261	TTCTTCTCTACAACGAGTCCCCAAC	CTRG_01086: -183~-159
HJL2262	CAATCTCTGCTTTCCAACCTTTTCT	CTRG_01086: +997~+970
HJL2267	CTTGAGTGTCGTTCCACTTGAATTG	CTRG_01098: -157~-133
HJL2268	CTGGTGGTGGTAGTTTTCATTGC	CTRG_01098: +1172~+1150
HJL2269	AACTTCTTCACCAAACCAACACAAG	CTRG_01438: +66~+90
HJL2270	AATCCATCGGGAACAAATTGC	CTRG_01438: +1537~+1517
HJL2271	TCAACATACTCATTGGGGAAACATG	CTRG_01541: -267~-243
HJL2272	TGCGTTACCTGAAGTGAAGATGG	CTRG_01541: +975~+953
HJL2273	ATCACATCGAGCACACCCAGTTC	CTRG_01646: -40~-18
HJL2274	CACCAACCAGTATCCCTGTCTTTC	CTRG_01646: +1389~+1366
HJL2275	CGACTCCCTACCAAAATGTCAGG	CTRG_01725: -15~+8
HJL2276	CCGTTGAGGGTGAACCATCC	CTRG_01725: +1238~+1219
HJL2277	GAGTATGAAGATGAGGTTACCAAGG	CTRG_02087: -872~-803
HJL2278	AATGGCTCAACGAAGGTTCAG	CTRG_02087: +385~+365
HJL2279	GATCCAAGACAAATCCAAGC	CTRG_02120: +550~+569
HJL2280	CAGGCTCTTTGTTCAGTGG	CTRG_02120: +2304~+2286
HJL2281	AATACAAAAAATCAAATCAAGAAAC	CTRG_02120: +1640~+1616
HJL2282	GTTGTGGAATAAGTTGCTTGTCTC	CTRG_02120: +2547~+2570
HJL2283	ACCACGAGCAATTCAGAAAGAG	CTRG_02120: +3834~+3813

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL2284	GTACCAATAAGACATCCTGCTTC	CTRG_02103: +347~+325
HJL2285	CTCTGATAAATTAAAAAGGAACAGC	CTRG_02103: -1158~-1134
HJL2286	AACACTTGTAGTGATCGACCTC	CTRG_02103: -733~-712
HJL2287	AAGCCTTGGTTGTGTGTGATAC	CTRG_02104: +1000~+1021
HJL2288	TTGTTTCTGTGAAGGTGATGAAC	CTRG_02104: +2074~+2052
HJL2289	TGTTCAACCACCAAACCATGATTAG	CTRG_01078: -17~+8
HJL2290	AGAATGTCTTCCCTCAACCCAAG	CTRG_01078: +1267~+1245
HJL2293	CAATGTTGATCTCCTCATCCACG	CTRG_01080: +1385~+1407
HJL2294	ACTTGCCAATGTGTTGCTTATGG	CTRG_01080: +2903~+2881
HJL2295	GCCCAGTGATCCTGAAGAAGATTAC	CTRG_01082: +3~+27
HJL2296	CAATGATAGGGTGTTGCGTCG	CTRG_01082: +1396~+1376
HJL2297	AAGGTCACCACTACAACCATCAAAG	CTRG_01082: +94~+118
HJL2298	ATTTCTGACAGGTTCCAGCATATTG	CTRG_01082: +1350~+1326
HJL2303	CTATGTTGGGTAGGGGGAATTATGC	CTRG_02123: -409~-385
HJL2304	CAAAGCAAGAGCAATCAAAGACACC	CTRG_02123: +1045~+1012
HJL2305	CTCTTCATCGTACTCTTCCATCC	CTRG_02123: +967~+989
HJL2306	TTCGGATATTCCAAACAAGGAC	CTRG_02123: +2440~+2419
HJL2307	AAGGAGACTAAGGAGATGTTTGC	CTRG_01076: +622~+644
HJL2308	CACATCCTGTCTGTCAAACTCTC	CTRG_01076: +2259~+2237
HJL2309	TTGTTGAACGATGAAGAAGACG	CTRG_01076: +724~+745
HJL2310	AAACGTCTCCAAATCACATTCC	CTRG_01076: +1390~+1369
HJL2313	TTGAGGGATTGGATATTCAAGCATC	CTRG_01745: +86~+110
HJL2314	TGATTGAGCCTTGTTCCTAGAAGC	CTRG_01745: +1365~+1342
HJL2315	ATTACATTTTTGTCACGACCACACC	CTRG_01756: -69~-45
HJL2316	CTTTTGTCTTCTTTCAGCGGCTT	CTRG_01756: +1143~+1121
HJL2317	CATCCAATTCGTCTGTGAATGATTC	CTRG_01765: +95~+119
HJL2318	ACGACGGTGTTCCATTTTCTTG	CTRG_01765: +1478~+1457
HJL2319	ACTAAAATCACAACAGCAGGGTTTG	CTRG_01775: -73~-49
HJL2320	TTTCTTCATGGTGCTCATCTTGG	CTRG_01775: +1267~+1245
HJL2321	GTGTGAGAGAGCGAATAGTTTGACG	CTRG_01785: -106~-82

Table 3. Continued

Name	Sequences (5' to 3')	Position
HJL2322	GCAAGGTAGTTTTCAAACAGTCACG	CTRG_01785: +1237~+1213
HJL2323	TCTTCACACTCACGCATATACCATG	CTRG_01793: -22~+3
HJL2324	AGTTTTACTGATTGGGTTGATGACG	CTRG_01793: +1554~+1530
HJL2325	AGAGAAGAACCACAAGTCCGAGAAG	CTRG_01804: +12~+36
HJL2326	CGATACCTCCATCTTCGAGTGC	CTRG_01804: +1513~+1492
HJL2329	AAGAAAATATCACGCCCTTGAGTTC	CTRG_01796: +167~+191
HJL2330	GTAACCAGCATCTTCTTCATTGTCC	CTRG_01796: +1683~+1659
HJL2331	ATGCCACAACTGATTCTTTCCAATC	CTRG_01798: +158~+182
HJL2332	TTGACTTCCCCATTTACATCGTCT	CTRG_01798: +1586~+1563
HJL2333	CAGTTGAAGGTCAAACCCCTGC	CTRG_01799: +116~+137
HJL2334	CATTCTGGACCACAGTTTTCGTTC	CTRG_01799: +1406~+1383
HJL2335	TTGAATCACTTGGTGGTGGAATAAG	CTRG_01800: +197~+221
HJL2336	TTTTACCGTCATCCTTGTGGTTG	CTRG_01800: +1551~+1529
HJL2337	ATGCTGATGAAGAATTACGTGAAGC	CTRG_01801: +263~+287
HJL2338	AACAATTCACCCATTCTCCACATC	CTRG_01801: +1756~+1733
HJL2339	AAAGAGAAGCTAAGAAAGCCGCTGC	CTRG_01803: +47~+71
HJL2340	GCTTCACCACTCAAACCATAACG	CTRG_01803: +1520~+1498
HJL2341	AAATGCTCTTGTGGTTCCGAAG	CTRG_01802: +364~+385
HJL2342	CTGGACGAGATAAGAGTCGAACTC	CTRG_01802: +1313~+1290
HJL2343	TGGACTAAGAGTCGTGGAGAAGATG	CTRG_01802: +1066~+1090
HJL2344	CCAAATGGTATCGGTAAATCCACT	CTRG_01802: +2057~+2034
HJL2345	ACCATTAAGGAAATCTTGGGAGTGG	CTRG_01803: +861~+837
HJL2346	AAATGCCTTTCCATGCTAGATCG	CTRG_01803: -596~-574
HJL2347	ATGGATTTCAATAAGGTGGATTTAC	CTRG_01803: +377~+353
HJL2348	GTTTTTGATGATAATGACGAGGTGG	CTRG_01804: -751~-727
HJL2349	CTTGGTCTCGATCTGGCATAAAAC	CTRG_01804: +335~+312
HJL2350	TATTTGTGGGGAGTTGATGATG	CTRG_02104: -1219~-1198
HJL2351	CACCAATAGATTTACCATGAACAG	CTRG_02104: +403~+380
HJL2352	TAGACGATTTTGTAAGGTTGGAAGG	CTRG_02104: -561~-537
HJL2353	AGGAGGAAGAAGTGTGAGAAATC	CTRG_02104: -205~-227
HJL2322	GCAAGGTAGTTTTCAAACAGTCACG	CTRG_01785: +1237~+1213

Table 4. Summary of broth microdilution data with 5FC/DMSO

Strain	Parental strain[b]	Position of SNP in *FCY2*			MIC (μg/ml) in each experiment[d]		
		-69 (145)	273 (486)	1518[c] (1731)	1	2	3
YM020291		T/G	T/G	T/C	0.5	0.5	0.5
YLO417	2291	T/T	T/T	T/T	64	64	≥64
YLO415	2291	G/G	G/G	C/C	0.25	0.25	0.5
YLO454	2291	T/G	T/T	T/C	2	2	1
YLO455	2291	T/G	T/T	T/C	-	2	-
YLO456	2291	T/G	G/G	T/C	0.5	0.5	1
YLO457	2291	T/G	G/G	T/C	-	0.5	-
YLO466	2291	T/T	T/G	T/C	-	-	≥64
YLO467	2291	T/T	T/G	T/C	-	-	≥64
YLO468	2291	G/G	T/G	T/C	-	-	0.5
YLO469	2291	G/G	T/G	T/C	-	-	0.25
YLO470	2291	T/G	T/G	T/T	-	-	0.5
YLO471	2291	T/G	T/G	T/T	-	-	0.5
YLO472	2291	T/G	T/G	C/C	-	-	0.5
YLO473	2291	T/G	T/G	C/C	-	-	0.5
YM060800		T/G	T/G	T/C	0.5	0.5	-
YLO440	6800	T/T	T/T	T/T	-	≥64	-
YLO441	6800	T/T	T/T	T/T	≥64	≥64	-
YLO447	6800	G/G	G/G	C/C	0.5	0.25	-
YLO448	6800	G/G	G/G	C/C	-	0.25	-
YLO6[a]					-	16	16
YLO7[a]					-	0.125	0.25
YLO12[a]					-	1	1

[a] YLO12 (*Candida albicans* ATCC90028, 0.5~2.0 μg/ml) was used as a reference strain, and YLO6 (*Candida krusei* ATCC6258, 4.0~16.0 μg/ml) and YLO7 (*Candida parapsilosis* ATCC22019, 0.12~0.5 μg/ml) were incorporated as quality control strains.

[b] 2291, YM020291; 6800, YM060800

[c] The position of this SNP is at 201 nt downstream of the stop codon of the *FCY2* gene.

[d] 5FC susceptibility was determined using 96-well broth microdilution assays. 5FC was dissolved in DMSO. –, not done.

Table 5. Summary of broth microdilution data with 5FC/H_2O

Strain	Parental strain[b]	Position of SNP in *FCY2*			MIC (μg/ml) in each experiment[d]			
		-69 (145)	273 (486)	1518[c] (1731)	4	5-1	5-2	5-3
YM020291		T/G	T/G	T/C	0.125	0.25	0.25	0.5
YLO417	2291	T/T	T/T	T/T	≥64	≥64	32	64
YLO418	2291	T/T	T/T	T/T	32	≥64	32	≥64
YLO415	2291	G/G	G/G	C/C	0.125	0.25	0.25	0.25
YLO454	2291	T/G	T/T	T/C	0.125	2	0.125	1
YLO456	2291	T/G	G/G	T/C	0.125	1	0.5	0.5
YLO466	2291	T/T	T/G	T/C	64	≥64	≥64	64
YLO467	2291	T/T	T/G	T/C	64	≥64	≥64	≥64
YLO468	2291	G/G	T/G	T/C	0.125	0.5	0.5	0.5
YLO469	2291	G/G	T/G	T/C	0.125	-	-	-
YLO470	2291	T/G	T/G	T/T	0.125	0.25	0.5	2
YLO472	2291	T/G	T/G	C/C	0.125	0.25	0.5	0.5
YLO6[a]					16	8	16	8
YLO7[a]					2	0.125	0.125	0.125
YLO12[a]					1	1	1	0.5

[a] YLO12 (*Candida albicans* ATCC90028, 0.5~2.0 μg/ml) was used as a reference strain, and YLO6 (*Candida krusei* ATCC6258, 4.0~16.0 μg/ml) and YLO7 (*Candida parapsilosis* ATCC22019, 0.12~0.5 μg/ml) were incorporated as quality control strains.

[b] 2291, YM020291; 6800, YM060800

[c] The position of this SNP is at 201 nt downstream of the stop codon of the *FCY2* gene.

[d] 5FC susceptibility was determined using 96-well broth microdilution assays. 5FC was dissolved in sterile distilled water. –, not done.

Table 6. Summary of 5FC susceptibility testing

Strain	Position of SNP in *FCY2*			5FC MIC (μg/ml)		Susceptibility[c]
	-69 (145)	273 (486)	1518[a] (1731)	Broth microdilution[b]	E-test[b]	
YM020291	T/G	T/G	T/C	≤0.5	0.16	S
YLO417	T/T	T/T	T/T	≥32	>32	R
YLO418	T/T	T/T	T/T	≥32	>32	R
YLO415	G/G	G/G	C/C	≤0.5	0.008	S
YLO454	T/G	T/T	T/C	≤2	0.094	S
YLO456	T/G	G/G	T/C	≤1	0.023	S
YLO466	T/T	T/G	T/C	≥64	>32	R
YLO467	T/T	T/G	T/C	≥64	>32	R
YLO468	G/G	T/G	T/C	≤0.5	0.023	S
YLO470	T/G	T/G	T/T	≤2	0.023	S
YLO472	T/G	T/G	C/C	≤0.5	0.023	S

[a] The position of this SNP is at 201 nt downstream of the stop codon of the *FCY2* gene.

[b] MIC values of broth microdilution are from at least three measurements including all the experiment data, while the values of *E*-test are a one-time experiment result.

[c] The interpretation of the susceptibility is based on the guidelines provided by NCCLS M27. S, susceptible to 5FC; R, resistant to 5FC.

Table 7. Phenotypes and genotypes of the 13 previously selected clinical isolates

Group[a]	Name	DST[b]	SNP site in *FCY2*		MICs of 5FC (μg/ml) by broth microdilution[c]		
			145	201	P	D1	D2
A	YM020291	155	T/G	G/G	0.5	>64	>64
	YM020715	160	G/G	A/G	0.25	64	64
	YM060088	188	T/G	G/G	<0.125	64	16
	YM060800	200	T/G	G/G	1	>64	64
	YM020112	-	T/G	G/G	0.25	64	64
	YM020347	-	T/G	G/G	<0.125	32	64
	YM020743	-	T/G	G/G	0.5	>64	>64
B	YM060369	139	G/G	G/G	0.25	>64	64
	YM060210	184	G/G	G/G	2	>64	ND
	YM020438	-	G/G	G/G	0.5	8	64
	YM060616	-	G/G	G/G	0.25	8	>64
C	YM060173	140	G/G	G/G	<0.125	0.5	0.5
	YM020055	-	G/G	G/G	0.25	2	4

[a] While group A includes strains carrying a nonsense mutation in *FCY2*, group B and C include strains exhibiting no noticeable nucleotide change that is involved in the 5FC resistance. Furthermore, group B includes strains whose derivatives are resistant or intermediate susceptible to 5FC, whereas group C are those generating 5FC susceptible progeny.

[b] The DST is based on the studies of Chou *et al.* (2007) and Li *et al.* (2009). –, not done.

[c] The MIC values are from Chia-Chen Wu (2009). P, parental strain; D1, derivative-1; D2, derivative-2; ND, no derivative.

Table 8. Phenotypes and genotypes of the remaining 22 clinical isolates (out of 35 pairs)

Group[a]	Name	DST[b]	SNP site in *FCY2*		MICs of 5FC (μg/ml) by broth microdilution[c]		
			145	201	P	D1	D2
D	YM060481	27	T/G	G/G	<0.125	64	64
	YM020311	90	T/G	G/G	0.5	64	64
	YM020693	90	T/G	G/G	<0.125	32	>64
	YM060299	134	T/G	G/G	0.25	>64	64
	YM060507	134	T/G	G/G	0.5	64	>64
	YM060508	134	T/G	G/G	1	>64	>64
	YM060512	134	T/G	G/G	<0.125	16	ND
	YM020274	153	T/G	G/G	0.25	64	64
	YM060330	186	T/G	G/G	<0.125	>64	>64
	YM060371	187	T/G	G/G	0.125	64	>64
	YM060146	188	T/G	G/G	<0.125	64	64
	YM060379	200	T/G	G/G	0.25	>64	64
	YM060325	201	T/G	G/G	1	64	64
	YM060565	202	T/G	G/G	0.5	>64	>64
	YM060237	-	T/G	G/G	0.25	64	32
	YM060300	-	T/G	G/G	0.5	>64	>64
	YM061047	-	T/G	G/G	<0.125	>64	>64
E	YM060097	149	G/G	G/G	<0.125	64	32
	YM060051	195	G/G	G/G	<0.125	1	<0.125
	YM020671	-	G/G	G/G	<0.125	>64	>64
	YM060071	-	G/G	G/G	<0.125	1	0.5
	YM060075	-	G/G	G/G	<0.125	32	ND

[a] While group D includes strains carrying a nonsense mutation in *FCY2*, group E includes strains exhibiting no noticeable nucleotide change that is involved in the 5FC resistance.

[b] The DST is based on the studies of Chou *et al.* (2007) and Li *et al.* (2009). –, not done.

[c] The MIC values are from Chia-Chen Wu (2009). P, parental strain; D1, derivative-1; D2, derivative-2; ND, no derivative.

Table 9. Genotype of the clinical isolates

Group[a]	Name	DST[b]	SNP in *FCY2*	
			145	201
F	YM060607	-	T/G	G/G
G	YM060547	98	G/G	G/G
	YM060647	98	G/G	G/G
	YM060509	140	G/G	G/G
	YM060828	140	G/G	G/G
	YM020649	156	G/G	G/G
	YM060559	183	G/G	G/G
	YM060302	185	G/G	G/G
	YM020367	-	G/G	G/G

[a] Group F is the 5FC resistant clinical isolate; group G includes the super-susceptible isolates.

[b] The DST is based on the studies of Chou *et al.* (2007) and Li *et al.* (2009). -, not done.

Table 10. Classification and association of *FCY2* genotype, MLST and PFGE of the clinical isolates

Group[a]	DST[b]	No.[c]	*ICL1*	*MDR1*	*SAPT2*	*SAPT4*	*XYR1*	*ZWF1a*	MLST clonal cluster[b]	PFGE type[b]	Clinical isolate[d]
1 (A/G)	160	1	5	1	1	34	1	21	-	V-1	YM020715
2 (T/G)	134	4	3	7	1	6	52	4	1-5	-	YM060299, YM060507 YM060508, YM060512
	90	2	3	7	1	6	22	4	1-3	III-4	YM020311, YM020693
	27	1	3	7	1	6	6	4	-	-	YM060481
	153	1	3	7	1	6	60	4	1-2	III-1	YM020274
	200	2	3	65	1	6	52	4	-	-	**YM060800**, YM060379
	155	1	3	7	3	6	22	4	1-4	III-5	**YM020291**
	201	1	3	7	4	6	75	10	-	-	YM060325
	202	1	10	7	4	6	22	4	-	-	YM060565
	188	2	1	33	2	30	69	25	-	-	YM060088, YM060146
	186	1	1	63	2	30	69	2	-	-	YM060330
	187	1	1	63	2	30	69	25	-	-	YM060371
	-	6							-	-	YM020112, YM020347 YM020743, YM060237 YM060300, YM061047
3 (G/G)	184	1	1	4	12	23	36	9	-	-	YM060210
	139	1	1	4	22	23	36	9	3-2	-	YM060369
	149	1	1	44	3	7	58	3	-	-	YM060097
	-	4							-	-	YM020438, YM060616 YM020671, YM060075
4 (G/G)	140	9	1	3	3	17	54	3	2-2	-	YM060173, YM060509 YM060828
	98	5	1	3	3	17	9	3	2-4	-	YM060547, YM060647
	183	1	1	20	12	17	68	3	-	-	YM060559
	185	1	1	62	2	23	5	1	-	-	YM060302
	195	1	1	9	1	10	6	7	-	-	YM060051
	149	1	1	44	3	7	58	3			
	others	26							-	-	
	-	21							-	-	YM020367, YM020055 YM060071
5 (G/G)	-	1							-	-	YM060607

[a] Based on MIC values, group1-3 include the clinical isolates that can generate 5FC resistant progeny, group 4 includes the super-susceptible clinical isolates, and group 5 is the 5FC resistant clinical isolate.

