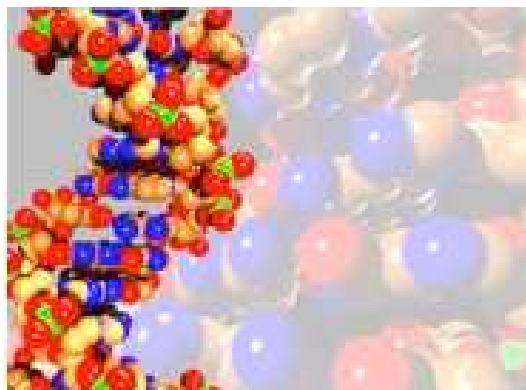


**Maestría en Biotecnología
Facultad de Ciencias
Universidad de la República**

“Diversidad molecular entre cepas industriales de *Saccharomyces cerevisiae*”

Lic. Sandra Jubany González



**Directora de Tesis: Dra. Carina Gaggero
Director Académico: Dr. Francisco Carrau**

**Julio, 2012
Departamento de Biología Molecular
Instituto de Investigaciones Biológicas Clemente Estable
Montevideo, Uruguay**

**A mis padres,
Teresa y Roberto**

Agradecimientos

Esta Tesis de Maestría fue realizada gracias a la financiación del proyecto PDT32/06 del Programa para el Desarrollo Tecnológico, DINACYT, Uruguay. La finalización de este trabajo ha sido posible gracias a la beca POS_2009_899 otorgada por la ANII (Agencia Nacional de Investigación y Desarrollo), Uruguay.

Quiero agradecer a mi Directora de Tesis Ph.D. Carina Gaggero y a mi Director Académico Dr. Francisco Carrau por su dedicación como docentes y por la oportunidad que me dieron de trabajar en su equipo. En especial quiero agradecer a Carina por su confianza, por su comprensión, por su paciencia, por la cantidad de tiempo que ha dedicado animándome a terminar con este proyecto y sobre todo agradecerle su amistad. Todas las expresiones de gratitud que puedo dar en estas líneas quedan pequeñas en comparación a todo lo que he recibido de Carina.

Agradezco a la Q. F. Laura Acevedo y al Dr. Antonio de Freitas por su docencia y por el apoyo que me brindaron en la realización de la pasantía en Laboratorios Microsules Uruguay. Agradezco a la Dra. Ivanna Tomasco por su docencia y por la ayuda con todo lo relacionado a los Microsatélites.

A todos mis compañeros del Departamento de Biología Molecular del Instituto de Investigaciones Biológicas Clemente Estable les agradezco todos los buenos momentos compartidos, gracias por proporcionarme su apoyo. A Mario le agradezco su disposición y ayuda con las figuras.

A todo el equipo de la sección Enología de la Cátedra de Ciencia y Tecnología de Alimentos de la Facultad de Química le agradezco las horas de trabajo, capacitación y congresos compartidos.

A mis amigas y amigos, que gracias a Dios son muchos, les agradezco todo lo compartido, las enseñanzas, la contención y el cariño recibido.

Finalmente, quiero agradecer de forma especial el apoyo de mi familia, a mis hermanos, cuñadas, sobrinos y primos que me acompañaron y me apoyaron a continuar en momentos muy difíciles. Agradezco muy especialmente a mis padres por la educación y el amor que me brindaron.

A todos muchas gracias de todo corazón.

Abreviaturas

A: Adenina

ADN: Ácido desoxirribonucleico

APS: Persulfato de amonio (ammonium persulphate)

C: Citosina

dNTPs: desoxinucleótidos

G: Guanina

fwd: oligonucleótido forward

Ho: Heterocigosidad observada

ORF: marco de lectura abierto (Open ReadinG Frame)

pb: pares de bases

PCR: Reacción en Cadena de la Polimerasa (Polymerase Chain Reaction)

PIC: Contenido de Información Polimórfica (Polymorphism Information Content)

rev: oligonucleótido reverse

SNP: Polimorfismo de un solo nucleótido (Single Nucleotide Polymorphism)

SSR: Repeticiones de Secuencias Simples (Simple SequenCe Repeats)

TEMED: NNN´N´-tetrametiletlenodiamina

T: Timina

INDICE

RESUMEN	1
1. INTRODUCCIÓN	2
1.1. Importancia de <i>S. cerevisiae</i> en las fermentaciones industriales	2
1.2. Diversidad genética de <i>S. Cerevisiae</i>	4
1.2.1. Microsatélites o SSRs (Simple Sequence Repeats)	5
1.2.2. SNPs (Single Nucleotide Polymorphism)	9
1.3. Genoma de <i>Saccharomyces cerevisiae</i>	9
1.4. Importancia del regulador FLO8	10
2. OBJETIVOS	11
2.1. Objetivo general	11
2.2. Objetivos específicos	11
3. MATERIALES Y MÉTODOS	12
3.1. Cepas de <i>Saccharomyces cerevisiae</i> y medios de cultivo	12
3.2. Extracción de ADN genómico	12
3.3. Diseño de cebadores	13
3.4. Amplificación de Microsatélites	14
3.5. Electroforesis en geles de secuencia	14
3.5.1. Preparación de muestras	14
3.5.2. Tratamiento y ensamblaje de vidrios	15
3.5.3. Preparación de solución de acrilamida para gel de pequeña escala	16
3.5.4. Preparación de solución de acrilamida para gel de gran escala	16
3.5.5. Armado de gel desnaturalizante de poliacrilamida al 5% y 7M urea	16
3.5.6. Condiciones de electroforesis	17
3.6. Visualización de microsatélites por tinción con nitrato de plata	18
3.7. Construcción de escaleras de alelos para SSR polimórficos	18
3.8. Amplificación para el análisis de SNPs en <i>FLO8</i>	19
3.8.1. Amplificación de una región del ORF de <i>FLO8</i>	19
3.8.2. Amplificación de una parte del promotor de <i>FLO8</i>	19

4. RESULTADOS	20
4.1. Criterios utilizados para la selección de loci de SSRs.	20
4.2. Análisis de microsatélites	26
4.2.1. Ensayo de todos los SSRs con una población de referencia inicial	26
4.2.2. Ensayo en gran escala con los SSRs más polimórficos	30
4.3. SNPs en <i>FLO8</i>	38
4.3.1. SNPs en una región del ORF de <i>FLO8</i>	38
4.3.2. SNPs en una parte del promotor de <i>FLO8</i>	42
5. DISCUSIÓN	45
6. CONCLUSIONES Y PERSPECTIVAS FUTURAS	51
7. BIBLIOGRAFÍA	52
8. ANEXOS	58
Anexo I. Tabla con las 120 cepas de <i>S. cerevisiae</i> de nuestra colección	58
Anexo II. Escaleras alelicas y tablas con las frecuencias alelicas de los loci TTA_XIII, TG_VI, GT_X, C4_XV, YOR267C, YGL013C, YGL028C y AT_X	61
Anexo III. Secuencia del Cromosoma V de <i>S. cerevisiae</i> desde la coordenada 375206 a 378176	70
Anexo IV. Jubany S, Tomasco I, Ponce de León I, Medina K, Carrau F, Arrambide N, Naya H, Gaggero C. 2008. Toward a global database for the molecular typing of <i>Saccharomyces cerevisiae</i> strains. FEMS Yeast Res. 8:472-484.	
Anexo V. Mercado L, Jubany S, Gaggero C, Masuelli R, Combina M. 2010. Molecular Relationships Between <i>Saccharomyces cerevisiae</i> Strains Involved in Winemaking from Mendoza, Argentina. Curr Microbiol 61:506–514.	

RESUMEN

La caracterización e identificación de subespecies de *Saccharomyces cerevisiae* es de gran interés en el área industrial, clínica y ambiental. En particular, en la producción de vinos de alta calidad, hoy en día es práctica usual la inoculación de cultivos puros de cepas seleccionadas para asegurar un proceso de fermentación reproducible. Diferentes subespecies de *S. cerevisiae* dan como resultado vinos de calidades muy diferentes. El análisis de microsatélites o SSR (Simple Sequences Repeats) con un alto grado de polimorfismo ha sido propuesto extensamente como método de tipificación molecular de eucariotas a nivel de subespecie. La diversidad molecular entre cepas de *S. cerevisiae* también puede deberse a polimorfismos de un solo nucleótido (SNP, Single Nucleotide Polymorphisms).

En este trabajo se tipificó una colección de 120 cepas de *S. cerevisiae* constituida por 48 cepas nativas aisladas de uva o mosto, 50 cepas comerciales de vino, 12 cepas productoras de H₂S, 1 cepa comercial de Sake, 6 cepas comerciales de panificación y 3 cepas de laboratorio mediante el uso de microsatélites. A partir de la información de dominio público de los SSR presentes en el genoma de *S. cerevisiae*, se ensayaron 33 loci con el fin de encontrar el conjunto mínimo de loci polimórficos que permitiera diferenciar las subespecies de *S. cerevisiae* en nuestra colección. Se seleccionaron los nueve loci de SSRs más polimórficos y se construyeron escaleras de alelos para cada locus seleccionado con el fin de lograr una correcta asignación de alelos. El análisis de los loci TTA_XIII, TG_VI, GT_X, C4_XV, YPL009C, YOR267C, YGL013C, YGL028C y AT_X permitió la discriminación de 93 cepas de las 120 cepas de nuestra colección de levaduras.

La floculación es una característica muy importante de las cepas industriales de *S. cerevisiae*. En este trabajo se analizó la presencia de SNPs en una región del gen regulador *FLO8*, clave en la capacidad de floculación y formación de pseudohifas.

Para la industria biotecnológica de *S. cerevisiae* es crucial poder asociar marcadores moleculares con rasgos complejos que dependen de múltiples genes y de sus variantes alélicas. Este trabajo permitió la creación de una base de datos (<http://www.pasteur.edu.uy/yeast/>) que en un futuro dará información sobre la biodiversidad molecular de cepas de *S. cerevisiae* provenientes de diferentes orígenes geográficos, de diversas aplicaciones tecnológicas (alimentos, probióticos, bioetanol), de aislamientos clínicos y de muestras ambientales.

1. INTRODUCCIÓN.

Desde la antigüedad las levaduras de la especie *Saccharomyces cerevisiae* han sido empleadas en numerosos procesos fermentativos. Dentro de la industria alimentaria se han utilizado en la producción de alimentos como el pan y en bebidas fermentadas como el vino, cerveza y sake. Recientemente, ciertas cepas de *S. cerevisiae* están siendo usadas como complemento alimenticio (debido a su contenido en vitaminas, minerales y proteínas) y como agentes probióticos en el tratamiento de diarreas crónicas o recurrentes con el fin de reestablecer el equilibrio en la flora intestinal (Posteraro et al, 2005).

S. cerevisiae, es considerada como el modelo eucariota unicelular por excelencia. Su facilidad para cultivarla, manipularla y el amplio conocimiento de sus características biológicas y genéticas ha hecho de esta levadura una herramienta fundamental de estudio en el campo de la biotecnología, donde igual que en el sector agroalimentario tiene el estatus de levadura "GRAS" (Generally Regarded As Safe). También se utilizan cepas de *S. cerevisiae* para la producción de combustibles alternativos como el bioetanol. Sin embargo, algunas cepas resultan ser patógenos oportunistas asociados a enfermedades en pacientes inmunocomprometidos (Piarroux et al., 1999).

Actualmente resulta imprescindible la caracterización e identificación de subespecies de *S. cerevisiae* de interés industrial, clínico y ambiental. La tipificación a nivel de subespecie es de relevancia desde el punto de vista industrial, debido a que muchos grupos de levaduras forman parte de la microflora natural de alimentos y bebidas fermentadas y/o participan en el proceso de obtención de éstos. De ahí la necesidad de poseer métodos de identificación rápidos, precisos y reproducibles, que puedan ser aplicados al control de la calidad en la industria, con el fin de asegurar que la cepa de partida, la que conduce el proceso y la que rinde el producto final, sea la misma.

1.1. Importancia de *S. cerevisiae* en las fermentaciones industriales.

Desde el inicio de la civilización humana, existe una relación estrecha entre el hombre y la actividad fermentativa de los microorganismos. Aunque las bacterias y los hongos filamentosos participan en la producción de alimentos y bebidas fermentadas, en una fermentación alcohólica las levaduras son los microorganismos predominantes. Por ejemplo, la fermentación vínica es un proceso complejo en el que está involucrada

una sucesión de diferentes especies de levaduras. Las primeras evidencias de producción de vino se remontan al VI milenio AC en la antigua Mesopotamia. Durante milenios, el proceso de vinificación se consideró una propiedad intrínseca del mosto, hasta que en el 1863 Louis Pasteur demostró su dependencia de la levadura. *S. cerevisiae* es prácticamente la única especie de la microflora presente en el mosto capaz de finalizar la fermentación vínica. Debido a la baja representación de esta especie en el mosto inicial, el proceso puede demorar varios días en comenzar o en algunos casos el inicio no se produce. Es por esta razón que se introdujo el concepto de inocular cultivos puros de levadura para iniciar fermentaciones vínicas (Pretorius, 2000). Actualmente existen dos tendencias en enología, aquella que confía en la microflora presente en la uva y en las instalaciones de la bodega para realizar la fermentación de forma espontánea y aquella que prefiere una fermentación más controlada utilizando inóculos comerciales.

En la producción de vinos, la calidad del mismo está estrechamente relacionada con la ecología microbiana de la fermentación. Para obtener vinos finos de muy buena calidad, todas las etapas involucradas en la transformación de la uva en vino deben estar optimizadas de manera de lograr un proceso de fermentación reproducible. La especie de levadura responsable de la mayor parte de la transformación del azúcar del mosto de uva en etanol es *Saccharomyces cerevisiae*. (Pretorius, 2000). Además del etanol y CO₂, el metabolismo de *S. cerevisiae* durante la fermentación alcohólica genera una gran cantidad de productos, entre ellos el glicerol, el ácido acético, el ácido succínico y el ácido láctico. Por otro lado, las propiedades aromáticas de los vinos se ven muy afectadas por la producción de alcoholes superiores (alcoholes con más de dos carbonos), por ésteres y otros compuestos volátiles (Antonelli et al., 1999).

Diferentes subespecies de *S. cerevisiae* dan como resultado vinos de calidades muy diferentes y por lo tanto, las cepas seleccionadas para ser empleadas en el proceso de fermentación deben presentar características enológicas óptimas. El proceso de aislamiento y selección de cepas es largo y tedioso, requiere la inversión de tiempo y dinero e implica desde ensayos bioquímicos hasta ensayos de fermentación a escala piloto. Las cepas enológicas comerciales hoy disponibles son el resultado de trabajos de selección a partir de cepas nativas. La selección de cepas implica elegir entre aquellas cuyo comportamiento fisiológico se adapte mejor a las propiedades enológicas requeridas, destacándose entre ellas: buen rendimiento en la

transformación de los azúcares del mosto en etanol, buena velocidad de fermentación, tolerancia al etanol, ausencia de producción de sulfuro de hidrógeno (H₂S). La producción de H₂S es una característica sensorial indeseable en las industrias del vino, la cerveza y el sake. La formación de este compuesto es consecuencia de su posición como producto final de la cadena de reducción del sulfato pero también por ser el producto metabólico intermedio en la síntesis de los aminoácidos azufrados.

Hoy en día es práctica usual la inoculación de cultivos puros de cepas seleccionadas para asegurar un comienzo apropiado de la fermentación y un proceso de fermentación reproducible. También se tiende a preparar estos inóculos a partir de cepas seleccionadas de la misma región de donde procede la uva, o sea, se tiende a utilizar cepas adaptadas a la ecología y características del mosto. En Uruguay se han seleccionado cepas nativas que, a diferencia de las cepas comerciales disponibles, son capaces de impartir aromas deseables a un vino en condiciones de baja demanda de nitrógeno (González Techera et al., 2001). De ahí la necesidad de poder diferenciar con certeza a nivel de subespecie aquellas cepas seleccionadas empleadas para preparar inóculos puros.

1.2. Diversidad genética de *S. cerevisiae*

Las técnicas convencionales empleadas para la caracterización de levaduras consisten en tests morfológicos y bioquímicos que además de ser complejos, laboriosos y consumir mucho tiempo, tienen un uso muy limitado y en la mayoría de los casos no permiten la identificación de las subespecies que se encuentran presentes en el mosto. Por esta razón se han desarrollado y aplicado diversos métodos moleculares como por ejemplo el análisis de cariotipo por electroforesis en gel de campo pulsado, análisis de restricción del ADN mitocondrial y el análisis de secuencias de ADN repetitivas interdelta (Schuller et al, 2004) para diferenciar cepas de *S. cerevisiae*. Estos métodos superan las limitaciones existentes con los métodos convencionales, aunque el grado de polimorfismo, la rapidez de la técnica y el poder de discriminación varían según el método empleado y en muchos casos se requiere de una combinación de técnicas para lograr el grado de discriminación buscado.

La situación ideal sería contar con un único método molecular que permitiera diferenciar todas las cepas en estudio. Un buen marcador molecular debe presentar

una buena distribución a lo largo del genoma, alto grado de polimorfismo, y la técnica para analizarlo debe ser simple, rápida y reproducible en diferentes laboratorios (Cheng & Crittenden 1994). En el 2001, nuestro grupo de trabajo propuso diferenciar cepas nativas de *S. cerevisiae* de interés enológico mediante el uso de los marcadores moleculares denominados microsatélites (González Techera et al., 2001). Esta técnica molecular es el método de elección para diferenciar individuos dentro de una misma especie eucariota debido a que los microsatélites presentan un alto grado de polimorfismo. Han sido muy utilizados en humanos para estudios de paternidad (Helminen et al., 1988), en medicina forense (Hagelberg et al., 1991) y en otros organismos como método de tipificación molecular a nivel de subespecie (Bretagne et al., 1997; Bowers et al., 1999).

1.2.1. Microsatélites o SSRs (Simple Sequence Repeats)

Los microsatélites o SSRs (del inglés "Simple Sequence Repeats") son regiones de ADN con secuencias cortas de dos a seis pares de bases repetidas en tandem cierto número de veces, como por ejemplo, (AT)_n o (CTT) (Charlesworth et al, 1994). En función de su pureza, entendiéndose como las repeticiones sin interrumpir de un mismo motivo, los SSR pueden ser:

- a) "perfectos", cuando consisten de un solo tipo de repetición, ejemplo (AT)₁₅.
- b) "imperfectos", cuando la región microsatélite presenta interrupciones de nucleótidos que no se ajustan al patrón de repetición, por ejemplo (CT)₄CTA(CT)₆.
- c) "compuestos", cuando se presentan adyacentes dos o más regiones microsatélites de diferente patrón de repetición, por ejemplo (CT)₅(GT)₁₀ (Jarne & Lagoda, 1996).

En estos marcadores moleculares, la diferencia entre individuos está dada por la variación en el número de motivos de repetición para un locus de microsatélite dado. El número de motivos repetidos para un locus dado, en individuos de una misma especie, es hipervariable debido a que la probabilidad de que sucedan errores durante la replicación del ADN en esas zonas repetidas es alta. Los microsatélites presentan una tasa de mutación del orden de 10^{-5} por generación (Strand *et al.*, 1993). Los patrones de mutación observados parecen corresponderse bien con un modelo de pequeños cambios en las unidades de repetición, vía un mecanismo de deslizamiento

de las hebras ("strand-slippage") sobre la secuencia repetida durante el proceso de extensión de la replicación del ADN (Levinson & Gutman, 1987; Strand *et al.*, 1993; Schlötterer & Tautz, 1992). El fenómeno del deslizamiento provoca la formación de un bucle del tamaño de uno o varios motivos de repetición, respetando la complementariedad de las bases, en la hebra molde o en la sintetizada. Se altera así, durante la síntesis de la hebra en formación, el número total de secuencias de repetición, obteniéndose una hebra más larga si el bucle se formó en la hebra sintetizada y una más corta si se formó en la hebra molde (ver Fig. 1).

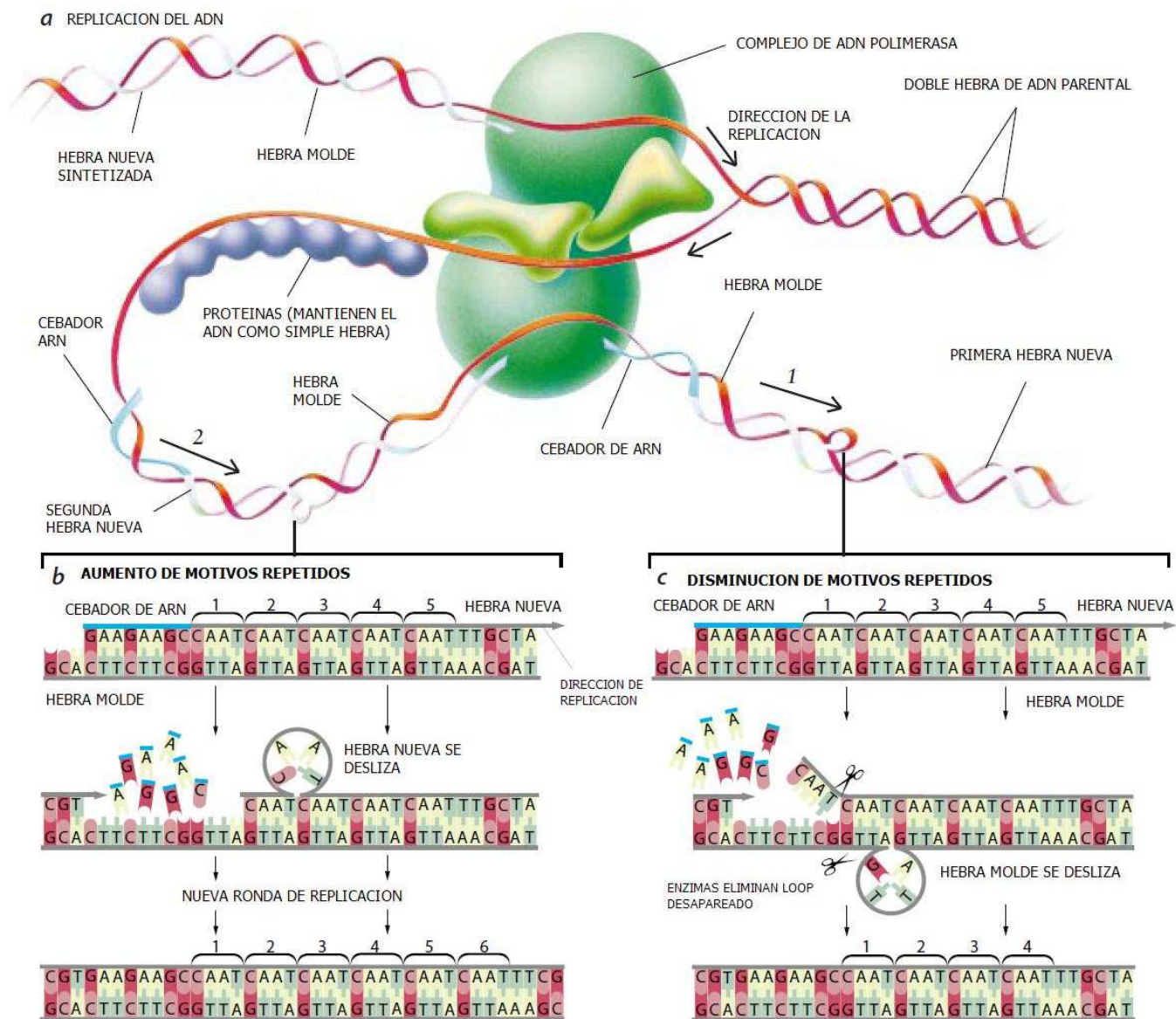


Figura 1. Modelo propuesto de pequeños cambios en las unidades de repetición de los microsatélites mediante un mecanismo de deslizamiento de las hebras ("strand slippage") sobre la secuencia repetida durante el proceso de extensión de la replicación del ADN (Moxon & Wills, 1999).

El análisis molecular por microsatélites consiste en determinar los tamaños de las moléculas de ADN obtenidas mediante la técnica de la reacción en cadena de polimerasa (PCR, "Polymerase Chain Reaction") desarrollada por Mullis *et al.* (1986). Esta técnica se basa en una reacción enzimática catalizada por una ADN polimerasa, que permite la amplificación «*in vitro*» de grandes cantidades de una región particular del ADN, a partir de una pequeña cantidad original de ADN molde. Esta enzima requiere, para la síntesis, de un par de oligos denominados «*primers*» o cebadores, cuyas secuencias son complementarias a la de las regiones flanqueantes 5' y 3' del segmento particular de ADN que se aspira amplificar. Las secuencias de ADN que contienen microsatélites están flanqueadas por regiones conservadas a partir de las cuales se pueden diseñar cebadores específicos para amplificar el locus en estudio. Normalmente al amplificar un microsatélite no se obtiene una única banda por alelo sino un cierto patrón de bandas característico; a la banda principal le acompañan otras bandas secundarias conocidas como bandas extras (denominadas "stuttering"), que son el resultado del fenómeno conocido como *slippage* (deslizamiento) de la *Taq* DNA polimerasa que ocurre cuando se amplifica ADN repetitivo. Además, durante la PCR se producen otros dos artefactos: la adición por parte de la *Taq* DNA polimerasa de una adenina en el extremo 3' en un cierto porcentaje de moléculas y la amplificación diferencial que favorece a los alelos más cortos, factores éstos que también contribuyen a dificultar la estimación correcta de los alelos presentes. La ocurrencia de estos artefactos dificulta la lectura de los geles. Alargando el tiempo de elongación final de la PCR, se logra que la *Taq* DNA polimerasa adicione adenina en los extremos 3' del 100% de las moléculas. Dependiendo del tamaño del producto de PCR y de su secuencia en las condiciones desnaturalizantes de la corrida electroforética también es posible que en algunos casos se separen las dos hebras de ADN y se visualicen como bandas diferentes.

En cada genotipo la amplificación de loci específicos revela la presencia de uno o más alelos por locus, según el grado de ploidía de la especie y según sea el individuo heterocigota u homocigota. La identificación del individuo (huella digital o "fingerprint") queda descrita por el conjunto total de alelos obtenido a partir del estudio de un número significativo de microsatélites polimórficos, llamado patrón alélico. La comparación de patrones alélicos mediante coeficientes de similitud permite establecer grupos de individuos con características génicas similares. En lo que respecta a las ventajas se puede señalar que los microsatélites son muy informativos,

presentan herencia codominante. Es una técnica simple y muy sensible pues requiere de poca cantidad de ADN y no necesariamente de alta calidad.

1.2.2. SNPs (Single Nucleotide Polymorphism)

La diversidad molecular entre cepas de *S. cerevisiae* también puede deberse a cambios puntuales o polimorfismos de un solo nucleótido (SNP, Single Nucleotide Polymorphisms). Este tipo de polimorfismo corresponde a cambios puntuales, por ejemplo sustituciones, deleciones o inserciones de un solo nucleótido. Aunque muchos de los SNPs se localizan en regiones no codificantes, un número importante de estas mutaciones están presentes en sitios codificantes y en humanos se han asociado a enfermedades o cambios en la expresión fenotípica. Su alta frecuencia en el genoma y su baja tasa de mutación, hacen que se constituyan como elementos deseables para la construcción de mapas genéticos. A partir del análisis estadístico de los patrones y frecuencias de SNPs en individuos de una misma especie es posible en ciertos casos asociar uno o varios SNPs a una característica fenotípica.

1.3. Genoma de *Saccharomyces cerevisiae*.

En 1996, el genoma de *Saccharomyces cerevisiae* fue secuenciado completamente (Goffeau et al. 1996). La célula haploide tiene un genoma total de 13.5 Mb, contenido en 16 cromosomas cuyos tamaños van de 230 a 2,350 Kb y presentan un total de 6,368 marcos de lectura abiertos (ORF, Open ReadinG Frame). A diferencia de los genomas de organismos multicelulares, el genoma de la levadura es muy compacto dado que las regiones intergénicas son pequeñas y únicamente un 4% de los genes codificantes poseen intrones. Aproximadamente el 70% del ADN cromosómico corresponde a secuencias codificantes, encontrándose en promedio un gen por cada 2 Kb de genoma.

El conocimiento del genoma completo de *S. cerevisiae* (<http://www.yeastgenome.org/>) ha permitido detectar la presencia de microsatélites (Field & Wills, 1998; Richard et al., 1999, Katti et al., 2001) y ha sido posible diseñar oligonucleótidos en las regiones conservadas que flanquean los repetidos de secuencia simple. En la página Web <http://www.ncl-india.org/ssr/> se encuentra toda la información de los microsatélites encontrados en el genoma de *S. cerevisiae*. Los

microsatélites se encuentran distribuidos a lo largo de todo el genoma de *S. cerevisiae* tanto en regiones codificantes como no codificantes. Mediante el uso de un programa de búsqueda bioinformática se encontraron 383 repetidos de dinucleótidos y 604 repetidos de trinucleótidos en el genoma. De estos últimos, hay 362 trinucleótidos que se encuentran dentro de marcos de lectura abiertos (Katti et al, 2001). A diferencia de lo observado en otros genomas eucariotas completamente secuenciados, en *S. cerevisiae* la mayoría de los repetidos son trinucleótidos (Katti et al, 2001). Las repeticiones de trinucleótidos son especialmente abundantes en genes reguladores relacionados con la transcripción y transducción de señales, mientras que son infrecuentes en genes que codifican proteínas estructurales. Estas observaciones apoyan el posible rol de este tipo de microsatélites en la regulación de la transcripción génica (Young et al, 2000).

1.4. Importancia del regulador *FLO8*

La capacidad de flocular es una característica deseada en las cepas de uso industrial. La floculación es un proceso reversible en el cual las células se adhieren formando estructuras multicelulares (denominadas flóculos) que pueden apreciarse a simple vista y que sedimentan rápidamente en cultivo líquido. En la levadura *S. cerevisiae*, este proceso implica la existencia de unas proteínas tipo lectinas que se denominan floculinas. Estas proteínas se localizan en la pared de las células y se unen selectivamente a los residuos manosa presentes en las paredes de células adyacentes, siendo necesaria la presencia de iones calcio para que tenga lugar dicha interacción (Verstrepen et al., 2003). Las floculinas son proteínas codificadas por los denominados genes *FLO*. *S. cerevisiae* tiene cinco miembros conocidos de la familia génica *FLO*: *FLO1*, *FLO5*, *FLO9*, *FLO10* y *FLO11*. Los cuatro primeros genes comparten una elevada identidad de secuencia y además se localizan en regiones subteloméricas del genoma (Teunissen et al., 1993; Bidard et al., 1994; Teunissen and Steensma, 1995). En la mayoría de las cepas de levadura de laboratorio los genes *FLO1*, *5*, *9* y *10* están silenciados. Cuando se expresan ectópicamente dan lugar a diferentes fenotipos adhesivos, como son floculación o crecimiento en forma de pseudohifas (Guo et al., 2000). De esta familia el gen *FLO11* es el único que se localiza cerca del centrómero y la regulación de su expresión es compleja e integra, a través de diversas vías de señalización, múltiples señales. La proteína Flo11p promueve la adhesión célula-célula y célula-superficie y es necesaria para que las células se mantengan adheridas

una a otras en un crecimiento filamentoso y también para que se adhieran e invadan el agar (Lo & Dranginis, 1996; Guo et al., 2000). Flo8p regula la expresión de muchas proteínas entre las cuales se encuentran las floculinas FLO1, FLO9 y FLO11 (Kobayashi et al., 1999).

FLO8 (YER109c, cromosoma V) codifica para un factor de transcripción de 798 aminoácidos necesario para la floculación, el crecimiento filamentoso de diploides y el crecimiento invasivo de haploides. Estudios realizados con la cepa secuenciada S288C muestran que la misma presenta una mutación en el gen *flo8* (Liu et al, 1996) que da lugar a un codón stop y por lo tanto no expresa una proteína funcional. A nivel internacional se han presentado varias patentes en relación a *FLO8* y su importancia en la formación de biofilms, metabolismo secundario, floculación en procesos industriales, etc. (ver www.freepatentsonline.com). Hasta el momento no se ha reportado una búsqueda de polimorfismos presentes en *FLO8* en diferentes cepas de *S. cerevisiae*. La asociación de un determinado polimorfismo en *FLO8* con una mayor capacidad de floculación podría ser una estrategia útil a la hora de seleccionar cepas industriales, por todo esto planteamos estudiar la presencia de SNPs en el gen regulador *Flo8*.

2. OBJETIVOS.

2.1. Objetivo general.

En este trabajo se plantea desarrollar un método de tipificación molecular de cepas de *S. cerevisiae*.

2.2. Objetivos específicos.

-Encontrar el conjunto mínimo de loci de microsatélites polimórficos que nos permita diferenciar todas las subespecies de *S. cerevisiae* nativas uruguayas *entre sí* y distinguirlas de todas las cepas comerciales presentes en nuestra colección.

-Analizar la presencia de SNPs en un gen regulador clave en la capacidad de floculación de las cepas industriales.

3. MATERIALES Y MÉTODOS.

3.1. Cepas de *Saccharomyces cerevisiae* y medios de cultivo.

Se trabajó con una colección de 120 cepas de *S. cerevisiae* (ver detalle de cepas en el Anexo I) constituida por 48 cepas nativas aisladas de uva o mosto [15 cepas uruguayas (URU), 28 cepas argentinas (ARG) y 5 de la colección de UCDavis (UCD)], 50 cepas comerciales de vino (COM), 12 cepas productoras de H₂S (H₂S), 1 cepa comercial de Sake (SAKE), 6 cepas comerciales de panificación (PAN) y 3 cepas de laboratorio (LAB). Como cepas controles se utilizaron las 3 cepas de laboratorio: AB972 y S288c (haploides), ambas utilizadas en el proyecto de secuenciación del genoma total de *S. cerevisiae*; BY4743, diploide, utilizada en el proyecto de delección de cada uno de los marcos de lectura abiertos del genoma de *S. cerevisiae*.

Los cultivos se realizaron en medio YPD (10g/L de extracto de levadura, 20g/L de peptona y 20g/L de glucosa) a una temperatura de incubación de 30°C con o sin el agregado de agar (20 g/L). Para mantenimiento a largo plazo se guardaron cultivos líquidos de cada cepa en glicerol (c.f. 20%) a -80 C, tal como recomienda Sherman, 1991.

3.2. Extracción de ADN genómico.

Para obtener el ADN genómico de las cepas de *S. cerevisiae* se realizó extracción de ADN en duplicado (a partir de cultivos líquidos iniciados de una colonia aislada o de mogollón) para cada cepa por un método rápido tal como se describe en González-Techera et al., 2001. El pellet correspondiente a 10⁹ células (comienzo fase estacionaria) fue lavado con agua estéril y resuspendido en 0,4 mL de buffer breaking (2% Triton X-100, 1% dodecilsulfato de sodio, 10 mM NaCl, 10 mM Tris, y 1 mM EDTA, pH 8). Las células fueron homogeneizadas por agitación a alta velocidad durante 3 minutos con 0,3 g de perlas de vidrio (Sigma G9268) en presencia de 0,4 mL de fenol pH 8. Luego se agregó 0.4 mL de TE (10 mM Tris pH 8 y 1 mM EDTA pH 8) y se agitó brevemente. Después de centrifugar a 4°C, la fase acuosa fue removida cuidadosamente. El ADN fue precipitado con etanol, centrifugado y resuspendido en TE. La estimación de la cantidad y calidad del ADN obtenido se evaluó en geles de agarosa 0.7 % teñidos con bromuro de etidio. Los ADNs se guardaron como stocks

concentrados (aproximadamente 100-400 ng/μl) y también diluidos a concentración de uso para PCR (aproximadamente 2-4 ng/μl) en freezer de -20°C.

3.3. Diseño de cebadores

Para el diseño de los oligonucleótidos se utilizó el programa informático Primer3 (Rozen and Skaletsky, 2000). El diseño de los cebadores que fueron utilizados en el análisis de microsatélites se realizó teniendo en cuenta las regiones conservadas flanqueantes de cada locus de SSR. El programa Primer3 considera diferentes variables a tener en cuenta, por ejemplo, la temperatura de "annealing", el tamaño del producto de amplificación y la máxima complementariedad en el extremo 3' del cebador. Para la máxima complementariedad en el extremo 3' se le exigió al programa el diseño de cebadores que tuvieran una calificación de 0,00, esto significa que no sucede alineamiento entre los dos cebadores en dicho extremo. Además, la presencia en el extremo 3' del primer de bases G y C ayuda a una unión específica por lo que este criterio se tuvo en cuenta en el diseño pero teniendo en cuenta que en las últimas 5 bases del extremo 3' no hubiera más de 3 bases G o C.

El trabajo a gran escala requiere de la optimización de todas las etapas involucradas para facilitar el trabajo experimental. El diseño de los cebadores para el análisis de SSRs se realizó de manera que todos tuvieran la misma temperatura de "annealing" y que los productos de amplificación se acotaran a un rango determinado de tamaño. La temperatura de "annealing" elegida para todas las reacciones de PCR fue de 55°C y los tamaños de los productos de PCR se eligieron entre 200-250 pb. Para los loci TTA_XIII y YOR267C se diseñaron nuevos cebadores (diferentes a los reportados previamente) para mejorar los productos de amplificación, buscando lograr productos de menor tamaño y con poco "stuttering".

El programa informático Primer3 también fue utilizado para diseñar los cebadores usados en el análisis de SNPs. Se diseñaron un par de primers (FLO8 prom fwd y FLO8 prom rev) para amplificar una región de 600 pb del promotor de FLO8 y se diseñaron los primers Flo8-fwd y Flo8-rev para el análisis de SNPs de una región de 300 pb del ORF de FLO8.

3.4. Amplificación de microsatélites.

La mezcla de reacción fue de 20 µl finales conteniendo: 10-20 ng de ADN genómico, 1X buffer PCR sin Mg^{+2} , 2,5 mM de $MgCl_2$, 200 µM de cada dNTP, 0,5 µM de cada oligonucleótido Forward y Reverse y 1 U de *Taq* ADN polimerasa. El programa utilizado consistió en: 5 minutos a 94 °C; 30 ciclos de (30 segundos a 94 °C, 30 segundos a 55 °C y 1 minuto a 72 °C); terminando con 5 minutos a 72 °C.

Para disminuir el excesivo stuttering, los loci de SSR: TG_VI, GT_X y C4_XV fueron amplificados con una enzima fiel. La mezcla de reacción fue de 20 µl finales conteniendo: 10-20 ng de ADN genómico, 1X buffer PCR sin Mg^{+2} , 4 mM de $MgSO_4$, 200 µM de cada dNTP, 0,5 µM de cada oligonucleótido Forward y Reverse y 0,5 U de *Pfu* ADN polimerasa (Fermentas). El programa utilizado consistió en: 2 minutos a 95 °C; 25 ciclos de (30 segundos a 95 °C, 30 segundos a 55 °C y 1 minuto a 70 °C); terminando con 5 minutos a 72 °C.

Para cada uno de los pares de oligonucleótidos diseñados se optimizaron las condiciones de amplificación de manera de obtener un producto único con la cepa haploide secuenciada de referencia. Los productos de amplificación se evaluaron en geles de agarosa 2% teñidos con bromuro de etidio para confirmar tamaño y para determinar la concentración apropiada a utilizar en el análisis en geles de secuencia.

La amplificación de los ADNs en gran escala se realizó en cajas de PCR de 96 pocillos y utilizando pipetas automáticas multicanal.

3.5. Electroforesis en geles de secuencia.

3.5.1. Preparación de muestras

Los productos de PCR se diluyeron a 1/3 en solución de corrida (10mM NaOH, 95% formamida, 0,005% azul de bromofenol, 0,05% Xylen cyanol) de manera que la cantidad de producto de PCR que se carga en cada pocillo del gel sea aproximadamente de 2-10 ng y esté presente en un volumen final de 3-4 µL. La mezcla se desnaturalizó a 94 °C por 4 minutos e inmediatamente se colocó en hielo. El gel se cargó lo más rápidamente posible para evitar que bajara la temperatura.

3.5.2. Tratamiento y ensamblaje de vidrios.

Los vidrios fueron lavados por ambos lados con detergente neutro diluido y abundante agua común. Se enjuagaron con agua de osmosis inversa, luego agua ultrapura tipo Milli-Q y se dejaron secar a temperatura ambiente.

El gel de secuencia de 0,4 mm de espesor debe quedar adherido al vidrio más corto, de manera de poder realizar todas las etapas de teñido con plata. Para ello se trata el vidrio corto con una solución que asegura el pegado del gel al mismo y también se trata el vidrio largo con otra solución que asegura que el gel se desprenda del mismo.

- Preparación del vidrio corto (en campana de gases):

Se limpió 3 veces la cara donde irá adherido el gel, con 2 mL de etanol 95% cada vez y con papel suave (Kleenex). Se mezcló la solución de 0,5% ácido acético en etanol 95% con *Silane* y se repartió la mezcla por la cara del vidrio en la que se adherirá el gel, utilizando un papel suave y movimientos que aseguren una distribución pareja por todo el vidrio.

Para el gel con peine de 24 pocillos, tamaño vidrio corto: 33,5 cm x 19,5 cm se preparó una mezcla de 0,44 mL (1,32 µL de *Silane* en 0,44 mL de solución de 0,5% ácido acético en etanol 95%).

Para el gel con peine de 62 pocillos, tamaño vidrio corto: 39 cm x 33 cm se preparó una mezcla de 1 mL (2 µL de *Silane* en 1 mL de solución de 0,5% ácido acético en etanol 95%).

Se dejó secar durante 10 minutos y se removió el exceso lavando 3 veces con 2 mL de etanol 95% cada vez.

- Preparación del vidrio largo (en campana de gases):

Se limpió 3 veces la cara donde será tratado el vidrio con 2 mL de etanol 95% cada vez y con papel suave. Se repartió la cantidad necesaria de *Sigmacote* sobre el vidrio utilizando un papel suave, de manera de obtener una distribución pareja por todo el vidrio.

Para el gel con peine de 24 pocillos, tamaño vidrio largo: 36 cm x 19,5 cm se aplicó 1,4 mL de *Sigmacote*.

Para el gel con peine de 62 pocillos, tamaño vidrio largo: 42 cm x 33 cm se aplicó 2,7 mL de *Sigmacote*.

Se dejó secar durante 10 minutos y se retiró el exceso con papel suave.

- Ensamblaje de los vidrios:

Se colocó en la mesada el vidrio largo con la cara tratada con Sigmacote hacia arriba. Se limpiaron los separadores con etanol 95% y se colocaron en los bordes laterales del vidrio largo. Se colocó el vidrio corto (cara tratada con silano hacia abajo) sobre el vidrio largo. Cuidando que los vidrios coincidan en su parte inferior (quedando al mismo nivel) se sellaron los lados y el borde inferior de los vidrios con una sola tira de cinta 3M amarilla de buena calidad. Los vidrios se colocaron en el soporte de goma que viene con el equipo Life Technologies Modelo SA o Modelo S2.

3.5.3. Preparación de solución de acrilamida para gel de pequeña escala (peine de 24 pocillos)

En un vaso de plástico Nalgene de 100 mL con pico vertedor se agregó 6,25 mL de solución stock poliacrilamida al 40% (Page Plus Concentrate, Amresco), 10 mL de TBE 5x (0,09 M Tris-borato, 0,002 M EDTA pH 8,5), 21 g de urea y 10 mL de agua ultrapura de tipo Milli-Q. Se disolvió la urea con agitador magnético y luego se llevó hasta volumen final de 50 mL. Se agregó 300 µL de APS 10% y 30 µL de TEMED agitando suavemente.

3.5.4. Preparación de solución de acrilamida para gel de gran escala (peine de 62 pocillos)

En un vaso de plástico Nalgene de 100 mL con pico vertedor se agregó 10 mL de solución stock poliacrilamida al 40%, 16 mL de TBE 5x, 33,6 g de urea y 16 mL de agua ultrapura de tipo Milli-Q. Se disolvió la urea con agitador magnético y luego se llevó hasta volumen final de 80 mL. Se agregó 400 µL de APS 10% y 40 µL de TEMED agitando suavemente.

3.5.5. Armado de gel desnaturalizante de poliacrilamida al 5% y 7M urea.

La solución de poliacrilamida al 5% y 7M urea se volcó entre los dos vidrios previamente ensamblados manteniendo un flujo constante y lento para impedir que se formen burbujas. Luego se insertó entre los dos vidrios el peine de manera invertida en la parte superior y se ajustó con ganchos metálicos. Esta zona se cubrió con Rolopac y se dejó polimerizar toda la noche. La polimerización se verificó en el resto de líquido que quedó en el vaso.

La cara exterior de los vidrios se limpió con agua y papel suave y los vidrios se colocaron adecuadamente en la cámara de electroforesis. Se llenó con buffer de

corrida 0,5x TBE y se verificó que no existieran fugas. Se retiró el peine y el espacio que estaba ocupado por este se enjuagó con 0,5x TBE a presión con una jeringa de manera de retirar la poliacrilamida sobrante. Se colocó el peine con los dientes hacia abajo con mucho cuidado hasta que los dientes se introdujeron ligeramente en el gel (apenas un par de mm). Se pegó un termómetro en la cara exterior del vidrio (en el centro) y se conectó el equipo de electroforesis a una fuente de poder de alto voltaje a una potencia constante de 55 watt (aproximadamente 30 mA y 1750 voltios) para precalentar el gel hasta alcanzar los 50°C (aproximadamente 1 hora).

3.5.6. Condiciones de electroforesis

Antes de cargar el gel con las muestras, se limpiaron los pocillos con buffer 0,5x TBE con una jeringa a presión para eliminar la urea. Luego de cargar las muestras, la corrida se realizó durante 10 minutos a 60 watt para subir rápidamente la temperatura a 50°C y luego se bajó a 55 watt, manteniendo potencia constante. La electroforesis se corre el tiempo necesario según el tamaño de los ADNs que se están separando. Para los productos de PCR de 200 a 300 pb el tiempo de corrida fue de 1 hora y 30 minutos. Las medidas fueron tomadas desde la banda hasta el borde inferior del gel. El Xylene Cyanol corre a 4,5 cm del borde inferior del gel. En todos los geles se incluyó el marcador de peso molecular de 10 pares de bases de Invitrogen (100 ng por pocillo). La banda de 200 pb del marcador corre hasta 6,5 cm del borde inferior del gel y la banda de 300 pb corre hasta 13,5 cm de borde inferior del gel.

En las pruebas preliminares de pequeña escala, cuando se estaba evaluando el grado de polimorfismo de distintos SSRs, en algunos casos un mismo gel se cargó dos o tres veces consecutivas con productos de PCR de tamaños diferentes para ahorrar tiempo y dinero.

Cuando finalizó la corrida, se desmontaron los vidrios y se colocaron sobre la mesada de manera que el vidrio corto quedara hacia abajo. Se removió el vidrio largo cuidadosamente. Para ello, se retiró uno de los separadores y se insertó entre los vidrios una punta de metal haciendo presión hacia arriba para que el vidrio largo se separara del vidrio corto que tiene pegado el gel.

3.6. Visualización de microsatélites por tinción con nitrato de plata.

La visualización de las bandas del ADN amplificado se obtuvo por tinción con nitrato de plata, basado en el protocolo del Kit Promega Silver Sequence Staining Reagents. La tinción con AgNO₃ consta de varias etapas cuyos tiempos de incubación son críticos para la obtención de buenos resultados. La preparación de las soluciones para la tinción y los pasos de lavado de los geles requieren de agua de excelente calidad, se utilizó agua ultrapura de tipo Milli-Q. El procedimiento de tinción consistió en sumergir el gel en diferentes soluciones en el siguiente orden, agitando suavemente en agitador orbital:

- fijación con ácido acético 10% por 20 minutos
- 3 lavados con agua Milli-Q durante 2 minutos cada vez
- tinción con solución de nitrato de plata 0,1 % y formaldehído 0,05 % por 30 minutos
- un lavado con agua Milli-Q por 5-10 segundos (este tiempo es crítico)
- revelado con carbonato de sodio 3%, formaldehído 0,05% y tiosulfato de sodio 0,02% (preparado el mismo día de uso y previamente mantenido a 4°C) por 2-5 minutos hasta que aparezcan las bandas
- stop con ácido acético 10% por 5 minutos y
- lavado final con agua Milli-Q por 3 minutos.

Estas soluciones (excepto la solución de revelado) pueden ser reutilizadas un máximo de 3-5 veces. El gel se secó al aire por 24 horas y se obtuvo una imagen por scanner o foto digital en formato JPEG.

Una vez obtenida la foto del gel para cada locus analizado, se procedió al estudio de los alelos obtenidos. Se determinaron rigurosamente las relaciones de tamaño entre cada genotipo, así como entre los alelos de un mismo genotipo.

3.7. Construcción de escaleras de alelos para SSR polimórficos.

Las escaleras de alelos se construyeron mezclando diferentes productos de PCR que fueron obtenidos a partir de varias cepas que representan a los diferentes alelos observados para cada SSR particular (ver tablas en el anexo II). La mezcla de ADNs fue precipitada con etanol, centrifugada y el pellet resuspendido en TE. Se diluyeron a 1/3 en solución de corrida (10mM NaOH, 95% formamida, 0,005% azul de

bromofenol, 0,05% Xylen cyanol) de manera de obtener aproximadamente de 2-5 ng de cada alelo mezclado en el pool y que estuviera presente en un volumen final de 4 μ L.

3.8. Amplificación para el análisis de SNPs en FLO8.

3.8.1. Amplificación de una región del ORF de FLO8.

Una región de 300 pb de YER109C se amplificó con los cebadores: Flo8-fwd 5'-CAGCAGCCTTTGCTCAAGATG-3' y Flo8-rev 5'-GTTCTGCATCGTGTGTAGCCTTG-3'. Los ADNs fueron amplificados en un volumen final de reacción de 20 μ L conteniendo 10-20 ng de DNA, 200 mM de cada dNTP, 2 μ l de 10X de PCR buffer menos Mg, 0.5U Pfu ADN polimerasa (Fermentas), 4mM MgSO₄ y 10 pmol de oligos forward y reverse. El programa utilizado consistió en: 2 minutos a 95 °C, 35 ciclos de (30 segundos a 95 °C, 30 segundos en 55 °C, 1 minuto a 70 °C) y 5 minutos a 72 °C. El producto de amplificación fue evaluado en gel de agarosa 1,5 % teñido con bromuro de etidio. Los productos de PCR fueron purificados y ambas hebras fueron secuenciadas en Macrogen (dna.macrogen.com).

3.8.2. Amplificación de una parte del promotor de FLO8.

Se diseñaron los siguientes primers: FLO8 prom fwd 5'-CCTTCTCCTGCACATTCTTG- 3' y FLO8 prom rev 5'-TGAATCTGGATACGAACTATTCACT-3' para amplificar una región de 600 pb del promotor de FLO8 (de 378174 a 377575 pb del Crom. V incluyendo las 36 primeras bases del ORF, ver anexo III). La mezcla de reacción fue de 20 μ l finales conteniendo: 10-20 ng de ADN genómico, 1X buffer PCR sin Mg, 2,5 mM de MgCl₂, 200 μ M de cada dNTP, 0,5 μ M de cada oligonucleótido Forward y Reverse y 1 U de *Taq* ADN polimerasa. El programa utilizado consistió en: 5 minutos a 95 °C; 30 ciclos de (1 minuto a 94 °C, 2 minutos a 52 °C y 1 minuto a 72 °C); terminando con 10 minutos a 72 °C. Una alícuota del producto de PCR fue evaluada en gel de agarosa 1,5 % teñido con bromuro de etidio. Los productos de PCR fueron purificados y ambas hebras fueron secuenciadas en Macrogen, Korea.

4. RESULTADOS.

4.1. Criterios utilizados para la selección de loci de SSRs.

A partir de la información que se encuentra publicada en la base de datos <http://www.ncl-india.org/ssr/> (Katti et al, 2001), se analizaron los 362 loci de repetidos de trinucleótidos presentes en marcos de lectura abiertos y 383 loci de dinucleótidos del genoma de *S. cerevisiae* para el diseño de cebadores. Los loci de SSRs fueron seleccionados de acuerdo a los siguientes criterios: un SSR por cada cromosoma, presencia de más de 7 repetidos perfectos, repetidos conteniendo C y/o Gs, localización del SSR dentro o cerca de genes de interés para la industria vitivinícola y loci de SSRs con antecedentes de alto grado de polimorfismo. En las tablas 1, 2 y 3 se detallan los primers diseñados para los 33 loci de SSRs seleccionados, algunos de ellos están localizados dentro o cerca de genes de interés. Los genes de interés se definieron teniendo en cuenta las características enológicas que se consideran para la selección de cepas de interés enológico. Los loci YER109C/Flo8 e YHR211W/Flo5 están ubicados dentro de genes vinculados a la floculación (sedimentación celular) (Teunissen et al., 1995). Algunos loci de SSRs (YGL 013C, YJL162C, TCA_XII, CTA_XIII, CAG_XIV y GAA_GTA_XIV) se encuentran cerca de genes vinculados a la biosíntesis de ergosterol (ver tabla 1, 2 y 3) y relacionados a la tolerancia al etanol (Wu et al., 2006). Otros loci (AT_X, CAG_II y el TG_VI) se encuentran cerca de genes que participan en la asimilación de sulfato (vinculado a la indeseada producción de ácido sulfhídrico). El locus AT_X está ubicado a 1,5 Kb de YJR010W/Met3, el SSR CAG_II a 1Kb de YBR213W/Met8 y el microsatélite TG_VI se encuentra a 3 Kb del gen YFR030W/Met10 que codifica la subunidad alfa de la reductasa sulfito, que es responsable de la conversión de sulfito a sulfuro (Hansen y Kielland Brandt, 1996). El locus YOR373W/NUD1 se encuentra a 7 Kb de ATF1 una alcohol acetil transferasa relacionada a la formación de ésteres (Verstrepen et al., 2003) y el SSR YPL026C/SKS1 se encuentra en una quinasa involucrada a la adaptación a bajas concentraciones de glucosa (Vagnoli & Bisson, 1998).

Se incluyeron dos SSRs (C4_XV y GT_X) que fueron propuestos como muy polimórficos por Legras y colaboradores (2005) y tres loci (TTA_XIII, YOR267C e YPL009C) previamente propuestos en el 2001 por nuestro grupo (González-Techera et al., 2001). TTA_XIII y YOR267C también fueron confirmados como polimórficos por Legras y colaboradores (2005). Se seleccionaron tres loci (YDR289C, YKL172W y

YLR177W) que han sido utilizados para discriminar entre cepas empleadas como agentes probióticos de cepas clínicas (Hennequin et al., 2001; Malgoire et al., 2005).

Los loci de SSR fueron nombrados teniendo en cuenta el nombre del marco de lectura abierto (ORF) propuesto por www.yeastgenome.org, si el motivo repetido se encuentra en una región codificante. Si el motivo de repetición está ubicado en una región intergénica del cromosoma, se nombró con el motivo repetido y el número de cromosoma, por ejemplo TTA_XIII. En el caso de C4_XV que es un SSR formado por un motivo compuesto (con repeticiones de diferentes tripletes) ubicado en el cromosoma XV, se conservó el nombre propuesto por Legras y colaboradores (2005).

Tabla 1. Detalle de 21 pares de primers que fueron diseñados para analizar microsatélites de trinucleótidos ubicados en ORFs.

Locus	Crom. (hebra)	Motivo repetido	primer forward 5'–3'	Primer reverse 5' –3'	PCR (pb)	Ubicación/ (Referencia)
YAL035W/FUN12	I (W)	(GAA)17	CGCAAGAATCCACCGCAAGC	ACAACAAATTGGGGAACGAA	206	
YBR017C/KAP104	II (C)	(GAT)12	GGAAATGGAGAGGATGCAGA	TGGGAACGCTATATCCATCA	162	
YCR067C/SED4	III (C)	(TCT)12	CCACCACAGAAGATGAGCAA	ATTCAGTGGAAGCGATGGTT	236	
YDR289C/RTT103	IV (C)	(GAT)15	TATGAAGTGGGGGATGGAGA	CCGGTATGTCATTAGCGTCA	268	(3)
YER109C/ FLO8	V (C)	(CAG)10	GCAATGGCAACGAATAGTGA	GGGTTTTGGCCTTTGTCTCT	233	
YGL013C/ PDR1	VII (C)	(AAC)12	ATGCTCAGCAGCAAATACCC	GATAAACGTCGCTCCACAGG	228	Ubicado a 0,5 Kb del gen Erg4
YGL028C/SCW11	VII (C)	Complejo1	GGCAACACTCAAAGTCAGCA	GCTGTTTCCTGTGTGGATGA	245	
YHR143W/DSE2	VIII (W)	(TCT)9	ATCATCTGCCCCTCTTTCCT	GCATGGATGTCGTTGCTCTA	236	
YHR211W/FLO5	VIII (W)	(TCT)7	CATCCGTCATTCCAACCAGTA	AAGCAGTTTCCTGGCCTGT	174	
YIL156W/UBP7	IX (W)	(CAG)10	GGTGGGTACGAAAAATGGAA	GCGGAAAAAGAGTCCTGAGA	161	
YJL162C	X (C)	(GAA)7	GCCGGTAAAATGGATACGAA	GCTGCTGTTGCTGTTGTTGT	169	Ubicado a 8 Kb del gen Erg20
YKL172W/EBP2	XI (W)	(GAA)12	TCAAAGCCAAGCCCTTTCTA	GGCCAATTTCTCAAGATCCA	190	(3)

Complejo1: (SerSerSerThr)9

Continuación

Locus	Crom. (hebra)	Motivo repetido	primer forward 5'–3'	Primer reverse 5' –3'	PCR (pb)	(Referencias)
YLR177W	XII (W)	(CAG)15	TGATGAACACTTCGCCTAACA	CGGAAGGAGCAGTGTATGGT	258	
YMR135C/GID8	XIII (C)	(CAG)8	CTGGAGGTGCTGGAAGAGAC	GATGTTCTCCTGTTGATGGTGA	171	
YNL298W/CLA4	XIV (W)	(CAG)10	CCAAATACCAAATCGCCTGT	GGTCCTTGATGGTGGTATGG	275	
YOR373W/NUD1	XV (W)	(AAT)20	TCTGATGACAGCGATTCTGG	TGAGAGGACGACATGGATGA	260	
YOR267C	XV (C)	(CAA)20	ATGACTGCAGCAATGAATCG	TCCTCTGTGCTGTTGACTCG	239	(1), (2)
YOR382W/FIT2	XV (W)	(TCC)11	ACCAAGACCGTCACTCAAGC	ACTTGCTCCTTGGAATGCAG	117	
YOL138C	XV (C)	(GAT)14	ACGTCGTTAAAAACGCAACC	CTCGTCAAACCTCCCTCGAAC	230	
YPL009C/ <u>SC8132</u>	XVI (C)	Complejo 2	CTGCTCAACTTGTGATGGGTTTTGG	CCTCGTTACTATCGTCTTCATCTTGC	207	(1), (2)
YPL026C/SKS1	XVI (C)	(CAA)13	CGAGCGAAGTTTACATGACG	ATCACTCTCCGGTTCAGCAT	210	

Complejo 2: (GAA)7GAG(GAA)16

Las referencias son las siguientes: (1) Field & Wills, 1998; (2) Gonzalez Techera et al, 2001; (3) Hennequin et al, 2001; (4) Legras et al, 2005

Tabla 2. Detalle de los pares de primers que fueron diseñados para analizar siete microsatélites de trinucleótidos ubicados en regiones intergénicas.

Locus	Crom. (hebra)	Posición	Motivo repetido	primer forward 5'–3'	Primer reverse 5' –3'	PCR (pb)	Ubicación/ (Referencias)
CAG_II	II (W)	649297	(CAG)12	TATGTCCTCAAATGGCCTGA	CTGCCATTCTTGTTCTGTG	225	Ubicado a 1Kb del gen Met8
TCA_XII	XII (W)	252909	(TCA)11	CTCATTCGTCGGTGTAGCAG	GCTGGAAAACGACTCAAAGC	190	Ubicado a 1Kb del gen Erg3
CTA_XIII	XIII W)	303998	(CTA)7	ATAGGTACGCGGACTGATGG	GTGCCCTAGCAATAGCACAA	158	Ubicado a 1 Kb del gen Erg5
TTA_XIII	XIII (W)	86941	(TTA)35	AAGCGTAAGCAATGGTGTAGA	AAGCCTCTTCAAGCATGACC	238	(1), (2)
GAG_XIV	XIV (W)	701520	(GAG)8	CACGGGTTTCTCTAGCCTCA	GCTGGAATCATTGGAAGCAG	210	Ubicado a 0,4 Kb del gen Erg19
GAA_GTA_XIV	XIV (W)	106056	(GAA)7 (GTA)8	CCTGCGTAGAGGACGTTTCTG	GCACTCATCCTCTTCCTCGT	209	Ubicado a 3 Kb del gen Erg24
C4_XV	XV (W)	110775	Complejo *	GAGAAAAATGCTGTTTATTCTGACC	CCTCCGGGACGTGAAATAAC	229	(4)

Complejo*. (ATA)5 CTA (ATA)9 (GTA)7 ATA (ATG)7

Las referencias son las siguientes: (1) Field & Wills, 1998; (2) Gonzalez Techera et al, 2001; (3) Hennequin et al, 2001; (4) Legras et al, 2005.

Tabla 3. Detalle de los pares de primers que fueron diseñados para analizar cinco SSR de dinucleótidos ubicados en regiones intergénicas.

Locus	Crom. (hebra)	Posición	Motivo repetido	primer forward 5'–3'	Primer reverse 5' –3'	PCR (pb)	Ubicación/ (Referencias)
TG_VI	VI (W)	210332	(TG)37	GACACAATAGCAATGGCCTTCA	ATATTGGCACCAGAAAGTGTCG	198	Ubicado a 3 Kb del gen MET10 (4)
AG_VIII	VIII (W)	123922	(AG)16	GCAAGGGCAAACAGGATAGAC	AACCTCTGCGGAAATCTCAA	268	Ubicado a 2 Kb del gen Erg11 (4)
GT_X	X (¿)	518944	(GT)20	TGCGCAGCTTAGTATGACCA	GATGGGCTTTCACTCCACTT	274	
AT_X	X (W)	455748	(AT)19	GCAACCTCCAAAGGAAATCA	CGACAGAGAGAGACCCAAGC	206	Ubicado a 135 pb del gen Met3
AG_XII	(W)	757467	(AG)17	GGAAAAGAGATGGATAGCATACG	ATCCAAAACCCAGCATGAAG	211	

Crom. (hebra): Cromosoma y hebra en que se ubica el locus de SSR

Posición: ubicación en el cromosoma

PCR (pb): Tamaño del fragmento esperado en la cepa secuenciada (S288C o AB972)

Las referencias son las siguientes: (1) Field & Wills, 1998; (2) Gonzalez Techera et al, 2001; (3) Hennequin et al, 2001; (4) Legras et al, 2005.

Para los loci TG_VI, GT_X y C4_XV se obtuvieron perfiles confusos con muchas bandas extras. Para solucionarlo estos SSRs problemáticos se amplificaron con la polimerasa fiel *Pfu* que disminuye mucho el “stuttering”.

Para nuestro trabajo a gran escala fue fundamental aplicar buenas prácticas de laboratorio y contar con estrategias específicas para minimizar la posibilidad de contaminación (Lo and Chan, 2006). Una de las estrategias es la separación física de las diferentes etapas en la técnica de PCR. Se establecieron diferentes áreas, un área pre-PCR donde se encuentra una cabina (con luz ultravioleta) donde se realizaron los mix con los reactivos y con la enzima. Se estableció un área exclusiva para manipular las muestras de ADN y en donde se encuentra el termociclador. Por último, se dedicó un área post-PCR para la manipulación de los productos de PCR donde se contaba con el instrumental para la realización de las electroforesis. Se trabajó con material exclusivo para cada área, por ejemplo, se dedicaron pipetas automáticas, tips estériles, tubos de PCR estériles, guantes, gradillas y recipientes para el hielo para cada área. Se evitó el transporte de materiales de un área a otra, por ejemplo el material utilizado en el área de post-PCR nunca se ingresaba al sector de pre-PCR ni al área de manipulación de muestras de ADN. En todas las reacciones de PCR se realizaron los controles sin ADN dando resultados negativos.

4.2. Análisis de microsatélites

4.2.1. Ensayo de todos los SSRs con una población de referencia inicial.

Como estrategia de trabajo, en primera instancia se analizaron un total de 33 loci de microsatélites en una población de referencia inicial (subconjunto de cepas representativo de nuestra colección) formada por 23 cepas de diferentes orígenes y aplicaciones. Esta población de referencia está integrada por las cepas secuenciadas AB972 y S288C, por un subgrupo de cepas nativas uruguayas (CP 863, CP 873, CP 874, CP 881, CP 882, CP KU1, FQU 02/16, UY-04), por una cepa nativa argentina (ARG-4), por 10 cepas comerciales (AWRI 796, Levuline CHP, Montrachet 522, Enoferm QA23, Lalvin BM45, Lalvin ICV-D254, UCD VEN 529, UCD VEN 713, UCD VEN 781, UCD VEN 820) que fueron seleccionadas en diferentes orígenes geográficos y por 2 cepas con otro uso industrial (Panther de panificación y UCD VEN 612 de sake).

Los valores absolutos para los tamaños de los productos de amplificación para la cepa secuenciada AB972 y S288C de referencia (ver tablas 1, 2 y 3) se calcularon a partir de la información del genoma secuenciado de *S. cerevisiae* contenida en <http://www.yeastgenome.org/>. Se demostró que se obtiene el mismo producto de amplificación a partir de ADN de mogollón y de colonia aislada, se optimizaron todas las etapas de amplificación y separación en geles para luego continuar con el análisis en gran escala. Para la evaluación del grado de polimorfismo de los distintos SSRs, en algunos casos en un mismo gel se cargaron a diferentes tiempos varios loci de SSRs con productos de PCR de tamaños diferentes. En la figura 2 se observa la foto de un gel de poliacrilamida donde se corrieron tres loci de microsatélites que fueron analizados en las 23 cepas de referencia. En el mismo se observa la presencia de tres alelos para el locus YHR143W, cuatro alelos para YOL138C y en cambio para el locus YGL013C se observan 11 alelos y 7 cepas heterocigotas. En la tabla 4 se detallan el nombre del locus de SSR, el valor H_o (Heterocigocidad observada) para cada locus y el número total de alelos obtenidos en las 23 cepas de referencia para los 33 loci de SSRs analizados. Para pasar al análisis en gran escala se seleccionaron aquellos loci de microsatélites que presentaban un mayor grado de polimorfismo y mayor porcentaje de cepas heterocigotas dentro del subconjunto de referencia.