[b] The DST, MLST and PFGE are based on the studies of Chou *et al.* (2007) and Li *et al.* (2009). –, not done.

[c] No., the numbers of the clinical isolates belong to the corresponding DST.

[d] The clinical isolates represent the corresponding DST and their genotypes (at position 145 and 201) has been identified by sequencing. Furthermore, for strains that are underlined their entire *FCY2* coding region and the flanking region have been sequenced.

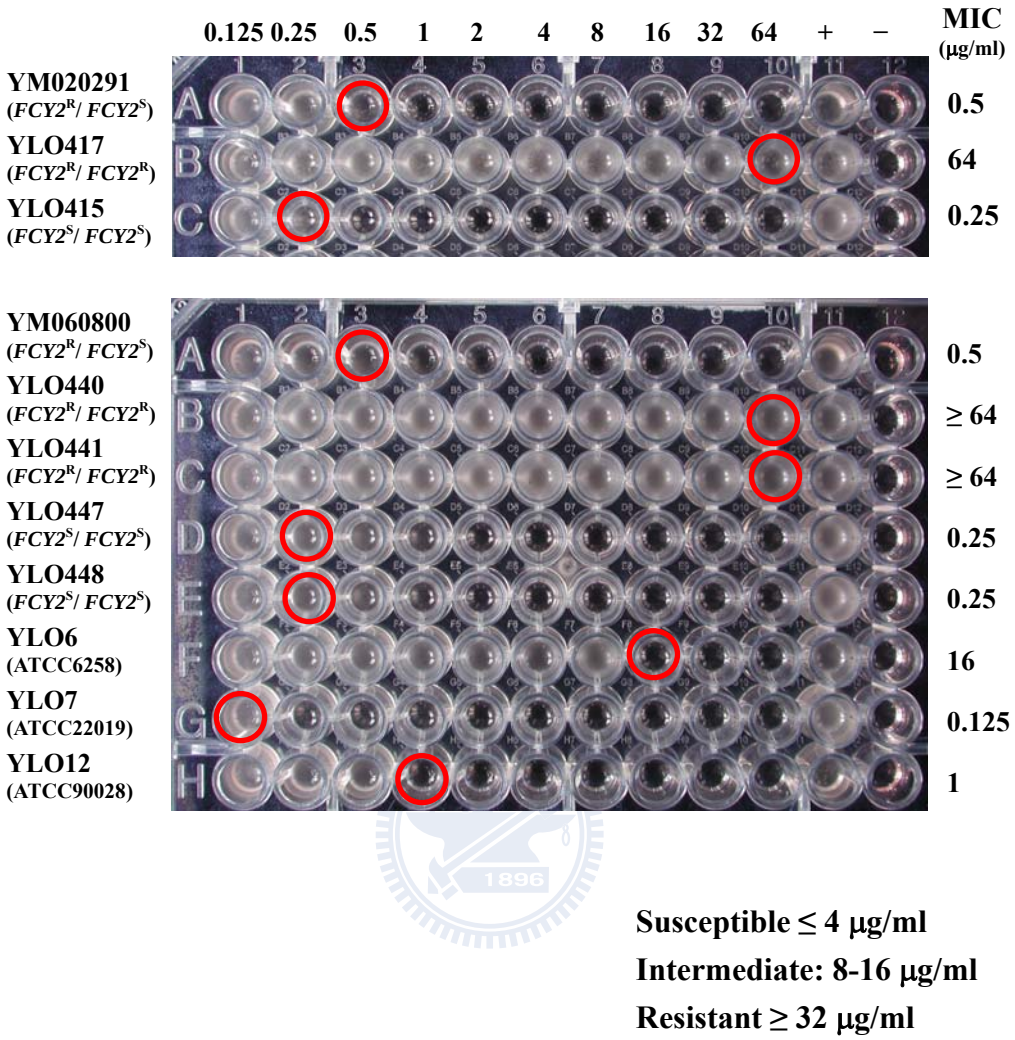


Figure 1. 5FC susceptibility testing of two clinical isolates and their homozygous mutants by both microdilution. The susceptibility of *C. tropicalis* to 5FC was determined by broth microdilution; the strains listed on the left are as follows: one clinical isolate YM020291, and its isogenic mutants YLO417 and YLO415; another clinical isolate YM060800, and its isogenic mutants YLO440, YLO441, YLO447 and YLO448. YLO417, YLO440 and YLO441 are *FCY2*^{R}/ *FCY2*^{R} homozygous mutants, while YLO415, YLO 447 and YLO448 are *FCY2*^{S}/ *FCY2*^{S} homozygous mutants. Drug concentrations are increased by two times from 0.125 to 64 μg/ml and note on the top of each column. "+" indicates the control for cellular growth; "-" is the sterility control. MICs are showed on the right. YLO12 (*Candida albicans* ATCC90028, 0.5~2.0 μg/ml) was used as a reference strain, and YLO6 (*Candida krusei* ATCC6258, 4.0~16.0 μg/ml) and YLO7 (*Candida parapsilosis* ATCC22019, 0.12~0.5 μg/ml) were incorporated as quality control strains.

A

Strain (Allele)	YM020291 ($FCY2^R$)	YM020291 ($FCY2^S$)	YM020367	YM020649	YM060302	YM060559
SNP at position -224 (-11) in *FCY2*	C	T	C/C	C/C	C/C	T/T

B

(-65) **-224 (-11)** (15)

```
FCY2R     TATCATTCACATTAGATATAGATAGATCCACCAAACTTATACATACACATATACACATAACACAATGGCTGATTACGAA
FCY2S     TATCATTCACATTAGATATAGATAGATCCACCAAACTTATACATACACATATATACATAACACAATGGCTGATTACGAA
YM020367  TATCATTCACATTAGATATAGATAGATCCACCAAACTTATACATACACATATACACATAACACAATGGCTGATTACGAA
YM020649  TATCATTCACATTAGATATAGATAGATCCACCAAACTTATACATACACATATACACATAACACAATGGCTGATTACGAA
YM060302  TATCATTCACATTAGATATAGATAGATCCACCAAACTTATACATACACATATACACATAACACAATGGCTGATTACGAA
YM060559  TATCATTCACATTAGATATAGATAGATCCACCAAACTTATACATACACATATATACATAACACAATGGCTGATTACGAA
```

C

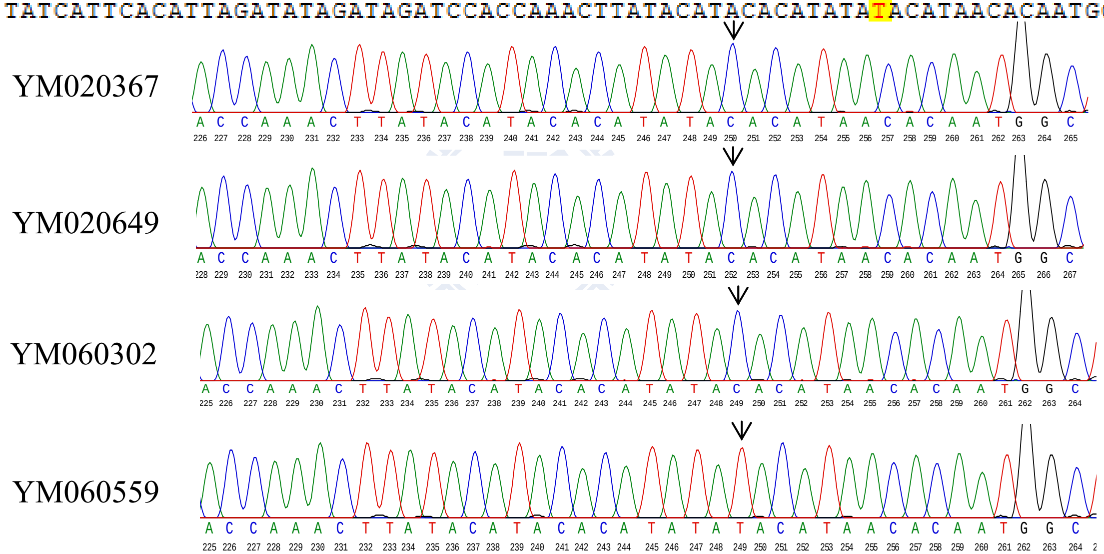
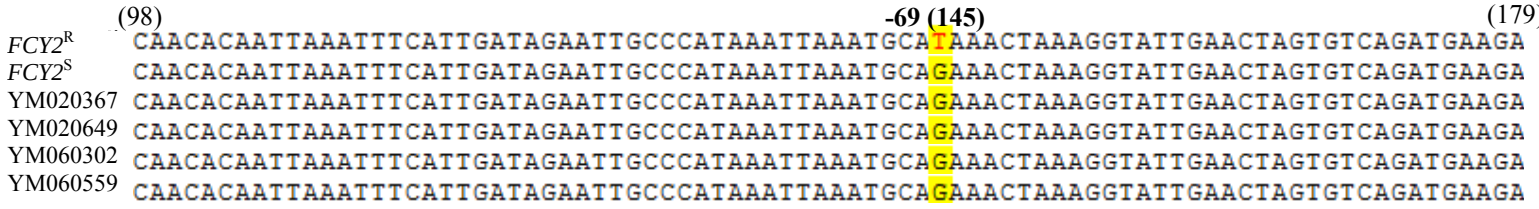


Figure 2. The SNPs at position -224 (-11) in *FCY2* promoter region among super-susceptible strains. (A) The SNPs at position -224 (-11) in *FCY2* promoter region of four super-susceptible strains YM020367, YM020649, YM060302 and YM060559 were compared to the $FCY2^R$ and $FCY2^S$ in YM020291. (B) It shows nucleotide sequence alignments of the promoter region of *FCY2* gene among the super-susceptible strains with the $FCY2^R$ and $FCY2^S$ in YM020291. The nucleotides boxed in yellow represent the SNPs at position -224 (-11). (C) DNA sequence profiles of the super-susceptible strains correspond to the region in (B). The SNPs at position -224 (-11) are indicated by arrows. $FCY2^R$ and $FCY2^S$ allele in YM020291.

A

Strain (Allele)	YM020291 ($FCY2^R$)	YM020291 ($FCY2^S$)	YM020367	YM020649	YM060302	YM060559
SNP at position -69 (145) in *FCY2*	T	G	G/G	G/G	G/G	G/G

B



C

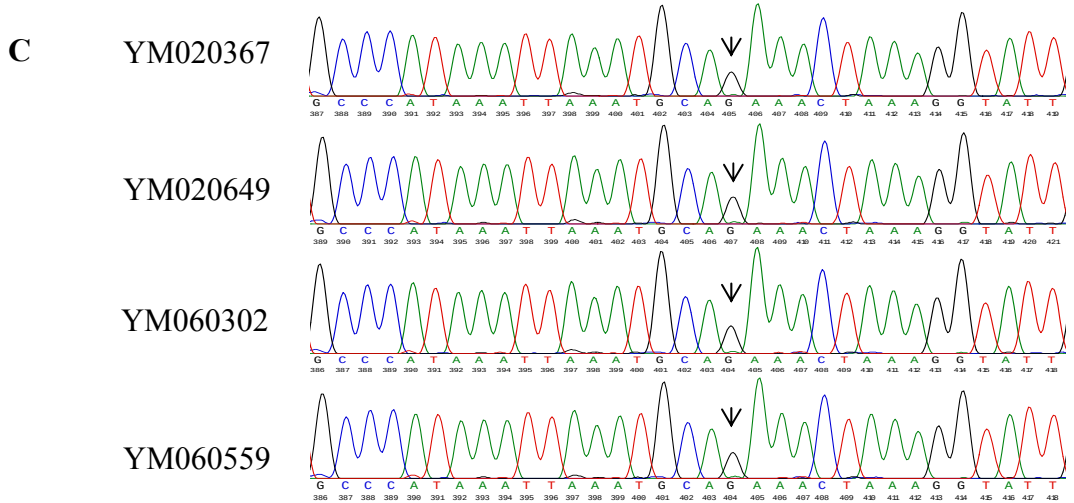
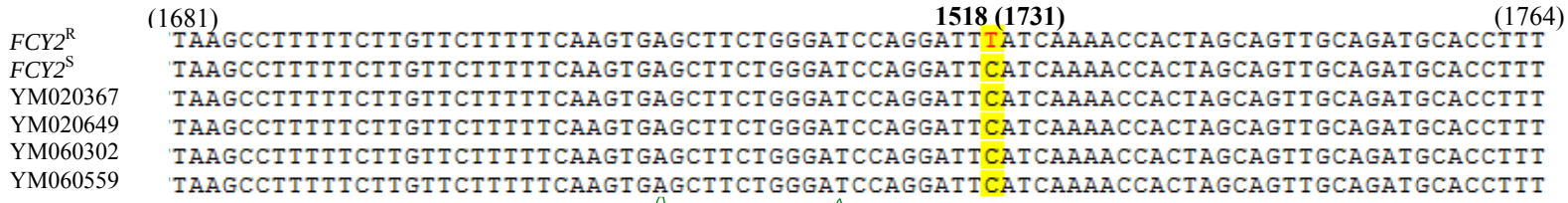


Figure 3. The SNPs at position -69 (145) in *FCY2* promoter region among super-susceptible strains. (A) The SNPs at position -69 (145) in *FCY2* promoter region of four super-susceptible strains YM020367, YM020649, YM060302 and YM060559 were compared to the $FCY2^R$ and $FCY2^S$ in YM020291. (B) It shows nucleotide sequence alignments of the promoter region of *FCY2* gene among the super-susceptible strains with the $FCY2^R$ and $FCY2^S$ in YM020291. The nucleotides boxed in yellow represent the SNPs at position -69 (145). (C) DNA sequence profiles of the super-susceptible strains correspond to the region in (B). The SNPs at position -69 (145) are indicated by arrows. $FCY2^R$ and $FCY2^S$ allele in YM020291.

A

Strain (Allele)	YM020291 ($FCY2^R$)	YM020291 ($FCY2^S$)	YM020367	YM020649	YM060302	YM060559
SNP at position 1518 (1731) in *FCY2*	T	C	C/C	C/C	C/C	C/C

B



C

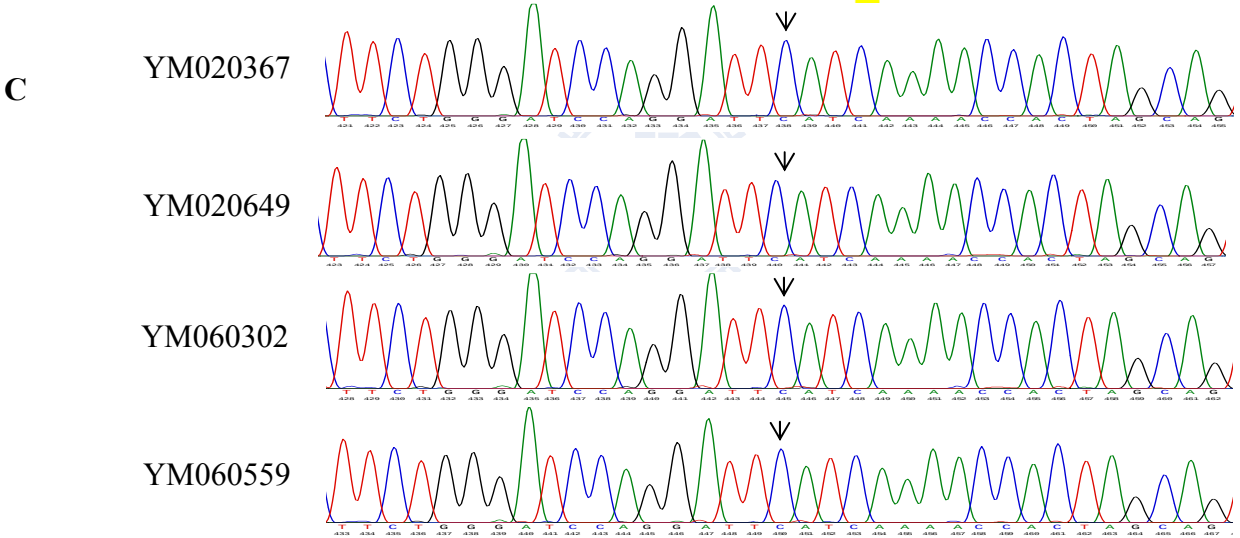


FIG. 4. The SNP in the 3' UTR of *FCY2* gene among hyper-susceptible clinical isolates. (A) The SNPs at position 1518 (1731) in the *FCY2* 3'UTR of four hyper-susceptible strains YM020367, YM020649, YM060302 and YM060559 were compared to the $FCY2^R$ allele and $FCY2^S$ allele in YM020291. (B) It shows nucleotide sequence alignments of the 3'UTR of *FCY2* gene among the hyper-susceptible strains with $FCY2^R$ and $FCY2^S$ in YM020291. The nucleotides boxed in yellow represent the SNPs at position 1518 (1731). (C) DNA sequence profiles of the hyper-susceptible strains correspond to the region in (B). The polymorphic nucleotides at position 1518 (1731) are indicated by the arrow. $FCY2^R$ and $FCY2^S$ allele in YM020291.

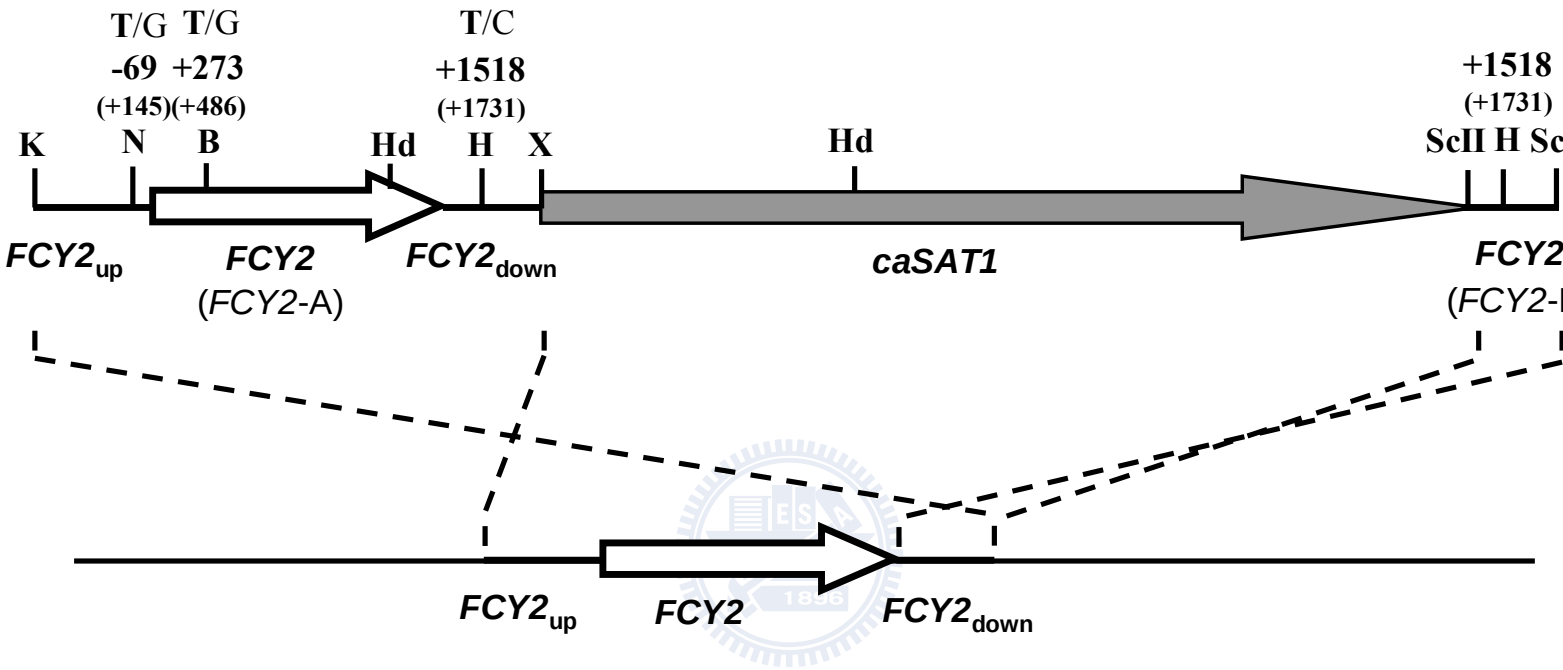


FIG. 5. The schematic representation of *FCY2* knockin with *SAT1* cassette. It depicts the structure of the recombinant DNA construct (top) used for the integration of different *FCY2* single mutant alleles into the chromosome of the clinical isolate YM020291 (bottom), in which one of the alleles was randomly replaced. The polymorphic nucleotides and their locations are given on the top of their corresponding enzyme sites which were used to select for the correct constructs. The *FCY2* coding region is represented by the white arrow and the upstream ($FCY2_{up}$) and downstream regions ($FCY2_{down}$) by the solid lines. The *FCY2* ORF, its upstream and downstream regions altogether is designated as *FCY2*-A, while $FCY2_{down}$ downstream of the *SAT1* cassette is *FCY2*-B. The *caSAT1* selection marker is indicated by the grey arrow. Only relevant restriction sites are shown: K, *Kpn*I; N, *Nsi*I; B, *Bbs*I; Hd, *Hind*III; H, *Hin*fI, *Xho*I; ScI, *Sac*II; ScI, *Sac*I.