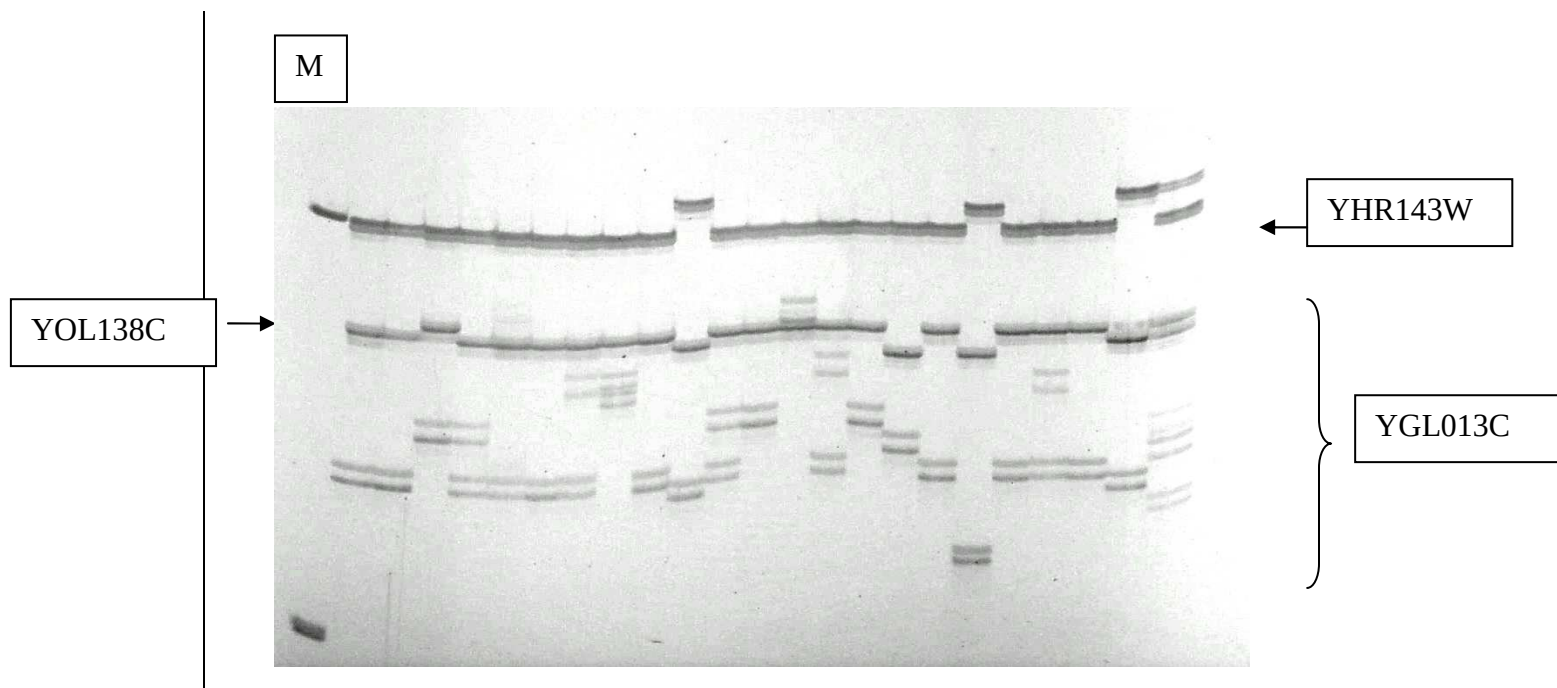


Figura 2. Análisis del grado de polimorfismo de tres loci de microsatélites: YHR143W, YOL138C e YGL013C en las 23 cepas de referencia. Foto de gel de poliacrilamida donde se sembraron sucesivamente las muestras correspondientes a los tres loci. En primer lugar se sembraron los productos de PCR correspondientes al locus YGL013C, luego YOL138C y por último YHR143W. En el primer pocillo (M) marcador de 100 pb (Invitrogen) que se sembró conjuntamente con las muestras correspondientes al locus YGL013C.

Tabla 4. Grado de polimorfismo de los 33 loci de SSRs ensayados en las 23 cepas de referencia

Locus	Nro. alelos	Ho
YHR211W/FLO5	1	0
YOR382W/FIT2	1	0
CTA_XIII	1	0
YBR017C_KAP104	2	0
TCA_XII	2	0
GAG_XIV	2	0.04
GAA_GTA_XIV	2	0.09
YDR289C/RTT103	3	0.04
YHR143W/DSE2	3	0.04
YER109C/FLO8	3	0.09
YNL298W/CLA4	3	0.09
AG_VIII	3	0.09
YKL172W/EBP2	3	0.13
YPL026C/SKS1	3	0.13
YIL156W/UBP7	4	0.04
YJL162C/JJJ2	4	0.04
AG_XII	4	0.04
YOL138C	4	0.04
YMR135C/SED4	4	0.09
CAG_II	4	0.09
YAL035W/FUN12	6	0.09
YLR177W	7	0.04
YOR373W/NUD1	7	0.17
YCR067C/SED4	8	0.17
YGL028C/SCW11	7	0.26
AT_X	7	0.30
C4_XV	9	0.13
GT_X	9	0.22
YGL013C/PDR1	11	0.30
TTA_XIII	13	0.26
YOR267C/HRK1	14	0.30
TG_VI	14	0.35
YPL009C	15	0.39

4.2.2. Ensayo en gran escala con los SSRs más polimórficos.

Se seleccionaron 9 loci de SSR (TTA_XIII, TG_VI, GT_X, C4_XV, YPL009C, YOR267C, YGL013C, YGL028C y AT_X) para ser amplificados en las 120 cepas de la colección. Este trabajo nos ha permitido adquirir experiencia en la técnica de PCR a gran escala y establecer un laboratorio con la capacidad de amplificar 96 muestras a la vez. Se obtuvieron muy buenos resultados gracias a la aplicación de las estrategias anteriormente mencionadas y a que se puso a punto tanto las concentraciones de ADN como las condiciones de PCR. En la figura 3 se muestra en un mismo gel la amplificación de las 96 muestras diferentes para un locus de SSR. Se utilizaron pipetas automáticas multicanal para cargar los geles de agarosa y de secuencia que se prepararon utilizando peines especialmente diseñados para respetar la distancia de las puntas de la pipeta multicanal. Resulta fundamental utilizar las puntas recomendadas para la pipeta multicanal para asegurar que los volúmenes pipeteados sean los mismos en las ocho puntas de la pipeta. Si se utilizan puntas inadecuadas los errores de pipeteo pueden resultar inadmisibles.

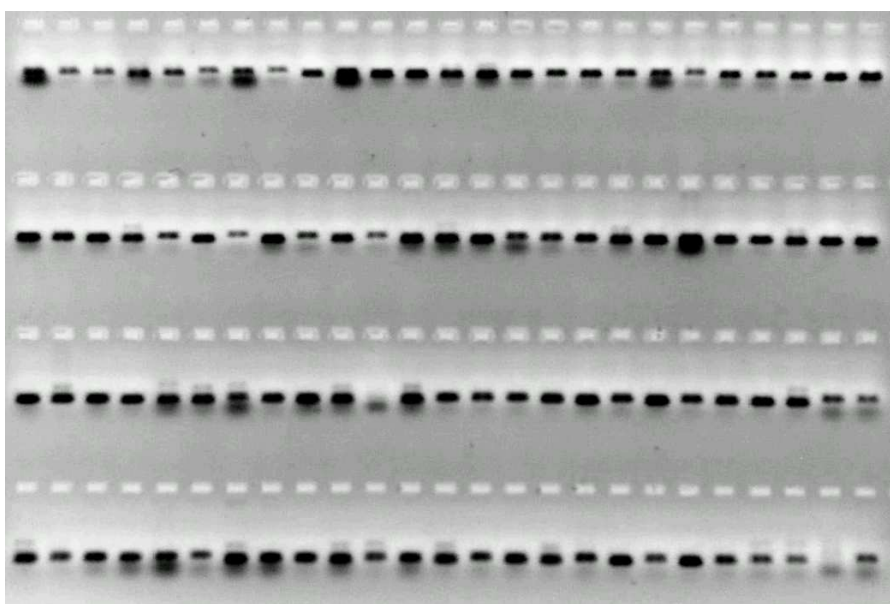


Figura 3. Productos de amplificación de un locus de SSR en 96 muestras. Foto de gel de agarosa 2% donde se separaron los productos de amplificación de 96 ADNs para un mismo locus de SSR.

Para confirmar la correcta asignación de alelos se construyeron escaleras de alelos para cada locus seleccionado mezclando diferentes productos de PCR que representan a todos o casi todos los alelos observados en nuestra colección de cepas. El tamaño de los alelos en las cepas analizadas se determinó por medición directa en el gel de poliacrilamida de la distancia de la banda con respecto a la banda correspondiente al alelo de la cepa secuenciada. Los alelos observados se expresaron como la diferencia en el número de repeticiones con respecto a la cepa de referencia secuenciada. De esta manera, por ejemplo, se le asignó el alelo +4 a la cepa en estudio que presenta cuatro repeticiones más que la cepa de referencia secuenciada o el alelo -7 si se observa que el tamaño del alelo presenta siete repeticiones menos que la cepa de referencia (ver figura 4). Los diferentes genotipos observados no son igualmente probables debido a que la formación de uno u otro genotipo es función directa de las frecuencias alélicas dentro de la población, por lo tanto cada genotipo presenta una probabilidad diferente de aparición en la población. En la tabla 5 se detallan las frecuencias alélicas para el locus YPL009C. En el anexo II se incluyen las fotos de las restantes escaleras de alelos construidas en este trabajo y las frecuencias alélicas observadas para cada locus.

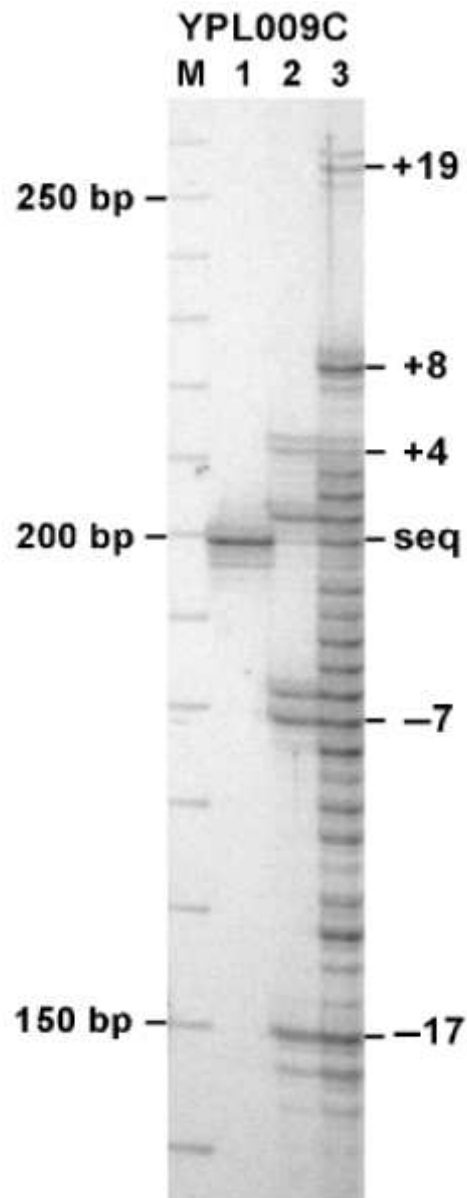


Figura 4. Escalera de alelos para el locus YPL009C.

Foto de un gel de secuencia con la escalera de alelos construida para el locus YPL009C. M es el marcador molecular de ADN de 10 pb. 1. Producto de PCR para la cepa secuenciada (alelo seq) en el locus YPL009C 2. Escalera de alelos parcial 3. Los productos de amplificación que representan a todos los alelos observados en este locus.

Tabla 5. Frecuencias alélicas para el locus YPL009C.

Se muestran los alelos observados, sus frecuencias alélicas y los ADNs utilizados para construir la escalera de alelos para el locus YPL009C.

YPL009C		
Alelo	Frecuencia	DNAs (allele ladders)
-18	0,012	Saf-instant
-17	0,049	CP 863
-15	0,004	FQU 02/13
-14	0,085	CP 873
-13	0,045	ARG-12
-11	0,004	UCD VEN 612
-10	0,004	UCD VEN 535
-9	0,012	Saf-instant
-8	0,012	FQU 02/13
-7	0,049	Levuline CHP
-6	0,024	UCD VEN 612
-5	0,012	FC 9
-4	0,041	AWRI 796
-3	0,057	Levuline CHP
-2	0,024	FQU 02/16
Seq	0,016	AB972
+1	0,285	Saf-instant
+2	0,138	<u>UCD VEN 781</u>
+3	0,081	UCD VEN 535
+4	0,033	ARG-17
+8	0,008	UCD VEN 932
+19	0,004	<u>UCD VEN 781</u>

Tal como era de esperar, cuando se amplificaron los SSRs seleccionados en un número mayor de ADN's diferentes se detectaron nuevos alelos. Los resultados de los nueve loci de SSR seleccionados para todas las cepas de la colección se encuentran en la base de datos www.pasteur.edu.uy/yeast que fue construida por investigadores del Institut Pasteur de Montevideo. Se obtuvieron un total de 184 alelos diferentes en las 120 cepas analizadas, la mayoría de las cepas de vinos fueron diploides y 11 cepas resultaron ser aneuploides (Uvaferm CM, RFV, Na33, Lalvin ICV-D80, UCV VEN612, Panther, Aladdins, Saf-instant, Oetker, Fleischmann y Direma). Todas las cepas comerciales de panificación y la cepa utilizada para la producción de sake tienen de tres a cuatro alelos para la mayoría de los loci. El análisis de las 120 cepas demostró que los nueve loci evaluados presentaron de 7 a 34 alelos. Un resumen de los resultados se presenta en la Tabla 6. En la misma se detalla para cada locus, el número de alelos, el número de alelos exclusivos, el alelo máximo, el alelo mínimo, Ho (Heterocigocidad observada) y PIC (contenido de información polimórfica). El parámetro PIC es frecuentemente utilizado para medir la capacidad de los microsatélites para discriminar entre diferentes individuos, se calcula mediante una fórmula que fue desarrollada por Botstein y colaboradores (1980) y sus valores pueden variar entre 0 y 1, indicando un mayor nivel de polimorfismo o variación cuando el valor es más cercano a 1. De acuerdo a esto, el poder de discriminación encontrado para estos microsatélites fue alto, mayor a 0,75 con la excepción de YGL028C y AT_X. Podemos decir que siete de los nueve loci seleccionados en este trabajo resultaron ser muy informativos a la hora de detectar variabilidad genética en las cepas de *S. cerevisiae*. Los nueve loci de SSR permiten la discriminación de 93 cepas de las 120 cepas analizadas. Se observaron 50 alelos exclusivos que permiten identificar 29 cepas. Una lista detallada de los alelos exclusivos y las cepas identificadas se incluye en la tabla 7.

Tabla 6. Resumen de los resultados obtenidos luego de analizar los ADNs de las 120 cepas de la colección con los nueve loci de SSRs seleccionados.

	TTA_XIII	TG_VI	GT_X	C4_XV	YPL009C	YOR267C	YGL013C	YGL028C	AT_X
Nro. Alelos	30	34	21	24	22	22	16	7	8
Nro. Alelos exclusivos	8	8	5	10	5	7	5	0	2
Exclusivos/Total	0.27	0.24	0.24	0.42	0.23	0.32	0.31	-	0.25
Alelo Máximo	+12	+34	+24	+34	+19	+41	+14	+3	+2
Alelo mínimo	-23	-25	-9	-3	-18	-9	-5	-3	-8
Ho	0.35	0.45	0.43	0.27	0.45	0.23	0.40	0.28	0.29
PIC	0.929	0.925	0.903	0.887	0.864	0.862	0.769	0.577	0.444

Tabla 7. Detalle de los 50 alelos exclusivos observados en el análisis a gran escala.

Cepa	Locus	Alelo
ARG-17	TG_VI	-12
ARG-26	TG_VI	-1
ARG-28	TTA_XIII	+12
ARG-3	TTA_XIII	+2
Fermol Bouquet	YGL013C	+7
UCD VEN 175	YOR267C	+11
UCD VEN 713	YOR267C	+5
UCD VEN 781	YPL009C	+19
UCD VEN 813	GT_X	+11
UCD VEN 932	YPL009C	+8
UCD VEN 952	TG_VI	-14
UCD VEN 941	YGL013C	+6
Lalvin Bourgoblanc CY 3079	TTA_XIII	-4
W	YOR267C	+21
Lalvin ICV-D80	TTA_XIII	-20
ICV-K1 Marquée	TTA_XIII	+7
UCD VEN 587	AT_X	-4
FQU 02/3	YOR267C	+16
ARG-22	GT_X	+9
ARG-22	C4_XV	+12
ARG-8	TTA_XIII	-22
ARG-8	YOR267C	+31
Uvaferm 299	TG_VI	+14
Uvaferm 299	TG_VI	+34
AWRI 350	TTA_XIII	+3
AWRI 350	TG_VI	+9
UCD VEN 765	C4_XV	+17
UCD VEN 765	C4_XV	+25
Lalvin ICV-D21	TTA_XIII	-1
Lalvin ICV-D21	TG_VI	+8
Lalvin ICV-D21	C4_XV	+1
UCD VEN 535	GT_X	+24
UCD VEN 535	C4_XV	+23
UCD VEN 535	YPL009C	-10
Panther	C4_XV	+32
Panther	TG_VI	-20
Panther	GT_X	-1
UY-04	GT_X	-3
UY-04	C4_XV	-2
UY-04	YGL013C	+14
FQU 02/13	C4_XV	+13
FQU 02/13	C4_XV	+18
FQU 02/13	YGL013C	-1
FQU 02/13	YPL009C	-15
UCD VEN 612	YOR267C	+35
UCD VEN 612	YOR267C	+41
UCD VEN 612	C4_XV	+28
UCD VEN 612	AT_X	+2
UCD VEN 612	YGL013C	-5
UCD VEN 612	YPL009C	-11

Los loci YGL013C, YGL028C, AT_X y TG_VI se encuentran dentro o a menos de 3 Kb de genes con interés enológico. El locus TG_VI resultó ser uno de los más polimórficos en nuestra colección de cepas, con 34 alelos diferentes y un 45 % de cepas heterocigotas (ver tabla 6). Las cepas AB972 y S288C se pudieron diferenciar en este locus, ya que la cepa S288C presenta 1 repetido más que la cepa AB972. En la secuenciación del genoma de *S. cerevisiae* (versión 63) para algunos cromosomas se utilizó la cepa AB972 y en la mayoría la cepa S288C. Debido a que el cromosoma VI fue secuenciado con la cepa AB972, le adjudicamos el alelo seq a esta cepa y el alelo +1 a la cepa S288C.

De los nueve SSR analizados a gran escala, el que presentó menos polimorfismo fue el locus AT_X que se encuentra a 135 pb del gen MET3/YJR010W. Este gen codifica una ATP sulfurilasa, la enzima que cataliza el primer paso de la activación intracelular de sulfato, esta enzima es esencial para la reducción de la asimilación de sulfato a sulfuro (Cherest et al., 1985). Las doce cepas productoras de H₂S (ver en tabla de anexo I, H₂S Producing) (Linderholm et al., 2006) presentes en nuestra colección muestran el mismo alelo en homocigosis para AT_X. También hay un alelo predominante en homocigosis para TG_VI en este subgrupo de cepas. Sin embargo, los alelos presentes en las doce cepas productoras de H₂S para estos dos loci próximos a genes MET no son exclusivos de este subgrupo y están presentes en otras cepas de nuestra colección.

El locus YGL013C está a 0,5 Kb de ERG4/YGL012W, pero en la hebra opuesta. ERG4 codifica la enzima esterol C-24 (28) reductasa, que cataliza el paso final en la biosíntesis de ergosterol (Zweytick et al., 2000). El motivo repetido localizado en SCW11/YGL028C se considera un minisatélite, en lugar de un microsatélite, ya que el motivo tiene 12 bases que codifican para los aminoácidos SSST. SCW11 codifica para una proteína de la pared celular con similitud a glucanasas.

4.3. SNPs en FLO8

4.3.1. SNPs en una región del ORF de *FLO8*.

La cepa de referencia del genoma (AB972 o S288C) y la mayoría de las cepas de laboratorio tienen una mutación nula en el gen *flo8* que da lugar a un fenotipo no floculante (Liu et al., 1996). Si el primer nucleótido de la secuencia codificante es el número 1, esta mutación consiste en la presencia de una A en la posición 425 de *Flo8*. Se diseñaron primers para amplificar una región de 300 pb que abarca este sitio conocido con mutación nula (ver secuencia en Figura 6) y se secuenciaron ambas hebras de ADN en un subgrupo de ocho cepas nativas de vino de Uruguay (CP 873, CP874, CP881, CP 882, FQU 02/16, UY-04, CP KU1, FQU 99/5), la cepa comercial Montrachet 522, AB972 (cepa de referencia haploide), BY4743 (cepa diploide utilizada en el proyecto de delección de cada uno de los marcos de lectura abiertos en *S. cerevisiae*) y la cepa comercial muy floculante AWRI350. Se encontraron cinco SNPs en los 300 pb de la región codificante del gen *flo8* y los polimorfismos observados se muestran en la Figura 5. Los SNPs fueron todos inspeccionados visualmente en ambas hebras para asegurarnos de que reflejan los polimorfismos reales y no errores de secuenciación. Los ADNs de todas las cepas de levadura de vino evaluadas presentaron los mismos SNPs en la posición 425 con una G que resulta en un codón de Trp. En la posición 334 se encontró la presencia de una A para las cepas CP 873, CP 874, CP 881, CP 882, FQU 02/16, UY-04 y la cepa comercial AWRI 350. Las cepas CP KU1, FQU 99/5 y Montrachet 522 fueron heterocigotas A/G en la posición 334 resultando en un codón para los aminoácidos Val o Ile. En la posición 306 se encontró que la cepa FQU 99/5 es heterocigota T/C que resulta en el mismo codón His que la cepa secuenciada y en la posición 315 se encontró que la cepa CP KU1 es heterocigota G/A codificando el mismo codón Arg que la cepa secuenciada. Las cepas heterocigotas se definen cuando en el cromatograma aparecen dos picos de colores diferentes y de una altura similar en la misma posición de la secuencia. Los aminoácidos Val e Ile son ambos no polares, mientras que Asn y Ser son aminoácidos con cadenas laterales cargadas polares.

Figura 5. SNPs en ORF de FLO8/YER109C en relación a las cepas secuenciadas AB972 y S288C. Se detallan los cinco SNPs encontrados en la posición del ORF: YER109C_306, YER109C_315, YER109C_334, YER109C_425 y YER109C_449.

Posición SNP en ORF:	YER109C_306	Posición SNP en Genoma:	ChrV_377305
-----------------------------	-------------	--------------------------------	-------------

Cepa	AB972 y S288C	FQU 99/5
Nucleótido	C	T/C
Codón	CAC	CAT/CAC
Aminoácido	His	His

Posición SNP en ORF:	YER109C_315	Posición SNP en Genoma:	ChrV_377296
-----------------------------	-------------	--------------------------------	-------------

Cepa	AB972 y S288C	CP KU1
Nucleótido	A	G/A
Codón	AGA	AGG/AGA
Aminoácido	Arg	Arg

Posición SNP en ORF:	YER109C_449	Posición SNP en Genoma:	ChrV_377162
-----------------------------	-------------	--------------------------------	-------------

Cepa	AB972 y S288C	CP 882
Nucleótido	A	A/G
Codón	AAT	AAT/AGT
Aminoácido	Asn	Asn/Ser

Posición SNP en ORF:	YER109C_334	Posición SNP en Genoma:	ChrV_377277
-----------------------------	-------------	--------------------------------	-------------

Cepa	AB972 y S288C	CP 873	CP 874	CP 881	CP 882	FQU 02/16	UY-04	AWRI350	CP KU1	FQU 99/5	Montrachet 522
Nucleótido	G	A	A	A	A	A	A	A	G/A	G/A	G/A
Codón	GTC	ATC	ATC	ATC	ATC	ATC	ATC	ATC	GTC/ATC	GTC/ATC	GTC/ATC
Aminoácido	Val	Ile	Ile	Ile	Ile	Ile	Ile	Ile	Val/Ile	Val/Ile	Val/Ile

Posición SNP en ORF:	YER109C_425	Posición SNP en Genoma:	ChrV_377186
-----------------------------	-------------	--------------------------------	-------------

Cepa	AB972 y S288C	CP 873	CP 874	CP 881	CP 882	FQU 02/16	UY-04	AWRI350	CP KU1	FQU 99/5	Montrachet 522
Nucleótido	A	G	G	G	G	G	G	G	G	G	G
Codón	TAG	TGG	TGG	TGG	TGG	TGG	TGG	TGG	TGG	TGG	TGG
Aminoácido	Stop	Trp	Trp	Trp	Trp	Trp	Trp	Trp	Trp	Trp	Trp

ATG	AGT	TAT	AAA	GTG	AAT	AGT	TCG	TAT	CCA	GAT	TCA	ATT	CCT	CCC	ACG	GAA	CAA
M	S	Y	K	V	N	S	S	Y	P	D	S	I	P	P	T	E	Q
CCG	TAC	ATG	GCA	AGC	CAG	TAT	AAA	CAA	GAT	TTG	CAG	AGT	AAT	ATT	GCA	ATG	GCA
P	Y	M	A	S	Q	Y	K	Q	D	L	Q	S	N	I	A	M	A
ACG	AAT	AGT	GAA	CAG	CAG	CGA	CAA	CAA	CAG	CAG	CAG	CAG	CAA	CAG	CAG	CAA	CAG
T	N	S	E	Q	Q	R	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q
CAG	TGG	ATA	AAT	CAA	CCT	ACG	GCG	GAA	AAT	TCG	GAT	TTG	AAG	GAA	AAA	ATG	AAC
Q	W	I	N	Q	P	T	A	E	N	S	D	L	K	E	K	M	N
TGC	AAG	AAT	ACG	CTC	AAT	GAG	TAC	ATA	TTT	GAC	TTT	CTT	ACG	AAG	TCG	TCT	TTG
C	K	N	T	L	N	E	Y	I	F	D	F	L	T	K	S	S	L
AAA	AAC	ACT	GCA	GCA	GCC	TTT	GCT	CAA	GAT	GCG	CAC	CTA	GAT	AGA	GAC	AAA	GGC
K	N	T	A	A	A	F	A	Q	D	A	H	L	D	R	D	K	G
CAA	AAC	CCA	GTC	GAC	GGA	CCC	AAA	TCT	AAA	GAA	AAC	AAT	GGT	AAC	CAG	AAT	ACG
Q	N	P	V	D	G	P	K	S	K	E	N	N	G	N	Q	N	T
TTC	TCG	AAG	GTA	GTA	GAT	ACA	CCT	CAA	GGC	TTT	TTG	TAT	GAA	TGG	TAG	CAA	ATA
F	S	K	V	V	D	T	P	Q	G	F	L	Y	E	W	*	Q	I
TTC	TGG	GAC	ATC	TTT	AAT	ACC	AGT	TCT	TCC	AGA	GGT	GGC	TCA	GAG	TTC	GCT	CAG
F	W	D	I	F	N	T	S	S	S	R	G	G	S	E	F	A	Q
CAA	TAT	TAT	CAA	CTA	GTT	CTT	CAA	GAA	CAA	AGG	CAG	GAA	CAA	ATA	TAT	AGA	AGC
Q	Y	Y	Q	L	V	L	Q	E	Q	R	Q	E	Q	I	Y	R	S
TTG	GCT	GTT	CAT	GCG	GCA	AGG	CTA	CAA	CAC	GAT	GCA	GAA	C				
L	A	V	H	A	A	R	L	Q	H	D	A	E					

Figura 6. Secuencia nucleotídica y aminoacídica de FLO8/YER109C en la cepa secuenciada S288C. Se detallan en color los cinco SNPs encontrados en la posición del ORF: YER109C_306 en amarillo, YER109C_315 en verde azulado, YER109C_334 en amarillo oscuro, YER109C_425 en rosado y YER109C_449 en azul. Los cebadores se indican en celeste.

4.3.2. SNPs en una parte del promotor de *FLO8*.

Se diseñaron primers para amplificar una región de 600 pb del promotor YER109C/Flo8 (ver secuencia en anexo III) y se secuenciaron ambas hebras de ADN en las cepas nativas CP 873, CP874, CP881, CP 882, CP KU1, FQU 02/16, UY-04, FQU 99/5 y en las cepas comerciales Montrachet 522 y AWRI350. Todas las cepas dieron igual a la cepa secuenciada S288C excepto en las cepas CP 873, CP KU1 y CP 882. Un alineamiento entre las secuencias de estas tres cepas y la cepa secuenciada S288C se muestra en la Figura 7. Se encontraron seis polimorfismos de SNPs en las posiciones ChrV_378060, ChrV_378032, ChrV_377995, ChrV_377955, ChrV_377799 y ChrV_377637 que se detallan en la Tabla 8. En la posición ChrV_378032 y ChrV_377799 no se observan diferencias en el alineamiento de secuencias de la Figura 7 pero cuando se analizaron los cromatogramas correspondientes de la secuenciación con el oligo fw y con el oligo rev se encontraron que CP882 y CP KU1 son heterocigotas T/G en la posición ChrV_378032 y la cepa CP882 es heterocigota T/C en la posición ChrV_377799 (ver tabla 8). La cepa CP 882 es la que presenta mayor polimorfismo, presentando heterocigocidad en los cinco SNPs de los seis encontrados en este análisis. La cepa CP 873 fue la única en mostrar polimorfismo en la posición ChrV_378060 presentando una C en lugar de la G de la cepa secuenciada.

Tabla 8. SNPs en promotor de FLO8/YER109C. Se detallan los nucleótidos observados y su ubicación en el cromosoma V de los seis SNPs encontrados en las cepas CP 873, CP KU1 y CP 882.

Cepa	Coordenadas en el Cromosoma V:					
	ChrV_378060	ChrV_378032	ChrV_377995	ChrV_377955	ChrV_377799	ChrV_377637
S288C	G	T	T	G	T	T
CP873	C	T	T	G	T	T
CPKU1	G	T/G	C/T	G	T	T
CP882	G	T/G	T/C	A/G	T/C	C/T

Figura 7. Alineamiento de las secuencias analizadas en la región del promotor de Flo8 en las cepas S288C, CP KU1, CP 873 y CP 882. Se indica con colores la ubicación de los seis SNPs encontrados, en color amarillo se indica la posición ChrV_378060, en rosado la coordenada ChrV_378032, en verde claro la ubicación ChrV_377995, en rojo ChrV_377955, en verde oscuro ChrV_377799 y en celeste la posición ChrV_377637.

S288C-FL08pr	TACCTTCTCCTGCACATTCTTGTGATAAGAATGTGAAAAATTTTTTGCTGTATTTCCAGT	60
CPKU1-FL08pr_fw	-----NNNNNNNANTTTTGCTGTATTTCCAGT	27
CP873-FL08pr_fw	-----TTTTGCTGTATTTCCAGT	18
CP882-FL08pr_fw	-----NNATNNNANNTTTGCTGTATTTCCAGT	28
	* * * * *	
S288C-FL08pr	TCTAATGCTGGCTCTAGTAGTAACAAAAATAGAAAATGCCTCGAATAAGTAGCCCTGGGA	120
CPKU1-FL08pr_fw	TCTAATGCTGGCTCTAGTAGTAACAAAAATAGAAAATGCCTCGAATAAGTAGCCCTGGGA	87
CP873-FL08pr_fw	TCTAATGCTGGCTCTAGTAGTAACAAAAATAGAAAATGCCTCGAATAAGTAGCCCTGGGA	78
CP882-FL08pr_fw	TCTAATGCTGGCTCTAGTAGTAACAAAAATAGAAAATGCCTCGAATAAGTAGCCCTGGGA	88
	* * * * *	
S288C-FL08pr	GTGGGATACTGAGGAATTAACGTTTATGCAGGCGTGTGCTTGCCTCCACATGCATGCAT	180
CPKU1-FL08pr_fw	GTGGGATACTGAGGAATTAACGTTTATGCAGGCGTGTGCTTGCCTCCACATGCATGCAT	147
CP873-FL08pr_fw	GTGGGATACTGAGGAATTAACGTTTATGCAGGCGTGTGCTTGCCTCCACATGCATGCAT	138
CP882-FL08pr_fw	GTGGGATACTGAGGAATTAACGTTTATGCAGGCGTGTGCTTGCCTCCACATGCATGCAT	148
	* * * * *	
S288C-FL08pr	ATCAATAGTGGACGCAACAGGGACCACAGTTCAACTGGCCAGGTCCATTGTTGTGTTTG	240
CPKU1-FL08pr_fw	ACCAATAGTGGACGCAACAGGGACCACAGTTCAACTGGCCAGGTCCATTGTTGTGTTTG	207
CP873-FL08pr_fw	ATCAATAGTGGACGCAACAGGGACCACAGTTCAACTGGCCAGGTCCATTGTTGTGTTTG	198
CP882-FL08pr_fw	ATCAATAGTGGACGCAACAGGGACCACAGTTCAACTGGCCAGGTCCATTGTTGTGTTTG	208
	* * * * *	
S288C-FL08pr	CCAACGAGTGTATAGTGCATGAAATCGCGCCTTCTCGGCTTCGGACTCTTTTACGAGGGT	300
CPKU1-FL08pr_fw	CCAACGAGTGTATAGTGCATGAAATCGCGCCTTCTCGGCTTCGGACTCTTTTACGAGGGT	267
CP873-FL08pr_fw	CCAACGAGTGTATAGTGCATGAAATCGCGCCTTCTCGGCTTCGGACTCTTTTACGAGGGT	258
CP882-FL08pr_fw	CCAACGAGTGTATAGTGCATGAAATCGCGCCTTCTCGGCTTCGGACTCTTTTACGAGGGT	268
	* * * * *	

S288C-FL08pr	CCGGAAGAGCGTGGGAAGACAACGAAGAAAGAATGGGATCACGATGAAGTTGTAGAGGGT	360
CPKU1-FL08pr_fw	CCGGAAGAGCGTGGGAAGACAACGAAGAAAGAATGGGATCACGATGAAGTTGTAGAGGGT	327
CP873-FL08pr_fw	CCGGAAGAGCGTGGGAAGACAACGAAGAAAGAATGGGATCACGATGAAGTTGTAGAGGGT	318
CP882-FL08pr_fw	CCGGAAGAGCGTGGGAAGACAACGAAGAAAGAATGGGATCACGATGAAGTTGTAGAGGGT	328

S288C-FL08pr	TGGTTTGAGACCGGTACGATAAAAATTCATAGAATACAGATTGAAAAAGTGACCATTTTT	420
CPKU1-FL08pr_fw	TGGTTTGAGACCGGTACGATAAAAATTCATAGAATACAGATTGAAAAAGTGACCATTTTT	387
CP873-FL08pr_fw	TGGTTTGAGACCGGTACGATAAAAATTCATAGAATACAGATTGAAAAAGTGACCATTTTT	378
CP882-FL08pr_fw	TGGTTTGAGACCGGTACGATAAAAATTCATAGAATACAGATTGAAAAAGTGACCATTTTT	388

S288C-FL08pr	TACTCCTGTTCAAGCGCATTTGCTTTGATACCATTTTTGTTTGCCGAAGACACGGTGAGTT	480
CPKU1-FL08pr_fw	TACTCCTGTTCAAGCGCATTTGCTTTGATACCATTTTTGTTTGCCGAAGACACGGTGAGTT	447
CP873-FL08pr_fw	TACTCCTGTTCAAGCGCATTTGCTTTGATACCATTTTTGTTTGCCGAAGACACGGTGAGTT	438
CP882-FL08pr_fw	TACTCCTGTTCAAGCGCATTTGCTTTGATACCATTTTTGTTTGCCGAAGACACGGTGAGTT	448

S288C-FL08pr	GACGTTAGTAAGTCACTGAGGCTATAAAAAATAAACACGAAGACGTTTATAGACATAAA	540
CPKU1-FL08pr_fw	GACGTTAGTAAGTCACTGAGGCTATAAAAAATAAACACGAAGACGTTTATAGACATAAA	507
CP873-FL08pr_fw	GACGTTAGTAAGTCACTGAGGCTATAAAAAATAAACACGAAGACGTTTATAGACATAAA	498
CP882-FL08pr_fw	GACGTTAGTAAGTCACTGAGGCTATAAAAAATAAACACGAAGACGTTTATAGACATAAA	508

S288C-FL08pr	AAAGAGGAAACGCATTCCGTGGTAGAATGAGTTATAAAGTGAATAGT-----TCGTAT-	593
CPKU1-FL08pr_fw	AAAGAGGAAACGCATTCCGTGGTAGAATGAGTTATAAAGTGAATAGN-----T-----	555
CP873-FL08pr_fw	AAAGAGGAAACGCATTCCGTGGTAGAATGAGTTATAAAGTGAATAGTGACGAATCATATA	558
CP882-FL08pr_fw	AAAGAGGAAACGCATTCCGTGGTAGAATGAGTTATAAAGTGAATAGTC-----	556

5. DISCUSIÓN.

En una fermentación controlada se añade al mosto una cepa seleccionada en cantidad suficiente de masa celular para que se imponga al resto de la microflora, de manera que sea la principal responsable del proceso fermentativo. El uso de levadura seca activa (LSA) requiere de controles fiables, sensibles, rápidos y económicos. Es necesaria la identificación inequívoca de que la cepa que se ha multiplicado y que se va a inocular es la cepa deseada. La técnica a elegir dependerá del costo económico, equipamiento, rapidez, poder de discriminación, la sensibilidad y la fiabilidad. En el 2001, González-Techera y colaboradores propusieron diferenciar cepas de interés enológico empleando la técnica molecular de microsatélites (González-Techera et al, 2001). En el presente trabajo se planteó encontrar el conjunto mínimo de loci de microsatélites polimórficos que permitiera diferenciar las 120 cepas de *S. cerevisiae* presentes en nuestra colección. En primera instancia se analizaron 33 loci de microsatélites en 23 cepas de la colección, nueve loci resultaron ser altamente polimórficos y los restantes presentaron una escasa o nula variabilidad. Entre los loci más polimórficos se encuentran TTA_XIII, TG_VI, GT_X, C4_XV, YPL009C y YOR267C que también fueron propuestos por Legras y colaboradores (2007) por su alto poder de discriminación. Los tres loci de microsatélites (YDR289C, YKL172W y YLR177W) que fueron usados para discriminar entre cepas empleadas como agentes probióticos de cepas clínicas (Hennequin et al, 2001; Malgoire et al., 2005) se descartaron por presentar solamente 3 alelos en el caso de los loci YDR289C y YKL172W y en el locus YLR177W se encontraron 7 alelos pero con un bajo porcentaje de heterocigotas en la población de referencia. No se conoce aun la explicación científica de porqué algunos loci de microsatélites son muy polimórficos y otros no. Por eso se requiere de un trabajo empírico previo para determinar cual es el conjunto de SSRs más polimórfico en la población en estudio. El nivel de polimorfismo depende de la población en estudio, cuanto mas grande sea la población más validez tendrán las estimaciones de las frecuencias alélicas.

En nuestro análisis a gran escala algunas cepas no se pudieron discriminar. Analizando la historia de estas cepas vemos que en el mercado existen cepas que son comercializadas por distintas empresas con diferente nombre tratándose en realidad de la misma cepa. En muchos casos también sucede que al pasar una cepa de un laboratorio a otro se le cambia el nombre, de ahí la importancia de conocer la historia de cada cepa. En el caso de estas cepas que no logramos discriminar solamente

podemos afirmar que con este conjunto de microsatélites no encontramos diferencias y podemos calcular (considerando las frecuencias alélicas de los alelos observados) cual es la probabilidad de que se trate de la misma cepa. En cambio podemos afirmar que cepas que presentan alelos diferentes son cepas diferentes.

Para establecer el fingerprinting de cada cepa, en primera instancia se planteó expresar los resultados como el número de repetidos presentes en el locus amplificado con respecto a la cepa secuenciada y no como valores absolutos (en pares de bases del producto de amplificación). De esta manera, los resultados serían independientes de los primers utilizados y permitiría el intercambio de resultados entre laboratorios. Para esto se tiene que definir el número de repeticiones para el locus considerado en la secuencia de referencia. Si el locus de SSR se encuentra dentro de una región codificante, se cuentan todos los motivos consecutivos, que codifican para el mismo aminoácido. En otros loci la definición del número de repetidos no es clara, por ejemplo, en el caso del locus TTA_XIII, el número de repetidos es 35 si el motivo de repetición es TTA o 36 si el motivo de repetición considerado es TAT. Esta dificultad en definir el número de repetidos en la secuencia de referencia nos llevó a considerar expresar los resultados como la diferencia en el número de repeticiones con respecto a la cepa de referencia secuenciada.

Los resultados del análisis a gran escala de los nueve microsatélites más polimórficos forman parte de una base de datos en la que a futuro se podrían incorporar los datos de más cepas de diferentes orígenes tecnológicos. Dado el constante avance de la tecnología y la disponibilidad de equipos nuevos, para garantizar la viabilidad de las bases de datos es importante que la interpretación de los datos producidos sea independiente del equipo usado para su obtención. Para la construcción de base de datos confiables es importante la reproducibilidad de los resultados entre laboratorios.

La tipificación de cepas mediante el uso de microsatélites se basa en la separación electroforética de fragmentos de ADN de diferentes longitudes que representa un patrón de bandas característico. El intercambio de datos de microsatélites entre laboratorios es difícil debido al uso de diferentes equipos y diferentes condiciones de electroforesis lo que conduce a la alteración de la evaluación de las longitudes de los fragmentos de PCR. La comparación de los perfiles de microsatélites obtenidos en diferentes laboratorios fue estudiado por This y colaboradores (This et al, 2004) en un esfuerzo para desarrollar un método estandarizado para la identificación de los cultivares de vid. Diez laboratorios

analizaron 46 cultivares de uva en seis loci de SSR. Todos los laboratorios utilizaron los mismos ADNs y los mismos primers pero las enzimas y los programas de PCR fueron diferentes. La comparación de los tamaños absolutos de los alelos no fue posible porque se encontraron discrepancias de hasta 3 pb. Los datos fueron comparables cuando todos los laboratorios utilizaron una escalera de alelos relativamente completa para cada uno de los seis loci de microsatélites. Más recientemente, García-Hermoso y colaboradores (2010) realizaron un estudio de estandarización de la tipificación por microsatélites de *Candida albicans* (García-Hermoso et al, 2010). Estos investigadores evaluaron la reproducibilidad entre seis laboratorios desarrollando una escalera de alelos que se compone de mezclas de fragmentos de referencia bien caracterizadas que incluye los alelos más frecuentes para un marcador de microsatélite polimórfico y que actúan como punto de referencia en la corrida electroforética. Los resultados entre laboratorios fueron discrepantes cuando se expresaron en tamaños de los fragmentos. A pesar de las variaciones en las determinaciones de tamaño de los fragmentos, los alelos fueron correctamente asignados gracias a la referencia interna de la escalera de alelos. Esto permitió que los resultados fueran comparables entre laboratorios. Las escaleras de alelos se presentan como un concepto sencillo y de aplicación universal para la normalización de la tipificación molecular mediante microsatélites y proporciona resultados confiables para la construcción de bases de datos.

Al publicar los resultados del presente trabajo (ver publicación en Anexo IV) se propuso un método de estandarización similar basado en la inclusión del ADN de la cepa de *S. cerevisiae* secuenciada de referencia (S288C) y la utilización de escaleras alélicas internas para expresar de manera inequívoca los resultados como la diferencia en el número de repeticiones en relación con la cepa de referencia.

Nuestro grupo de trabajo ofrece la distribución de los ADNs seleccionados, en lugar de escaleras alélicas, para permitir el uso de diferentes plataformas (manual o automática), diferentes cebadores y uso de reacciones de multiplex. Con esta propuesta cada laboratorio puede crear su propia escalera de alelos, el único requisito previo para asegurar la validación de los resultados es que el mismo ADN de referencia se incluya junto con tres a cuatro ADNs seleccionados (que deberían ser los mismos para todos los laboratorios). El intercambio de ADN impide la confusión asociada con los nombres dados a las cepas.

También en este trabajo (Jubany et al., 2008) se realizó un análisis filogenético con los resultados obtenidos con los nueve loci de microsatélites seleccionados. En el dendograma de la Fig 2 de Jubany et al., 2008, se observa la gran diversidad genética

de las cepas nativas uruguayas (URU) que aparecen distribuidas en todo el dendograma conjuntamente con cepas de diferentes orígenes geográficos. En cambio, se observa que las cepas nativas argentinas están filogenéticamente relacionadas y aparecen agrupadas (ver conjunto de cepas ARG).

Las variaciones en minisatélites han sido propuestas como un método para caracterizar las cepas de levadura de vino (Marinangeli et al., 2004). Aunque el poder de discriminación no es alto, pueden ser útiles marcadores porque la amplificación es robusta y en general la diferencia de tamaños entre los alelos es grande y se puede determinar en geles de agarosa en lugar de acrilamida (Howell et al, 2004). En el minisatélite localizado en YGL028C/SCW11 se observaron sólo siete alelos en nuestra colección de cepas y el alelo de mayor tamaño (12 motivos repetidos que codifican para los aminoácidos SSST) no se encontró en homocigosis. El bajo nivel de polimorfismo puede reflejar una restricción funcional para las posibles variantes de la proteína.

En la actualidad la industria biotecnológica de *S. cerevisiae* se interesa en los rasgos complejos que dependen de múltiples genes y de sus variantes alélicas. La floculación es una característica muy importante de las cepas industriales de *S. cerevisiae*, que facilita las operaciones de clarificación del producto y recuperación de levadura. El estudio de SNPs fue restringido a FLo8 por su papel regulador en la floculación. Se encontraron cinco SNPs en unos 300 pb de la región codificante de Flo8 y seis SNPs en una región de 600 pb del promotor de FLo8. En el 2005, Deutschbauer y Davis encontraron entre cepas diferentes de *S. cerevisiae* un polimorfismo, en promedio, cada 173 pares de bases (Deutschbauer & Davis, 2005). En la base de datos yeastgenome (ver <http://www.yeastgenome.org/cgi-bin/FUNGI/alignment.pl?locus=YER109C>) reportan el alineamiento de secuencias proteicas de Flo8 realizado en un conjunto de 24 cepas de *S. cerevisiae* incluyendo la cepa secuenciada S288C. En la posición del ORF YER109C_306 y en YER109C_315 todas las cepas presentan el mismo aminoácido que la cepa secuenciada. Además, en la posición YER109C_425 se reporta Triptófano en todas las cepas con excepción de la cepa secuenciada que presenta el codón stop y en la posición YER109C_334 reportan la presencia de Valina o Isoleucina. Estos datos coinciden con los aminoácidos asociados a los SNPs encontrados en este trabajo con excepción de la posición YER109C_449 de Flo8 que se describe la presencia de Asparagina en las 24 cepas del alineamiento a diferencia de nuestro trabajo que encontramos una variante heterocigota en la cepa CP 882 que presenta Asparagina/Serina.

La posición en el cromosoma V de los SNPs encontrados en este trabajo se reportan considerando las coordenadas del gen YER109C/Flo8 ChrV 377610-375211 de la versión 63 del genoma de *S. cerevisiae* al igual que lo hace la base de datos YSB: Yeast SNPs Browser (http://gbrowse.princeton.edu/cgi-bin/gbrowse/yeast_strains_snps/ se). En esta base de datos muestran la variación genómica a nivel de SNPs encontrada en una colección de 63 cepas de *S. cerevisiae* de diferentes nichos ecológicos y orígenes geográficos (Schacherer et al, 2009). Comparando nuestros resultados con los SNPs reportados en esta base de datos encontramos que las seis variantes encontradas en el promotor de Flo8 no han sido descritas. Considerando la región del ORF de Flo8 analizada en este trabajo encontramos que estas dos bases de datos no son concordantes con los datos reportados para las cepas M22, YJM269 y CLIB324. En la base de datos de yeastgenome se reporta para estas tres cepas la presencia del SNP ubicado en la posición del ORF YER109C_334 que se asocia a la presencia de Valina o Isoleucina en cambio la presencia de dicho SNP no se describe en la base de datos YSB: Yeast SNPs Browser. A partir de febrero de 2011 las coordenadas para Flo8 han cambiado a partir de que se constató que había una inserción de cuatro bases corriente arriba de Flo8 y es por esta razón que actualmente en yeastgenome.org se encuentra la versión 64 con las coordenadas ChrV 377614-375215 para YER109C/Flo8. Decidimos usar la versión 63 del genoma para poder comparar nuestros datos con los reportados por la base de datos YSB: Yeast SNPs Browser.

Nuestro grupo de trabajo no pudo establecer una correlación directa entre los SNPs encontrados en Flo8, su crecimiento en forma de pseudohifas y la floculación (Jubany et al, 2008, en anexo IV). La cepa CP 882 que presenta mayor polimorfismo en los SNPs encontrados en este trabajo tiene la característica de flocular muy bien (González-Techera al, 2001) pero muestra un nivel bajo de formación de pseudohifas en placas con medio SLAD (Synthetic Low-Ammonia Dextrose) (Jubany et al, 2008). En trabajos futuros habría que introducir las distintas variantes de Flo8 en un mismo contexto genético y analizar como influyen en los fenotipos de floculación. En los programas de mejoramiento de cepas industriales se prefiere partir de una cepa que tiene buena performance a nivel de fermentador (algo que depende de la interacción de muchos genes) y a esa cepa introducirle ciertos cambios puntuales que por ejemplo mejoren su capacidad de floculación y aseguren una buena decantación al terminar el proceso de fermentación (Nevoigt et al, 2008). Los microsatélites ubicados dentro o

cerca de genes de interés serían marcadores moleculares ideales en un programa de mejoramiento de cepas industriales.

En los últimos años se ha propuesto utilizar el método MLST (Multi Locus Sequence Typing) como método de discriminación de cepas de levaduras de *Candida albicans* (Bougnoux et al, 2002; Tavanti et al, 2003). Este método consiste en determinar las variaciones en un conjunto definido de genes seleccionados, la identificación es de forma directa por secuenciación del ADN y comparando los SNPs entre cepas para esos genes. La secuencia de ADN es un dato objetivo fácilmente intercambiable entre diferentes laboratorios. En el 2006, Ayoub y colaboradores analizaron 84 cepas de *S. cerevisiae* (incluyendo cepas aisladas de bodegas tradicionales en el Líbano, cepas comerciales de vino, cepas aislados de Asia y cepa secuenciada) por el método MLST y por microsatélites utilizando los loci TTA_XIII, TG_VI, GT_X, C4_XV, YPL009C y YOR267C. El esquema de MLST que utilizaron presentó un menor poder de discriminación que los seis loci de SSRs.

Al momento de plantearnos una tipificación por SNPs y SSRs debemos tener en cuenta que los microsatélites dan información bien definida sobre los niveles de ploidía. Muchas cepas industriales son aneuploides como lo hemos observado en este trabajo. La comparación de la variación genética entre poblaciones requiere un gran número de SNPs con respecto a los microsatélites, normalmente los loci de SSR tienen muchos alelos (34 alelos en el loci TG_VI se encontraron en este trabajo) en comparación a los SNPs que normalmente son bialélicos. A futuro la tipificación de un gran número de cepas con muchos SSR polimórficos y/o SNPs en muchos genes permitirá asociar estos marcadores moleculares a rasgos fenotípicos complejos.

6. CONCLUSIONES Y PERSPECTIVAS FUTURAS

En este trabajo se seleccionó un conjunto de microsatélites polimórficos y se puso a punto el análisis de microsatélites a gran escala.

Los resultados de este trabajo permitieron la creación de una base de datos en la que es posible agregar datos de nuevas cepas y así en un futuro dará información sobre biodiversidad molecular de cepas de *S. cerevisiae* provenientes de diferentes orígenes geográficos, de diversas aplicaciones tecnológicas (alimentos, probióticos, bioetanol), de aislamientos clínicos y de muestras ambientales. Será posible la discriminación e identificación de posibles sinónimos en las colecciones de cepas, cálculo de las frecuencias alélicas, combinación genotípica de los alelos y asociación de los marcadores moleculares para caracteres de interés.

Esta tecnología de microsatélites se podría utilizar para discriminar cepas aisladas para otras aplicaciones industriales como por ejemplo producción de bioetanol.

Aislando cepas en distintas etapas de una fermentación espontánea y analizando sus ADNs con microsatélites es posible saber que cepa es la que predomina y llega al final de la fermentación. También comparando los datos obtenidos con la base de datos podemos saber si las cepas de *S. cerevisiae* aisladas en fermentaciones espontáneas son diferentes a las cepas nativas previamente incorporadas a la base de datos.

Con respecto al efecto de los polimorfismos encontrados en Flo8 habría que incorporar las diferentes variantes en un mismo contexto genético para poder evaluar su impacto en la floculación o en su capacidad de formar pseudohifas. Es interesante resaltar que la formación de pseudohifas es un proceso que se induce en condiciones de escasez de fuentes de nitrógeno y todas las cepas nativas uruguayas fueron seleccionadas en un medio con bajo contenido de nitrógeno.

Por último, consideramos que sería necesario seguir buscando nuevos loci polimórficos de microsatélites en *S. cerevisiae* y analizar cientos o miles de cepas para poder llegar a estimaciones de frecuencias alélicas en poblaciones de cepas de diferentes orígenes geográficos.

7. BIBLIOGRAFÍA.

Antonelli A, Castellari L, Zambonelli C, Carnacini A. (1999) Yeast influence on volatile composition of wines. J Agric Food Chem. 47:1139-1144.

Ayoub MJ, Legras JL, Saliba R, Gaillardin C. (2006) Application of Multi Locus Sequence Typing to the analysis of the biodiversity of indigenous *Saccharomyces cerevisiae* wine yeasts from Lebanon. J Appl Microbiol. 100:699-711.

Bidard F, Blondin B, Dequin S, Vezinhet F, Barre P. (1994) Cloning and analysis of a *FLO5* flocculation gene from *S. cerevisiae*. Curr Genet. 25:196-201.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 32(3):314-31.

Bougnoux ME, Morand S, d'Enfert C. (2002) Usefulness of multilocus sequence typing for characterization of clinical isolates of *Candida albicans*. J Clin Microbiol. 40:1290-1297.

Bowers J, Boursiquot JM, This P, Chu K, Johansson H, Meredith C. (1999) Historical Genetics: The Parentage of Chardonnay, Gamay, and Other Wine Grapes of Northeastern France. Science 285:1562-1565.

Bretagne S, Costa JM, Besmond C, Carsique R, Calderone R. (1997) Microsatellite polymorphism in the promoter sequence of the elongation factor 3 gene of *Candida albicans* as the basis for a typing system. J Clin Microbiol. 35:1777-1780.

Charlesworth B, Sniegowski P, Stephan W. (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371(6494):215-20.

Cheng HH, Crittenden LB. (1994) Microsatellite markers for genetic mapping in the chicken. Poult Sci. 73(4):539-46.

Cherest H, Nguyen NT, Surdin-Kerjan Y. (1985) Transcriptional regulation of the *MET3* gene of *Saccharomyces cerevisiae*. Gene. 34:269-281.

Deutschbauer AM, Davis RW. (2005) Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat Genet.* 37:1333-1340.

Field D, Wills C. (1998) Abundant microsatellite polymorphism in *Saccharomyces cerevisiae*, and the different distributions of microsatellites in eight prokaryotes and *S. cerevisiae*, result from strong mutation pressures and a variety of selective forces. *Proc Natl Acad Sci U S A.* 95:1647-1652.

Garcia-Hermoso D, MacCallum DM, Lott TJ, Sampaio P, Serna MJ, Grenouillet F, Klaassen CH, Bretagne S. (2010) Multicenter collaborative study for standardization of *Candida albicans* genotyping using a polymorphic microsatellite marker. *J Clin Microbiol.* 48:2578-2581.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG. (1996) Life with 6000 genes. *Science.* 274:546- 567.

González Techera A, Jubany S, Carrau FM, Gaggero C. (2001) Differentiation of industrial wine yeast strains using microsatellite markers. *Lett Appl Microbiol.* 33:71-75.

Guo B, Styles CA, Feng Q, Fink GR. (2000) A *Saccharomyces* gene family involved in invasive growth, cell-cell adhesion, and mating. *Proc Natl Acad Sci U S A.* 97:12158-12163.

Hagelberg E, Gray IC, Jeffreys AJ. (1991) Identification of the skeletal remains of a murder victim by DNA analysis. *Nature.* 352:427-479.

Hansen J, Kielland-Brandt MC. (1996) Inactivation of *MET10* in brewer's yeast specifically increases SO₂ formation during beer production. *Nat Biotechnol.* 14:1587-1591.

Helminen P, Ehnholm C, Lokki ML, Jeffreys A, Peltonen L. (1988) Application of DNA "fingerprints" to paternity determinations. *Lancet.* 1:574-576.

Hennequin C, Thierry A, Richard GF, Lecointre G, Nguyen HV, Gaillardin C, Dujon B. (2001) Microsatellite typing as a new tool for identification of *Saccharomyces cerevisiae* strains. *J Clin Microbiol.* 39:551-559.

- Howell KS, Bartowsky EJ, Fleet GH, Henschke PA. (2004) Microsatellite PCR profiling of *Saccharomyces cerevisiae* strains during wine fermentation. Lett Appl Microbiol. 38:315-320.
- Jarne P, Lagoda PJ. (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol. 11(10):424-9.
- Jubany S, Tomasco I, Ponce de León I, Medina K, Carrau F, Arrambide N, Naya H, Gaggero C. (2008) Toward a global database for the molecular typing of *Saccharomyces cerevisiae* strains. FEMS Yeast Res. 8:472-484.
- Katti MV, Ranjekar PK, Gupta VS. (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol Biol Evol. 18:1161-1167.
- Kobayashi O, Yoshimoto H, Sone H. (1999) Analysis of the genes activated by the *FLO8* gene in *Saccharomyces cerevisiae*. Curr Genet. 36:256-261.
- Legras JL, Merdinoglu D, Cornuet JM, Karst F. (2007) Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. Mol Ecol. 16:2091-2102.
- Legras JL, Ruh O, Merdinoglu D, Karst F. (2005) Selection of hypervariable microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains. Int J Food Microbiol. 102:73-83.
- Levinson G, Gutman GA. (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol. 4:203-221.
- Linderholm AL, Olineka TL, Hong Y, Bisson LF. (2006) Allele diversity among genes of the sulfate reduction pathway in wine strains of *Saccharomyces cerevisiae*. Am. J. Enol. Vitic. 57:431-440.
- Liu H, Styles CA, Fink GR. (1996) *Saccharomyces cerevisiae* S288C has a mutation in *FLO8*, a gene required for filamentous growth. Genetics. 144:967-978.
- Lo WS, Dranginis AM. (1996) *FLO11*, a yeast gene related to the *STA* genes, encodes a novel cell surface flocculin. J Bacteriol. 178:7144-7151.

Lo YM, Chan KC. (2006) Setting up a polymerase chain reaction laboratory. *Methods Mol Biol.* 336:11-18.

Malgoire JY, Bertout S, Renaud F, Bastide JM, Mallié M. (2005) Typing of *Saccharomyces cerevisiae* clinical strains by using microsatellite sequence polymorphism. *J Clin Microbiol.* 43:1133-1137.

Marinangeli P, Angelozzi D, Ciani M, Clementi F, Mannazzu I. (2004) Minisatellites in *Saccharomyces cerevisiae* genes encoding cell wall proteins: a new way towards wine strain characterisation. *FEMS Yeast Res.* 4:427-435.

Moxon ER, Wills C. (1999) DNA microsatellites: agents of evolution? *Sci Am.* 280:94-99.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 51 Pt 1:263-273.

Nevoigt E. (2008) Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev.* 72:379-412.

Piarroux R, Millon L, Bardonnnet K, Vagner O, Koenig H. (1999) Are live *Saccharomyces* yeasts harmful to patients? *Lancet.* 353:1851-1852.

Posteraro B, Sanguinetti M, Romano L, Torelli R, Novarese L, Fadda G. (2005) Molecular tools for differentiating probiotic and clinical strains of *Saccharomyces cerevisiae*. *Int J Food Microbiol.* 103:295-304.

Pretorius IS. (2000) Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. *Yeast.* 16:675-729.

Richard GF, Hennequin C, Thierry A, Dujon B. (1999) Trinucleotide repeats and other microsatellites in yeasts. *Res Microbiol.* 150:589-602.

Rozen S, Skaletsky H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol.* 132:365-386.

Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L. (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. Nature. 458:342-345.

Schlötterer C, Tautz D. (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res. 20:211-215.

Schuller D, Valero E, Dequin S, Casal M. (2004) Survey of molecular methods for the typing of wine yeast strains. FEMS Microbiol Lett. 231:19-26.

Sherman F. (1991) Getting started with yeast. En: Guide to yeast genetics and molecular biology. Ed. Guthrie C & Fink GR. Methods in Enzymology. Vol. 194. Academic Press, Inc.

Strand M, Prolla TA, Liskay RM, Petes TD. (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature. 365:274-276.

Tavanti A, Gow NA, Senesi S, Maiden MC, Odds FC. (2003) Optimization and validation of multilocus sequence typing for *Candida albicans*. J Clin Microbiol. 41:3765-3776.

Teunissen AW, Holub E, van der Hucht J, van den Berg JA, Steensma HY. (1993) Sequence of the open reading frame of the *FLO1* gene from *Saccharomyces cerevisiae*. Yeast. 9:423-427.

Teunissen AW, Steensma HY. (1995) Review: the dominant flocculation genes of *Saccharomyces cerevisiae* constitute a new subtelomeric gene family. Yeast. 11:1001-1013.

Teunissen AW, van den Berg JA, Steensma HY. (1995) Localization of the dominant flocculation genes FLO5 and FLO8 of *Saccharomyces cerevisiae*. Yeast. 11(8):735-45.

This P, Jung A, Boccacci P, Borrego J, Botta R, Costantini L, Crespan M, Dangl GS, Eisenheld C, Ferreira-Monteiro F, Grando S, Ibáñez J, Lacombe T, Laucou V, Magalhães R, Meredith CP, Milani N, Peterlunger E, Regner F, Zulini L, Maul E. (2004) Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theor Appl Genet. 109:1448-1458.

Vagnoli P, Visón LF. (1998) The SKS1 gene of *Saccharomyces cerevisiae* is required for long-term adaptation of *snf3* null strains to low glucose. *Yeast*. 15;14(4):359-69.

Verstrepen KJ, Derdelinckx G, Verachtert H, Delvaux FR. (2003) Yeast flocculation: what brewers should know. *Appl Microbiol Biotechnol*. 61:197-205.

Verstrepen KJ, Van Laere SD, Vanderhaegen BM, Derdelinckx G, Dufour JP, Pretorius IS, Winderickx J, Thevelein JM, Delvaux FR (2003) Expression levels of the yeast alcohol acetyltransferase genes ATF1, Lg-ATF1, and ATF2 control the formation of a broad range of volatile esters. *Appl Environ Microbiol*. 69(9):5228-37.

Young ET, Sloan JS, Van Riper K. (2000) Trinucleotide repeats are clustered in regulatory genes in *Saccharomyces cerevisiae*. *Genetics*. 154:1053-1068.

Wu H, Zheng X, Araki Y, Sahara H, Takagi H, Shimoi H. (2006) Global gene expression analysis of yeast cells during sake brewing. *Appl Environ Microbiol*. 72(11):7353-8.

Zweytick D, Hrastnik C, Kohlwein SD, Daum G. (2000) Biochemical characterization and subcellular localization of the sterol C-24(28) reductase, Erg4p, from the yeast *Saccharomyces cerevisiae*. *FEBS Lett*. 470:83-87.

8. ANEXOS

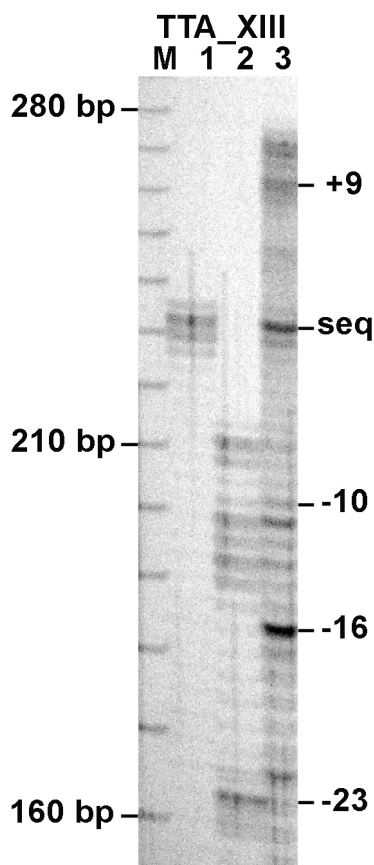
Anexo I. Tabla con las 120 cepas de *S. cerevisiae* de nuestra colección.