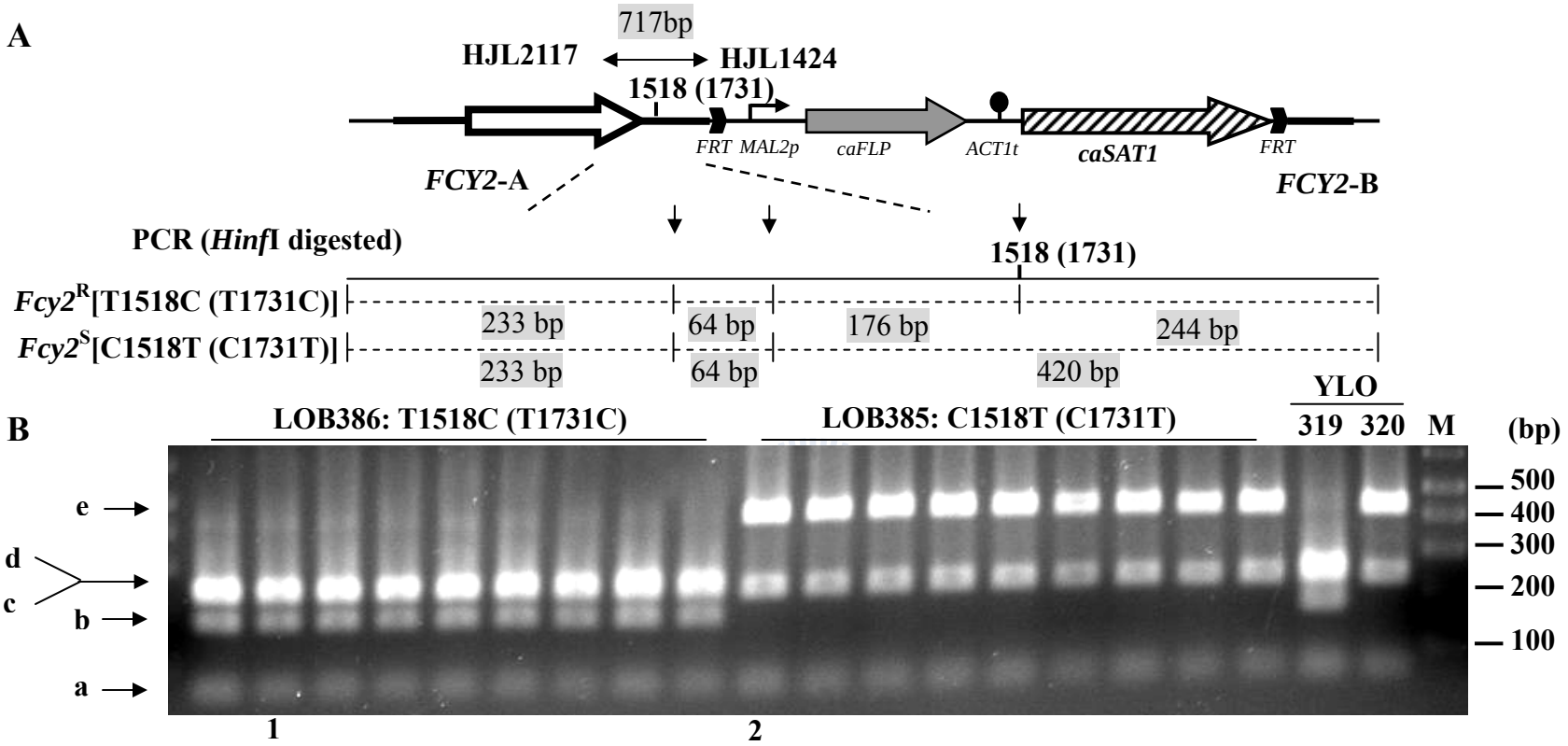


FIG. 6. Verification of plasmids carrying *FCY2*^{S}[C1518T (C1731T)] or *FCY2*^{R}[T1518C (T1731C)] allele. (A) Schematic representation of the relevant regions of the *SAT1* flipper, *FCY2*-A, *FCY2*-B and the point mutation site in plasmids LOB385 and LOB386. The generations of the mutations were confirmed by *Hinf*I-digested PCR products of plasmids. The sizes of the PCR products and the digested DNA fragments are boxed in grey. Arrows represent the potential enzyme recognition sites of *Hinf*I (5'-GANTC-3') on the PCR products. The dotted lines separately stand for the two distinct PCR products. (B) Evaluation of plasmids carrying *Fcy2*^{R}[T1518C (T1731C)] or *Fcy2*^{S}[C1518T (C1731T)] allele. The digested PCR products were analyzed by 1.5% agarose gel electrophoresis. The sizes of each digested fragment are corresponded to (A). For control purpose, YLO319 and YLO320 were used in parallel. The sizes of marker (M) are given on the right side of the photo. Lane 1 and 2 (below the photo) separately correspond to LOB386 and LOB385. a to e indicate the sizes of each band. a: < 100 bp; b: close to 200 bp; c and d: refer to 200~300 bp; e: refer to 400~500 bp.

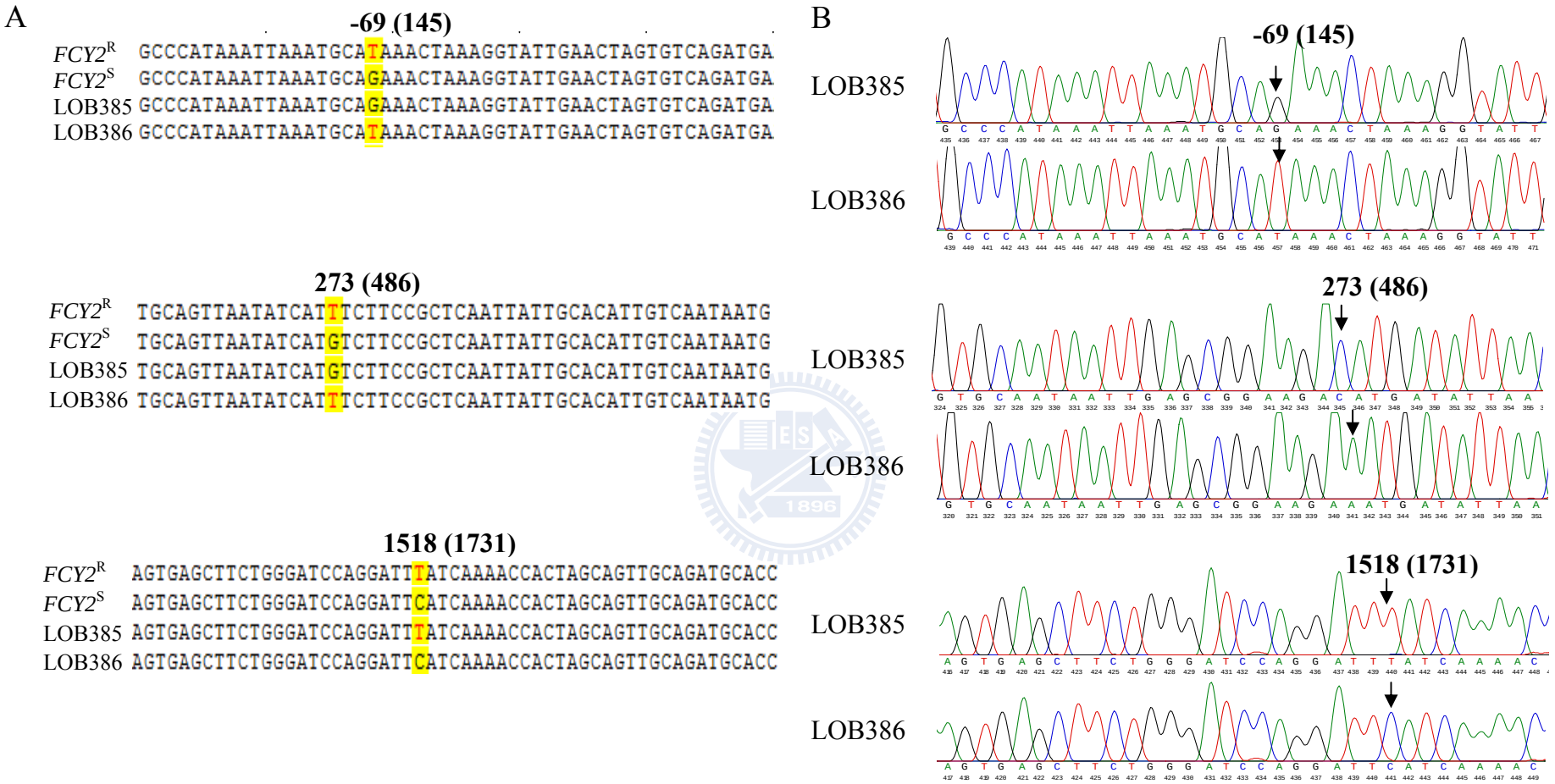


FIG. 7. Nucleotide sequences of *FCY2* gene in LOB385 and LOB386. (A) It shows nucleotide sequence alignments of the entire *FCY2* gene in the plasmids LOB385 and LOB386 with the *FCY2*^{R} allele and *FCY2*^{S} allele in YM020291. The nucleotide boxed in yellow represents the SNPs at position -69 (145), 273 (486) or 1518 (1731). (B) DNA sequence profiles of the plasmids LOB385 and LOB386 correspond to the region (A). The polymorphic nucleotides at position -69 (145), 273 (486) and 1518 (1731) are indicated by the arrows. *FCY2*^{R} and *FCY2*^{S} allele in YM020291.

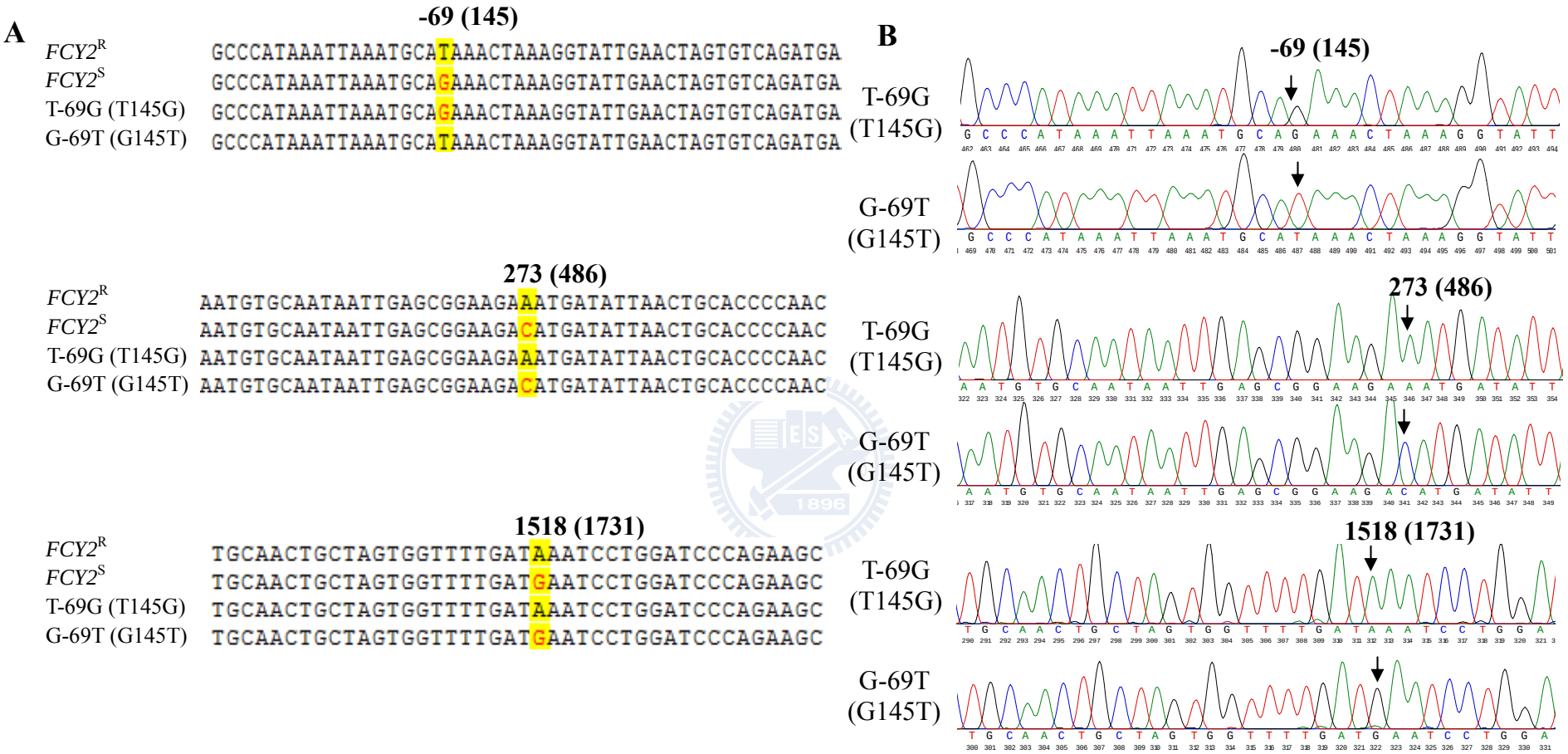


FIG. 8. Nucleotide sequences of plasmids carrying either *FCY2*^{S}[G-69T (G145T)] or *FCY2*^{R} [T-69G (T145G)] allele. (A) It shows nucleotide sequence alignments of the *FCY2*-A in pGEMT easy cloning vector with the *FCY2*^{R} allele and *FCY2*^{S} allele in YM020291. The nucleotide boxed in yellow represents the SNPs at position -69 (145), 273 (486) or 1518 (1731). (B) DNA sequence profiles of the plasmids correspond to the region in (A). The polymorphic nucleotides at position -69 (145), 273 (486) and 1518 (1731) are indicated by the arrows. FCY2-S, *FCY2*^{S} allele in YM020291; FCY2-R, *FCY2*^{R} allele in YM020291; T145G, *FCY2*^{R} [T-69G (T145G)] allele; G145T, *FCY2*^{S}[G-69T (G145T)]

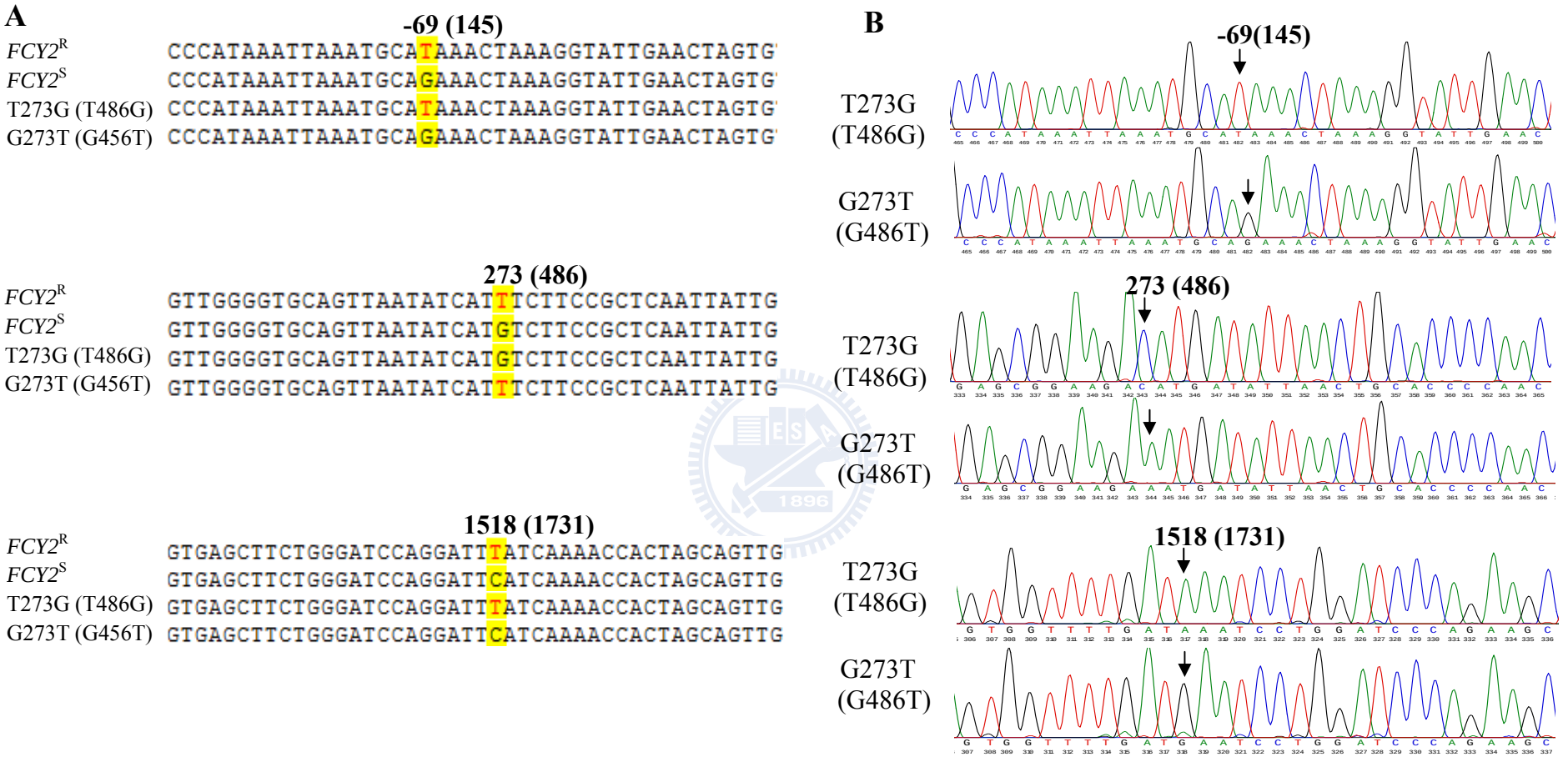


FIG. 9. Nucleotide sequences of plasmids carrying either *FCY2*^{S}[G273T (G486T)] or *FCY2*^{R} [T273G (T486G)] allele. (A) It shows nucleotide sequence alignments of the *FCY2*-A in the pGEMT easy cloning vector with the *FCY2*^{R} allele and *FCY2*^{S} allele in YM020291. The nucleotide boxed in yellow represents the SNPs at position -69 (145), 273 (486) or 1518 (1731). (B) DNA sequence profiles of the plasmids correspond to the region in (A). The polymorphic nucleotides at position -69 (145), 273 (486) and 1518 (1731) are indicated by the arrows. FCY2-S, *FCY2*^{S} allele in YM020291; FCY2-R, *FCY2*^{R} allele in YM020291; T486G, *FCY2*^{R} [T273G (T486G)] allele; G486T, *FCY2*^{S}[G273T (G486T)]