Dendogram	Name	Origin	Application
URU-1	<u>CP 863</u>	Uruguay	Wine/native
URU-2	<u>CP 873</u>	Uruguay	Wine/native
URU-3	<u>CP 874</u>	Uruguay	Wine/native
URU-4	<u>CP 881</u>	Uruguay	Wine/native
URU-5	<u>CP 882</u>	Uruguay	Wine/native
COM-1	<u>AWRI 796</u>	South Africa	Wine/commercial
COM-2	<u>Levuline CHP</u>	France/Champagne	Wine/commercial
COM-3	<u>Montrachet 522</u>	USA/California	Wine/commercial
URU-6	<u>CP KU1</u>	Uruguay	Wine/native
LAB-1	<u>AB972</u>	Laboratory	Laboratory
ARG-1	ARG-1	Argentina	Wine/native
ARG-2	ARG-2	Argentina	Wine/native
ARG-3	ARG-3	Argentina	Wine/native
ARG-4	<u>ARG-4</u>	Argentina	Wine/native
ARG-5	ARG-5	Argentina	Wine/native
ARG-6	ARG-6	Argentina	Wine/native
ARG-7	ARG-7	Argentina	Wine/native
ARG-8	ARG-8	Argentina	Wine/native
ARG-9	ARG-9	Argentina	Wine/native
ARG-10	ARG-10	Argentina	Wine/native
ARG-11	ARG-11	Argentina	Wine/native
ARG-12	ARG-12	Argentina	Wine/native
ARG-13	ARG-13	Argentina	Wine/native
ARG-14	ARG-14	Argentina	Wine/native
ARG-15	ARG-15	Argentina	Wine/native
ARG-16	ARG-16	Argentina	Wine/native
ARG-17	ARG-17	Argentina	Wine/native
ARG-18	ARG-18	Argentina	Wine/native
ARG-19	ARG-19	Argentina	Wine/native
ARG-20	ARG-20	Argentina	Wine/native
ARG-21	ARG-21	Argentina	Wine/native
ARG-22	ARG-22	Argentina	Wine/native
ARG-23	ARG-23	Argentina	Wine/native
ARG-24	ARG-24	Argentina	Wine/native
ARG-25	ARG-25	Argentina	Wine/native
ARG-26	ARG-26	Argentina	Wine/native
ARG-27	ARG-27	Argentina	Wine/native
ARG-28	ARG-28	Argentina	Wine/native
URU-7	<u>FQU 02/16</u>	Uruguay	Wine/native
URU-8	FQU 02/3	Uruguay	Wine/native
URU-9	FQU 02/11	Uruguay	Wine/native
URU-10	FQU 02/5	Uruguay	Wine/native
URU-11	<u>UY-04</u>	Uruguay	Wine/native
URU-12	FQU 02/13	Uruguay	Wine/native
URU-13	FQU 02/4	Uruguay	Wine/native
URU-14	FQU 99/5	Uruguay	Wine/native
URU-15	FQU 96/8	Uruguay	Wine/native
COM-4	Uvaferm 299	France/Beaujolais	Wine/commercial
COM-5	Uvaferm CM	USA/California	Wine/commercial
COM-6	AWRI 350	Australian	Wine/commercial

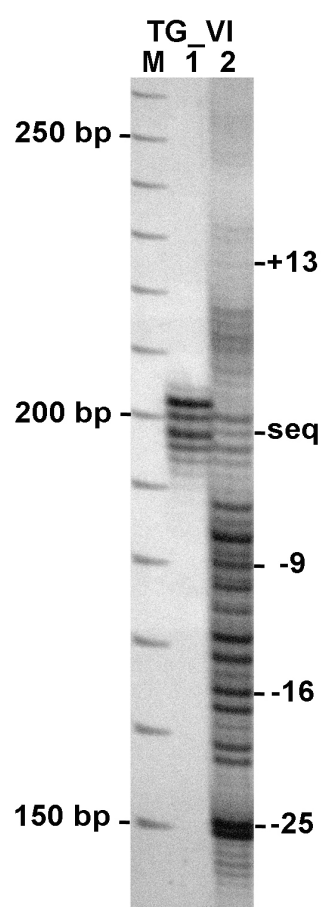
COM-7	RFV	Unknown	Wine/commercial
COM-8	RFA	Unknown	Wine/commercial
COM-9	S. bayanus	Unknown	Wine/commercial
COM-10	Lallem Laffort F15	France/Bordeaux	Wine/commercial
COM-11	Laffort Actiflora C	Unknown	Wine/commercial
COM-12	ALG 111	France/Burgundy	Wine/commercial
COM-13	Fermol Bouquet	France	Wine/commercial
COM-14	Merlot Cepage	France/Bordeaux	Wine/commercial
COM-15	Lalvin EC-1118	France/Champagne	Wine/commercial
COM-16	Lalvin ICV-D21	France/Pic Saint Loup.	Wine/commercial
COM-17	<u>Enoferm QA23</u>	Portugal/Vinhos Verdes.	Wine/commercial
COM-18	QD 145	Portugal/Dao	Wine/commercial
COM-19	<u>Lalvin BM45</u>	Italy/Brunello di monracino.	Wine/commercial
COM-20	Burgundy (BGY)	France	Wine/commercial
COM-21	AMH(Assmanshausen)	Germany	Wine/commercial
COM-22	Simi White	France	Wine/commercial
COM-23	Na33	Spain/Olite	Wine/commercial
COM-24	Enoferm L2226	France/Côtes du Rhône.	Wine/commercial
	Lalvin Bourgo blanc CY		
COM-25	3079	France	Wine/commercial
COM-26	W	Unknown	Wine/commercial
COM-27	T 18	Unknown	Wine/commercial
COM-28	R 7	France/Champagne	Wine/commercial
COM-29	Lalvin ICV-D80	France/Côte Rôtie.	Wine/commercial
COM-30	FC 9	Unknown	Wine/commercial
COM-31		France/Vallée du Rhône	Wine/commercial
COM-32	C (<i>S. cerevisiae bayanus</i>)	Unknown	Wine/commercial
COM-33	ICV-K1 Marquée	France/Montpellier	Wine/commercial
COM-34	DV 10	France/Champagne	Wine/commercial
COM-35	Lalvin 71B	France/Narbonne	Wine/commercial
UCD-1	UCD VEN 49	USA	Wine/native
COM-36	UCD VEN 51	France/Burgundy	Wine/commercial
COM-37	UCD VEN 175	Italy/Sicily	Wine/commercial
COM-38	<u>UCD VEN 529</u>	Germany/Steinberg	Wine/commercial
UCD-2	UCD VEN 535	USA	Wine/native
UCD-3	UCD VEN 587	USA/California	Wine/native
SAKE	<u>UCD VEN 612</u>	Japan/Kurashi	Sake/commercial
COM-39	<u>UCD VEN 713</u>	USA	Wine/commercial
COM-40	UCD VEN 765	Australian	Wine/commercial
COM-41	UCD VEN 777	France	Wine/commercial
COM-42	<u>UCD VEN 781</u>	Switzerland	Wine/commercial
COM-43	UCD VEN 813	France/Epernay II	Wine/commercial
COM-44	UCD VEN 819	France	Wine/commercial
COM-45	<u>UCD VEN 820</u>	South Africa	Wine/commercial
COM-46	UCD VEN 904	France	Wine/commercial
COM-47	UCD VEN 905	France	Wine/commercial
H2S-1	UCD VEN 932	Italy	H2S Producing
H2S-2	UCD VEN 934	Italy	H2S Producing
H2S-3	UCD VEN 935	Italy	H2S Producing
H2S-4	UCD VEN 937	Italy	H2S Producing
H2S-5	UCD VEN 939	Italy	H2S Producing
H2S-6	UCD VEN 940	Italy	H2S Producing
H2S-7	UCD VEN 941	Italy	H2S Producing
H2S-8	UCD VEN 949	Italy	H2S Producing
H2S-9	UCD VEN 950	Italy	H2S Producing

H2S-10	UCD VEN 951	Italy	H2S Producing
H2S-11	UCD VEN 952	Italy	H2S Producing
H2S-12	UCD VEN 957	Italy	H2S Producing
COM-48	UCD VEN 969	Spain/Valencia	Wine/commercial
COM-49	UCD VEN 2031	France	Wine/commercial
COM-50	UCD VEN 2032	France	Wine/commercial
UCD-4	UCD VEN 2118	USA/California	Wine/native
UCD-5	UCD VEN 2120	USA/California	Wine/native
LAB-2	UCD VEN 2516 (BY4743)	Laboratory	Laboratory
LAB-3	<u>UCD VEN 2517 (S288C)</u>	Laboratory	Laboratory
PAN-1	<u>Panther</u>	Turkey	Bread/commercial
PAN-2	Aladdin's	Turkey	Bread/commercial
PAN-3	Saf-instant	Argentina	Bread/commercial
PAN-4	Oetker	Brazil	Bread/commercial
PAN-5	Fleischmann	Brazil	Bread/commercial
PAN-6	Direma	Brazil	Bread/commercial

Anexo II. Escaleras alelicas construidas en este trabajo y tablas con las frecuencias alelicas de los loci TTA_XIII, TG_VI, GT_X, C4_XV, YOR267C, YGL013C, YGL028C y AT_X. Se muestran las fotos de los geles de poliacrilamida con las escaleras de alelos y se detalla en las tablas la frecuencia observada para cada alelo en la población y el ADN utilizado para la construcción de las escaleras de alelos.

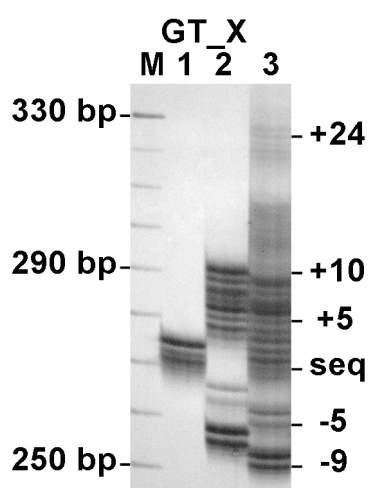


TTA_XIII		
Alelo	Frecuencia	DNAs (allele ladders)
-23	0,024	Uvaferm 299
-22	0,008	ARG-8
-20	0,008	Lalvin ICV-D80
-19	0,012	UCD VEN 2118
-18	0,045	UY-04
-17	0,024	ARG-3
-16	0,008	FQU 02/13
-14	0,016	Uvaferm 299
-13	0,012	FC 9
-12	0,073	UCD VEN 765
-11	0,045	Uvaferm 299
-10	0,151	Lalvin Bourgoblanc CY 3079
-9	0,029	UCD VEN 529
-8	0,090	Uvaferm 299
-7	0,020	UCD VEN 957
-6	0,012	UCD VEN 587
-5	0,012	FQU 02/13
-4	0,004	Lalvin Bourgoblanc CY 3079
-3	0,073	Burgundy (BGY)
-2	0,041	Lalvin ICV-D21
-1	0,004	Lalvin ICV-D21
Seq	0,045	AB972
+1	0,049	ARG-7
+2	0,004	ARG-3
+3	0,008	AWRI 350
+4	0,094	ARG-11
+6	0,053	CP 863
+7	0,008	ICV-K1 Marquée
+9	0,016	ARG-22
+12	0,008	ARG-28

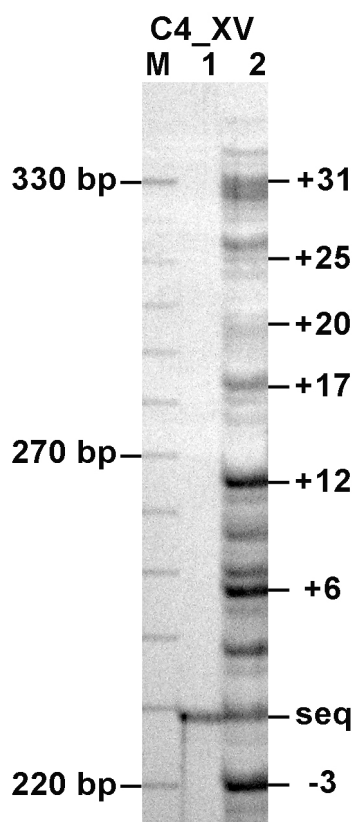


TG VI

Alelo	Frecuencia	DNAs (allele ladders)
-25	0,177	UCD VEN 957
-24	0,1	Panther
-23	0,076	CP 882
-20	0,004	Panther
-19	0,056	CP 02/11
-17	0,036	ALG 111
-16	0,088	Burgundy (BGY)
-15	0,044	Na33
-14	0,008	UCD VEN 952
-13	0,032	Lallemand Laffort F15
-12	0,008	ARG-17
-11	0,036	AWRI 350
-10	0,016	Simi White
-9	0,008	CP 882
-8	0,028	Oetker
-7	0,036	ARG-22
-6	0,02	CP 874
-5	0,016	Oetker
-4	0,028	UY-04
-3	0,008	Direma
-2	0,02	CP 873
-1	0,008	ARG-26
seq	0,04	AB972
+1	0,016	UCD VEN 2517
+2	0,012	ARG-28
+4	0,008	Lalvin ICV-D80
+5	0,012	Panther
+6	0,016	ARG-9
+7	0,012	AWRI 350
+8	0,004	Lalvin ICV-D21
+9	0,004	AWRI 350
+14	0,004	Uvaferm 299
+28	0,012	Oetker
+34	0,004	Uvaferm 299

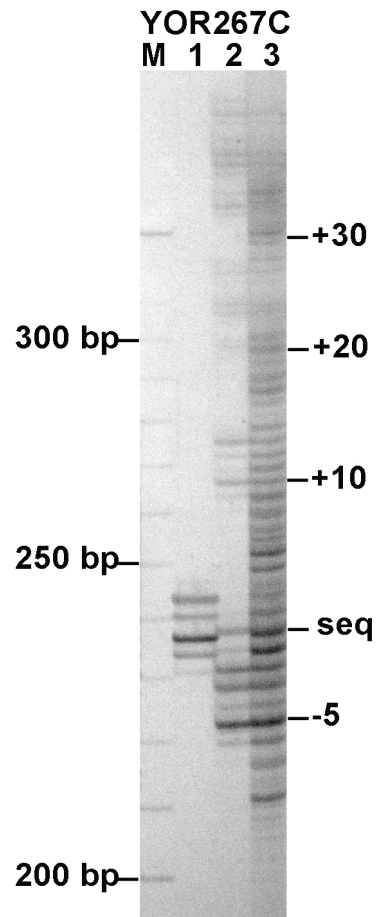


GT_X		
Alelo	Frecuencia	DNAs (allele ladders)
-9	0,099	<u>CP 863</u>
-7	0,049	FQU 02/13
-6	0,165	<u>Panther</u>
-5	0,008	FQU 02/13
-4	0,049	ARG-8
-3	0,008	<u>UY-04</u>
-2	0,033	<u>UCD VEN 612</u>
-1	0,004	<u>Panther</u>
Seq	0,045	Uvaferm 299
+1	0,012	Uvaferm 299
+3	0,041	UCD VEN 957
+4	0,144	<u>Montrachet 522</u>
+5	0,066	CP 873
+6	0,049	Uvaferm 299
+7	0,111	UCD VEN 813
+8	0,062	ARG-14
+9	0,008	ARG-22
+10	0,016	UCD VEN 587
+11	0,004	UCD VEN 813
+14	0,021	<u>Montrachet 522</u>
+24	0,004	UCD VEN 535



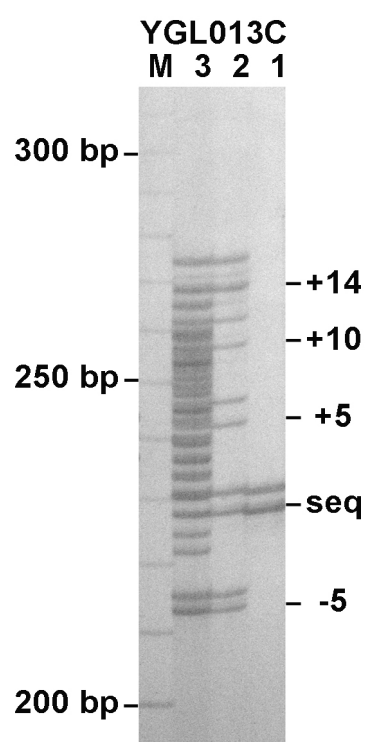
C4_XV

Alelo	Frecuencia	DNAs (allele ladders)
-3	0,017	W
-2	0,008	UY-04
Seq	0,017	UCD VEN 2517
+1	0,008	Lalvin ICV-D21
+3	0,121	ARG-1
+4	0,105	Montrachet 522
+5	0,159	UCD VEN 535
+6	0,109	CP 873
+7	0,159	Lalvin 71B
+8	0,084	CP 874
+9	0,013	ARG-26
+10	0,021	ARG-1
+12	0,008	ARG-22
+13	0,004	FQU 02/13
+17	0,004	UCD VEN 765
+18	0,004	FQU 02/13
+20	0,017	Panther
+21	0,092	Panther
+23	0,004	UCD VEN 535
+25	0,004	UCD VEN 765
+28	0,008	UCD VEN 612
+31	0,017	Lalvin 71B
+32	0,004	Panther
+34	0,013	Uvaferm 299



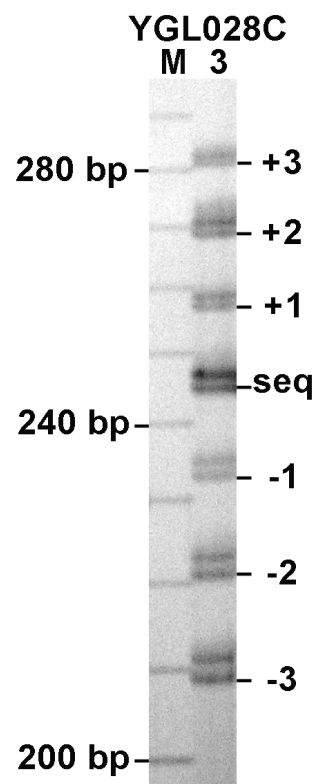
YOR267C

Alelo	Frecuencia	DNAs (allele ladders)
-9	0,0126	UCD VEN 535
-6	0,0168	<u>Panther</u>
-5	0,2647	UCD VEN 175
-2	0,0378	ARG-10
-1	0,0588	Na33
seq	0,0168	<u>AB972</u>
+4	0,0462	UCD VEN 957
+5	0,0084	<u>UCD VEN 713</u>
+6	0,1134	<u>AWRI 796</u>
+9	0,0924	<u>CP 874</u>
+10	0,1429	ARG-3
+11	0,0042	UCD VEN 175
+16	0,0042	FQU 02/3
+17	0,0924	<u>Montrachet 522</u>
+18	0,0126	ARG-26
+21	0,0084	W
+24	0,0084	Direma
+29	0,0168	Direma
+31	0,0084	ARG-8
+32	0,0252	Enoferm L2226
+35	0,0042	<u>UCD VEN 612</u>
+41	0,0042	<u>UCD VEN 612</u>



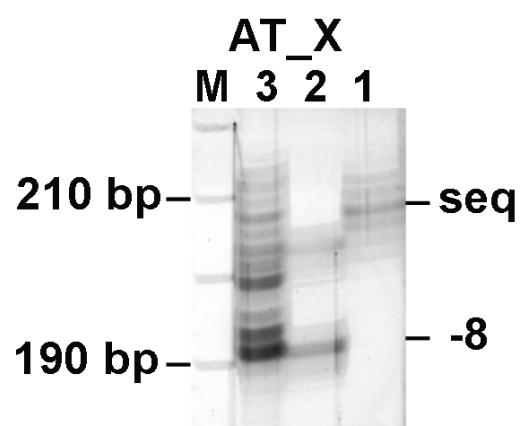
YGL013C

Alelo	Frecuencia	DNAs (allele ladders)
-5	0,008	UCD VEN 612
-2	0,017	Uvaferm 299
-1	0,004	FQU 02/13
Seq	0,025	UCD VEN 2517
+1	0,326	FQU 02/13
+2	0,025	Uvaferm 299
+3	0,033	ARG-10
+4	0,021	UCD VEN 2120
+5	0,289	Uvaferm 299
+6	0,008	UCD VEN 941
+7	0,004	Fermol Bouquet
+8	0,033	ARG-17
+9	0,099	Fermol Bouquet
+10	0,045	ARG-23
+11	0,054	UCD VEN 937
+14	0,008	UY-04



YGL028C

Alelo	Frecuencia	DNAs (allele ladders)
-3	0,012	ARG-11
-2	0,132	ARG-12
-1	0,504	ARG-27
Seq	0,302	UCD VEN 612
+1	0,012	UCD VEN 612
+2	0,025	Merlot Cepage
+3	0,012	ARG-27



AT_X

Alelo	Frecuencia	DNAs (allele ladders)
-8	0,725	<u>CP 873</u>
-7	0,063	<u>UCD VEN 612</u>
-6	0,046	Lalvin 71B
-4	0,008	UCD VEN 587
-3	0,063	<u>UCD VEN 713</u>
-2	0,029	<u>CP 863</u>
Seq	0,063	AB972
+2	0,004	<u>UCD VEN 612</u>

Anexo III. Secuencia del Cromosoma V de *S. cerevisiae* desde la coordenada 375206 a 378176 que contiene las regiones analizadas para la presencia de SNPs en el promotor y en el ORF de FLO8/YER109C.

En verde se señalan los cebadores usados para analizar la presencia de SNPs en la región del promotor de FLO8/YER109C que va desde 378174 a 377575 e incluye los 36 primeras bases del ORF. El comienzo del ORF es en la ubicación 377610 (en rojo se indica el codon de inicio ATG del ORF). En celeste se indican los cebadores utilizados para estudiar la presencia de SNPs en la región codificante de FLO8/YER109C que contiene la mutación nula en la posición 425 del ORF (se indica con color rosa).

```
TACCTTCTCCTGCACATTCTTG TGATAAGAATGTGAAAAATTTTTTGCTGTATTTCCAGT
TCTAATGCTGGCTCTAGTAGTAACAAAAATAGAAAAATGCCTCGAATAAGTAGCCCTGGGA
GTGGGATACTGAGGAATTAACGTTTATGCAGGCGTGTGCTTGCCTCCACATGCATGCAT
ATCAATAGTGACGCAACAGGGACCACAGTTCAACTGGCCAGGGTCCATTGTTGTGTTTG
CCAACGAGTGTATAGTGCATGAAATCGCGCCTTCTCGGCTTCGGACTCTTTTACGAGGGT
CCGGAAGAGCGTGGGAAGACAACGAAGAAAGAATGGGATCACGATGAAGTTGTAGAGGGT
TGGTTTGAGACCGGTACTGATAAAATTCATAGAATACAGATTGAAAAAGTGACCATTTTT
TACTCCTGTTCAAGCGCATTTTGCTTTGATACCATTTTTGTTTGCCGAAGACACGGTGAGTT
GACGTTAGTAAGTCACTGAGGCTATAAAAAATAAACACGAAGACGTTTATAGACATAAAT
AAAGAGGAAACGCATTCCGTGGTAGAATGAGTTATAAAGTGAATAGTTCGTATCCAGATT
CAATTCTCCACGGAACAACCGTACATGGCAAGCCAGTATAAAACAAGATTTGCAGAGTA
ATATTGCAATGGCAACGAATAGTGAACAGCAGCGACAACAACAGCAGCAGCAGCAACAGC
AGCAACAGCAGTGGATAAATCAACCTACGGCGGAAAATTTCGGATTTGAAGGAAAAAATGA
ACTGCAAGAATACGCTCAATGAGTACATATTTGACTTTCTTACGAAGTCGTCTTTGAAAA
ACACTGCAGCAGCCTTTGCTCAAGATGCGCACCTAGATAGAGACAAAGGCCAAAACCCAG
TCGACGGACCCAAATCTAAAGAAAACAATGGTAACCAGAATACGTTCTCGAAGGTAGTAG
ATACACCTCAAGGCTTTTTGTATGAATGGTAGCAAAATATTCTGGGACATCTTTAATACCA
GTTCTTCCAGAGGTGGCTCAGAGTTTCGCTCAGCAATATTATCAACTAGTTCTTCAAGAAC
AAAGGCAGGAACAAATATATAGAAGCTTGGCTGTTTCATGCGGCAAGGCTACAACACGATG
CAGAACGAAGAGGGGAATATAGTAACGAGGACATAGACCCCATGCACTTGGCTGCTATGA
TGCTAGGAAATCCTATGGCACCTGCGGTTCAAATGCGCAATGTTAATATGAACCCTATAC
CAATTCCTATGGTTGGTAACCCTATCGTTAATAATTTTTCCATTCCACCATAACAATAATG
CAAACCCACGACTGGTGCAACTGCTGTTGCTCCACAGCGCCGCTTCCGGCGATTTTA
CAAATGTAGGGCCAAACCCAGAAATCGGAGTCAAAACGTTACTGGCTGGCCAGTCTATAAT
ATCCAATGCAACCCCACTACGGAAAAATCCAGTGGGAAACCCGTGTAACAATAATACCACAA
ATAATACAACATAAACAATCTCCAGTGAACCAACCTAAAAGTTTAAAAACTATGCATT
CAACAGATAAAACCAATAATGTCCCGACGTCAAAATCTACAAGAAGTAGATCTGCAACCT
CAAAAGCGAAGGGTAAAGTTAAAGCCGGTCTAGTGGCTAAGAGACGAAGAAAAAATAATA
CCGCTACAGTTTCCGCGGGATCGACGAACGCTTGTTCGCCAAATATTACCACACCAGGCT
CAACAACAAGTGAACCCGCTATGGTAGGTTCAAGAGTAAATAAGACTCCAAGATCAGATA
TTGCTACTAACTTCCGCAATCAAGCAATAATATTTGGCGAGGAAGATATTTATTCTAATT
CCAAATCTAGCCCATCGTTGGATGGAGCATCACCTTCCGCTTTAGCTTCTAAACAGCCCA
CAAAGGTAAGGAAAAATACAAAAAAGGCATCCACCTCAGCTTTTCCAGTAGAGTCTACGA
ATAAACTCGGTGGCAACAGCGTGGTGACAGGTAAAAAGCGCAGTCCCCCTAACACTAGAG
TGTCGAGGAGGAAATCCACTCCTTCTGTTATTCTGAATGCTGATGCCACTAAGGATGAGA
ATAATATGTTAAGAACATTCTCGAATACTATTGCTCCGAATATTCAATCCGCTCCGCCCA
CTAAACTGCGAATTCTCTCCCTTTTCCAGGTATAAATTTGGGAAGTTTCAACAAGCCGG
CTGTATCCAGTCCATTATCTTCAGTGACAGAGAGTTGCTTCGATCCAGAAAGTGGCAAGA
TTGCCGGAAGAATGGACCCAAGCGAGCAGTAAACTCAAAGTTTCGGCATCATCCCCAT
TAAGCATAGCAACACCTCGGTCTGGTGACGCTCAGAAGCAAAGAAGTTCTAAGGTACCAG
GAAACGTGTTTATAAGCCGCCACATGGGTTTTCAACCACCAATTTGAATATTACTTTAA
AGAACCTGAAAAATACTCACTTACAGAAATAACAGTATCCCAAGAATTGCCGAATGGGG
GAAACATACTGGAGGCGCAAGTAGGCAATGATTCAAGAAGTAGTAAAGGCAATCGTAACA
CATTATCTACTCCAGAGGAAAAAAGCCGAGTAGTAATAATCAAGGATATGATTTTGACG
CCCTCAAAAATTCAAGTTCTTTGTTGTTTCTAATCAAGCTTATGCTTCTAACAATAGAA
CACCAAACGAGAATTCAAATGTTGCTGATGAAACCTCTGCATCTACAAATAGTGGCGATA
ATGATAACACATTAATTCAGCCCTCATCCAATGTGGGTACAACCTTTGGGTCTCAGCAAA
CCAGTACTAATGAAAAATCAGAAATGTACACTCTCAGAACTTGAAGTTTGGGAATATTGGTA
TGTTTGAAGACCAAGGACCGGATTACGATCTCAATTTACTGGATACAAATGAAAATGATT
TCAATTTTATTAATTGGGAAGGCTGAAGTCA
```

Anexo IV.

Jubany S, Tomasco I, Ponce de León I, Medina K, Carrau F, Arrambide N, Naya H, Gaggero C. 2008. Toward a global database for the molecular typing of *Saccharomyces cerevisiae* strains. FEMS Yeast Res. 8:472-484.

Anexo V.

Mercado L, Jubany S, Gaggero C, Masuelli R, Combina M. 2010. Molecular Relationships Between *Saccharomyces cerevisiae* Strains Involved in Winemaking from Mendoza, Argentina. Curr Microbiol 61:506–514

RESEARCH ARTICLE

Toward a global database for the molecular typing of *Saccharomyces cerevisiae* strains

Sandra Jubany¹, Ivanna Tomasco², Inés Ponce de León¹, Karina Medina³, Francisco Carrau³, Nicolás Arrambide⁴, Hugo Naya⁴ & Carina Gaggero¹

¹Departamento de Biología Molecular, Instituto de Investigaciones Biológicas Clemente Estable (IIBCE), Uruguay; ²Sección Evolución, Facultad de Ciencias, Udelar, Uruguay; ³Sección Enología, Facultad de Química, Udelar, Uruguay; and ⁴Unidad de Bioinformática, Institut Pasteur de Montevideo, Uruguay

Correspondence: Carina Gaggero, Departamento de Biología Molecular, IIBCE (Instituto de Investigaciones Biológicas Clemente Estable), Av. Italia 3318, 11600 Montevideo – Uruguay. Tel.: +5982 487 2605; fax: +5982 487 5461; e-mail: carina@iibce.edu.uy

Received 29 October 2007; revised 22 January 2008; accepted 22 January 2008.
First published online 22 February 2008.

DOI:10.1111/j.1567-1364.2008.00361.x

Editor: Isak Pretorius

Keywords

Saccharomyces cerevisiae; microsatellites; SNPs; FLO8.

Abstract

Most of the yeast strains used in fermented beverages and foods are classified as *Saccharomyces cerevisiae*. However, different strains are suitable for different fermentation processes. The purpose of this work is the proposal of a standardized methodology for the molecular genotyping of *S. cerevisiae* strains based on polymorphisms at microsatellite loci and/or single nucleotide polymorphisms (SNPs). Single nucleotide variants in the coding region of FLO8, a key regulator of flocculation and pseudohyphae formation, were analyzed in a subset of Uruguayan wine strains. Polymorphism analysis at nine microsatellite loci (selected from 33 loci tested) was performed in a collection of 120 strains, mostly wine strains, from different origins. From a total of 184 different alleles scored, 50 were exclusive alleles that could identify 29 strains. Four selected microsatellite loci are located within or near genes of putative enological interest. The Uruguayan strains are highly diverse and evenly distributed in the phylogenetic reconstructions, suggesting an evolutionary history previous to human use. The *Saccharomyces cerevisiae* Microsatellites and SNPs Genotyping Database is presented (www.pasteur.edu.uy/yeast). Comparison of standardized results from strains coming from different settings (industrial, clinical, environmental) will provide a reliable and growing source of information on the molecular biodiversity of *S. cerevisiae* strains.

Introduction

In his early studies on alcoholic fermentation, Louis Pasteur already reported that each type of fermentation requires a certain type of microorganism (Pasteur, 1866). Different strains of the budding yeast *Saccharomyces cerevisiae* have been used for centuries in baking, distilling, brewing and wine making. Phenotypic variation among wine yeast isolates was recognized by winemakers long before being appreciated by geneticists. Under the same enological conditions, a grape juice can result in a good or bad-quality wine, depending on the *S. cerevisiae* strain that completes the fermentation. Therefore, most modern winemakers inoculate grape must with a pure culture of a selected *S. cerevisiae* strain to ensure a reliable and predictable fermentation process. Among many attributes, selected strains must tolerate high ethanol concentrations, flocculate

at the end of the fermentation, confer desirable aromas and color, and must not produce undesirable by-products like hydrogen sulfide.

In a completely different application, live *Saccharomyces boulardii*, now considered a strain of *S. cerevisiae* (McCullough *et al.*, 1998), has been used as a nutritional supplement and also as a probiotic agent for reequilibration of the intestinal flora (McFarland & Bernasconi, 1993).

Although *S. cerevisiae* is not considered a pathogen in healthy individuals, it is increasingly isolated from immunocompromised patients. The use of live *Saccharomyces* in the treatment of diarrhea in Europe has been linked to yeast sepsis (Piarroux *et al.*, 1999). Clinical yeast isolates are able to grow at higher temperatures (41 °C) than laboratory strains, and this characteristic has been correlated with their survival in mice (McCusker *et al.*, 1994). In *S. cerevisiae*, the physiological response to nutrient deprivation (carbon,

nitrogen or amino acid deprivation) results in changes in morphology and cell surface characteristics, switching from spherical or ovoid cells to filaments of invasive pseudohyphae. For fungal pathogens, such as *Candida albicans*, this dimorphic transition is strongly correlated with invasion of host tissues and virulence (Sudbery *et al.*, 2004).

Recently, the budding yeast *S. cerevisiae* has attracted interest as an emerging model in ecological and evolutionary genetics. *Saccharomyces cerevisiae* occupies and flourishes in numerous habitats that are not necessarily associated with human activities. Insects and birds are considered important agents for the dispersal of yeasts. It is well recognized that natural *S. cerevisiae* isolates, which are generally prototrophs, exhibit very large genotypic and phenotypic diversity, resulting in wide variations in their secondary metabolite production. Metabolic parameters generally vary in a complex, continuous way that can be attributed to a typical polygenic determinism associated to quantitative trait loci (QTL) (for a review see Landry *et al.*, 2006). Three QTLs that control *S. cerevisiae* sporulation efficiency were mapped to single-nucleotide resolution showing that the interaction of a few genetic variants (SNPs) can have a profound phenotypic effect (Deutschbauer & Davis, 2005). Many traits of industrial strains (ethanol production and tolerance, flocculation, production of desirable or undesirable aromas, etc) or clinical isolates (invasive growth) depend on the expression of multiple loci of variable phenotypic contribution.

Traditional morphological and biochemical tests are of limited value in revealing the genetic diversity of *S. cerevisiae* strains. In 2001, polymorphism analysis of selected microsatellite loci was proposed as a very powerful and unique method to discriminate *S. cerevisiae* at the strain level (González Tejera *et al.*, 2001; Hennequin *et al.*, 2001). Microsatellites or simple sequence repeats (SSRs) consist of direct tandem repeats of a short DNA motif, usually < 10 bp, that are hypervariable in length as a result of DNA-replication errors (Strand *et al.*, 1993). Thus, microsatellites show a substantial level of polymorphism between individuals of the same species and are extensively used for paternity exclusion tests (Helminen *et al.*, 1988), forensic medicine (Hagelberg *et al.*, 1991) and the molecular typing of different eukaryotic organisms, including cultivars of *Vitis vinifera* (This *et al.*, 2004) and the pathogen yeast *C. albicans* (Botterel *et al.*, 2001). Methods for the analysis of either microsatellites or SNPs require a previous knowledge of the sequence under study. The choice of a suitable set of polymorphic loci is also crucial for both methods. Microsatellites are particularly suitable for the detection of polyploids and have a higher discrimination power than nucleotide sequence-based methods such as multilocus sequence typing (MLST), particularly when closely related strains are compared (Ayoub *et al.*, 2006). In contrast to

microsatellites, SNPs have a low rate of recurrent mutation, making them stable indicators of evolutionary history. SNPs are increasingly used for linkage and biodiversity studies in all kinds of organisms (Wang *et al.*, 1998; Pearson *et al.*, 2004; Clark *et al.*, 2007), including yeasts (Ben-Ari *et al.*, 2005; Aa *et al.*, 2006).

After the entire *S. cerevisiae* genome was publicly available (Goffeau *et al.*, 1996), different computer searches for short tandem repeats were conducted (Field & Wills, 1998; Katti *et al.*, 2001; Aishwarya *et al.*, 2007). Recently, several high-throughput microsatellite polymorphism analyses have been performed (Legras *et al.*, 2005, 2007; Schuller & Casal, 2007). Although these studies served to confirm the level of polymorphism of several microsatellite loci, they represent stand-alone efforts, and it is still not possible to compare results from different groups or calculate allelic frequencies. The genetic diversity of strains isolated from very different settings, e.g. clinical cases, environmental studies or technological applications, cannot be compared. In this work we report the molecular diversity of *S. cerevisiae* strains using two powerful and complementary tools for discriminating individuals within any eukaryotic species: microsatellites and SNPs. We propose a standardization method to report data from microsatellite polymorphism analysis, and we present a database that aims to collect and standardize data from different laboratories.

Materials and methods

Yeast strains and media

Saccharomyces cerevisiae AB972 and S288C, two strains used in the sequencing project, are haploid, MAT alpha strains, and were used as standard DNAs of known sequence. The diploid strain BY4743, used for the *S. cerevisiae* genome deletion project, was also included. Several of the hydrogen sulfide (H₂S)-producing strains have been studied in detail by Linderholm *et al.* (2006). The Argentine native yeast strains have been described by Mercado *et al.* (2007). The complete list of strains used in this work is included as supplementary material (supplementary Table S1). All yeast strains were grown on YPD [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) glucose] medium for DNA isolation. All yeast strains were routinely grown at 30 °C, except when tested for growth at 41 °C. Synthetic low-ammonia dextrose (SLAD) plates were prepared as described (Gimeno *et al.*, 1992).

Quick preparation of DNA template for PCR

The pellet corresponding to around 10⁹ cells (early stationary phase) was washed with sterile water and resuspended in 0.4 mL breaking buffer (2% Triton X-100, 1% sodium dodecylsulphate, 10 mM NaCl, 10 mM Tris, pH 8 and 1 mM EDTA, pH 8). The cells were homogenized by

vortexing at high speed for 3 min with 0.3 g glass beads (Sigma G9268) in the presence of 0.4 mL phenol pH 8. Then, 0.4 mL TE (10 mM Tris pH 8 and 1 mM EDTA pH 8) was added and vortexed briefly. After centrifugation at 4 °C, the aqueous phase was carefully removed. DNA was ethanol-precipitated, centrifuged and resuspended in TE. DNA concentration and quality was estimated in 0.7% agarose gels and diluted to c. 10–20 ng in 5 µL for PCR reactions.

DNA was isolated from early stationary phase cultures started either from one isolated colony or from a streak. In all cases, these duplicated DNAs showed exactly the same PCR profile for all SSRs tested.

Analysis of SSR loci

Description of the 33 microsatellite loci analyzed is included as supplementary material (supplementary Table S2). The specific pairs of primers used for the nine selected polymorphic loci are shown in Table 1.

PCR amplifications were performed in a Thermo PXE 0.2 Thermal Cycler. YPL009C, YOR267C, TTA_XIII, YGL013C, YGL028C and AT_X SSR loci were amplified in 20 µL reactions consisting of 10–20 ng DNA, 200 µM of each dNTP, 2 µL of 10X PCR buffer minus Mg, 1 U Taq DNA polymerase, 2.5 mM MgCl₂ and 10 pmol of forward and reverse primers. Amplification was performed as follows: 5 min at 94 °C, 30 cycles of (30 s at 94 °C, 30 s at 55 °C, 1 min at 72 °C) and 5 min at 72 °C. To avoid excessive stuttering, TG_VI, GT_X and C4_XV SSR loci were amplified in 20 µL reactions consisting of 10–20 ng DNA, 200 µM of each dNTP, 2 µL of 10X PCR buffer minus Mg, 0.5 U Pfu DNA polymerase (Fermentas), 4 mM MgSO₄ and 10 pmol of forward and reverse primers. Amplification was performed as follows: 2 min at 95 °C, 25 cycles of (30 s at 95 °C, 30 s at 55 °C, 1 min at 70 °C) and 5 min at 72 °C. Amplification was confirmed by running an aliquot of the PCR reaction product in 2% agarose gels. DNA concentration was then adjusted, and 1/3 volume of denaturing dye solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol) was added. One to four microlitres of this mixture were denatured and electrophoresed in a

sequencing gel [6% polyacrylamide gel electrophoresis-Plus acrylamide (Fermentas) plus 7 M urea] and then silver stained according to the Promega Silver Staining Kit. Product sizes were determined by comparison with the known size of the amplified DNA from the sequenced reference strains, the 10-bp DNA marker (Invitrogen), and by mixing several of the different alleles observed. In some cases, the intrinsic stuttering of the microsatellite served as an internal allele ladder.

Absolute values for DNA microsatellite markers were calculated from GenBank database information and are the following: 239 bp for locus YOR267C, 207 bp for locus YPL009C, 238 bp for locus TTA_XIII, 229 bp for locus C4_XV, 274 bp for locus GT_X, 198 bp for locus TG_VI, 245 bp for locus YGL028C, 206 bp for locus AT_X and 228 bp for locus YGL013C. Both AB972 and S288C strains gave the same amplification product size for all SSRs tested, except for locus TG_VI, where one repeat difference was observed between the two strains. In this case, the absolute value of 198 bp was assigned to the amplification product of strain AB972 because this was the strain used to sequence chromosome VI (see 'FAQs about *S. cerevisiae*' in www.yeastgenome.org)

SNPs in FLO8

Primers used to amplify part of the coding region of FLO8 were as follows: YER109C-fwd, 5'-GCATGGCAACGAATAGTGA-3'; FLO8-fwd, 5'-CAGCAGCCTTTTGC TCAAGATG-3'; Flo8-rev, 5'-GTTCTGCATCGTGTGTAG CCTTG-3'.

A region of YER109C was amplified with two pairs of primers: FLO8-fwd and Flo8-rev or YER109C-fwd and Flo8-rev.

DNAs were amplified in 20 µL reactions consisting of 10–20 ng DNA, 200 µM of each dNTP, 2 µL of 10X PCR buffer minus Mg, 0.5 U Pfu DNA polymerase (Fermentas), 4 mM MgSO₄ and 10 pmol of forward and reverse primers. Amplification was performed as follows: 2 min at 95 °C, 35 cycles of (30 s at 95 °C, 30 s at 55 °C, 1 min at 70 °C) and 5 min at 72 °C. Amplification was confirmed by running an

Table 1. Nucleotide sequence of the pairs of primers used to amplify the nine selected SSR loci

SSR loci	Forward primer 5'–3'	Reverse primer 5'–3'
YPL009C	CTGCTCAACTTGTGATGGGTTTTGG	CCTCGTTACTATCGTCTTCATCTTGC
YOR267C	ATGACTGCAGCAATGAATCG	TCCTCTGTGCTGTTGACTCG
TTA_XIII	AAGCGTAAGCAATGGTGTAGA	AAGCCTCTTCAAGCATGACC
TG_VI	GACACAATAGCAATGGCCTTCA	ATATTGGCACCAGAAAGTGTGC
GT_X	TGCGCAGCTTAGTATGACCA	GATGGGCTTTCACTCCACTT
C4_XV	GAGAAAAATGCTGTTTATTCTGACC	CCTCCGGGACGTGAAATAAC
YGL013C	ATGCTCAGCAGCAAATACCC	GATAAACGTCGCTCCACAGG
YGL028C	GGCAACACTCAAAGTCAGCA	GCTGTTTCTGTGTGGATGA
AT_X	GCAACCTCCAAAGGAAATCA	CGACAGAGAGAGACCCAAGC

aliquot of the PCR reaction product in 1.5% agarose gels. PCR products were purified and both strands were sequenced at Macrogen (dna.macrogen.com). Only SNPs-confirmed sequencing of both strands are reported. The presence of heterozygote DNAs for a certain SNP was confirmed by visually inspecting the superimposed presence of two peaks of different colors for that position in readings from both strands.

Phylogenetic reconstructions

Only strains with different genotypes were included in the phylogenetic reconstructions. In order to compare strains with different ploidy, we considered alleles at different loci as independent characters and scored them in a presence/absence (0/1) manner. We estimated the Jaccard coefficient between pairs of strains and grouped them using Neighbor-Joining and unweighted pair-group method with arithmetic average cluster (UPGMA) algorithms using the PAST 1.74 software (Hammer *et al.*, 2001), rooting the resulting tree by the midpoint method. This method has, at least, three main drawbacks: it does not give the same weight to all loci, it leaves out the correlation between loci and their alleles, and it does not reflect the processes that bring about the differences between strains.

Restricting the analysis to the 82 different diploid strains, mainly wine strains, we computed the Cavalli-Sforza's chord distance (Cavalli-Sforza & Edwards, 1967), which assumes that all differences between strains arise only from genetic drift. We applied the Neighbor Joining algorithm in 1000 Bootstrap pseudo replicates with PHYLIP 3.67 Phylogeny Inference Package (<http://evolution.gs.washington.edu/phylip.html>), considering the sake strain as the outgroup. In all cases, we visualized the resulting trees with MEGA 4 software (Tamura *et al.*, 2007).

Results

This molecular-typing work was performed with a collection of 120 *S. cerevisiae* strains: 48 native wine strains [15 from Uruguay (URU), 28 from Argentina (ARG), five from UC Davis collection (UCD)], 50 commercial wine strains (COM), 12 strains producing different levels of H₂S (H2S), one commercial strain for Sake (SAKE), six commercial bread strains (PAN) and three laboratory strains (LAB) (see supplementary material for more details).

At present, almost all commercial wine strains available require the addition of ammonium salts to avoid the production of undesirable aromas. However, there is a growing tendency to avoid this practice because the negative effects of excess nitrogen (residual undesirable compounds, wine contamination during ageing, etc) have been well documented (Bell & Henschke, 2005). Several of the Uruguayan native strains included in this work were selected

as low nitrogen demanding strains (González Techera *et al.*, 2001).

Screening for new polymorphic SSR loci

Based on the computer search for SSRs performed by Katti *et al.* (2001) (publicly available at <http://www.ncl-india.org/ssr/>) on the completely sequenced genome of *S. cerevisiae*, we chose and tested the polymorphism of several loci. We performed this initial screening with a representative subset of 23 strains used for different applications and representing different geographical origins (strains in bold and underlined in supplementary Table S1). The criteria for the selection of SSRs to be tested were the following: (1) at least one SSR per chromosome, (2) perfect motifs and long SSRs were preferred, (3) motifs rich in As or Ts that tend to give stuttering were avoided, (4) SSRs present within or near genes of putative enological interest were chosen. Loci previously reported as polymorphic were also included. Primers around 20 bp long were designed with the online PRIMER3 software so as to generate PCR product sizes within 200–250 bp and a unique annealing temperature for all PCR reactions of 55 °C. High-throughput PCR amplification and gel electrophoresis analysis was thus simplified. Criteria for naming the SSR loci were the following: (1) the normalized name of the ORF (according to www.yeastgenome.org) if the repeat was located in a coding region and (2) the motif and the chromosome number if the repeat was a perfect repeat on a noncoding region. C4_XV is a compound repeat (containing different triplet repeats) in chromosome XV and therefore we kept the name proposed by Legras *et al.* (2005). The precise position in base pairs, indicating the chromosome and strand, unambiguously identifies the SSR. Details of the 33 tested loci together with the number of alleles and heterozygosity observed in this subset of 23 strains are included as supplementary material (Supplementary Table S2). Only nine loci were polymorphic out of the 33 tested, within this reference strain population.

Polymorphism analysis at nine selected SSRs

Details of the nine selected SSRs that were analyzed with the collection of 120 strains are included in Table 2. Microsatellites that have been previously used by us or others (González Techera *et al.*, 2001; Legras *et al.*, 2005) were renamed following the standardized naming criteria described above. After obtaining the raw data in bp for the alleles present in the 120 strains for each SSR, we initially thought it would be easier to exchange results among different laboratories if alleles were expressed as number of repeats. In this way, the results would be independent of the primers used. However, when we started to carefully define the number of repeats present in the reference sequenced strain, in some cases the definition was clear cut, but in

Table 2. Name, definition and location of selected SSRs

SSR name	Other names	Chromosome/strand	Start position	End position	SSR present in sequenced strain
TTA_XIII	SCAAT1 SCPTSY7	XIII/Watson	86 953	87 057	(TTA) ₃₅
TG_VI	C5	VI/Watson	210 333	210 394	(TG) ₃₁
GT_X	C11	X/Watson	519 249	519 288	(GT) ₂₀
C4_XV	C4	XV/Watson	110 776	110 865	(ATA) ₅ CTA(ATA) ₉ (GTA) ₇ ATA(ATG) ₇ 30 triplets
YPL009C	SC8132X	XVI/Crick	536 776	536 705	(GAA) ₇ GAG(GAA) ₁₆ 24 triplets Glu
YOR267C	–	XV/Crick	822 985	822 926	(CAA) ₂₀
YGL013C	–	VII/Crick	469 276	469 247	AAT(AAC) ₈ AAT 10 triplets Asn
YGL028C	–	VII/Crick	442 254	442 147	(SerSerSerThr) ₉
AT_X	–	X/Watson	456 060	456 097	(AT) ₁₉

others it was uncertain. For instance, we assigned 30 triplets for the compound SSR C4_XV, but if this SSR is searched for using another search engine, like the one recently provided in EuMicroSatdb (Aishwarya *et al.*, 2007) (<http://www.veenuash.info/web/intro.htm>), the recognized repeats turn out to be different: (TAA)₉(TAG)₇. In the case of SSR TTA_XIII, the number of repeats is 35 if the repeated motif is TTA or 36 if the repeated motif is considered to be TAT. When the SSR was located within a coding region, we counted all the consecutive motifs that coded for the same amino acid.

In an effort to standardize the reporting of results independently of either the primers or the definition of the number of repeats present in the sequenced reference strain, we realized that the best option would be to report the difference in the number of repeats relative to the reference sequenced strain. Thus, the allele sizes were expressed as the difference in the number of repeats present in the locus under study relative to the sequenced reference strain, e.g. +4 if the strain under study showed four repeats more than the sequenced strain analyzed at that locus or –7 if seven repeats less than the reference strain were observed (see Fig. 1). The results for all the strains and loci analyzed can be found in the database (www.pasteur.edu.uy/yeast). To confirm the correct assignment of alleles, we constructed allele ladders for each locus by mixing together different amplified DNAs representing all or almost all of the alleles observed. Running these allele ladders in parallel with the samples under study allowed a precise estimation of the difference in the number of repeats relative to the sequenced reference DNA. An example of a typical allele ladder is shown in Fig. 1.

A total of 184 different alleles were scored, and 11 strains turned out to be aneuploids. All the commercial bread strains and the strain used for the production of sake showed three to four alleles for most loci. Most of the wine strains were diploids, and a few were aneuploid. A summary of the results is presented in Table 3. These nine SSR loci allowed the discrimination of 93 strains out of the 120 analyzed. There were 50 exclusive alleles that could identify 29 strains. The strain showing the highest number of exclusive alleles (6) was the sake strain, followed by URU12,

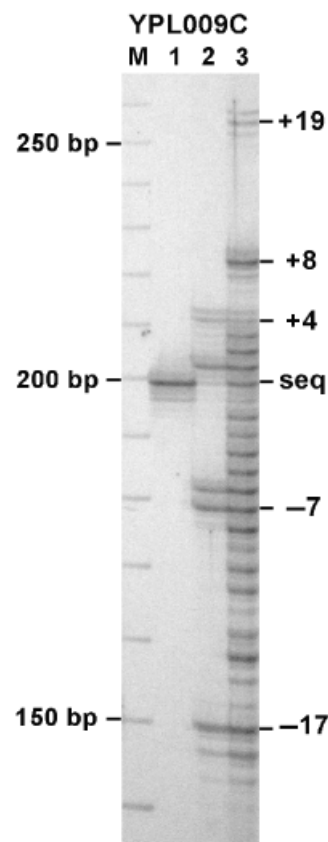


Fig. 1. Allelic diversity at microsatellite locus YPL009C. M stands for the 10-bp DNA molecular marker. (1) SC288C PCR product for this locus (seq), (2) partial allele ladder for locus YPL009C, (3) amplification products representing all the alleles observed at locus YPL009C, reported as the difference in the number of repeats relative to the sequenced reference DNA.

with four exclusive alleles. A detailed list of exclusive alleles and strains identified is included as supplementary material (Supplementary Table S3). All data obtained and the calculated allelic frequencies for the nine SSRs are included in the database. DNAs used for the allele ladders are available on request.

Table 3. Summary of results obtained analyzing 120 strains with nine selected SSRs

	TTA_XIII	TG_VI	GT_X	C4_XV	YPL009C	YOR267C	YGL013C	YGL028C	AT_X
# alleles	30	34	21	24	22	22	16	7	8
#excl.allele	8	8	5	10	5	7	5	0	2
Exc/total	0.27	0.24	0.24	0.42	0.23	0.32	0.31	0	0.25
Max. allele	+12	+34	+24	+34	+19	+41	+14	+3	+2
Min. allele	−23	−25	−9	−3	−18	−9	−5	−3	−8
Ho	0.35	0.45	0.43	0.27	0.45	0.23	0.40	0.28	0.29
PIC	0.929	0.925	0.903	0.887	0.864	0.862	0.769	0.577	0.444

alleles stands for the total number of alleles observed, #excl. allele is the total number of alleles present exclusively in certain strain, Exc/total is the fraction of exclusive alleles from the total observed alleles. Max and Min alleles are reported as the difference in the number of repeats relative to the sequenced reference strain. Ho is the observed heterozygosity and PIC the polymorphism information content.

Tree topologies obtained from UPGMA and Neighbor Joining algorithms, which compared all strains, gave similar results. The characteristic pattern of microsatellite data was evident, with deep branches and low definition at the basal nodes (see supplementary Fig. S1). The most differentiated strains are those used for bread, followed by sake and laboratory strains. However, in both cases, one wine strain was also differentiated from the rest. Besides, none of the categories of strains have a unique origin, except for a subgroup of H₂S producers.

Restricting the analysis to the 82 diploid strains reproduced the above pattern but also showed that two groups of Argentine strains are monophyletic (Fig. 2).

Four polymorphic SSRs are within or near genes of putative enological interest

The four polymorphic SSRs YGL013C, YGL028C, AT_X and TG_VI are located within or near genes of putative enological interest. We defined 'near' as <3 kb because on average 1 cM corresponds to 3 kb for *S. cerevisiae*.

YGL013C is next to *ERG4*/YGL012W, but on the opposite strand. *ERG4* encodes the enzyme sterol C-24(28) reductase, which catalyzes the final step in ergosterol biosynthesis (Zweytick *et al.*, 2000). Ergosterol is an essential component of yeast cells that maintains the integrity of the membrane. High concentrations of ergosterol have been correlated to high ethanol tolerance (Wu *et al.*, 2006). Testing the subgroup of Uruguayan native strains, which include representatives of high (>14%) and low (<12%) ethanol tolerance, we could not demonstrate a possible association of this molecular marker to ethanol tolerance (results not shown).

The repeat present within *SCW11*/YGL028C is considered a minisatellite, rather than a microsatellite, because the repeated motif has 12 bases. *SCW11* codes for a cell wall protein with similarity to glucanases. A null mutant in *SCW11* is viable but exhibits defects in separation after division and displays flocculant growth (Cappellaro *et al.*,

1998). The observed alleles in the population are shown in Fig. 3. We sequenced homozygote DNAs with three motifs less and two motifs more than the sequenced reference DNA and confirmed that the variation corresponded to changes in four codons coding for the amino acids SerSerSerThr; therefore, the repeated motif is best defined as SSST. Furthermore, we verified that in all sequenced DNAs, there were four different groups of codons coding for the same amino acid motif SSST:

- (1) TCG TCC TCT ACG
- (2) TCG TCC TCT ACT
- (3) TCT TCT TCT ACT
- (4) TCT TCC TCT ACT

Motif types (1) and (2) remained the same for all sequenced DNAs whereas variations in the number of motif types (3) and (4) accounted for the size differences observed.

The largest allele with 12 motifs (SSST) was not observed in homozygosis. Flocculation assays of strains carrying the different alleles observed in homozygosis did not result in evident phenotypic differences (results not shown).

The production of H₂S is an undesirable sensory characteristic in the wine, beer and sake industries. Yeast strain background has a strong influence on H₂S production in wine strains of *S. cerevisiae* (Spiropoulos *et al.*, 2000).

The TG_VI SSR (named as C5 by Legras *et al.*, 2005) is located at around 3 kb from *MET10*/YFR030W. *MET10* codes for the subunit alpha of assimilatory sulfite reductase, which is responsible for the conversion of sulfite into sulfide (Hansen & Kielland-Brandt, 1996). The AT_X SSR is located at 135 bp from *MET3*/YJR010W. *MET3* codes for ATP sulfurylase, the enzyme that catalyzes the primary step of intracellular sulfate activation, ATP sulfurylase is essential for the assimilatory reduction of sulfate to sulfide (Cherest *et al.*, 1985). All of the H₂S-producing strains (H₂S in the dendograms) show the same allele in homozygosis for AT_X SSR. There is also one predominant allele present in homozygosis for TG_VI in this subset of strains. However, the alleles present in these 12 strains for these two loci next

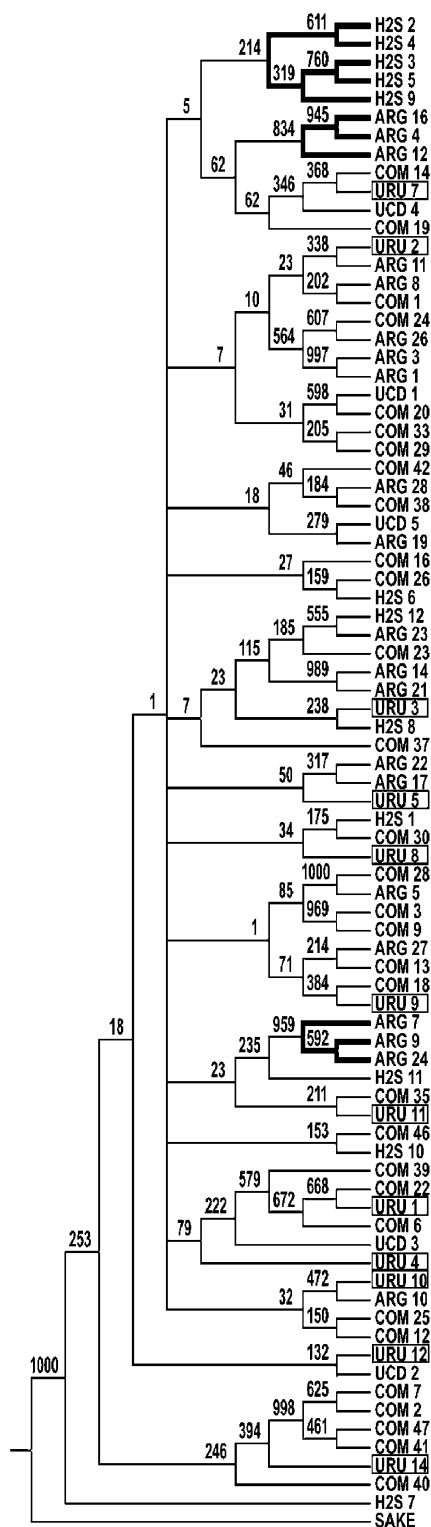


Fig. 2. Dendrogram of diploid *Saccharomyces cerevisiae* strains. Neighbour-joining tree constructed from the chord distance between yeast strains based on the polymorphism at nine microsatellite loci and rooted considering Sake strain as outgroup. Numbers above nodes are times of occurrence of each node after 1000 pseudoreplicates of bootstrap.

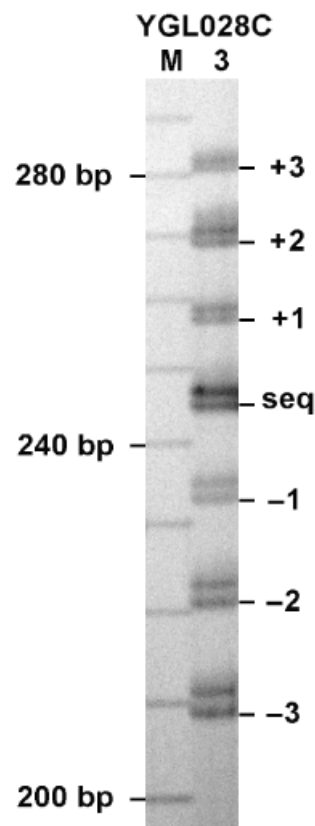


Fig. 3. Allelic diversity at microsatellite locus YGL028C. M stands for the 10-bp DNA molecular marker. (3) Amplification products representing all the alleles observed at locus YGL028C, reported as the difference in the number of repeats relative to the sequenced reference DNA (seq).

to *MET* genes are not exclusive to this subgroup and are also present in other strains of our collection.

SNPs in *FLO8*, a key regulator of flocculation and pseudohyphal growth

FLO8 is a key transcription factor required for flocculation and filamentous growth. The genome reference strain (AB972 or S288C) and most laboratory strains have a null mutation in this gene and a nonflocculant phenotype (Liu *et al.*, 1996). If the first nucleotide of the coding sequence is given number 1, this null mutation consists of the presence of an A in position 425 of *FLO8*, generating a stop codon at the amino acid level. We designed primers to amplify a 300-bp region encompassing this known null mutation site and sequenced both strands in a subset of Uruguayan native wine strains, AB972 (reference haploid strain), BY4743 (diploid strain used for the *S. cerevisiae* deletion project) and the very flocculant commercial strain AWRI 350. The polymorphisms observed are shown in Table 4.

Filamentous growth was tested on SLAD plates (see Fig. 4). Although pseudohyphae formation and agar invasion

Table 4. SNPs observed in *FLO8/YER109C*

Strains	SNP position in <i>FLO8/YER109C</i>	Nucleotide	Codon	Amino acid
AB972, BY4743	306	C	CAC	His
FQU 99/5	306	T/C	CAT/CAC	His
AB972, BY4743	315	A	AGA	Arg
CP KU1	315	G/A	AGG/AGA	Arg
AB972, BY4743	334	G	GTC	Val
CP 873, CP 874, CP 881, CP 882, FQU 02/16, UY-04, AWRI 350	334	A	ATC	Ile
CP KU1, FQU 99/5, M522	334	G/A	GTC/ATC	Val/Ile
AB972, BY4743	425	A	TAG	STOP
CP 873, CP 874, CP 881, CP 882, CP KU1, FQU 02/16, UY-04, FQU 99/5, AWRI 350, M522	425	G	TGG	Trp
AB972, BY4743	449	A	AAT	Asn
CP 882	449	A/G	AAT/AGT	Asn/Ser

The position in bp (number 1 is given to the first nucleotide of the coding sequence), nucleotide present and codon/amino acid resulting is indicated. Both alleles are indicated for heterozygote DNAs.

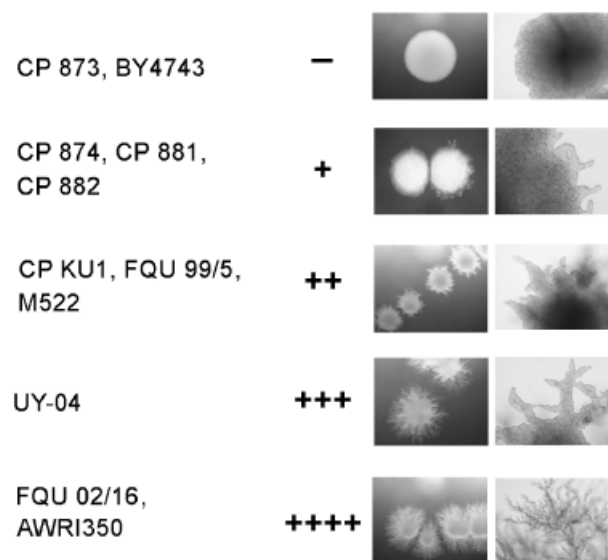


Fig. 4. Pseudohyphal growth on synthetic low-ammonia dextrose plates. Strains were streaked on SLAD media poured onto sterile microscopic slides and incubated at 30 °C for 3–4 days. Different levels of pseudohyphal growth are indicated: from no pseudohyphae (–) to formation of pseudohyphal mats (++++). Photographs were taken from representative colonies (left column, $\times 20$). A closer look at the colony borders is shown (right column, $\times 200$).

were evident in some cases, none of this subset of strains could grow on YPD at 41 °C, as has been reported for clinical isolates.

***Saccharomyces cerevisiae* microsatellites and SNPs genotyping database**

The *S. cerevisiae* microsatellites and SNP Genotyping Database was conceived to gather and search information of

microsatellites (or SSRs) and SNPs alleles from different strains. The aim of this interactive site is to provide a reliable and growing source of information on the biodiversity of *S. cerevisiae* strains isolated from different geographical origins (for applications as diverse as the production of wine, beer, bread, sake, bioethanol, probiotics, biofertilizers, etc), from clinical cases or from environmental samples. As more and more strains are characterized world-wide, it is imperative to provide a means to exchange and add up results from different laboratories. A convenient spreadsheet to transform raw data to differences in the number of repeats relative to the sequenced strain is provided. The inclusion of internal controls with reference DNAs will serve to validate results from different laboratories. The DNAs used to construct the allele ladders for each locus are available on request. If researchers add their results to this site, the database would serve the following purposes:

- (1) identification of new strains of *S. cerevisiae*,
- (2) identification of possible synonyms in strain collections,
- (3) identification of possible synonyms in probiotic and clinical isolates,
- (4) calculation of allelic frequencies,
- (5) genotypic combination of alleles,
- (6) assessment of geographical biodiversity and
- (7) association of molecular markers to traits of interest.

The database is organized with a user-friendly interface. Three main search entries are provided: by strain, by SSR or by SNP. Strains can be searched by their given names, geographical origin or application. In the SSR search, allele(s) observed for different loci can be entered, and all the matching strains will be displayed. If entered data is not found in the database, the user is prompted to submit it as new data subjected to validation. Only SNPs in *FLO8* (this work), *CYS4* and *MET6* (Linderholm *et al.*, 2006) for a

small subset of strains are included for the moment. Links to information in published literature or other websites are provided. The present database (Beta version) includes data from nine different SSRs and three different SNP loci. Data for these loci from as many different strains as possible are welcomed. In the future, we look forward to including data from other polymorphic SSRs and SNPs as well. The database will be continuously updated and improved, responding to the input and suggestions of the yeast community.

Discussion

Molecular evidence for the presence of *S. cerevisiae* in wine fermentation dates back to 3150 BC, from pottery jars found in the tomb of the King Scorpion I of Egypt. However, several lines of evidence suggest that the evolutionary history of *S. cerevisiae* strains is previous to human use and spans millions of years (Blair *et al.*, 2005; Landry *et al.*, 2006).