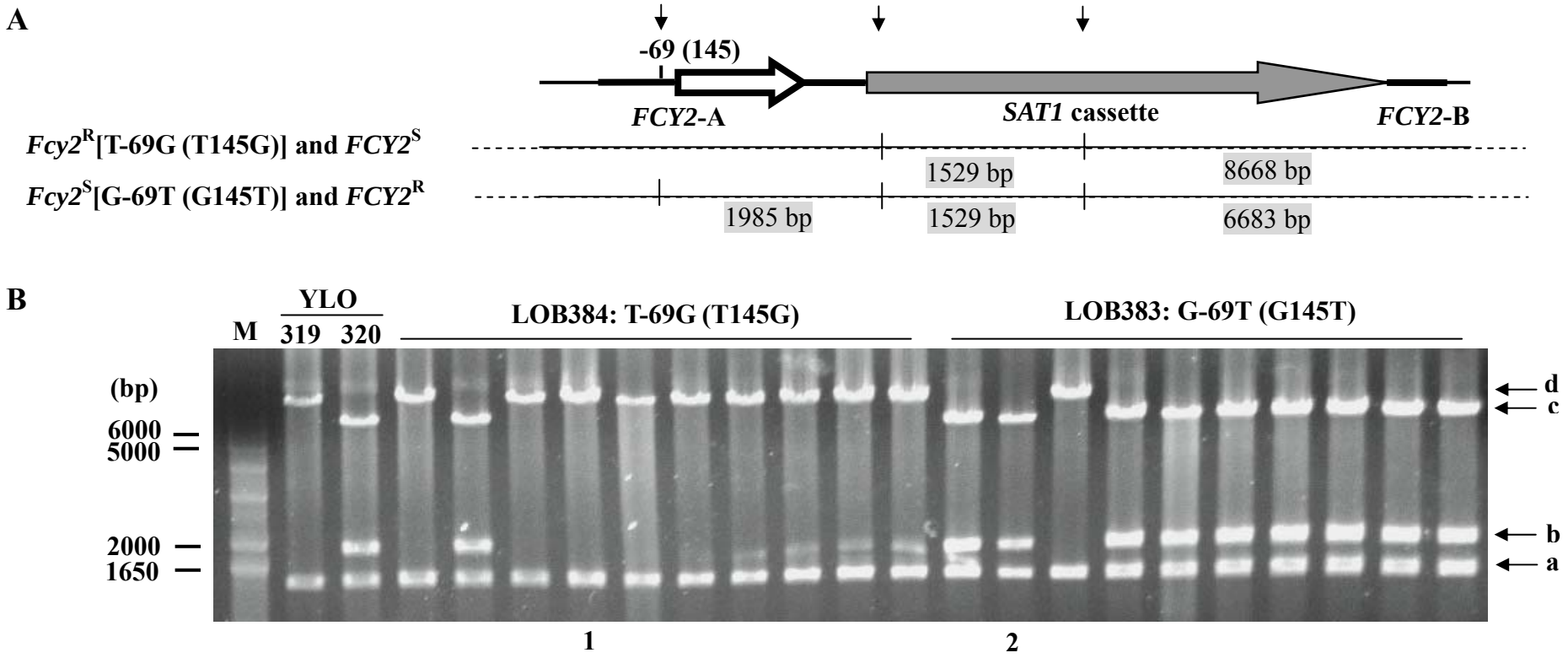


FIG. 10. Verification of plasmids carrying *FCY2*^{S}[G-69T (G145T]] or *FCY2*^{R} [T-69G (T145G)] allele. (A) Schematic representation of the relevant regions of the *SAT1* flipper, *FCY2*-A, *FCY2*-B and the point mutation site in plasmids LOB383 and LOB384. The generations of the mutations were confirmed by *Nsi*I-digested plasmids. The sizes of the digested DNA fragments are boxed in grey. Arrows represent the potential enzyme recognition sites of *Nsi*I (5'-ATGCAT-3') on the plasmids. (B) Evaluation of plasmids carrying *Fcy2*^{R}[T-69G (T145G)] or *Fcy2*^{S}[G-69T (G145T)] allele. The digested products were analyzed by 1% agarose gel electrophoresis. The sizes of each digested fragment are corresponded to (A). For control purpose, YLO319 and YLO320 were used in parallel. The sizes of marker (M) are given on the left side of the photo. Lane 1 and 2 (below the photo) separately corresponded to LOB384 and LOB383. a to d indicate the sizes of each band. a: close to 1650 bp; b: close to 2000 bp; c: close to 6000 bp; d: over 6000 bp.

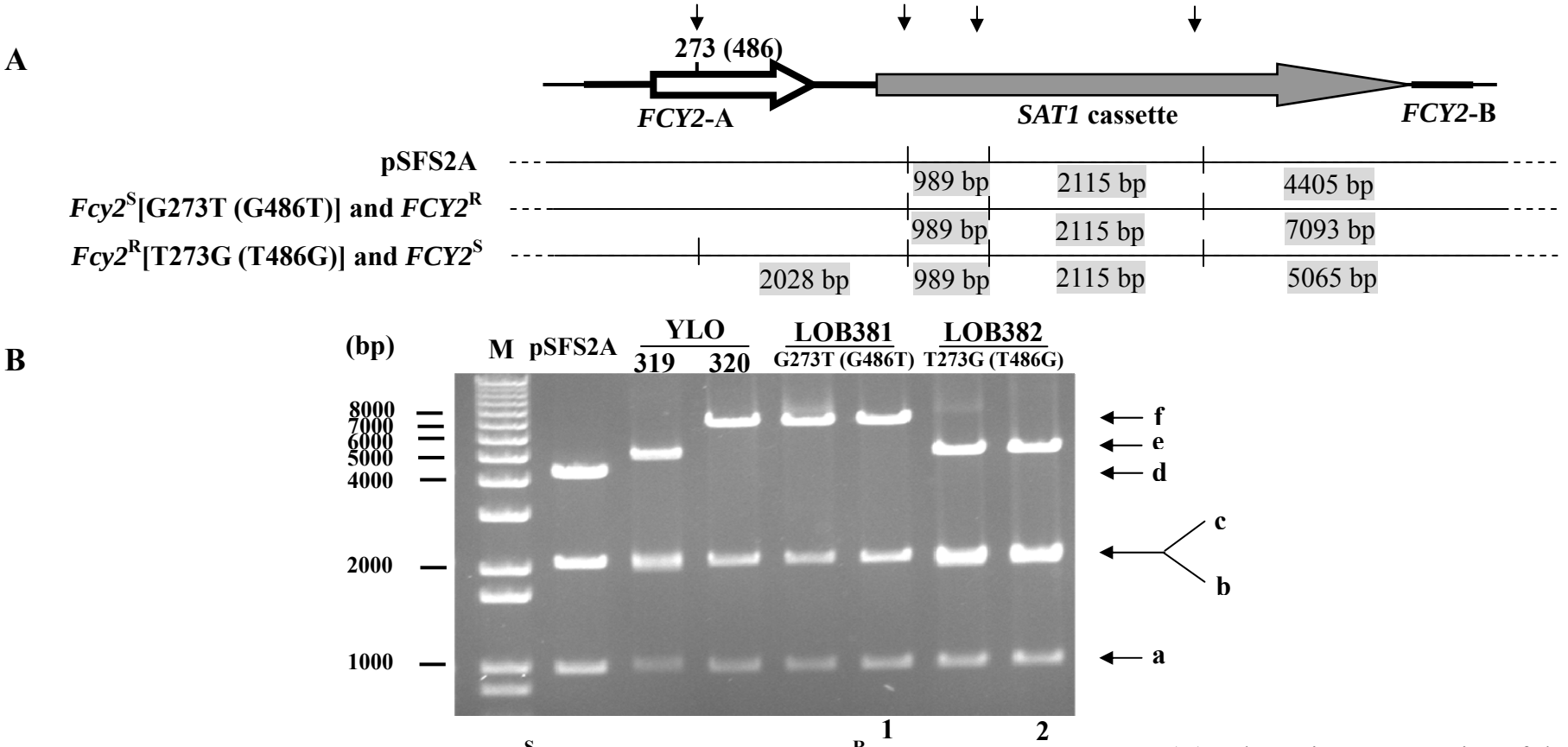


FIG. 11. Verification of plasmids carrying *FCY2*^{S}[G273T (G486T)] or *FCY2*^{R} [T273G (T486G)] allele. (A) Schematic representation of the relevant regions of the *SAT1* flipper, *FCY2*-A, *FCY2*-B and the point mutation site in plasmids LOB381 and LOB382. The generations of the mutations were confirmed by *Bbs*I-digested plasmids. The sizes of the digested DNA fragments are boxed in grey. Arrows represent the potential enzyme recognition sites of *Bbs*I (5'-GAAGAC -3') on the plasmids. (B) Evaluation of plasmids carrying *Fcy2*^{R}[T273G (T486G)] or *Fcy2*^{S}[G273T (G486T)] allele. The digested products were analyzed by 1% agarose gel electrophoresis. The sizes of each digested fragment are corresponded to (A). For control purpose, pSFS2A, YLO319 and YLO320 were used in parallel. The sizes of marker (M) are given on the left side of the photo. Lane 1 and 2 (below the photo) separately correspond to LOB381 and LOB382. a to f indicate the sizes of each band. a: close to 1000 bp; b and c: close to 2000 bp; d: refer to 4000~5000 bp; e: close to 5000 bp; f: refer to 7000~8000 bp

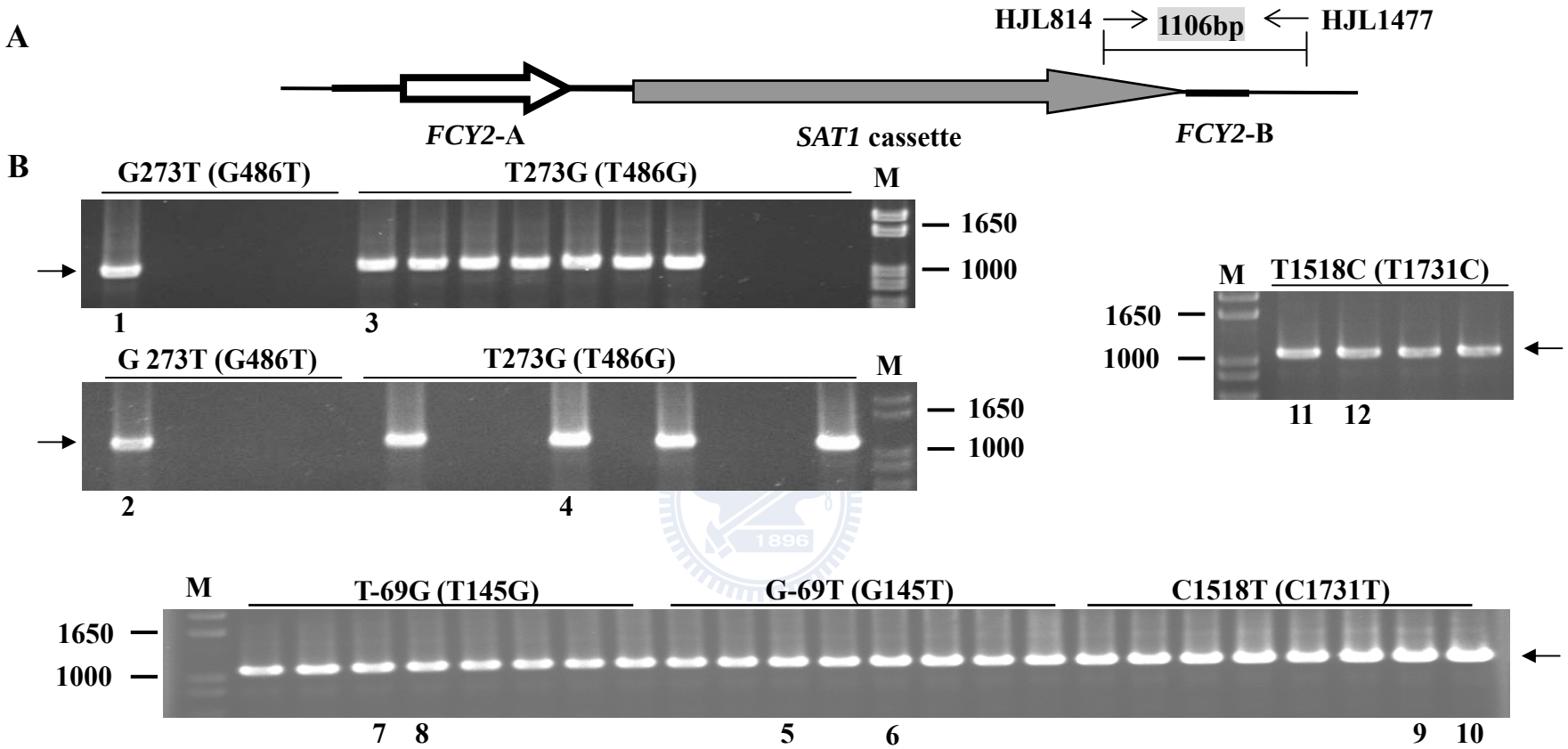


FIG. 12. Evaluation the constructs of the single mutation strains by colony PCR. (A) Schematic representation of the relevant positions of the primer pair HJL814 (located on the *SAT1* cassette) and HJL1477 (located at the downstream of the *FCY2*-B) on the chromosome of the single mutation strains. The integrations of each cassette from plasmids LOB381, LOB382, LOB383, LOB384, LOB385 and LOB386 were confirmed by PCR. The sizes of the PCR products are boxed in grey. (B) Verification of the correct integration of each cassette in single mutation strains. The PCR products were analyzed by 1% agarose gel electrophoresis. The size of PCR products is corresponded to (A), and the sizes of marker (M) are given on the right or left side of the photo. Lane 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 (below the photo) separately corresponded to mutants YLO454, YLO455, YLO456, YLO457, YLO466, YLO467, YLO468, YLO469, YLO470, YLO471, YLO472 and YLO473. Arrows indicate the size of the bands, which is close to 1000 bp.

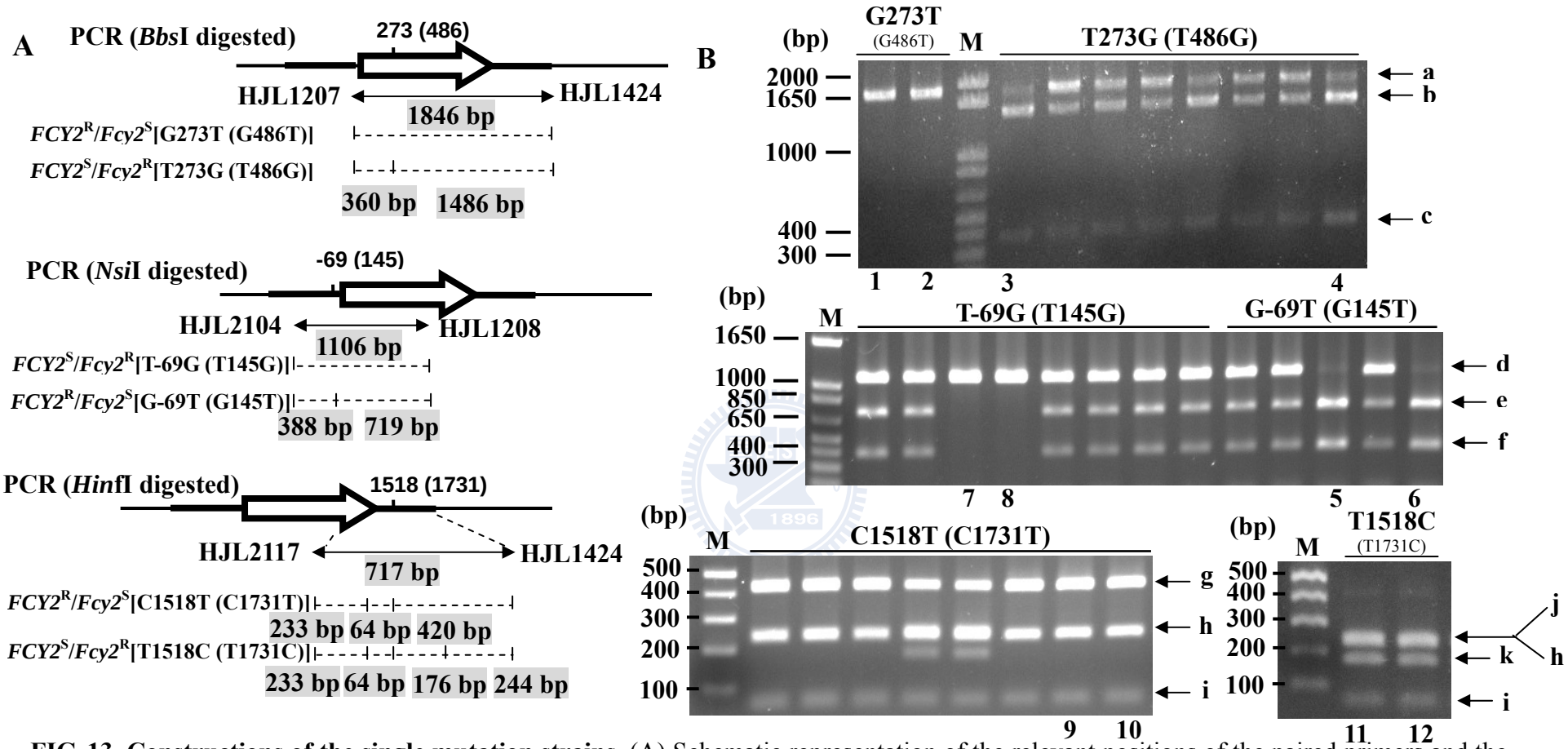
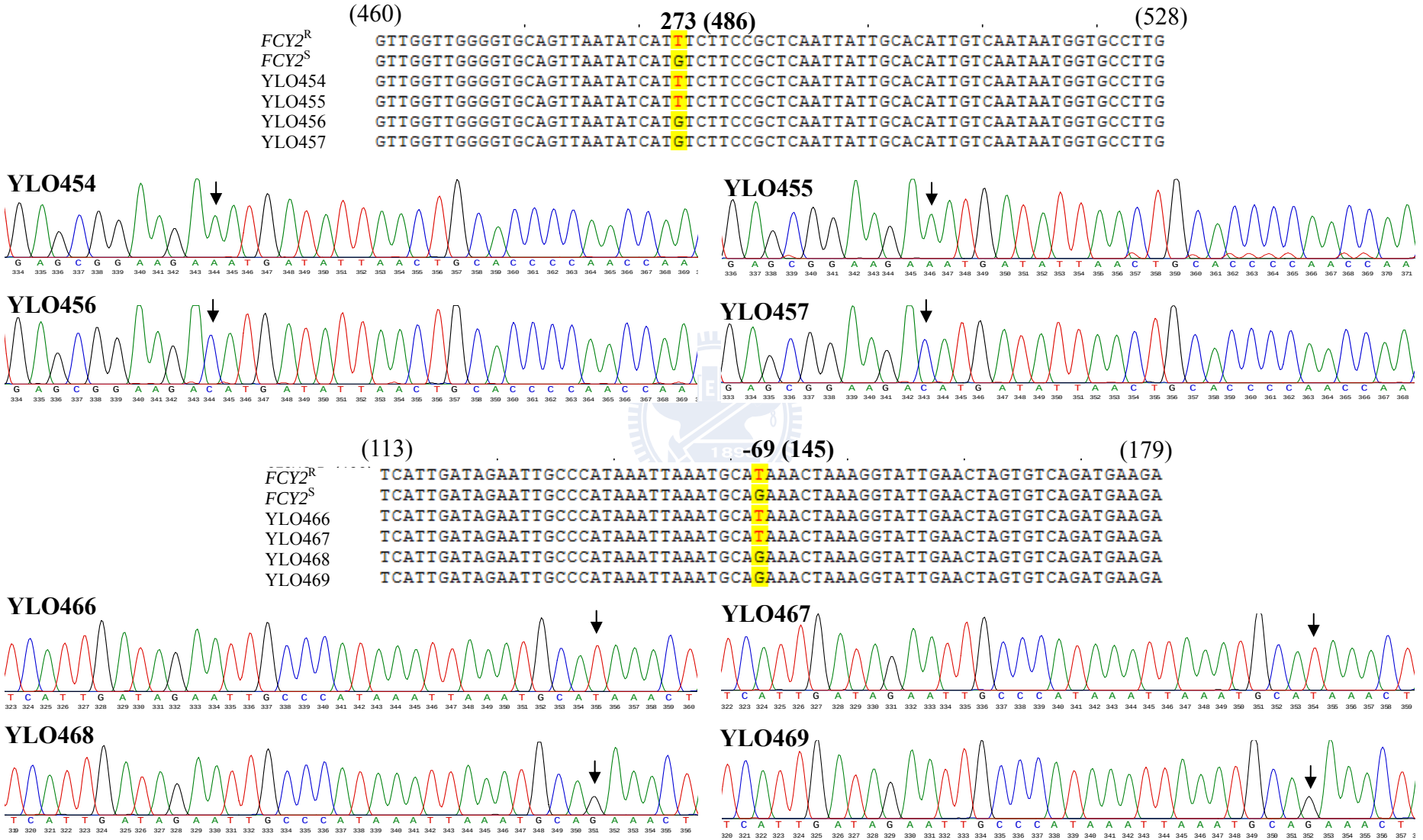