Although microsatellites have become extremely popular molecular markers, little is known about the role of microsatellites in genome organization, gene regulation, quantitative genetic variation and evolution of genes. Some SSRs are highly polymorphic while others show very little or no variation within individuals. In our initial screening, we discarded three SSRs (YDR289C, YKL172W and YLR177W), which have been claimed as useful to discriminate probiotic from clinical isolates (Hennequin *et al.*, 2001; Malgoire *et al.*, 2005) due to their low discriminative power in our analyzed population. In agreement with Legras *et al.* (2007), we have found that TTA_XIII, TG_VI, GT_X, C4_XV, YPL009C and YOR267C SSRs show the highest discriminative power. Polymorphisms in YGL013C, YGL028C and AT_X SSRs have not been reported previously. Our microsatellite polymorphism analysis also indicates that sake and bread yeast strains are highly differentiated from most wine strains, in accordance with their technological origins. Nevertheless, this conclusion is taken from a distance which has several drawbacks, as was pointed out in the Materials and methods section of this article. To be conclusive, a suitable distance to compare organisms with different ploidy would need to be developed.

On the other hand, phylogenetic trees from diploid strains suggest that, in general, native strains – Uruguayan (URU), Californian (UCD) and some Argentine (ARG) strains – have a multiple origin. In fact, this origin must have been previous to a possible migration of these strains to America. The American native strains may proceed from European strains or may have been recruited *in situ*. This second possibility would be contradictory to the hypothesis proposed by Legras *et al.* (2007). However, our analysis does not include the proper geographic representation of strains to demonstrate an *in situ* recruitment. If we considered

the American native strains as strains of ‘European origin’, our results would be compatible with the proposal of Legras *et al.* (2007).

The opposite situation is observed with a subset of H₂S producers and two groups of Argentine strains, between which a close relationship is suggested, with high bootstrap values for all of them (see Fig. 2). Different methods confirmed this relationship for the subgroup of H₂S producers. Interestingly, this result suggests a differentiation *in situ* of some Argentine strains.

In the *S. cerevisiae* genome, the majority of the genes containing intragenic minisatellites encode cell wall proteins. Variations in the number of intragenic repeats could provide the functional diversity of cell surface antigens that, in fungi and other pathogens, allows rapid adaptation to the environment and/or elusion of the host immune system. In the same genetic background, Verstrepen *et al.* (2005) could demonstrate that size variations in an intragenic minisatellite in *FLO1* create quantitative alterations in phenotypes like adhesion, flocculation or biofilm formation. Size variations for the minisatellite in *SCW11/YGL028C* were restricted to only seven alleles in this population, and the largest allele was not found in homozygosis. This limited variation may reflect a functional restriction for the possible protein variants. Variations in minisatellites have been proposed as a method to characterize wine yeast strains (Marinangeli *et al.*, 2004). Although the discrimination power of these loci is not high enough, they can be useful markers because amplification is robust and, in most cases, the allele size differences can be checked on agarose gels.

Analyzing DNAs from eight Uruguayan native strains we found five SNPs in a 300 bp coding region of *FLO8*. Deutschbauer & Davis (2005) found a polymorphism, on average, every 173 bp between DNAs from two strains with high divergence in sporulation efficiencies. These eight URU native strains studied are very diverse, as evidenced by their different ethanol tolerance, flocculation or pseudohyphae formation and confirmed by microsatellite and SNP molecular analysis.

DNAs from all the wine yeast strains tested (eight Uruguayan native and two commercial strains) showed the same two SNPs: a G at position 425 (resulting in a Trp codon) and an A or heterozygous A/G at position 334 (resulting in a Val or Ile codon). DNAs with SNPs in all the other YER109C positions (306, 315 and 449) coded for either different codons of the same amino acid or different amino acids of the same chemical groups. Val and Ile are both nonpolar amino acids, while Asn and Ser are amino acids with charged polar side chains.

No straightforward correlation could be established between pseudohyphal growth, flocculation and SNPs in *FLO8*. CP 882 is a highly flocculant strain when compared with CP 881 (González Techera *et al.*, 2001), but both show

very few pseudohyphae in SLAD plates. AWRI350 is a highly flocculant strain compared with FQU 02/16, and both show mats of pseudohyphae invading the agar. The three strains (CP KU1, FQU 99/5 and M522) with intermediate levels of pseudohyphal growth showed heterozygote DNAs in the *FLO8* region analyzed. Cell-cell adhesion (flocculation) and adhesion to abiotic surfaces or to tissues are properties of medical and industrial relevance. At least three different signaling pathways (some requiring the positive regulator Flo8) are activated in response to stress, nutrient limitation or small signaling molecules. Adhesion is a complex response controlled by integrated pathways working together (Verstrepen & Klis, 2006). A comparative genomics analysis within individuals of *S. cerevisiae* analyzed at many loci could eventually serve to pinpoint the crucial genotypical differences underlying complex phenotypes.

Nowadays, one of the greatest challenges for geneticists is the dissection of complex quantitative genetic variation into genes at the molecular level. Most traits of biotechnological interest in *S. cerevisiae* strains are complex traits that depend on multiple genes and their allelic variants. Codominant molecular markers like SSRs and SNPs are widely used for the molecular discrimination of individuals within eukaryotic species, for biodiversity studies, QTL mapping and linkage studies. In plants, molecular markers are used for marker-assisted introgression of favorable alleles in breeding programs (Andersen & Lübberstedt, 2003). In plant breeding, introgression is a common and effective practice to improve specific traits in an already good accession called 'elite'. Recently, this approach has been used to construct industrial yeast strains (Marullo *et al.*, 2007). Also, 260 SSR markers across the 16 yeast chromosomes were recently developed to discriminate two strains with extreme phenotypes and to genetically dissect the QTL regions responsible for ethanol tolerance in *S. cerevisiae* (Hu *et al.*, 2007). *Saccharomyces cerevisiae* provides an ideal framework for QTL analysis due to its high recombination rate, its richly annotated genome, and the fact that genes can be directly manipulated in their genomic context.

Sequencing projects of entire genomes of *Saccharomyces* yeast strains are in progress and will certainly add very valuable information about SNPs and probably will also serve to choose a set of suitable genes for MLST. However, this 'brute force' effort will not substitute or invalidate microsatellite typing. SSR typing is a cheap and accessible method that has the following unique features compared with SNPs (or MLST):

- (1) SSRs give clearcut information on ploidy levels. Many industrial strains are aneuploids or polyploids, and this has been associated with an adaptation mechanism (Querol *et al.*, 2003).
- (2) SSRs can be easily adapted to a simple method (using agarose gels) to monitor *S. cerevisiae* strains during

alcoholic fermentation (Howell *et al.*, 2004) and to detect the presence of *S. cerevisiae*-*Saccharomyces bayanus* hybrids (Masneuf-Pomarède *et al.*, 2007).

- (3) The reason why some SSRs are highly polymorphic while others are invariable is still an open question. Variation in the efficiency of DNA mismatch repair at different sites in the yeast genome has been proposed as a possible explanation (Hawk *et al.*, 2005). An assessment of SSR instability (an important phenomenon in cancer development) could also be a by-product of this database.

- (4) For closely related *S. cerevisiae* strains, MLST has proven to be less discriminatory than SSRs (Ayoub *et al.*, 2006).

- (5) Precise estimation and comparison of genetic variation among populations requires a large number of SNP relative to microsatellites because microsatellite loci typically have many alleles (more than 30 for *S. cerevisiae*), whereas two is the norm for SNP loci. Ascertainment bias in SNPs identification can also be a serious issue for studies of population structure since it has the potential to introduce systematic bias in estimates of variation within and among populations (Morin *et al.*, 2004).

There is an increasing need for standardization in the reporting of results from different laboratories as more *S. cerevisiae* strains and SSR markers are being tested. The discrimination power of the selected SSRs depends on the population of strains analyzed and, therefore, it would be very valuable information to be able to calculate allelic frequencies from strains coming from industrial, clinical or environmental settings. At present, it is not possible to extrapolate microsatellite data from different laboratories. Sizing with ladders, containing many or all of the observed alleles for a given SSR locus, is a common practice when analyzing human microsatellites and it certainly allows comparison of data after careful validation procedures (see 'Genetic Identity' at www.promega.com). The standard in humans is to report alleles as the absolute number of repeats. Only a small core set of loci have been selected and commercial kits providing premixed primers and allelic ladders are available. Because all users work with the same primers, these allelic ladders can be used to calibrate PCR product sizes to SSR repeat number for genotyping purposes (Butler, 2007). However, in some cases, there is still the need to reach a consensus on the definition of the core repeat structure to prevent confusion and allow a comparison of results between laboratories (see 'Comment on nomenclature for STR alleles and repeat structure' at <http://www.cstl.nist.gov/biotech/strbase/intro.htm>).

The comparability of microsatellite profiles obtained in different laboratories was exhaustively studied by This *et al.* (2004) in an effort to develop a standardized method for the identification of grape cultivars. Ten laboratories in seven countries analyzed 46 grape cultivars at six SSR loci,

and no effort was made to standardize equipment or protocols. All the participants used the same DNAs and the same primers but different enzymes, PCR programs and manual or automatic sequencing. The comparison of absolute allele sizes was impossible because there were discrepancies of up to 3 bp. The data coded as size differences between the smallest allele observed and the allele of the cultivar under study were more consistent but still not satisfactory. Even the automatic scoring of peak sizes can produce artificial shifts due to the simple algorithms used for automatically rounding up fragment sizes. Highest data consistency was obtained when every laboratory used the same selected cultivar-specific fragments as internal size standards, comprising a relatively complete allelic ladder for each of the six microsatellite loci.

We propose a similar standardization method based on the inclusion of the DNA from sequenced *S. cerevisiae* strains (AB972 and/or S288C) and the usefulness of internal allelic ladders to unequivocally express the results as the difference in the number of repeats relative to the sequenced reference strain. We offer the distribution of selected DNAs, rather than allele ladders, to allow for the use of different platforms (manual or automatic), different primers and/or multiplexing. Every laboratory can create its own allele ladder; the only prerequisite to ensure validation of results is that the same reference sequenced DNA is included plus three to four selected DNAs (which should be the same for all laboratories) that will cover the whole range of allele sizes. The exchange of DNAs prevents confusion associated with the names given to strains. The comparison of data from many strains and different SSRs will allow the selection of a minimal set of robust and highly polymorphic markers with clear fragment patterns. If researchers used the same pairs of primers giving robust amplification products, it would be possible to exchange allele ladders for manual or automatic platforms.

Although DNAs showing different microsatellite patterns surely correspond to different strains, if two or more DNAs show the same pattern, we can only say that we are unable to discriminate them with the analyzed microsatellite loci. Analysis of a higher number of polymorphic loci might be needed to discriminate closely related individuals. However, in some cases, analysis of the strain history reveals that the same strain was transported from one collection to another and different names were assigned to the same strain.

Confirmation of the presence of exclusive alleles for certain strains could eventually serve for identification purposes. A bar-coding method based on microsatellites could be used to identify industrial strains.

The future usefulness of the database presented here as a Beta version will depend on the scale of validated results incorporated by all researchers working with strains of *S. cerevisiae* from different settings. If the database is welcomed

by the *S. cerevisiae* community, regular updates and improvements will certainly be needed. Comparison of several hundreds or thousands of strains with a large number of polymorphic SSRs and/or SNPs in many genes will eventually allow an association of molecular markers to complex phenotypic traits.

Acknowledgements

This research work was funded by the Program for Technological Development PDT32/06, Dinacyt, Uruguay and Pedeciba Química, Uruguay. We acknowledge Linda Riles for sending AB972 yeast strain. We are very grateful to Lucy Joseph, Perrine Languet and Mariana Combina for sending us strains and information. We thank Mario Lalinde for visual art services.

References

- Aa E, Townsend JP, Adams RI, Nielsen KM & Taylor JW (2006) Population structure and gene evolution in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **6**: 702–715.
- Aishwarya V, Grover A & Sharma PC (2007) EuMicroSatdb: a database for microsatellites in the sequenced genomes of eukaryotes. *BMC Genomics* **8**: 225.
- Andersen JR & Lubberstedt T (2003) Functional markers in plants. *Trends Plant Sci* **8**: 554–560.
- Ayoub MJ, Legras JL, Saliba R & Gaillardin C (2006) Application of multi locus sequence typing to the analysis of the biodiversity of indigenous *Saccharomyces cerevisiae* wine yeasts from Lebanon. *J Appl Microbiol* **100**: 699–711.
- Bell SJ & Henschke PA (2005) Implications of nitrogen nutrition for grapes, fermentation and wine. *Aust J Grape Wine Res* **11**: 242–295.
- Ben-Ari G, Zenvirth D, Sherman A, Simchen G, Lavi U & Hillel J (2005) Application of SNPs for assessing biodiversity and phylogeny among yeast strains. *Heredity* **95**: 493–501.
- Blair JE, Shah P & Hedges SB (2005) Evolutionary sequence analysis of complete eukaryote genomes. *BMC Bioinform* **6**: 53.
- Botterel F, Desterke C, Costa C & Bretagne S (2001) Analysis of microsatellite markers of *Candida albicans* used for rapid typing. *J Clin Microbiol* **39**: 4076–4081.
- Butler JM (2007) Short tandem repeat typing technologies used in human identity testing. *Biotechniques* **43**: Sii–Sv.
- Cappellaro C, Mrsa V & Tanner W (1998) New potential cell wall glucanases of *Saccharomyces cerevisiae* and their involvement in mating. *J Bacteriol* **180**: 5030–5037.
- Cavalli-Sforza LL & Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution* **21**: 550–570.
- Cherest H, Nguyen NT & Surdin-Kerjan Y (1985) Transcriptional regulation of the *MET3* gene of *Saccharomyces cerevisiae*. *Gene* **34**: 269–281.
- Clark RM, Schweikert G, Toomajian C et al. (2007) Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* **317**: 338–342.

- Deutschbauer AM & Davis RW (2005) Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat Genet* **37**: 1333–1340.
- Field D & Wills C (1998) Abundant microsatellite polymorphism in *Saccharomyces cerevisiae*, and the different distributions of microsatellites in eight prokaryotes and *S. cerevisiae*, result from strong mutation pressures and a variety of selective forces. *Proc Natl Acad Sci USA* **95**: 1647–1652.
- Gimeno CJ, Ljungdahl PO, Styles CA & Fink GR (1992) Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* **68**: 1077–1090.
- Goffeau A, Barrell BG, Bussey H *et al.* (1996) Life with 6000 genes. *Science* **274**: 546, 563–567.
- González Techera A, Jubany S, Carrau FM & Gaggero C (2001) Differentiation of industrial wine yeast strains using microsatellite markers. *Lett Appl Microbiol* **33**: 71–75.
- Hagelberg E, Gray IC & Jeffreys AJ (1991) Identification of the skeletal remains of a murder victim by DNA analysis. *Nature* **352**: 427–429.
- Hammer Ø, Harper DAT & Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron* **4**: 9.
- Hansen J & Kielland-Brandt MC (1996) Inactivation of MET10 in brewer's yeast specifically increases SO₂ formation during beer production. *Nat Biotechnol* **14**: 1587–1591.
- Hawk JD, Stefanovic L, Boyer JC, Petes TD & Farber RA (2005) Variation in efficiency of DNA mismatch repair at different sites in the yeast genome. *Proc Natl Acad Sci USA* **102**: 8639–8643.
- Helminen P, Ehnholm C, Lokki ML, Jeffreys A & Peltonen L (1988) Application of DNA “fingerprints” to paternity determinations. *Lancet* **1**: 574–576.
- Hennequin C, Thierry A, Richard GF, Lecointre G, Nguyen HV, Gaillardin C & Dujon B (2001) Microsatellite typing as a new tool for identification of *Saccharomyces cerevisiae* strains. *J Clin Microbiol* **39**: 551–559.
- Howell KS, Bartowsky EJ, Fleet GH & Henschke PA (2004) Microsatellite PCR profiling of *Saccharomyces cerevisiae* strains during wine fermentation. *Lett Appl Microbiol* **38**: 315–320.
- Hu XH, Wang MH, Tan T, Li JR, Yang H, Leach L, Zhang RM & Luo ZW (2007) Genetic dissection of ethanol tolerance in the budding yeast *Saccharomyces cerevisiae*. *Genetics* **175**: 1479–1487.
- Katti MV, Ranjekar PK & Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol Biol Evol* **18**: 1161–1167.
- Landry CR, Townsend JP, Hartl DL & Cavalieri D (2006) Ecological and evolutionary genomics of *Saccharomyces cerevisiae*. *Mol Ecol* **15**: 575–591.
- Legras JL, Ruh O, Merdinoglu D & Karst F (2005) Selection of hypervariable microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains. *Int J Food Microbiol* **102**: 73–83.
- Legras JL, Merdinoglu D, Cornuet JM & Karst F (2007) Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* **16**: 2091–2102.
- Linderholm AL, Olineka TL, Hong Y & Bisson LF (2006) Allele diversity among genes of the sulfate reduction pathway in wine strains of *Saccharomyces cerevisiae*. *Am J Enol Vitic* **57**: 431–440.
- Liu H, Styles CA & Fink GR (1996) *Saccharomyces cerevisiae* S288C has a mutation in FLO8, a gene required for filamentous growth. *Genetics* **144**: 967–978.
- Malgoire JY, Bertout S, Renaud F, Bastide JM & Mallie M (2005) Typing of *Saccharomyces cerevisiae* clinical strains by using microsatellite sequence polymorphism. *J Clin Microbiol* **43**: 1133–1137.
- Marinangeli P, Angelozzi D, Ciani M, Clementi F & Mannazzu I (2004) Minisatellites in *Saccharomyces cerevisiae* genes encoding cell wall proteins: a new way towards wine strain characterisation. *FEMS Yeast Res* **4**: 427–435.
- Marullo P, Yvert G, Bely M, Aigle M & Dubourdieu D (2007) Efficient use of DNA molecular markers to construct industrial yeast strains. *FEMS Yeast Res* **7**: 1295–1306.
- Masneuf-Pomarède I, Le Jeune C, Durrens P, Lollier M, Aigle M & Dubourdieu D (2007) Molecular typing of wine yeast strains *Saccharomyces bayanus* var. *uvarum* using microsatellite markers. *Syst Appl Microbiol* **30**: 75–82.
- McCullough MJ, Clemons KV, McCusker JH & Stevens DA (1998) Species identification and virulence attributes of *Saccharomyces boulardii* (nom. inval.). *J Clin Microbiol* **36**: 2613–2617.
- McCusker JH, Clemons KV, Stevens DA & Davis RW (1994) *Saccharomyces cerevisiae* virulence phenotype as determined with CD-1 mice is associated with the ability to grow at 42 degrees C and form pseudohyphae. *Infect Immun* **62**: 5447–5455.
- McFarland LV & Bernasconi P (1993) *Saccharomyces boulardii*: a review of an innovative biotherapeutic agent. *Microb Ecol Health Dis* **343**: 171–172.
- Mercado L, Dalcero A, Masuelli R & Combina M (2007) Diversity of *Saccharomyces* strains on grapes and winery surfaces: analysis of their contribution to fermentative flora of Malbec wine from Mendoza (Argentina) during two consecutive years. *Food Microbiol* **24**: 403–412.
- Morin PA, Luikart G, Wayne RK *et al.* (2004) SNPs in ecology, evolution and conservation. *Trends Ecol Evol* **19**: 208–216.
- Pasteur L (1866) Etudes sur le vin. Paris, Impr. Imperiale.
- Pearson T, Busch JD, Ravel J *et al.* (2004) Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. *Proc Natl Acad Sci USA* **101**: 13536–13541.
- Piarroux R, Millon L, Bardonnnet K, Vagner O & Koenig H (1999) Are live *Saccharomyces* yeasts harmful to patients? *Lancet* **353**: 1851–1852.
- Querol A, Fernández-Espinar MT, del Olmo M & Barrio E (2003) Adaptive evolution of wine yeast. *Int J Food Microbiol* **86**: 3–10.

- Schuller D & Casal M (2007) The genetic structure of fermentative vineyard-associated *Saccharomyces cerevisiae* populations revealed by microsatellite analysis. *Antonie Van Leeuwenhoek* **91**: 137–150.
- Spiropoulos A, Tanaka J, Flerianos I & Bisson LF (2000) Characterization of hydrogen sulfide formation in commercial and natural wine isolates of *Saccharomyces*. *Am J Enol Vitic* **51**: 233–248.
- Strand M, Prolla TA, Liskay RM & Petes TD (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* **365**: 274–276.
- Sudbery P, Gow N & Berman J (2004) The distinct morphogenic states of *Candida albicans*. *Trends Microbiol* **12**: 317–324.
- Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **10**: 1093.
- This P, Jung A, Boccacci P *et al.* (2004) Development of a standard set of microsatellite reference alleles for identification of grape cultivars. *Theor Appl Genet* **109**: 1448–1458.
- Verstrepen KJ & Klis FM (2006) Flocculation, adhesion and biofilm formation in yeasts. *Mol Microbiol* **60**: 5–15.
- Verstrepen KJ, Jansen A, Lewitter F & Fink GR (2005) Intragenic tandem repeats generate functional variability. *Nat Genet* **37**: 986–990.
- Wang DG, Fan JB, Siao CJ *et al.* (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**: 1077–1082.
- Wu H, Zheng X, Araki Y, Sahara H, Takagi H & Shimoi H (2006) Global gene expression analysis of yeast cells during sake brewing. *Appl Environ Microbiol* **72**: 7353–7358.
- Zweyck D, Hrastnik C, Kohlwein SD & Daum G (2000) Biochemical characterization and subcellular localization of the sterol C-24(28) reductase, erg4p, from the yeast *Saccharomyces cerevisiae*. *FEBS Lett* **470**: 83–87.

Supplementary material

The following supplementary material is available for this article:

Table S1. Strain list.

Table S2. Initial screening.

Table S3. Exclusive alleles.

Figure S1. Dendrogram.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1567-1364.2008.00361.x> (This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Molecular Relationships Between *Saccharomyces cerevisiae* Strains Involved in Winemaking from Mendoza, Argentina

Current Microbiology

ISSN 0343-8651

Volume 61

Number 6

Curr Microbiol (2010)

61:506-514

DOI 10.1007/s00284-010-9645-

y

Current Microbiology

An International Journal

Volume 61
Number 6
December 2010

Endophytic Bacteria Improve Seedling Growth of Sunflower Under Water Stress, Produce Salicylic Acid, and Inhibit Growth of Pathogenic Fungi
G. Forchetti · O. Masciarelli · M.J. Izaguirre · S. Alemano · D. Alvarez · G. Abdala 485

Identification of a Novel Virulence-Related Gene in *Streptococcus suis* Type 2 Strains
H. Zhang · H. Fan · C. Lu 494

Adhesion Ability of *Campylobacter jejuni* to H-29 Cells Increases with the Augmentation of Oxidant Agent Concentration
N. Haddad · G. Mailart · A. Garlneau · F. Jugiau · M. Federighi · J.-M. Cappelletti 500

Molecular Relationships Between *Saccharomyces cerevisiae* Strains Involved in Winemaking from Mendoza, Argentina
L. Mercado · S. Jubany · C. Gaggero · R.W. Masuelli · M. Combarro 506

A Quantitative Real-time PCR Assay for Quantification of Viable *Listeria monocytogenes* Cells After Bacteriocin Injury in Food-First Insights
A. Cobo Molinos · H. Abouel · N. Ben Omar · M. Martinez-Canamero · A. Gálvez 515

Comparison of Real-Time PCR with Disk Diffusion, Agar Screen and E-test Methods for Detection of Methicillin-Resistant *Staphylococcus aureus*
L. Sharifi · M. Valdi · M.A. Tabatabaiefar · A. Karimi · M.R. Nafisi 520

PCR Detection of Thermophilic Spore-Forming Bacteria Involved in Canned Food Spoilage
S. Prevost · S. Andre · E. Remize 525

In Vitro Biofilm Activity of Non-*Candida albicans* *Candida* Species
S. Silva · M. Henriques · R. Oliveira · D. Williams · J. Azevedo 534

Characterization of Tunisian *Bacillus thuringiensis* Strains with Abundance of *kurstaki* Subspecies Harbouring Insecticidal Activities Against the Lepidopteran Insect *Ephestia kuehniella*
I. Saadoun · R. Al-Thani · F. Al-Saadi · N.B.-B. Hassan · L. Abdelkafi-Mesrahi · P. Schultze · S. Reus · S. Jassbi 541

Production and Characterization of N- and C-terminally Truncated Mtx2: a Mosquitocidal Toxin from *Bacillus sphaericus*
K. Phansinchel · P. Ratsat · T. Limvutteejeerat · B. Promdonkoy 549

Biosurfactants as Agents to Reduce Adhesion of Pathogenic Bacteria to Polystyrene Surfaces: Effect of Temperature and Hydrophobicity
A.E. Zeraik · M. Nitschke 554

Screening and Evaluation of Human Intestinal Lactobacilli for the Development of Novel Gastrointestinal Probiotics
P. Kott · R. Minder · I. Smidt · P. Hoff · K. Trautwein · R.-H. Mikkelsen · J. Shchepetova · K. Krogh-Andersen · H. Marcolla · L. Hammarström · M. Mikelsaar 560

***Bacillus Anthracis* Endospores Regulate Ornithine Decarboxylase and Inducible Nitric Oxide Synthase Through ERK1/2 and p38 Mitogen-Activated Protein Kinases**
S. Porasuphatana · G.-L. Cao · P. Tsai · F. Tavakkoli · T. Hsuwar · L. Baillie · A.S. Cross · P. Shapiro · G.M. Rosen 567

Characterization of Bacterial Isolates Collected from a Sheep Model of Osteointegration
D.L. Williams · R.D. Bloebaum · J.P. Beck · C.A. Petti 574

***Lactobacillus plantarum* 24, Isolated From the Manila Fruit (*Scolocarya birrea*), has Probiotic Properties and Harbors Genes Encoding the Production of Three Bacteriocins**
W.A. Prins · M. Botha · M. Botes · M. de Kwaadsteniet · A. Endo · L.M.T. Dicks 584

Risk Assessment of Genetically Modified Lactic Acid Bacteria Using the Concept of Substantial Equivalence
J.G. LeBlanc · D. Van Sinderen · J. Hugenoltz · J.-C. Plard · F. Sesma · G. Savoy de Giori 590

Utilization of Phenol and Naphthalene Affects Synthesis of Various Amino Acids in *Corynebacterium glutamicum*
S.Y. Lee · T.-H. Lee · S.-T. Chang · J.-S. Park · Y.-H. Kim · J. Min 596

Isolation and Characterization of Bacteriophages Infecting *Staphylococcus epidermidis*
D. Gutiérrez · B. Martínez · A. Rodríguez · P. García 601

Environmentally Safe Production of 7-ACA by Recombinant *Acetomonium chrysogenum*
Y. Liu · G. Gong · C. Zhu · B. Zhu · Y. Hu 609

ERRATUM
Erratum to: Biosorption of Copper by Cyanobacterial Bloom-Derived Biomass Harvested from the Eutrophic Lake Dianchi in China
K. Wang · G. Collica · R. De Philippis · Y. Liu · D. Li 615

Further articles can be found at www.springerlink.com

Instructions for Authors for Curr Microbiol are available at www.springer.com/00284

Curr Microbiol ISSN 0343-8651

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Molecular Relationships Between *Saccharomyces cerevisiae* Strains Involved in Winemaking from Mendoza, Argentina

Laura Mercado · Sandra Jubany · Carina Gaggero ·
Ricardo W. Masuelli · Mariana Combina

Received: 2 October 2009 / Accepted: 26 March 2010 / Published online: 21 April 2010
© Springer Science+Business Media, LLC 2010

Abstract Three molecular typing techniques were applied to assess the molecular relationships of *Saccharomyces cerevisiae* strains isolated from winery equipment, grapes, and spontaneous fermentation in a cellar located in “Zona Alta del Río Mendoza” (Argentina). In addition, commercial *Saccharomyces* strains widely used in this region were also included. Interdelta PCR typing, mtDNA restriction analysis, and microsatellite (SSR) genotyping were applied. Dendrograms were constructed based on similarity among different patterns of bands. The combination of the three techniques discriminated 34 strains among the 35 isolates. The results of this study show the complex relationships found at molecular level among the isolates that share the same ecological environment, i.e., the winemaking process. With a few exceptions, the yeast isolates were generally clustered in different ways, depending on the typing technique employed. Three clusters were conserved independently of the molecular method applied. These groups of yeasts always clustered

together and had high degree of similarity. Furthermore, the dendrograms mostly showed clusters combining strains from winery and fermentation simultaneously. Most of the commercial strains included in this study were clustered separately from the other isolates analyzed, and just a few of them grouped with the strains mainly isolated from spontaneous fermentation. Only one commercial strain was clustered repetitively with a noncommercial strain isolated from spontaneous fermentation in the three dendrograms. On the other hand, this study has demonstrated the importance of selecting an appropriate molecular method according to the main objectives of the research.

Introduction

The alcoholic fermentation of grape juice into wine by yeasts is an ecologically complex process [1]. Though many types of yeast species are associated with wine fermentation, *Saccharomyces cerevisiae* is the main species involved. Various studies on wine microbial ecology have elucidated the complex biodiversity involved in the development of the fermentation [2–7], establishing that *Saccharomyces* populations are integrated by multiple strains, even in inoculated fermentations. Consequently, it is important to have simple and appropriate methods that allow for discrimination at the strain level.

Numerous molecular techniques have been proposed to study *Saccharomyces* strain diversity, such as pulsed field electrophoresis [8], mitochondrial DNA (mtDNA) restriction analysis [9], and interdelta element PCR amplification [10, 11]. In recent years, the amplification of polymorphic microsatellite loci (SSRs, simple sequence repeats) has also been proposed as a powerful tool for *S. cerevisiae* strain differentiation [12–15].

L. Mercado · R. W. Masuelli · M. Combina (✉)
Instituto Nacional de Tecnología Agropecuaria (INTA),
EEA Mendoza and EEA La Consulta, San Martín 3853,
5507 Lujan de Cuyo, Mendoza, Argentina
e-mail: mcombina@mendoza.inta.gov.ar

R. W. Masuelli · M. Combina
Consejo Nacional de Investigaciones Científicas y Técnicas
(CONICET), Buenos Aires, Argentina

L. Mercado · R. W. Masuelli
Facultad de Ciencias Agrarias, Universidad Nacional de Cuyo,
Mendoza, Argentina

S. Jubany · C. Gaggero
Departamento de Biología Molecular, Instituto de
Investigaciones Biológicas Clemente Estable (IIBCE),
Montevideo, Uruguay

Several studies have used molecular characterization of *Saccharomyces* to examine their biodiversity and distribution in vineyards, the geographic origin of the native isolates, and the impact of the enological practices on autochthonous ecosystems [16–18].

The “Zona Alta del Río Mendoza” (ZARM) is an important Argentinean wine region, which has optimum conditions for the growth of Malbec grapes for the production of high-quality wines. Currently, little information is available about the yeast populations involved in the winemaking process carried out in this region [5, 19]. A previous study attempted to identify the native *Saccharomyces* strains associated with winery equipment, grapes, and spontaneous fermentation on Malbec grape must from a cellar located in the ZARM region, during two consecutive years. It was demonstrated that a stable and resident *Saccharomyces* microbiota was present in the winery, which exhibited a dynamic behavior across different seasons and between years. Moreover, low occurrence of *Saccharomyces* on grapes and their limited participation during fermentation was also observed. Fermentations showed strain sequential substitution, and about 30–60% of yeast population at the end of fermentation were from the population already present in the winery. In addition, some commercial yeast strains were found during fermentation and on winery equipment in low percentage [5].

Although these results showed the microbial ecology of alcoholic fermentation in an industrial winery from Mendoza, there is still a lack of information on the genetic relationships among these *Saccharomyces* native populations.

The aim of this study was to assess the genetic relationships among *Saccharomyces* native population isolated from winery equipment, grapes, and spontaneous fermentation in a cellar located in “Zona Alta del Río Mendoza” (Argentina). Three molecular typing techniques (interdelta amplification analysis, mtDNA restriction analysis, and SSR genotyping) were applied for the molecular characterization of *Saccharomyces* isolates. The molecular analysis included the comparison of the polymorphism obtained by each method and the determination of the genetic relationship among the strains from the same ecological origin and selected commercial strains.

Materials and Methods

Yeast Strains

Saccharomyces strains were previously isolated from winery equipment and spontaneous fermentation of Malbec grape must from ZARM wine region during two consecutive years [5]. Twenty-eight previously isolated strains

were selected as representatives of the diversity found on winery equipment and during fermentation (Table 1). Seven commercial strains originally isolated in France and widely used in the ZARM wine region were included in this study; they were named with an alphabetical code to preserve the company's confidentiality.