FIG. 13. Constructions of the single mutation strains. (A) Schematic representation of the relevant positions of the paired primers and the mutation sites at the *FCY2* locus. The generations of the mutations were confirmed by enzyme-digested PCR products. The sizes of PCR products and the digested DNA fragments are boxed in grey. (B) Evaluation of single mutants. The digested PCR products were analyzed by 1% or 2% agarose gel electrophoresis. The sizes of each digested fragment are corresponded to (A). The sizes of marker (M) are given on the left side of the photo. Lane 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 (below the photo) separately corresponded to mutants YLO454, YLO455, YLO456, YLO457, YLO466, YLO467, YLO468, YLO469, YLO470, YLO471, YLO472 and YLO473. a to k indicated the sizes of each band. a ~2000 bp; b ~1650 bp; c,f,g ~400 bp; d: ~1000 bp; e: refer to 650~850 bp; j,h: refer to 200~300 bp; i: <100 bp; k ~200 bp


(460)
273 (486)
(528)
FCY2R GTTGGTTGGGGTGCAGTTAATATCATTTCTTCCGCTCAATTATTGCACATTGTCAATAATGGTGCCTTG
FCY2S GTTGGTTGGGGTGCAGTTAATATCATGTCTTCCGCTCAATTATTGCACATTGTCAATAATGGTGCCTTG
YLO454 GTTGGTTGGGGTGCAGTTAATATCATTTCTTCCGCTCAATTATTGCACATTGTCAATAATGGTGCCTTG
YLO455 GTTGGTTGGGGTGCAGTTAATATCATTTCTTCCGCTCAATTATTGCACATTGTCAATAATGGTGCCTTG
YLO456 GTTGGTTGGGGTGCAGTTAATATCATGTCTTCCGCTCAATTATTGCACATTGTCAATAATGGTGCCTTG
YLO457 GTTGGTTGGGGTGCAGTTAATATCATGTCTTCCGCTCAATTATTGCACATTGTCAATAATGGTGCCTTG
YLO454
YLO455
YLO456
YLO457
(113)
-69 (145)
(179)
FCY2R TCATTGATAGAATTGCCCATAAATTAAATGCATAAACTAAAGGTATTGAACTAGTGTCAGATGAAGA
FCY2S TCATTGATAGAATTGCCCATAAATTAAATGCAGAAACTAAAGGTATTGAACTAGTGTCAGATGAAGA
YLO466 TCATTGATAGAATTGCCCATAAATTAAATGCATAAACTAAAGGTATTGAACTAGTGTCAGATGAAGA
YLO467 TCATTGATAGAATTGCCCATAAATTAAATGCATAAACTAAAGGTATTGAACTAGTGTCAGATGAAGA
YLO468 TCATTGATAGAATTGCCCATAAATTAAATGCAGAAACTAAAGGTATTGAACTAGTGTCAGATGAAGA
YLO469 TCATTGATAGAATTGCCCATAAATTAAATGCAGAAACTAAAGGTATTGAACTAGTGTCAGATGAAGA
YLO466
YLO467
YLO468
YLO469

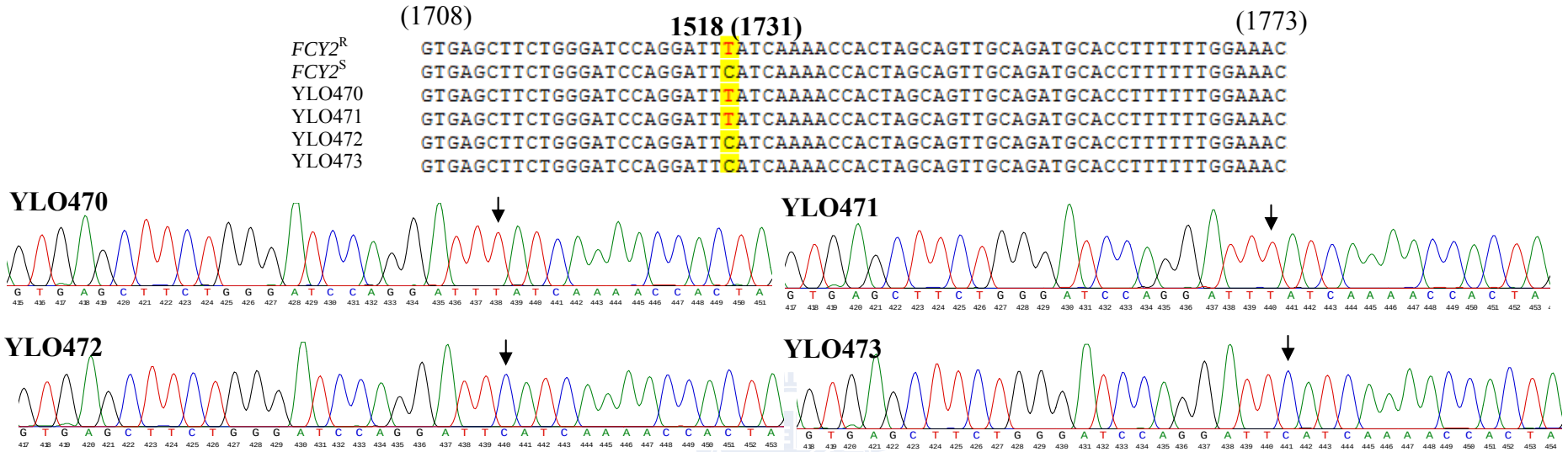


FIG. 14. Nucleotide sequences of *FCY2* gene in single mutation strains. Nucleotide sequence alignments of the DNA segments containing the SNPs at position -69 (145), 273 (486) or 1518 (1731) at the *FCY2* locus among various single mutation strains as well as the *FCY2*^{R} allele and *FCY2*^{S} allele in YM020291. The nucleotides boxed in yellow represent the SNPs at position -69 (145), 273 (486) or 1518 (1731). DNA sequencing profiles corresponding to each region are indicated bellow the aligned sequences. *FCY2*^{S} and *FCY2*^{R} allele in YM020291.

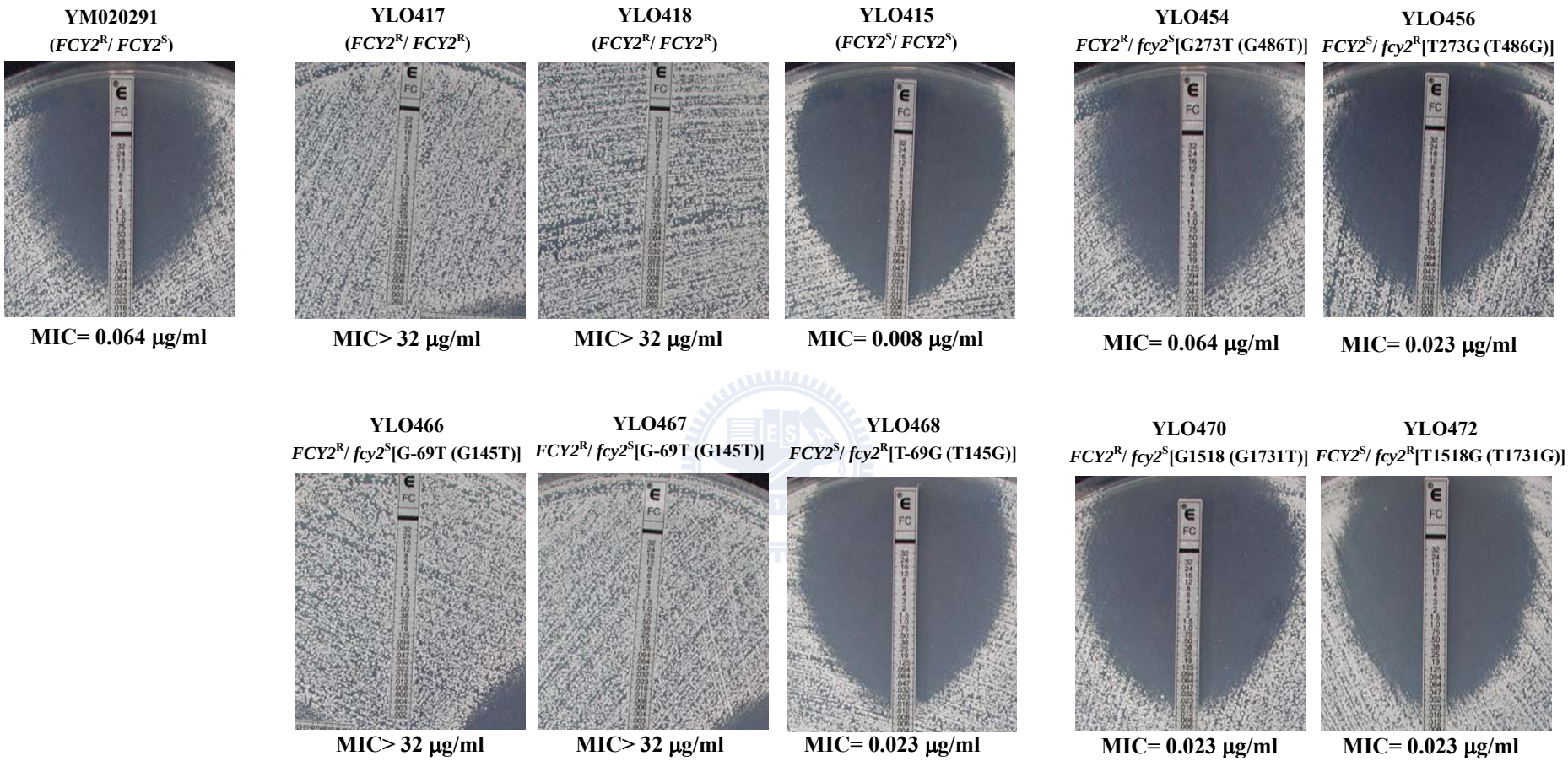


FIG. 15. E-test validation of broth microdilution data. The MICs of 5FC for *C. tropicalis* after 2 days of incubation at 35°C. E-test reading pattern for the parental strain, the homozygous mutant strains and the single mutation strains. The MICs are given bellow the photos and their values are based on the scale of 5FC concentrations imprinted on the E-test strip (in μg/ml).

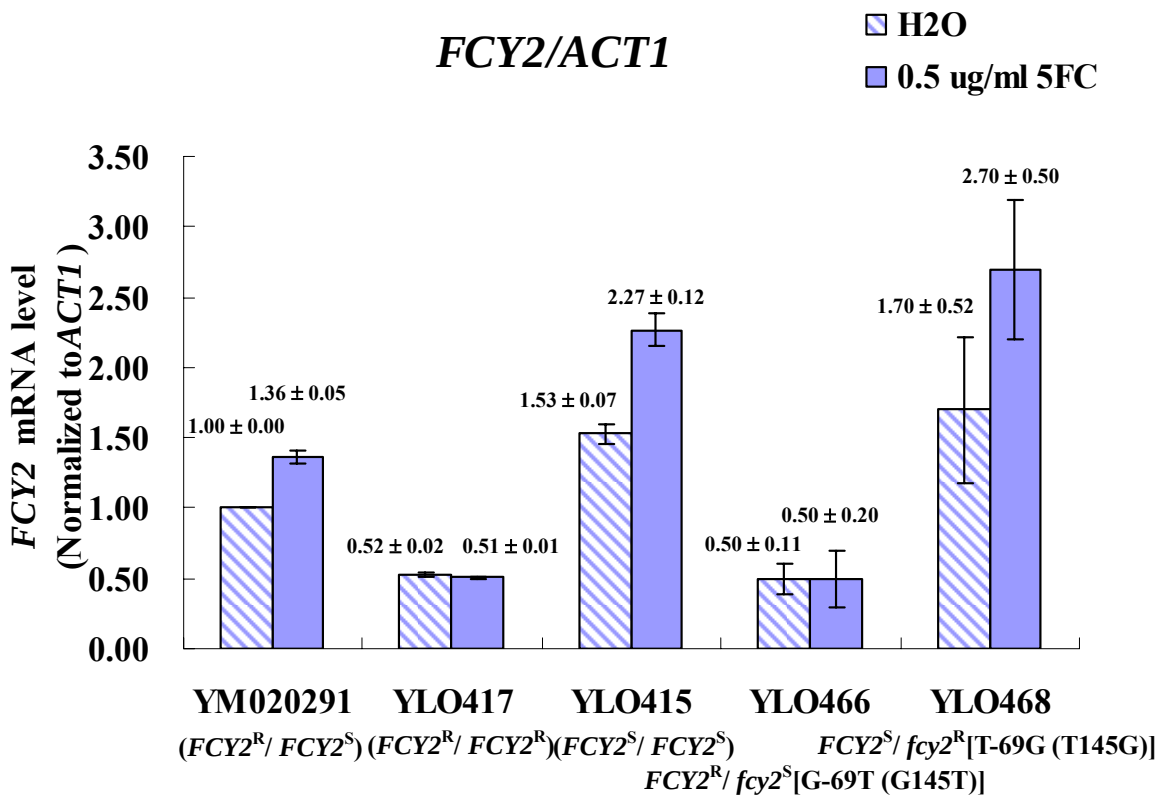


FIG. 16. *FCY2* mRNA levels among clinical isolate, its homozygous mutants and its single mutation strains by real-time quantitative PCR. Total RNA was isolated from cells with or without 5FC treatment. The reported values are the *FCY2* mRNA levels of each strain including homozygous resistant mutant YLO417, homozygous susceptible mutant YLO415, G-69T (G145T) single mutation strain YLO466 and T-69G (T145G) single mutation strain YLO468 compared to the clinical isolate YM020291, followed by the normalization using *ACT1* expression. The values are shown on the top of each bar in the histogram. The reported values are the means ± standard deviations. The results are derived from two independent experiments.

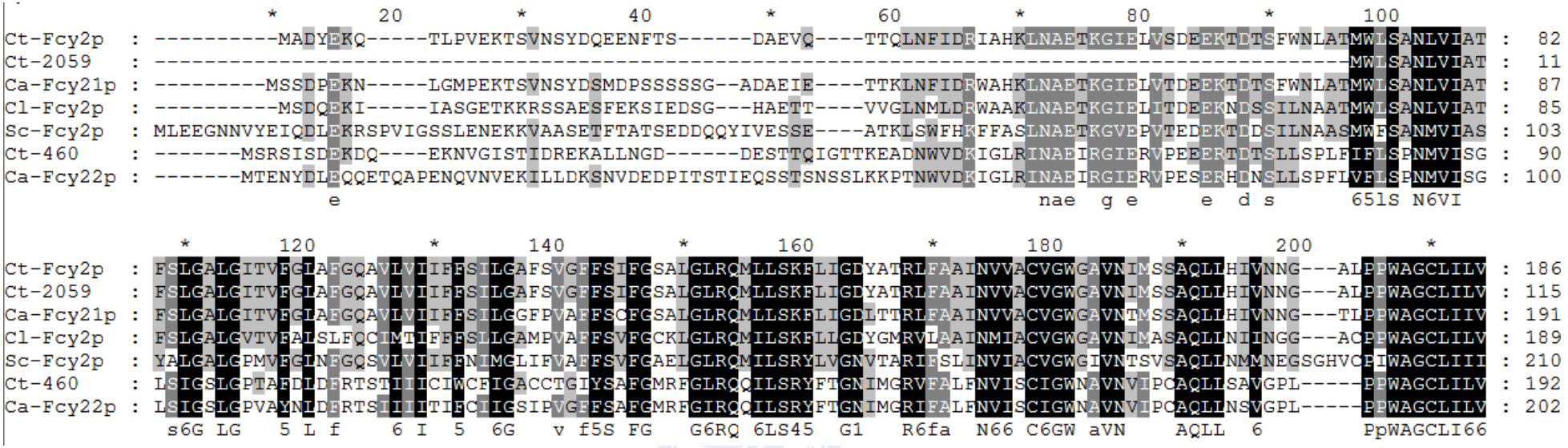


FIG. 17. Amino acid sequence alignments of the N terminal of the PCP. Ct-Fcy2p, deduced amino acid sequence from the *FCY2*^{S} of YM020291; Ct-2059 and Ct-460, Broad Institute *C. tropicalis* database, CTRG_02059 and CTRG_00460, respectively; Ca-Fcy21p and Ca-Fcy22p, Stanford *Candida* genome database, orf19.1357 and orf19.333, respectively; Sc-Fcy2p, GenBank accession no. X51751; Cl-Fcy2p, GenBank accession no. AY506866. Absolutely conserved residues are boxed in black, highly conserved residues are boxed in 50% grey and weakly conserved residues are boxed in 25% grey. Ct, *C. tropicalis*; Ca, *C. albicans*; Cl, *C. lusitaniae*; Sc, *S. cerevisiae*.

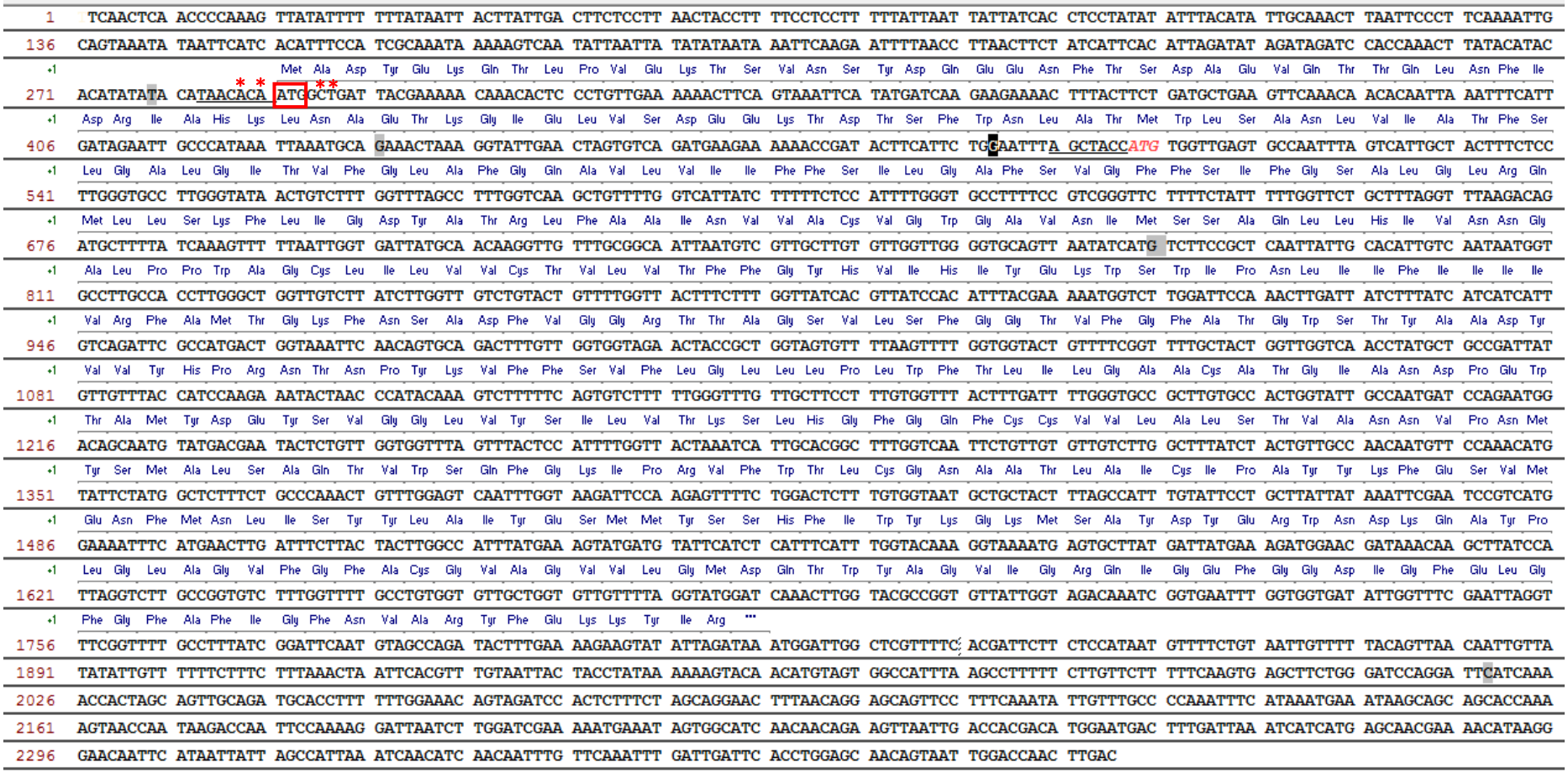


FIG. 18. Nucleotide sequence of a 2258 bp DNA segment containing the entire *FCY2* gene and the deduced amino acid sequence of the purine cytosine permease protein. The start coden identified by database is shown in italic font in red. Regions including seven bases preceding the newly defined start codon and the original one are underlined. Nucleotides around the start codon that are favorable for the initiation of translation are marked with asterisks. Polymorphic nucleotide at position -69 (145), position 273 (486) and position 1518 (1731) are shaded in grey. A *fcy2* G201A mutation leading to a nonsense mutation is shown in solid dark square. The proposed initiation codon is boxed in red.

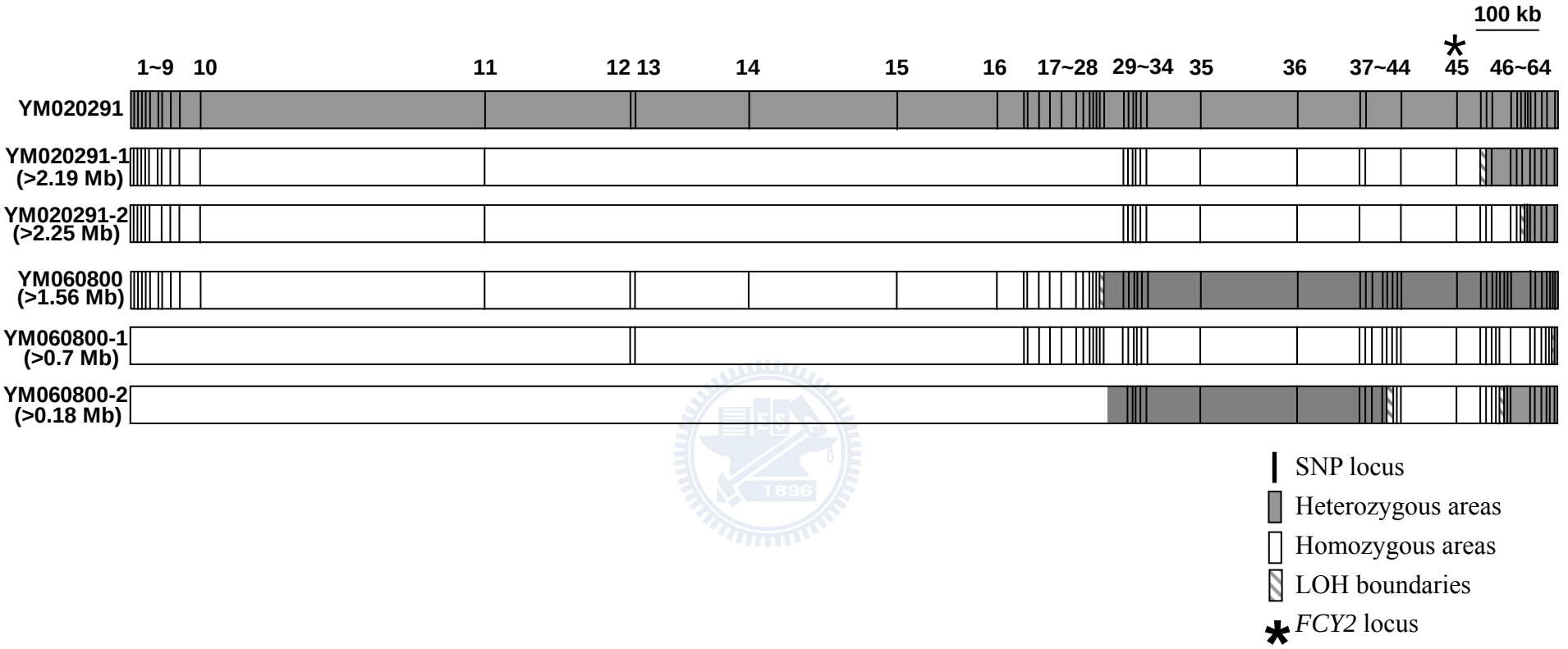


FIG. 19. SNP map of the chromosome carrying *FCY2* gene in clinical isolates, YM020291 and YM060800, and their derivatives. Chromosomal locations of SNP loci and LOH events of strains YM020291, YM060800, YM020291-1, YM020291-2, YM060800-1 and YM060800-2 (map modified from Diogo *et al*.) Numbers above chromosomes (a.k.a., supercontig 2, available on the database http://www.broadinstitute.org/annotation/genome/candida_group/FeatureSearch.html) correspond to SNP order in Figure 20. Regions containing SNPs in the parental strains YM020291 or YM060800 are identified (small vertical black lines) and the heterozygous (grey) and homozygous (white) areas in the corresponding regions of their derivatives are determined on the chromosome. Accordingly, LOH boundaries were narrowed down to a smaller range (hatched). The *FCY2* locus is marked with asterisk. The parentheses on the left side of the figure refer to the extent of the LOH events occurred in each strain.

SNP map location	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
SNP locus (CTRG)	1078	1080	1082	1083	1085	1098	1100	1112	1120	1138	1337	1438	1439	1541	1646	1725	1745	1755	1756	1765	1775
Location	2080	5273	8705	9596	11071	32314	35929	57814	73941	112101	569124	812195	813328	1028000	1238118	1403012	1443451	1461630	1463024	1480107	1495295
	3204	7900	9556	10228	11745	34611	37083	59109	75974	113703	570032	813307	815709	1030285	1239509	1404406	1448454	1462862	1464199	1481870	1496605
SNP number	6	1	1	1	1	1	1	1	3	1	1	2	1	2	2	2	2	2	3	5	4
YM020291	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het
YM020291-1	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	-	-	-	-	-	-	-	-	-	-
YM020291-2	Hom	Hom	Hom	Hom	Hom	-	Hom	Hom	Hom	Hom	Hom	-	-	-	-	-	-	-	-	-	-
YM060800	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom
YM060800-1	-	-	-	-	-	-	-	-	-	-	-	Hom	Hom	-	-	-	Hom	Hom	Hom	Hom	Hom
YM060800-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

SNP map location	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
SNP locus (CTRG)	1785	1793	1796	1798	1800	1802	1804	1817	1820	1822	1823	1837	1843	1886	1944	1991	1993	1999	2008	2009	2010
Location	1526667	1539505	1548599	1551511	1556896	1561090	1565480	1601857	1606071	1609875	1610712	1627871	1643554	1731285	1883328	1983786	1994160	2006268	2021331	2023676	2029381
	1529114	1541490	1550536	1553118	1557744	1562187	1572091	1602714	1607084	1610429	1611176	1629046	1645512	1732322	1885439	1985714	1995590	2006951	2023052	2026795	2030733
SNP number	1	3	5	3	(0,1,1)	1	(2,2,0)	4	(0,0,1)	1	(0,1,1)	3	1	1	1	(1,1,0)	2	1	1	8	10
YM020291	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Hom	-	-	-	-
YM020291-1	-	-	-	-	-	-	-	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	-	-	-	-
YM020291-2	-	-	-	-	-	-	-	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	-	-	-	-	-
YM060800	Hom	Hom	Hom	Hom	Hom	Hom	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het
YM060800-1	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom
YM060800-2	-	-	-	-	-	-	-	-	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Hom

SNP map location	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	
SNP locus (CTRG)	2011	2012	2059	2080	2082	2083	2084	2086	2087	2088	2089	2099	2103	2104	2105	2109	2114	2117	2119	2120	2121	2123	
Location	2032580 2033755	2034001 2036349	2136372 2137688	2191805 2194174	2195668 2199483	2201361 2203973	2204168 2206054	2207935 2208711	2209077 2209883	2209966 2211957	2212792 2214423	2239505 2241031	2251245 2252384	2253712 2254596	2255239 2256126	2265384 2268029	2275744 2277564	2286111 2288372	2293778 2295835	2296116 2297057	2297301 2297681	2299852 2303001	
SNP number	3	3	(1,5,1)	(5,4,0)	6	9	2	(2,1,0)	5	1	2	1	4	(5,0,0)	3	1	2	1	(2,10,0)	2	(0,1,3)	(0,4,1)	43
YM020291	-	Het	Het	Het	Het	Het	-	-	-	-	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	-	-	Het
YM020291-1	-	Hom	Hom	Hom	Het	Het	-	-	-	-	Het	Het	Het	-	-	Het	Het	Het	Het	Het	-	-	Het
YM020291-2	-	Hom	Hom	Hom	Hom	Hom	-	-	-	-	Hom	Hom	Hom	Hom	Het	Het	Het	Het	Het	Het	-	-	Het
YM060800	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Het	Hom	-	-	-	Het	Het	Het	Het	Het	Het	Het	Het
YM060800-1	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	-	-	Hom	Hom	-	-	-	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Het
YM060800-2	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Hom	Het	Het	Het	Hom	-	-	-	Het	Het	Het	Het	Het	-	Het	Het

*

FIG. 20. Genotypes of clinical isolates and their derivatives on the chromosome, containing *FCY2* gene. Characteristics of variant loci for the clinical isolates, YM020291 and YM060800, and their derivatives YM020291-1, YM020291-2, YM060800-1 and YM060800-2. The chromosome of each parental strain was screened for the presence of heterozygous SNPs by sequencing partial regions of genes spread on it. Two states are observed at each locus: a heterozygous state (Het, grey) where the alleles of a pair are heterozygous and a homozygous state (Hom, white) where the alleles of a pair are homozygous. Numbering corresponds to the locations of SNPs on the SNP map. CTRG is the gene locus number designated by the *Candida* database of Broad Institute (http://www.broadinstitute.org/annotation/genome/candida_group/FeatureSearch.html). Three numbers in the parentheses respectively stand for the amount of SNPs found in the upstream, ORF and downstream region of the gene. The *FCY2* locus is marked with asterisk. –, not done.

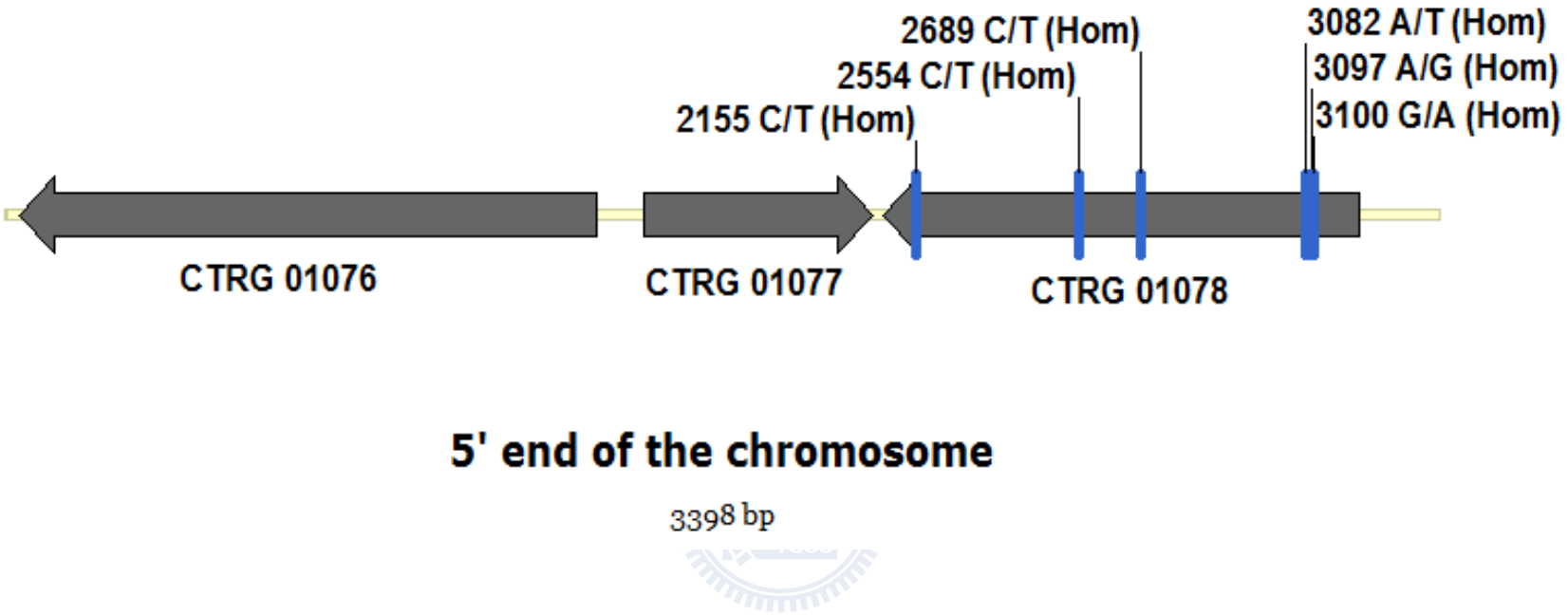


FIG. 21. 5' end of the chromosome in the clinical isolate YM060800 and the derivatives of YM020291. The map above depicts the 5' end of the chromosome, corresponding to nucleotides from position 1 to 3398 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in the clinical isolates YM060800 as well as the derivatives of YM020291 (YM020291-1 and YM020291-2). Each solid arrow stands for the annotated gene given by the Broad institute, and represents the orientation of the gene. The relevant position of SNP markers is indicated by the vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in the clinical isolate YM020291, and the status of each polymorphic site in YM060800, YM020291-1 and YM02091-2. A, adenine; T, thymine; C, cytosine; G, guanine. Hom, homozygous state.

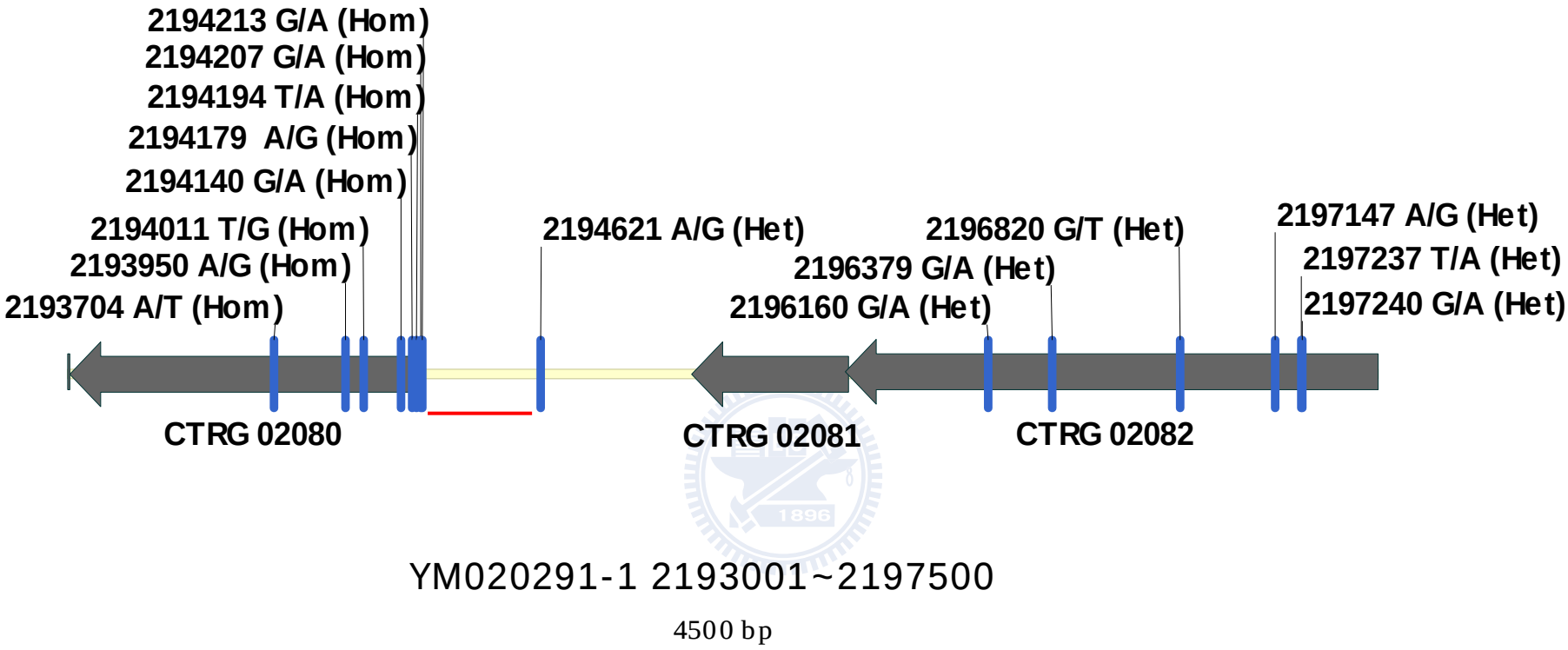


FIG. 22. The LOH boundary on the right end side of the chromosome downstream of the *FCY2* gene in YM020291-1. The map above depicts the right end side of the chromosome, corresponding to nucleotides from position 2193001 to 2197500 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in YM020291-1. Each solid arrow stands for the annotated gene given by the Broad institute, and shows the orientation of the gene. The relevant position of SNP markers is indicated by the thick vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in the parental strain YM020291, and the status of each polymorphic site in YM020291-1. The identified LOH boundary is underlined. A, adenine; T, thymine; C, cytosine; G, guanine; Hom, homozygous state; Het, heterozygous state.