Interdelta PCR Analysis

Total DNA extraction was performed as previously described by Hoffman and Winston [20]. Oligonucleotides primers delta 1 (5'-CAAAATTCACCTAT[A/T]TCTCA-3') and delta 2 (5'-GTGGATTTTATTCAACA-3'), or delta 12 (5'-TCAACAATGGAATCCCAAC-3') and delta 21 (5'-CATCTTAACACCGTATATGA-3') were used to amplify the total genomic DNA between the repeated interspersed delta sequences as previously described [10, 11]. PCR products were separated in 1.5% agarose gels in 0.5× TBE buffer. The molecular marker 100-bp DNA ladder (Promega, Madison, USA) was used as the molecular size standard. Electrophoresis gels were stained with ethidium bromide (5 µg ml⁻¹), visualized by UV transillumination, and processed using the Gel Doc 1000 Video Gel Documentation System (BioRad, Richmond, USA).

Mitochondrial DNA Restriction Fragment Length Polymorphism (mtDNA RFLP)

The total DNA was extracted using the Wizard Genomic DNA extraction kit (Promega, Madison, USA) according to the manufacturer's instructions. Digestions were performed with the restriction enzyme *Hinf*I as previously described [9]. Restriction fragments were separated by electrophoresis in 1% agarose gels, stained, and visualized as described earlier for PCR products.

Analysis of SSR Loci

Quick preparation of DNA template for PCR was carried out as described by Jubany et al. [13]. The set of six SSR loci used in this study (Table 2) were chosen from a previous study [13], in which a large group of SSR loci were evaluated. In that study, nine SSRs were selected and six of them showed the highest discriminative power between closely related native strains. SSRs were named following the standardized criteria recently described by Jubany et al. PCR amplifications and electrophoresis were performed as described by Jubany et al. [13].

Cluster Analysis

The patterns of bands obtained after gel electrophoresis were employed for the construction of presence/absence

Table 1 Summary of patterns obtained by all typing methods used on 28 *Saccharomyces* isolated at ZARM wine region and 7 commercial wine strains

Isolate number code	Origin of isolation	Year of isolation	Pattern assigned by each molecular method			
			PCR interdelta	“Improved” PCR interdelta	RFLP ADNmt	SSR
1	Fermentation tank ^a	2001	1	1	1	1
2	Pipe for filling of tank ^a	2002	1	2	1	1
3	Spontaneous fermentation ^a	2002	1	3	1	3
4	Pipe for must transport ^a	2002	2	4	1	4
5	Spontaneous fermentation	2001	3	5	2	5
6	Spontaneous fermentation ^a	2002	1	2	1	1
7	Spontaneous fermentation	2001	4	6	3	6
8	Fermentation tank	2001	4	7	4	7
9	Fermentation tank ^a	2001	4	6	3	8
10	Spontaneous fermentation	2002	5	8	5	9
11	Grape	2002	5	9	5	10
12	Pipe for filling of tank ^a	2002	6	10	6	11
13	Fermentation tank ^a	2002	6	11	7	11
14	Pipe for must transport ^a	2002	6	12	8	12
15	Reception hopper ^a	2002	6	13	8	12
16	Pipe for must transport ^a	2002	7	14	1	13
17	Crusher	2002	6	15	9	15
18	Fermentation tank	2002	6	10	10	11
19	Spontaneous fermentation	2002	6	10	11	15
20	Pipe for must transport ^a	2001	6	16	7	11
21	Fermentation tank ^a	2001	6	12	8	16
22	Crusher	2001	6	17	12	17
23	Spontaneous fermentation	2001	6	18	13	18
24	Fermentation tank	2001	6	6	3	19
25	Crusher ^a	2002	8	19	7	11
26	Fermentation tank	2002	8	20	14	20
27	Spontaneous fermentation	2002	8	21	15	21
28	Spontaneous fermentation	2002	8	22	16	22
29	Commercial strain A	–	9	23	17	23
30	Commercial strain B	–	10	24	18	24
31	Commercial strain C	–	11	25	19	25
32	Commercial strain D	–	12	26	20	26
33	Commercial strain É	–	13	27	21	27
34	Commercial strain F	–	14	28	22	28
35	Commercial strain G	–	15	29	2	29
Total of different patterns			15	29	22	29

^a Perennial strains defined by Mercado et al. [5]

matrix, taking into account the total number of different bands observed for each method. A total of 25 bands were obtained with improved interdelta PCR, 53 bands with mtDNA RFLP, and 82 fragments with the combination of the six SSR markers used. The Dice coefficient was calculated to estimate the similarity between the strains and a dendrogram was developed with the UPGMA method using NTSYSPC1.1 software (Exceter software, Setauket, USA). Cophenetic correlation as an estimator of goodness of fit of clustering was calculated by applying the same software package. The *S. cerevisiae* laboratory reference

strain, S288c, was used as the out-group for the development of the dendrograms.

Results

Molecular Strain Characterization

About 28 wine *Saccharomyces* isolates from a commercial winery from ZARM (Argentina) and 7 commercial yeast strains were analyzed by applying three different molecular

Table 2 Oligonucleotide primers used for the SSR loci amplification

SSR loci	Forward primer 5'–3'	Reverse primer 5'–3'
YPL009C	CTGCTCAACTTGTGATGGGTTTGG	CCTCGTTACTATCGTCTTCATCTTGC
YOR267C	ATGACTGCAGCAATGAATCG	TCCTCTGTGCTGTTGACTCG
TTA_XXIII	AAGCGTAAGCAATGGTGTAGA	AAGCCTCTTCAAGCATGACC
TG_VI	GACACAATAGCAATGGCCTTCA	ATATTGGCACCAGAAAGTGTCG
GT_X	TGCGCAGCTTAGTATGACCA	GATGGGCTTTCACTCCACTT
C4_XV	GAGAAAAATGCTGTTTATTCTGACC	CCTCCGGGACGTGAAATAAC

typing techniques. This group of yeasts was selected considering the results of a previous study developed in this wine region [5], in which the yeast populations analyzed were obtained from winery equipment, Malbec grapes, and musts during spontaneous fermentation. These latter samples were obtained at different stages of the process carried out in the same winery according to their own standard protocols, including the lack of inoculation with selected commercial strains, which is a current practice of this company for the production of red wines. Among the strains analyzed in this study, 13 yeasts were selected because they represent the “perennial” strains of the winery [5]. These yeasts were the isolates that were able to survive in the cellar, including the isolates found on winery equipment that were not used since the previous year, as well as those recovered from the winery in successive years. Some of these isolates were also found at the end of the spontaneous fermentation conducted in the same cellar (Table 1).

The results presented in Table 1 clearly indicate that according to the technique used, distinct levels of discrimination were obtained. Interdelta PCR with primers delta 1 and delta 2 was the least discriminative technique, distinguishing 15 patterns among the 35 strains analyzed (Table 1). When improved primers, delta 12 and delta 21, were used, the discrimination power was increased, allowing for the discrimination of 29 patterns of the *Saccharomyces* strains analyzed. On the other hand, the analysis of genetic variability of 35 *Saccharomyces* isolates using mtDNA RFLP allowed for the differentiation of 22 patterns. This technique showed a good level of polymorphism with an intermediate discriminating power, when compared with the other molecular markers used (Table 1).

The discrimination power obtained by combining the information from six SSR loci was very high, similar to that of the improved interdelta PCR (Table 1). Out of the 35 *Saccharomyces* isolates analyzed, 29 genotypes had a total of 87 alleles, ranging from 10 to 20 alleles and 12 to 24 genotypes per locus. However, the frequency of appearance of different alleles at each locus was very variable (Fig. 1). TG_VI and TTA_XIII showed a great

discriminating power, differentiating 20 and 17 genotypes, respectively. Some alleles were detected more frequently, being more characteristic of the population under study. On the other hand, alleles found less frequently could indicate genotypes rarely detected. In summary, the combination of results obtained with the three applied molecular typing techniques allowed for the differentiation of 34 strains among the 35 isolates. Only two yeast isolates showed identical profile with all the molecular methods applied, and they could be considered as the same strain (Table 1).

Molecular Relationships Among the Strains

Similarities based on the Dice coefficient and UPGMA clustering from improved interdelta PCR, mtDNA RFLP, and SSR are presented in Figs. 2, 3, and 4, respectively. The dendrograms showed different degrees of relationship among the strains, according to the molecular method used. Interdelta PCR showed a superior similitude among the strains (Fig. 2), mainly because the patterns obtained by this technique have bands shared between all the isolates. The lowest genetic similitude was obtained with SSR analysis, where the first dendrogram branching was produced at a similitude coefficient value lower than 0.1. The evaluated SSR loci showed a high polymorphism between the yeast isolates analyzed, where only few of them remain closely associated (Fig. 4).

The molecular methods applied reflect the polymorphism at two levels—nuclear genomic and mitochondrial genomic level. Some strains exhibited differences in each genome, by clustering together with the nuclear markers (interdelta PCR and SSR analyses) and separately with the mitochondrial marker (RFLP mtDNA), such as isolates 12, 18, and 25 (Figs. 2, 3, and 4).

With a few exceptions, the yeast isolates were generally clustered in different ways, depending on the typing technique employed. Three clusters were conserved independently of the molecular method applied (cluster 7, 9, and 24; cluster 1, 2, and 6; and cluster 14, 15, and 21) (Figs. 2, 3, and 4). These clusters were mainly integrated by isolates from the winery equipment, and mostly constituted a part

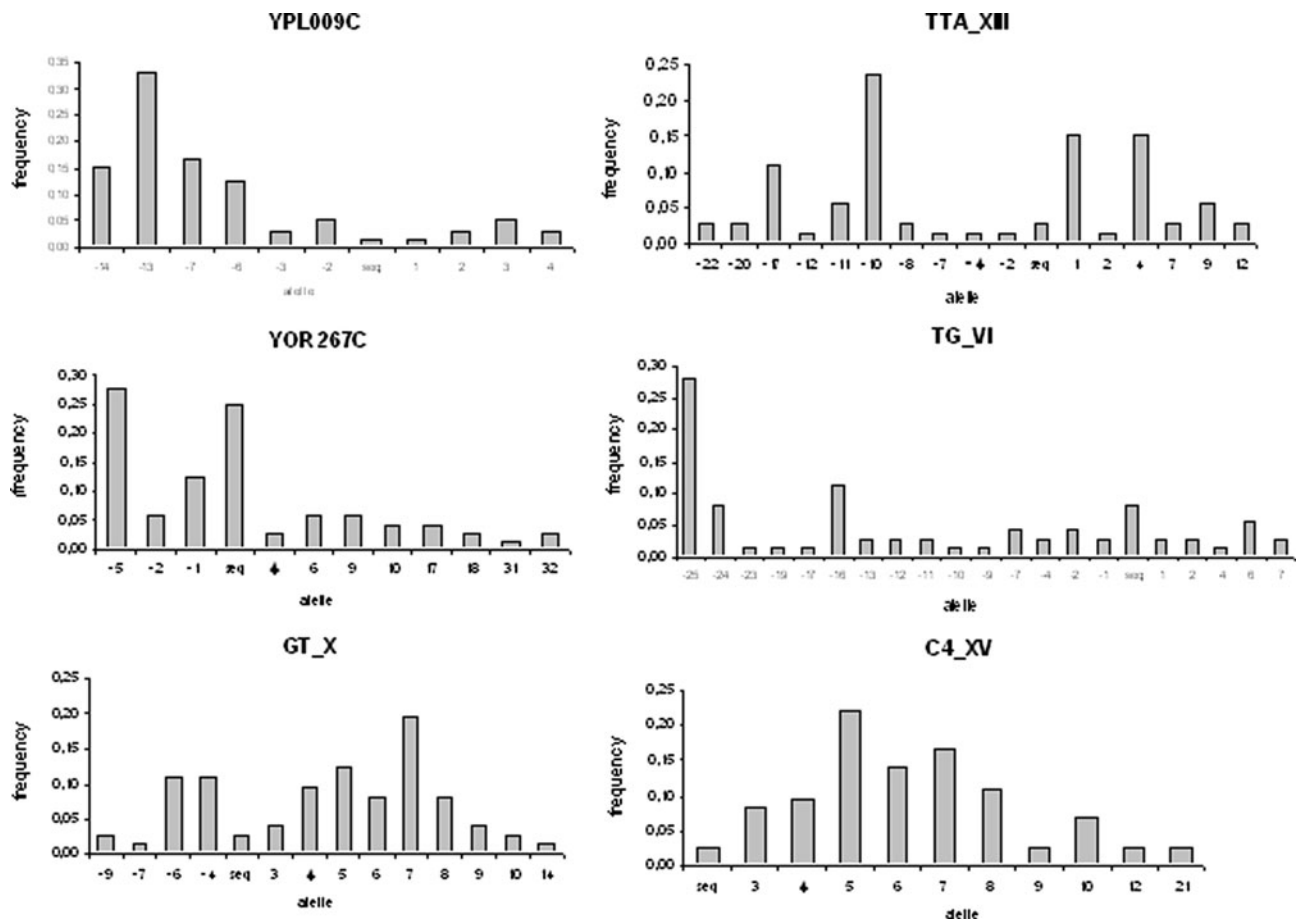


Fig. 1 Polymorphisms of alleles obtained for the six SSR loci on 28 native and 7 commercial *S. cerevisiae* wine strains. Alleles are expressed as + or – number of repeats relative to the sequenced (seq) strain

of the “perennial” population of the winery. Besides these strains, the clusters included some other strains that did not remain clustered with the three molecular markers. Furthermore, the dendrograms mostly showed clusters combining strains from winery and fermentation simultaneously (Figs. 2, 3, and 4).

Most of the commercial strains included in this study were clustered separately from the other isolates analyzed, and just a few of them grouped with the strains mainly isolated from spontaneous fermentation. Only the commercial strain 35 was clustered repetitively with a non-commercial strain, grouping with strain 5 isolated from spontaneous fermentation in the three dendrograms, and showing different degrees of similarity according to the method applied (Figs. 2, 3, and 4).

Discussion

In this study, different molecular typing methods have been applied to evaluate the genetic relationships between 28 *Saccharomyces* strains isolated in a previous study [5]. The

strains were obtained in the same cellar, and included isolates widely distributed on different winery equipment, those able to survive between vintages on the cellar surfaces, as well as those frequently found in spontaneous fermentation. Seven commercial strains widely used in this region, currently employed in this winery for the production of white wines, were also included in the analysis.

Recently, new genetic techniques have been developed to provide a differentiation of *Saccharomyces* strains based on single-nucleotide polymorphisms (SNPs), like multilocus sequence typing (MLS) [21, 22] or microarray techniques [23, 24]. Nevertheless, the molecular markers selected in this study still represent the simplest and most widely used techniques to study *Saccharomyces* biodiversity. They not only provide differentiation at the strain level but also give some information about the genetic relationship.

In the first approach, the discrimination power of the molecular techniques applied to differentiate closely related *Saccharomyces* strains was analyzed. Interdelta PCR with primers, delta 1 and delta 2, was the least discriminating method, as shown in previous studies [18], but the

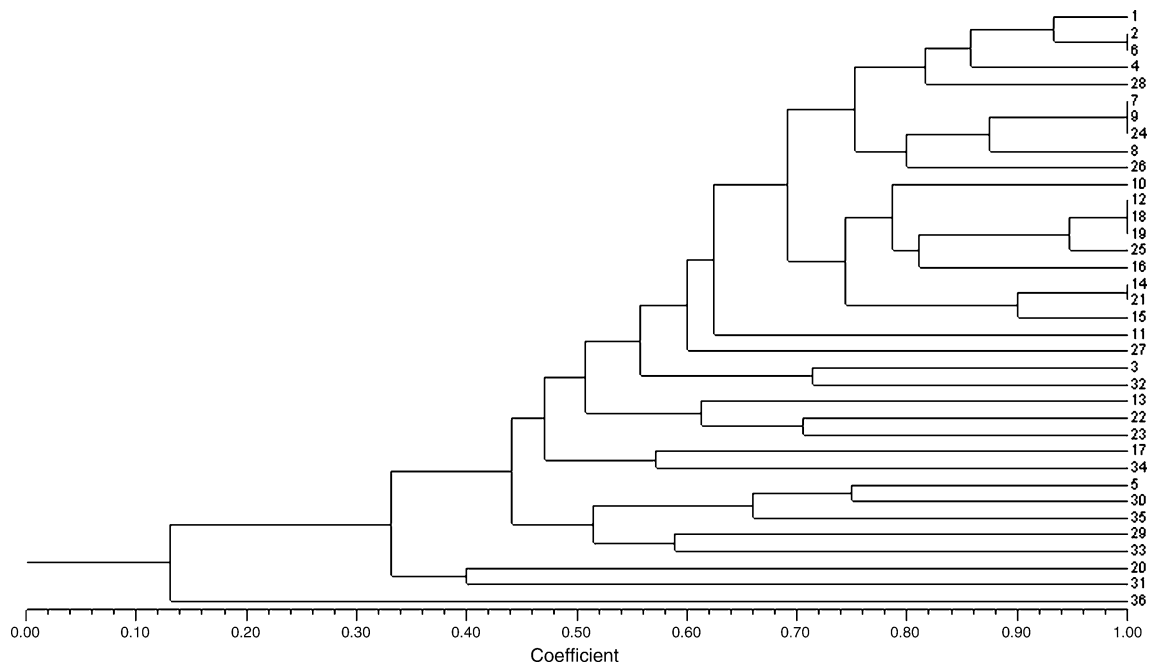


Fig. 2 Dendrogram showing molecular relationships based on improved PCR interdelta patterns of 28 wine *S. cerevisiae* isolates and 7 commercial selected strains. Cophenetic correlation $r = 0.88019$

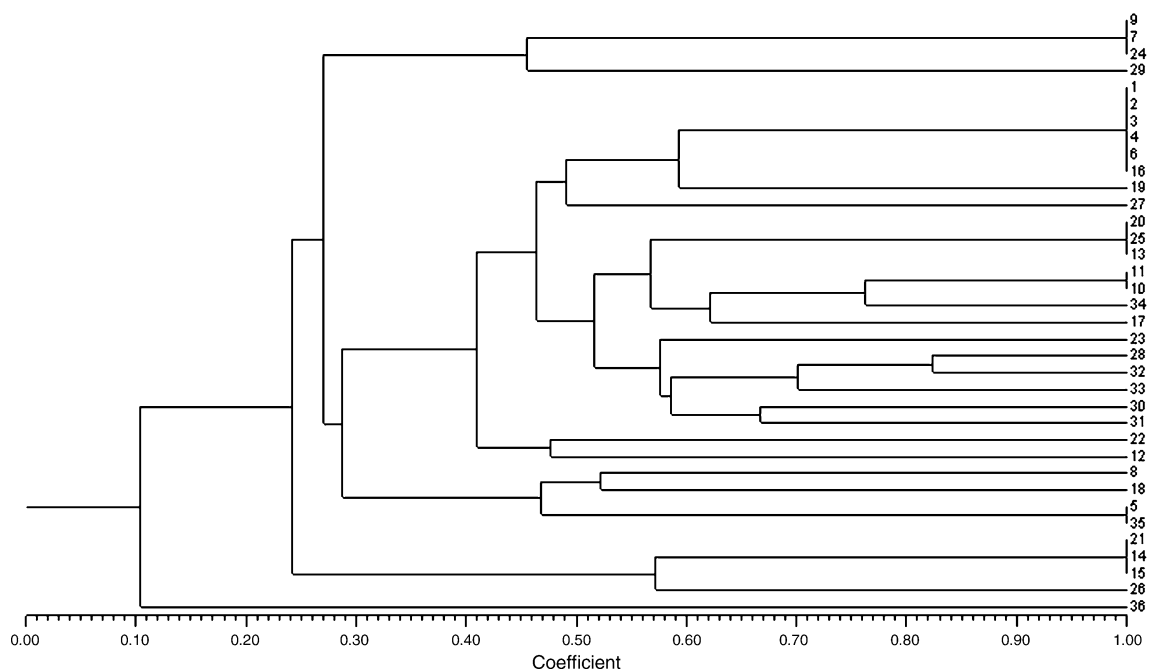


Fig. 3 Dendrogram showing molecular relationships based on RFLP mtDNA patterns of 28 wine *S. cerevisiae* isolates and 7 commercial selected strains. Cophenetic correlation $r = 0.85498$

differentiation efficiency was increased with the use of improved primers [10, 25].

The combination of the allele sizes from the six SSR loci showed a high degree of resolution, even though they were not adequate for the unequivocal characterization of

the present population of 35 strains. According to the observed results, the SSR technique was not equivalent to the mtDNA RFLP, and showed higher discrimination power. Furthermore, the same/different SSR amplification profile did not always corresponded to the same/different

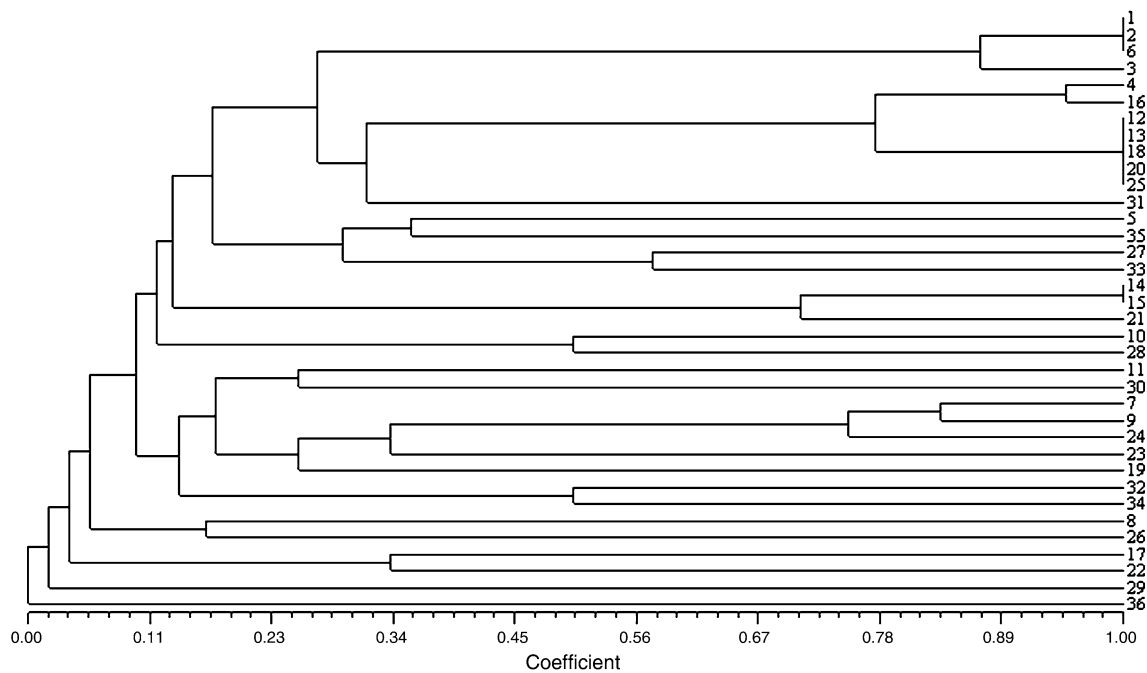


Fig. 4 Dendrogram showing molecular relationships based on SSR analysis patterns of 28 wine *S. cerevisiae* isolates and 7 commercial selected strains. Cophenetic correlation $r = 0.88078$

mtDNA RFLP patterns, as reported previously [18, 26]. In spite of this result, mtDNA RFLP analysis was adequate for the differentiation of yeast strains from the same ecosystem, because although some isolates were not discriminated by this technique, they remained closely related when the SSR method was used. Numerous authors have demonstrated that mtDNA RFLP analysis is an efficient technique to differentiate at the strain level [9, 18, 27, 28]. According to the results of this study, studies aiming to characterize the strains that are genetically closely related would need the inclusion of other SSR loci or the combination of SSR with other molecular typing techniques.

Molecular relationships among *Saccharomyces* strains from the same environment, represented by equipment, grapes, and musts spontaneously fermented from a winery located in the ZARM region, were evaluated, and complex relations were evidenced.

Only nine yeast isolates were repetitively grouped in three clusters. These groups comprised some of the “perennial” yeasts that were isolated from different winery equipment during the two vintages and were also found at the end of fermentation [5]. As shown in the dendrograms, these groups of yeasts always clustered together and had high degree of similarity. The isolates included in each of these clusters showed identical mitochondrial restriction patterns and the discrimination among them was achieved by molecular markers related to the nuclear repetitive sequences. It has been proposed that recombination

processes could be favored by the presence of repeated sequences that may increase the probability of ectopic recombination events that generate different molecular patterns [29]. It is presumed that some kind of change at the nuclear level occur during the permanence of yeasts in the winery. Such changes could allow for yeast adaptation to the winery environment and survival from one vintage to the following by improving their competitive traits and tolerance to stress. Similar mechanism has been proposed by Bond to explain the dynamic genome of lager yeasts [30]. The monophyletic origin of each of these clusters and its rare relationship with the evaluated commercial strains suggest their American origin and the recent occurrence of microevolutionary events.

On the other hand, the rest of the isolates from the winery and fermentation showed a random clustering according to the molecular marker applied; this result suggests some kind of change at nuclear level, which could occur at a different rate in the nucleus with respect to the mitochondria, during the yeast life cycle. Recent studies have demonstrated a low stability of the genome in wine yeast [31], which may be due to the high reorganization capacity of its genome by Ty-promoted translocation, mitotic recombination, and gene conversion [16, 29, 32]. Alternatively, it has been proposed that ethanol and acetaldehyde introduce breaks in the DNA, with a much higher mutation rate on the mitochondrial genome. This may be due to a higher efficiency of the yeast nuclear DNA repair

system compared with the mitochondrial system that lacks proofreading activity [16, 33].

In this study, the commercial strains of European origin were not restricted to a sole cluster, but were generally clustered separately or appeared related to the fermentation isolates. The existence of some genetic relation between a few “winemaking-related” strains with some of the commercial strains could support the hypothesis that some American native strains may proceed from European strains as it was recently suggested [13]. Conversely, it has been suggested that the existence of “genomic resemblance” among the native *Saccharomyces* and commercial strains, caused by centuries of positive selection during wine production, could lead to the same characteristics explored during the process of selection for obtaining commercial cultures [23]. These authors suggest that resemblance in phenotype is reflected in genotypic characteristics.

Results of this study show the complex relationships found at the molecular level among the yeast isolates that share the same ecological environment. Moreover, this study has revealed that beyond the abundant diversity observed, the yeasts share many genetic characteristics. On the other hand, this study has demonstrated the importance of selecting an appropriate molecular method according to the main objectives of the research.

Acknowledgments This study was supported by Viticulture Regional Project MZASJ O7 from the *Instituto Nacional de Tecnología Agropecuaria (INTA)* (and Project PDT 32/06 from Dinacyt, Uruguay). L.M. is a fellow of *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)* and *Programa de Doctorado en Ciencias Biológicas (PROBIOL)*, *Universidad Nacional de Cuyo*.

References

- Ribereau-Gayon P, Dubourdieu D, Donèche B, Lonvaud A (eds) (2006) Cytology, taxonomy and ecology of grape and wine yeast. In: Handbook of enology—the microbiology of wine and vinifications, 2nd edn, vol 1. John Wiley & Sons, Chichester, pp 1–49
- Agnolucci M, Scarano S, Santoro S et al (2007) Genetic and phenotypic diversity of autochthonous *Saccharomyces* spp. strains associated to natural fermentation of ‘Malvasia delle Lipari’. *Lett Appl Microbiol* 45:657–662
- Blanco P, Ramilo A, Cerdeira M, Orriols I (2006) Genetic diversity of wine *Saccharomyces cerevisiae* strains in an experimental winery from Galicia (NW Spain). *Antonie van Leeuwenhoek* 89:351–357
- Lopes C, Lavalley T, Querol A, Caballero A (2005) Combined use of killer biotype and mtDNA-RFLP patterns in a Patagonian wine *Saccharomyces cerevisiae* diversity study. *Antonie van Leeuwenhoek* 3:1–10
- Mercado L, Dalcero A, Masuelli R, Combina M (2007) Diversity of *Saccharomyces* strains on grapes and winery surfaces: Analysis of their contribution to fermentative flora of Malbec wine from Mendoza (Argentina) during two consecutive years. *Food Microbiol* 24:403–412
- Santamaría P, Garijo P, López R et al (2005) Analysis of yeast population during spontaneous alcoholic fermentation: effect of the age of the cellar and the practice of inoculation. *Int J Food Microbiol* 103:49–56
- Torija M, Rozes N, Poblet M et al (2001) Yeast population dynamics in spontaneous fermentations: comparison between two different wine-producing areas over a period of three years. *Antonie van Leeuwenhoek* 79:345–352
- Blondin B, Vezinhet F (1988) Identification de souches de levures oenologiques par leurs caryotypes obtenus en électrophorèse en champ pulsé. *Rev Fr Oenol* 115:7–11
- Querol A, Barrio F, Ramon D (1992) A comparative study of different methods of yeast strain characterization. *Syst Appl Microbiol* 15:439–446
- Légras J, Karst F (2003) Optimisation of interdelta for *Saccharomyces cerevisiae* strain characterization. *FEMS Microbiol Lett* 221:249–255
- Ness C, Lavalley F, Dubourdieu D et al (1993) Identification of yeast strains using the polymerase chain reaction. *J Sci Food Agric* 62:89–94
- Gallego F, Perez G, Martinez I, Hidalgo P (1998) Microsatellites obtained from database sequences are useful to characterize *Saccharomyces cerevisiae*. *Am J Enol Vitic* 49:350–351
- Jubany S, Tomasco I, Ponce de León I et al (2008) Toward a global database for the molecular typing of *Saccharomyces cerevisiae* strains. *FEMS Yeast Res* 8:472–484
- Légras J, Ruh O, Merdinoglu D, Karst F (2005) Selection of hypervariable microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains. *Int J Food Microbiol* 102:73–83
- Pérez M, Gallego F, Hidalgo P (2001) Evaluation of molecular techniques for the genetic characterization of *Saccharomyces cerevisiae* strains. *FEMS Microbiol Lett* 205:375–378
- Martínez C, Cosgaya P, Vásquez C et al (2007) High degree of polymorphism and geographic origin of wine yeast strains. *J Appl Microbiol* 103:2185–2195
- Schuller D, Alves H, Dequin S, Casal M (2005) Ecological survey of *Saccharomyces cerevisiae* strains from vineyards in the Vinho Verde Region of Portugal. *FEMS Microbiol Ecol* 51:167–177
- Schuller D, Valero E, Dequin S, Casal M (2004) Survey of molecular methods for the typing of wine yeast strains. *FEMS Microbiol Lett* 231:19–26
- Combina M, Elía A, Mercado L et al (2005) Dynamics of indigenous yeast populations during spontaneous fermentation of wines from Mendoza, Argentina. *Int J Food Microbiol* 99:237–243
- Hoffman C, Winston F (1987) A ten-minute DNA preparation from yeast efficiently release autonomous plasmids for transformation of *E. coli*. *Gene* 57:267–272
- Liti G, Carter D, Moses A et al (2009) Population genomics of domestic and wild yeast. *Nature* 458:337–341
- Vigentini I, Fracassetti D, Picozzi C, Foschino R (2009) Polymorphism of *Saccharomyces cerevisiae* involved in wine production. *Curr Microbiol* 58:211–218
- Carreto L, Eiriz M, Gomes A et al (2008) Comparative genomics of wild type yeast strains unveil important genome diversity. *BMC Genomics* 9:524
- Schacherer J, Shapiro J, Ruderfer D, Kruglyak L (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* 458:342–346
- Pramateftaki PV, Lanaridis P, Typas MA (2000) Molecular identification of wine yeast at species or strain level: a case study with strains from two vine-growing areas of Greece. *J Appl Microbiol* 89:236–248
- Schuller D, Casal M (2007) The genetic structure of fermentative vineyard-associated *Saccharomyces cerevisiae* populations

- revealed by microsatellite analysis. *Antonie van Leeuwenhoek* 91:137–150
27. Fernández-Espinar M, López V, Ramón D et al (2001) Study of the authenticity of commercial wine yeast strains by molecular techniques. *Int J Food Microbiol* 70:1–10
28. Guillamón JM, Barrio E, Querol A (1996) Characterization of wine yeast strains of the *Saccharomyces* genus on the basis of molecular markers: relationships between genetic distance and geographic or ecological origin. *Syst Appl Microbiol* 19:122–132
29. Puig S, Querol A, Barrio E, Pérez-Ortín JE (2000) Mitotic recombination and genetic changes in *Saccharomyces cerevisiae* during wine fermentation. *Appl Environ Microbiol* 66:2057–2061
30. Bond U (2009) Chapter 6 the genomes of lager yeasts. *Adv Appl Microbiol* 69:159–182
31. Pretorius IS (2000) Tailoring wine yeast for the new millenium: novel approaches to the ancient art of winemaking. *Yeast* 16: 1–55
32. Rachidi N, Barre P, Blondin B (1999) Multiple Ty-mediated chromosomal translocation lead to karyotype changes in a wine strain of *Saccharomyces cerevisiae*. *Mol Gen Genet* 261:841–850
33. Castrejón F, Codón A, Cubero B, Benítez T (2002) Acetaldehyde and ethanol are responsible for mitochondrial DNA (mtDNA) restriction fragment length polymorphism (RFLP). *Syst Appl Microbiol* 25:462–467