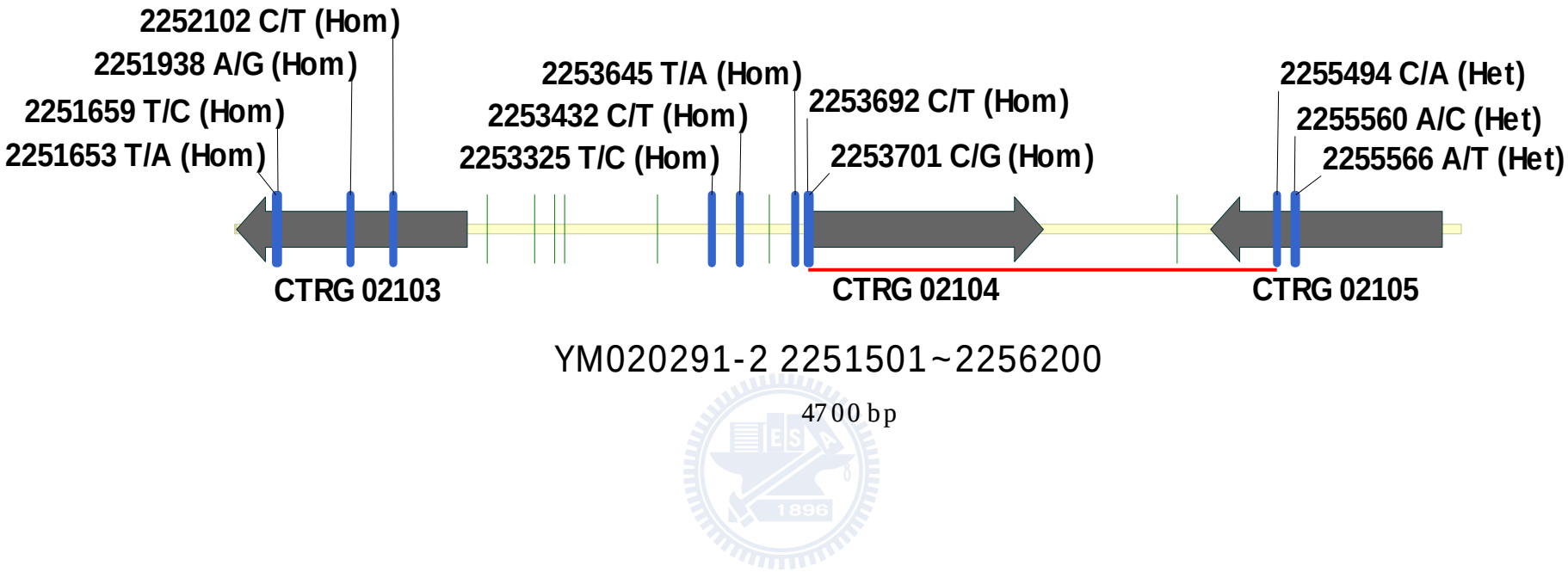


FIG. 23. The LOH boundary on the right end side of the chromosome downstream of the *FCY2* gene in YM020291-2. The map above depicts the right end side of the chromosome, corresponding to nucleotides from position 2251501 to 2256200 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in YM020291-2. Each solid arrow stands for the annotated gene given by the Broad institute, and shows the orientation of the gene. The relevant position of SNP markers is indicated by the thick vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in the parental strain YM020291, and the status of each polymorphic site in YM020291-2. Regions that possess poly A or poly T (≥10 bp) are indicated by thin vertical lines. The identified LOH boundary is underlined. A, adenine; T, thymine; C, cytosine; G, guanine; Hom, homozygous state; Het, heterozygous state.

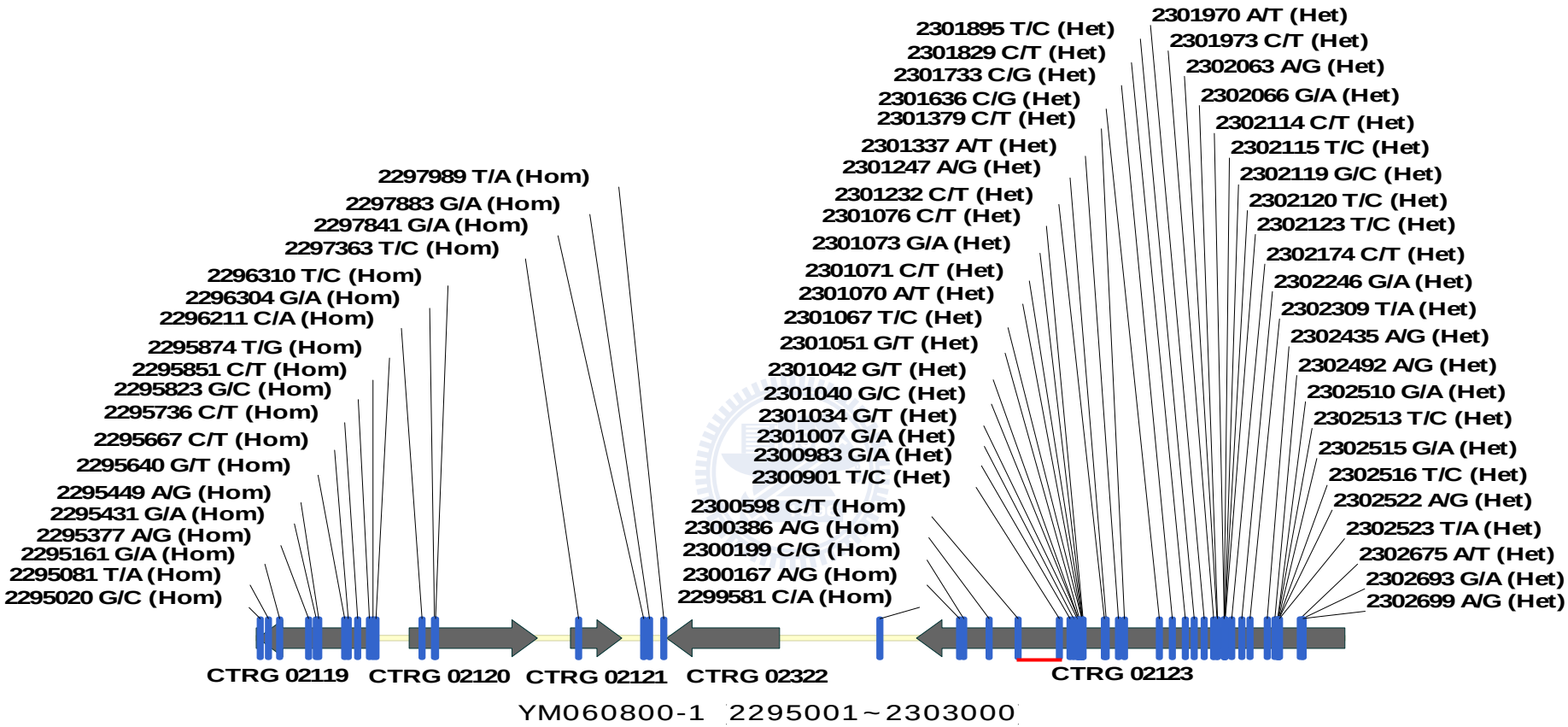


FIG. 24. The LOH boundary on the right end side of the chromosome downstream of he *FCY2* gene in YM060800-1. The map above depicts the right end side of the chromosome, corresponding to nucleotides from position 2295001 to 2303000 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in YM060800-1. Each solid arrow stands for the annotated gene given by the Broad institute, and shows the orientation of the gene. The relevant position of SNP markers is indicated by the thick vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in the parental strain YM060800, and the status of each polymorphic site in YM060800-1. The identified LOH boundary is underlined. A, adenine; T, thymine; C, cytosine; G, guanine; Hom, homozygous state; Het, heterozygous state.

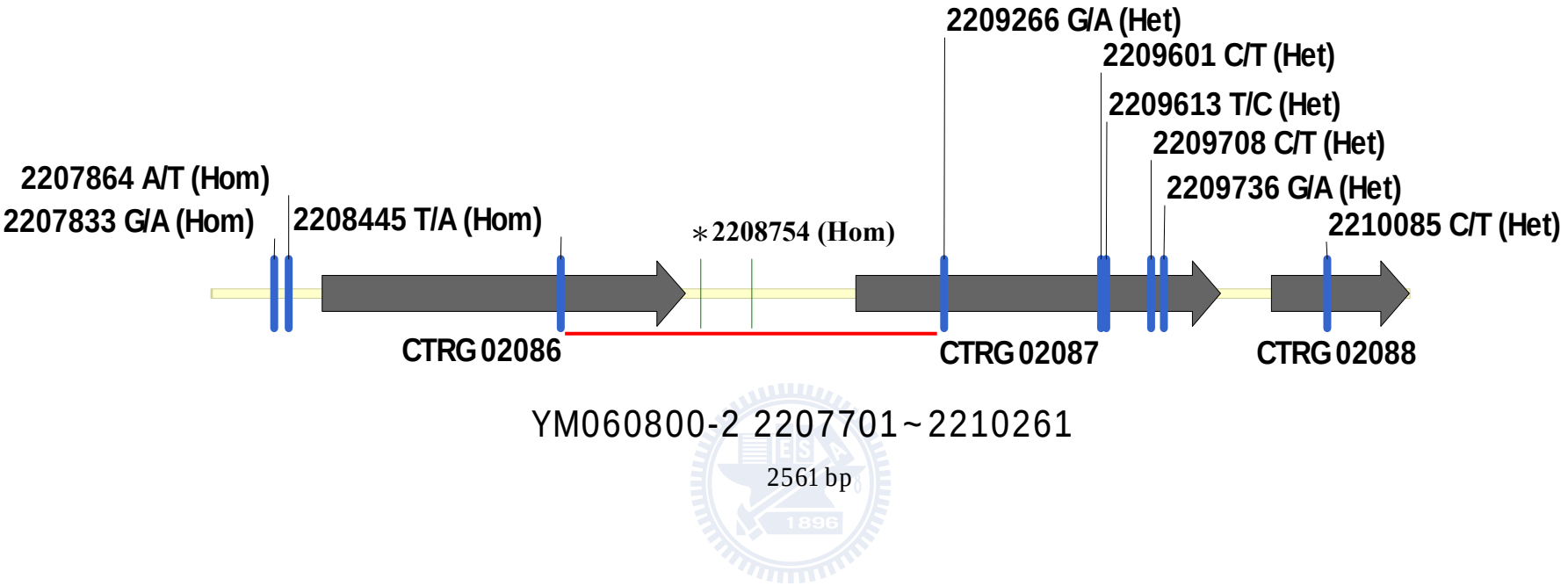


FIG. 25. The LOH boundary on the right end side of the chromosome downstream of the *FCY2* gene in YM060800-2. The map above depicts the right end side of the chromosome, corresponding to nucleotides from position 2207701 to 2210261 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in YM060800-2. Each solid arrow stands for the annotated gene given by the Broad institute, and shows the orientation of the gene. The relevant position of SNP markers is indicated by the thick vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in the parental strain YM060800, and the status of each polymorphic site in YM060800-2. Regions that possess poly A or poly T (≥10 bp) are indicated by thin vertical lines. The identified LOH boundary is underlined. The deletion polymorphism is marked with asterisk. A, adenine; T, thymine; C, cytosine; G, guanine; Hom, homozygous state; Het, heterozygous state.

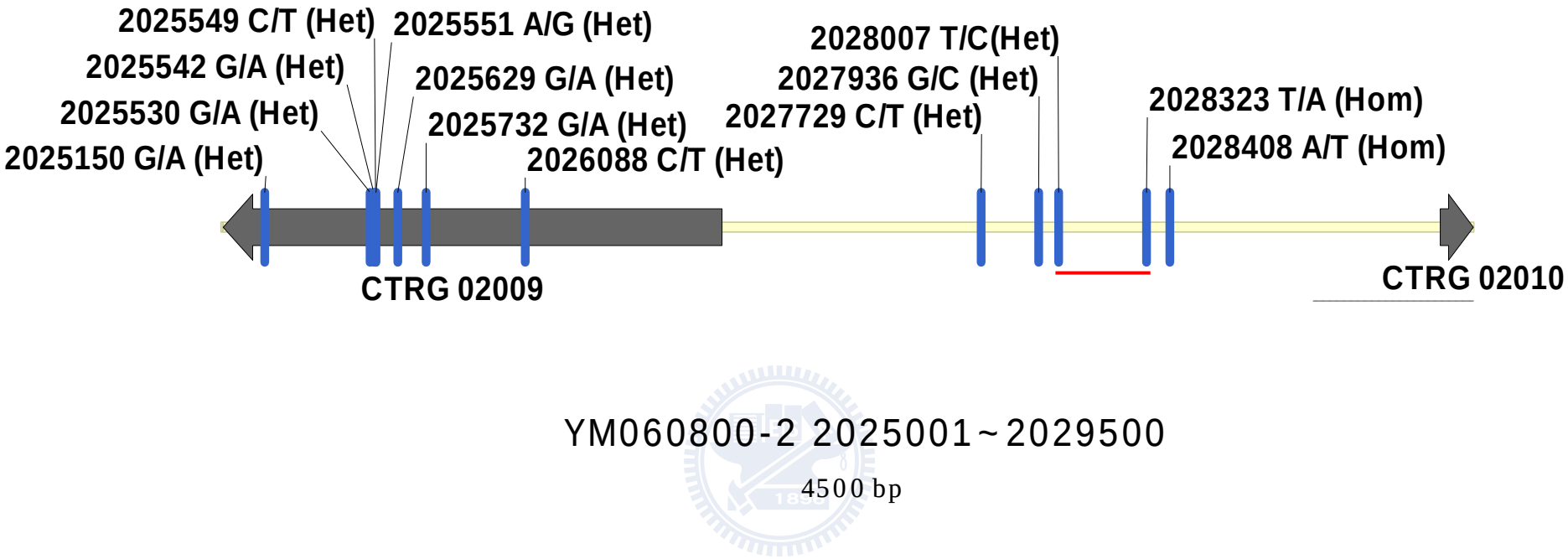


FIG. 26. The LOH boundary on the left end side of the chromosome preceding the *FCY2* gene in YM060800-2. The map above depicts the left end side of the chromosome, corresponding to nucleotides from position 2025001 to 2029500 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in YM060800-2. Each solid arrow stands for the annotated gene given by the Broad institute, and shows the orientation of the gene. The relevant position of SNP markers is indicated by the thick vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in the parental strain YM060800, and the status of each polymorphic site in YM060800-2. The region that was not checked by sequencing is shown in dot line. The identified LOH boundary is underlined. A, adenine; T, thymine; C, cytosine; G, guanine; Hom, homozygous state; Het, heterozygous state.

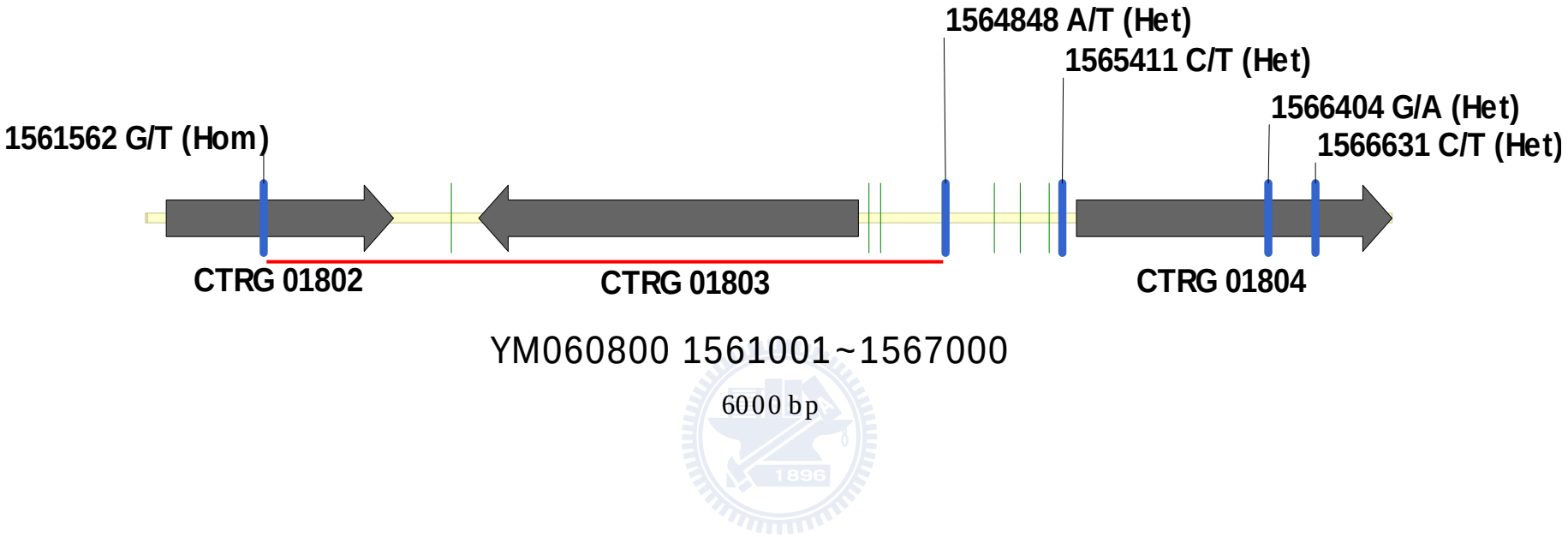


FIG. 27. The LOH boundary on the left end side of the chromosome downstream of the *FCY2* gene in YM060800. The map above depicts the left end side of the chromosome, corresponding to nucleotides from position 1561001 to 1567000 in the supercontig 2 available on the database (http://www.broadinstitute.org/annotation/genome/candida_group/Regions.html), in YM060800. Each solid arrow stands for the annotated gene given by the Broad institute, and shows the orientation of the gene. The relevant position of SNP markers is indicated by the thick vertical line, and the attached information is its location on the chromosome, the polymorphic nucleotides identified in another clinical isolate YM020291, and the status of each polymorphic site in YM060800. Regions that possess poly A or poly T (≥10 bp) are indicated by thin vertical lines. The identified LOH boundary is underlined. The identified LOH boundary is underlined. A, adenine; T, thymine; C, cytosine; G, guanine; Hom, homozygous state; Het, heterozygous state.

YM020291-1 (409 bp)

ctaatgattaaacgcgtcatttgaaaggtgtaactcagtaatcgcgtctgggtcaaatttataatacaacaaataaataaaaatacataaaacaaggtcgtgtacattggcatatattttccccaccaatgagctggtctact
caaaccaactgtatctatagactggactcaatcctgttgaaactcaaaaagtatgttcacattgtagatttaaaaagataagtttatttgtaacaatgacactgtcaaacagcatttccagaatgtacaccaaggactatgtta
gaccttgcaactacatgaactttacatgttcaataatgccatgtatgacgttattgatactatgtctctgatatggtaatcaagtactacttttccaagccaaaaatatccaatttttgtta

YM020291-2 (1794 bp)

cgacaaacataatgacttcaactaacttaaacggtactttcccattaattgctaaaggtaaagtcagagatatttatcaagttgatgaaaatactcttttatttgttgcaactgatagaatttcagcttatgatgttattatgtcaa
atggtattcctcaaaagggtaaaattttaactaaattatctgaattttggtttgatttcttacctattgaaaatcatttaattaaaggtgatattttcactaaatatccacaattagaaccatatcgtgatcaattagaaggtagatca
ttattagttcgtaaattgaaattaattccattagaagttattgttagaggttatattactggttccggttggaaagaatatactaaatctaaaactgttcatggtaaatctattggtgaaggaattcttgaatctcaaaaacttgaa
ccacctattttcactccttcaactaaagctgaacaaggtgaacatgatgaaaatattagtaaagaacaagctgataaaattgttggtaaagaattatgtgatagaattgaaaaaattgccattaatttatatacaaaagcaag
agattatgctgctaccaaaggaattattattgctgatactaaatttgaatttggtttagatgataataacaatattattcttgttgatgaagttttaactcctgattcttcaagattctggaatgctagtaaatatgaaattggtaaat
ctcaagaatcttatgataaacaatttttacgtgattggttaacttcaaatggtgttgctggtaaagatggtgttgccatgccagaagatattgcaaccaaaactaaagaaaaatatgttgaagcttatgaatgtttaactggtg
ataaatgggtcgattaagaaaacaaataactgaaaaagcaaccaattagataatacatgtaagaaaaaataatactatatataatgtaagtgaatgaacttgtagttgtaaaacaattgggctgcaagtttaagccttggtt
gtgtgtgatacgccgtaatagttataagcagacgtgtcagtatagatcggcaggcaatctttcttcctacacgtgtattaatccctgaagcataaggaatgtaaaatacgccttgttggaatttttggttttagcaactgttgc
tctggtatagataatctgaatataaagaatgttctcgcattgtggttgttgttttcccaacgaacccgagtggcttaagcaaagttgctgccttgtttgtttgtttgtcggccgtgtcgactcgttcctcgttcgagagaatcg
agaaagagatatatgaataacaacattcattcatttgtttgttgtgtttgatttgttattaatgtttatcaaggtaaatatatataactctctgtatacaaaaaaaaaaatactgttccccatttttttgaaaagttgacggggagg
aggattagattccttctatcgttaattgaaattcatagatgcgagggaggtaggggagggtgctagacaataattcttttcaaaatctaccaatataatccttcgcactaattgtagttgttgcataagattcttgatcattttc
catttgattcatatctgcaaatgaatagaaatcaatggcattagtatcaactttgtctttgaaataggagtttgaccagcataactcttcagtcgttttgcaagtggatggaacgataatgaagaaacaccagttgaagaag
tcgaagaacgaagattaccactaccactaccactaccaatagaatgg

YM060800-1 (304 bp)

catatcccaattgctcttttgttcgcaactgatcaaaacaaggttccttgatgactgtacagaacaagtctgtcaacacacgcaatttcaaatcatctgcatttggtgcgatctggatatagtactcaatgcacgaattgatat
ttgcttcatccttcaatgacatctcatagcggatcacctcctctggttcaaagacataaccttccaagtggaatttctcttcatcgtactcttccatccatggcttgatgtgttttatagaatcatttatggtgtctttgattgctctt
gctttggtaat

YM060800-2R (513 bp)

atagacgatcaattttaattttgttcttttcgtgtacaaaaaaactattttggcaacaatcccaaaaaaatgttcaaagatccttctctgaatgaataccaaaaaaaaaaaaccaacggctctataatacaatagaaaaatgtac
atgtagatttgttttgttttgttttgttttacaaccattcatattttgattttgtttatattcttcacaaatatcaaccactgttttttctgtttcattagtagagttctatatccttttttttctgatcaatcttgcaaccagttttttgggatcca
ccatatcaaaattctgagcaatgtccatgctttcatctacaattgatacattgataagtgaagctggtggtatggttgaacaaatacttcctttatccaataattctagcaagttaaatgggagtcggggttctttatcaagcac
agctgaagaaatgacaccaagtgcaagtattgaagaactttgtactgcagttgacgaattacctg

YM060800-2L (317 bp)

Tcatcaaatctgtcaattgacggaactctgtcttatattattgtataattattacccattcattataatttgacaagagagaaatttaaagagacaggaattaaagaaggtgtcaaatagaaagttcactacatttgacgtatta
ttttttagattcaatgcattgttttcccaatttgtcaatatccgctagctagaagccagaattataaacctatgttttcaacttatttttggttatcagtgtaagatacaatggaagtaagaaggagcaatgaattacggatcaag
aggaaagcttgggtatgggcaaatt

YM060800 (3287 bp)

ggaaattagttgagttcccaagaattaataatccagtatcattattaaaatcgagaaaaggtagatgtggagaatgggtgaattgttttatgttaatattacaagcattaataggtggaggtgatactgatagaatcagatac
gtttggaatcacgaagatcacgtttggtgtgaatattactccaatagttctaaaagatggatacatttagatccatgtgaagcagttgctgatgaaccattgttatattgtaataactggggcaagaagatgagttttgttattg
gattcaatcgtcattatattattgatttgagtgacaagtatataactcctgaaaaacaaataagtaaagcttcgattaccagtgaaagtaaagttaaatcagccatacattatatcaacagtaagaaacagtcagatttttaca
ggcaattaaaaattcaaatgaacgaagaacaggcattaattaaagtgtatgaggaaatcattgtacctcataatgctgaaacttgaagaaactggaaaaagaattaccatcaacaactacaagtgatttacctgttgga
agacaatctgggtctgcagagtggactaagagtcgtggagaagatggaaagtgagtgaggttttagaaactatgtagaagaaatataaatttcaatgaaatataattgatttactaaagggttaatattaccaatgaagat
tggtggtgaatctaacaagaccataaagccaccaacttttaaatttcaacaaacagggcgtgtggcgtagttggtagcgcgttcccttagcatgggaaaggtcatgagttcgactcttatctcgtccagttttcttttttttga
atttttaacagcttgctgataaagtaattacttgaatcgtcattgctttttttttttatatttaattgataatttaaacaagagtcaatgatatgtttatttccaagaaattctattaataaagagcaatgaatgaatataaaaataat
gataataataataataaaaaaagaacaactataatacaaccttgaagccaatgatttttgtaatccaaaatagaaccgtcccatcttgtagcattcttgtttcaataacaaaaatgtctttagcaactttatccaataatctg
aaatcgtgggagacgacaacaacaccaccattaaaggcgttgatggcatctgccaaagaatcaatagtagccaaatccaaaccgttagtaggttcgtccaataaaatcaagttaggagcttccaatgctaataatgca
aaaacgacacgagatctttgaccttcagataaagtagccatttgggaagtttgggcttcaccactcaaaccataacgacccaattgacctctccagtattgataatcttgtgaaatactagaaaatttatcacgaacgaatt
ctaatggggatttagtcaaatccaattggtctgcagaatgttgtgagtaaacacccaatttaatatgagtatgttgaataactctacccttttgtggttgtaatttaccttggaacaatttcaacaaagtggatttaccgatacc
atttggaccaacaagagcaactctagaatccatatcgataccaatatccaaatgttcatacaagttgtcttcttctttacctgagtaagagaaagacatatcgtcgaatgccaagactggaggtggcaatttttcaacttca
gggaatctgaaactgaagaccttatcaggaacaactggttgaatcaaaccatcggcttccattttatccaaaatcttttgtctagatttagcttgcttgaccaagttagcataagtaccagcagatgcaatgaatttcttgata
tgagcaatttcttcttgttgcttgttgtattgtttcatttggttagtttccaattcagctctggttttaacataggagtcgtagttaccaccatacatagttaaaactttcaatctcatatcaatcatgttagtacaaacaccattaagg
aaatcttgggagtgggagaccaagatcaaaattctgtcaaatctttttaagtattcttccaaccaaacacaagctgccaaatccaaatgggcagttgggtcatccaataacaataaggttggtttaacaaacaaagctttg
gccaaggcaacacgcattctccaaccaccagacatgtctttggttttttgttgatggtaactgagttgaaacccaaaccagtcaagataacagcagctctacttttcaaaagttgaagggtccatttcatcgatttttcgta
aagagcttccaaagctggacattctggaccttctttgacaatcaattcttcaaccaaatcttccaatctcttcatttcatgttcggcttctctgacaacatattccaaagcagaatattcagttgcttcagctggttcattcaaca
agtaaatatcaatatgttctgggattgggaattctctagcagcaatggatttcaataaggtggatttaccacaaccattttcacccaataaaccatatctacgaccataattcaattccaaagtagaatcttgaatcaaaactt
taccgtggaataaaagggaaacagaagacaatttaatatctctggaagtgaccaatgattccaaaacaccagtagtgactctgtctgacaaaccatgttcatcggtttgtaatttcaatttggcaatttgttgctcagcatct
tcaacttcatcttcagcagcatccttttcggcttgcttcttagttttcttggtagttttaccagcagcggcttttcttagcttctctttctaaacgcttttgttctctttagcttttgatgcagaaatagacatgactagttgttgcttact
atgaatatgaagagaagagaaatcaataaaaattttttttttttttcactcccacttccttcaataaggagaagggcgagagtaatagaaaaaaaaaagaaacaaagtacccgatgagtgaacagcttttatgcagtacc
caatataatggcgtttttttttaatagtctattgtttgcaacaccaagtttcatcatcagggtagtaccatacaaaggacatgaacaggttagaggaattgatcttattaagaaccttccaatgattaattgaattgtatgtgttttt
gatgataatgacgaggtggtaaaacaagaactattcaacagatgcacaaaattaaatattagaacgcagcaaccaagaatgtgtacagcagtttattcaataataattcacata

FIG. 28. Nucleotide sequences of the LOH boundaries. Nucleotide sequences of the LOH boundaries of five strains, YM020291-1, YM020291-2, YM060800-1, YM060800-2 and YM060800. The number in the parenthesis is the length of each defined LOH boundary. YM060800-2R, right border of the LOH boundary of the derivative YM060800-2; YM060800-2L, left border of the LOH boundary of the derivative YM060800-2.

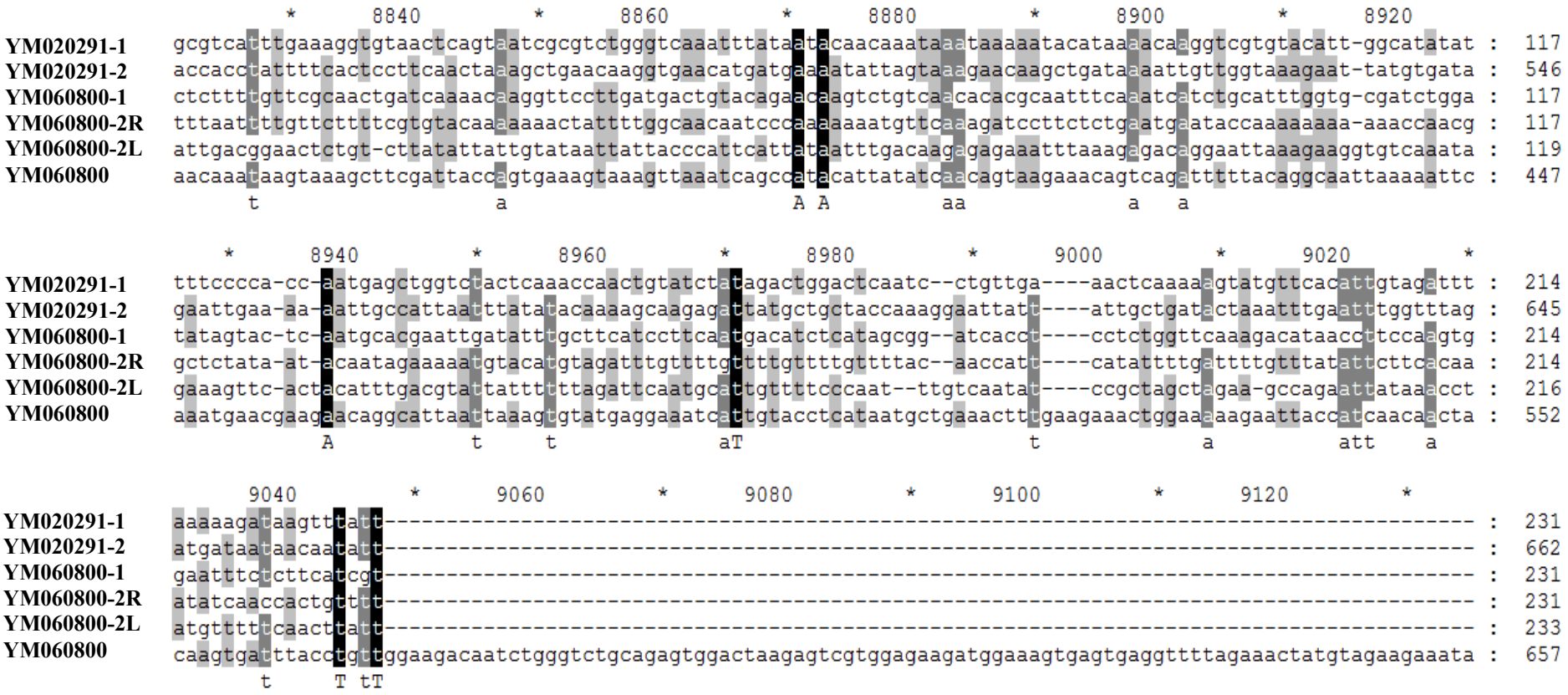


FIG. 29. Nucleotide sequence alignment of the LOH boundaries. Nucleotide sequence alignment of the LOH boundaries of five strains, YM020291-1, YM020291-2, YM060800-1, YM060800-2 and YM060800. YM060800-2R, right border of the LOH boundary of the derivative YM060800-2; YM060800-2L, left border of the LOH boundary of the derivative YM060800-2. Absolutely conserved residues are boxed in black, highly conserved residues are boxed in 50% grey and weakly conserved residues are boxed in 25% grey.

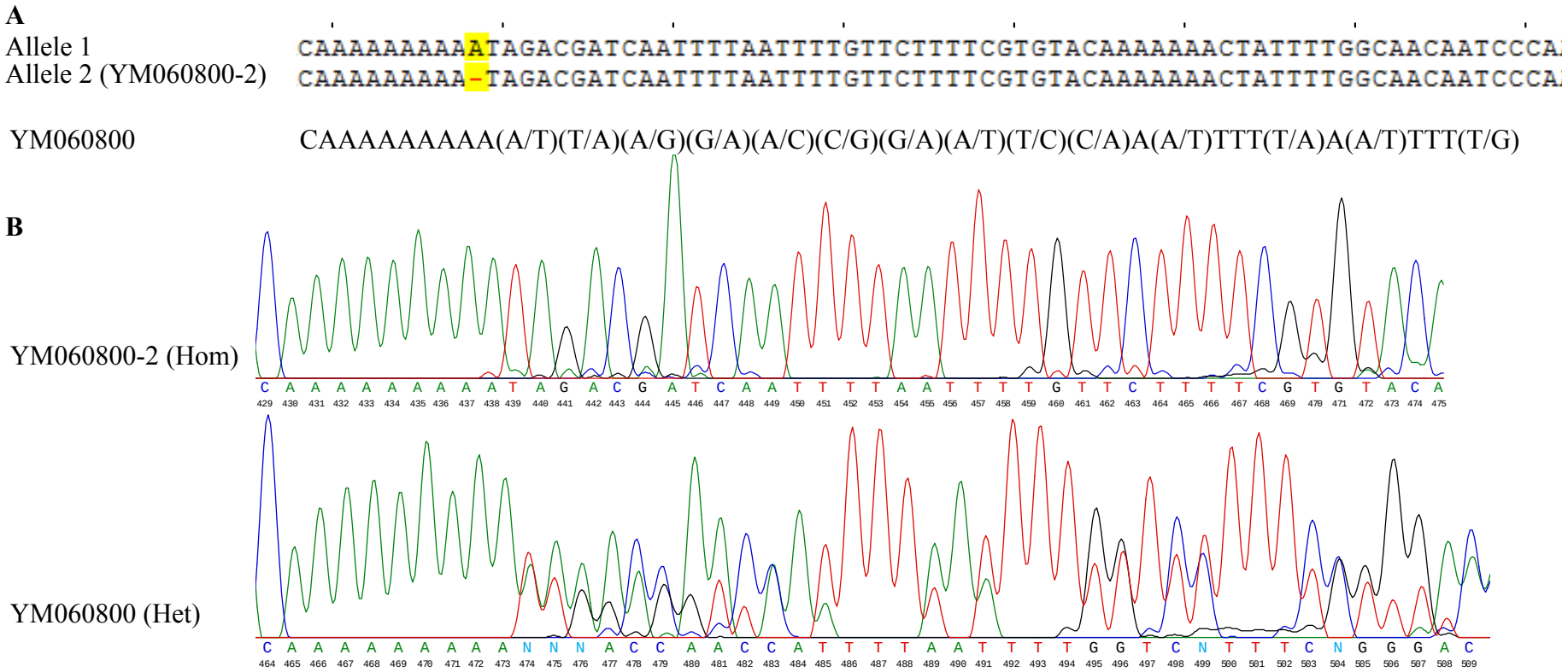
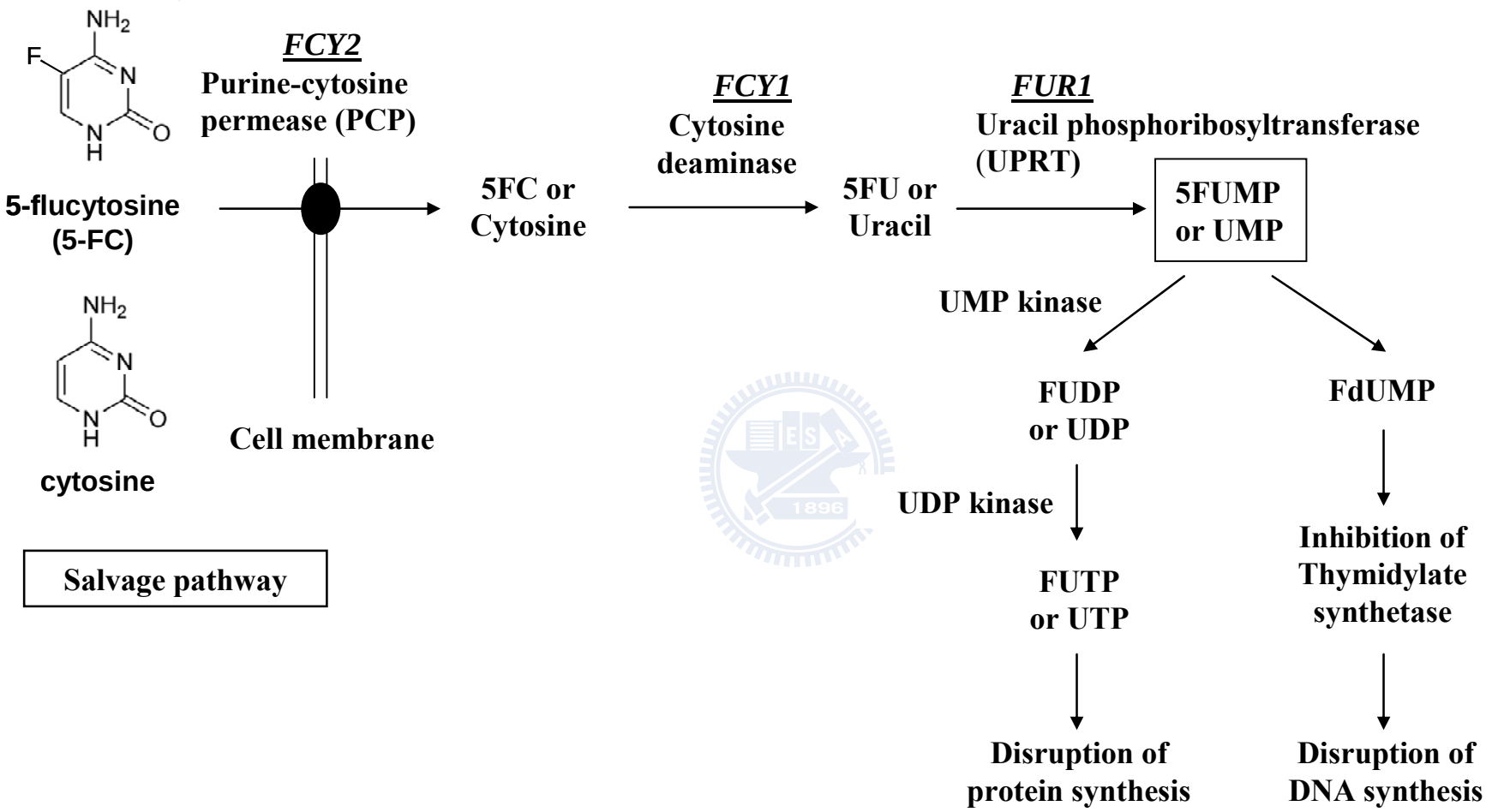


FIG. 30. A deletion polymorphism in YM060800. (A) The nucleotide sequences show the nucleotide sequences (from position 2208744 to 2208817 based on supercontig 2) at the downstream of the stop codon of CTRG_02086 in the clinical isolate YM060800 (allele 1 and allele 2), which differs by a deletion at position 2208754 (boxed in yellow). YM060800-2 is homozygous at this locus, exhibiting the sequence pattern of allele 2. Sequencing result of YM060800 suggesting the combination of two type sequences is given in third row. (B)Sequencing results of YM060800-2 and YM060800 correspond to the nucleotide sequences above. Hom, homozygosity; Het, heterozygosity.



Appendix 1. Metabolism of 5FC. 5FC metabolism is through pyrimidine biosynthetic pathway. 5FU, 5-fluorouracil; 5FUMP, 5-fluorouridine monophosphate; UMP, uridine monophosphate; FUDP, 5-fluorouridine diphosphate; UDP, uridine diphosphate; UTP, 5-fluorouridine triphosphate; UTP, uridine triphosphate; FdUMP, 2' deoxy- 5-fluorouridine monophosphate