FELIPE BARATHO BEATO

Physiology of *Saccharomyces cerevisiae* strains isolated from Brazilian biomes: new insights into biodiversity and industrial applications

A dissertation submitted to the Biotechnology Graduate Program of the University of São Paulo for the degree of Master in Science

Concentration area: Biotechnology

Supervisor: Andreas Karoly Gombert

Original version

São Paulo

2016

FELIPE BARATHO BEATO

Fisiologia de linhagens de *Saccharomyces cerevisiae* isoladas de biomas Brasileiros: novos olhares sobre a biodiversidade e aspectos relevantes para aplicações industriais

Dissertação apresentada ao Programa de Pós-Graduação Interunidades em Biotecnologia USP/ Instituto Butantan/ IPT, para obtenção do Titulo de Mestre em Biotecnologia

Área de concentração: Biotecnologia

Orientador: Andreas Karoly Gombert

Versão original

São Paulo

2016

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do Instituto de Ciências Biomédicas da Universidade de São Paulo

Beato, Felipe Baratho.
Fisiologia de linhagens de *Saccharomyces cerevisiae* isoladas de biomas brasileiros: novos olhares sobre a biodiversidade e aspectos relevantes para aplicações industriais / Felipe Baratho Beato. -- São Paulo, 2016.

Orientador: Prof. Dr. Andreas Karoly Gombert.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan. Área de concentração: Biotecnologia. Linha de pesquisa: Fisiologia e engenharia metabólica de leveduras.

Versão do título para o inglês: Physiology of *Saccharomyces cerevisiae* strains isolated from brazilian biomes: new insights into biodiversity and industrial applications.

1. Levedura 2. *Saccharomyces cerevisiae* 3. Etanol combustível 4. Tolerância a estresse 5. Biodiversidade 6. Biotecnologia industrial I. Gombert, Prof. Dr. Andreas Karoly II. Universidade de São Paulo. Instituto de Ciências Biomédicas. Programa de Pós-Graduação Interunidades em Biotecnologia USP/IPT/Instituto Butantan III. Título.

ICB/SBIB063/2016

UNIVERSIDADE DE SÃO PAULO

Programa de Pós-Graduação Interunidades em Biotecnologia

Universidade de São Paulo, Instituto Butantan, Instituto de Pesquisas Tecnológicas

---

Candidato(a): Felipe Baratho Beato.

Título da Dissertação: Fisiologia de linhagens de *Saccharomyces cerevisiae* isoladas de biomas brasileiros: novos olhares sobre a biodiversidade e aspectos relevantes para aplicações industriais.

Orientador(a): Prof. Dr. Andreas Karoly Gombert.

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado,**

em sessão pública realizada a ...../...../....., considerou

( ) Aprovado(a) ( ) Reprovado(a)

Examinador(a): Assinatura: .....
Nome: .....
Instituição: .....

Examinador(a): Assinatura: .....
Nome: .....
Instituição: .....

Presidente: Assinatura: .....
Nome: .....
Instituição: .....

To my parents, for the unconditional love and support.

Acknowledgments

I would like to thank my supervisor Andreas Gombert for accepting me as his MSc student, for the trust and all the support along my studies and especially for being an example and sharing with me your academic rigor, discipline and knowledge on yeast research.

I also would like to thank Jochen Forster for having me in your research group at DTU Biosustain, for all the support and supervision, and Basti Bergdahl for all the patience, shared knowledge and co-supervision during the six months I spent in Denmark. To all my colleagues at DTU Biosustain, Ólafur, Christof, Laura, Tomas, Vratislav, Stefan, Marie, Peter, Solvej, Klara, Paula, Gheorge and Kira, thanks for the scientific discussions, for the support and for making lunch time very funny.

A special thanks to Solvej for having me at your house, and to Ólafur for all the support with bureaucratic tasks we have performed, including countless visits to police departments in the Danish countryside.

I would like to thank also Ditte, David and José for the friendship. You were fundamental for helping me surviving the Danish winter.

In addition, I would like to thank Carlos Rosa for providing the indigenous strains used in this work as part of our research network within a FAPESP/FAPEMIG/VALE initiative (2010/51318-2), Luiz Carlos Basso for the help with the automated plate reader TECAN, Fariza for the help with the serum vial experiments, Eutimio Gustavo Fernández for the help with statistical analysis, Vanessa Martins da Silva and Vijayendran Raghavendran, for their help with the elemental analyses.

To the friends and colleagues at LEMEB, Bruno, José, Wesley, Vijay, Suellen, Pamela, Juliana, Rosa, Laura, Cristiane, Suzan, Marcus, Daniele, Aureliano, Priscila, Fátima and Allan (from LMQA), thanks for all the support and friendship during my work at UNICAMP.

I would like to thank Isabela for the patience, comprehension, all the support and for your unyielding devotion during the past nine years.

To my parents, Dimas and Vera, and to my brother, Thiago, thanks for the unconditional support and encouragement, and for always being there to me.

I would like to acknowledge, University of São Paulo, University of Campinas and Technical University of Denmark, for accepting me as their student and researcher, and CAPES and FAPESP for the financial support in Brazil (process# 2013/17859-4) and in Denmark (process# 2014/10366-5).

“Look deep into nature, and then you will understand everything better.”

Albert Einstein

Resumo

BEATO, F. B. **Physiology of *Saccharomyces cerevisiae* strains isolated from Brazilian biomes: new insights into biodiversity and industrial applications**. 2016. 142 f. Dissertação (Mestrado em Biotecnologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2016

A levedura *Saccharomyces cerevisiae* tem servido a humanidade há milhares de anos. Atualmente, além de ser utilizada como um organismo modelo para pesquisa científica, é um dos microrganismos mais utilizados na produção industrial de etanol e de outras moléculas de valor comercial. No Brasil, *S. cerevisiae* é amplamente empregada na produção de etanol de 1ª geração (1G). Neste processo, as condições do processo de fermentação podem ser estressantes para as leveduras. As linhagens de levedura comerciais, usadas anualmente para dar início ao processo industrial no período de safra da cana de açúcar, são naturalmente substituídas por linhagens selvagens, durante o período de operação. Além do fato de não se conhecer a origem dessas leveduras, o habitat natural da espécie *S. cerevisiae* ainda não foi completamente compreendido e caracterizado, e os estudos realizados para isolar indivíduos de ambientes naturais foram conduzidos principalmente na Europa e América do Norte. O estudo de linhagens selvagens obtidas a partir de habitats naturais tem o potencial de auxiliar na elucidação dessas questões, além de poder revelar linhagens com características promissoras para aplicações industriais. Nesse estudo, avaliamos as propriedades fisiológicas e genéticas de 14 linhagens selvagens de *S. cerevisiae* isoladas de Mata Atlântica e Cerrado brasileiros, em comparação com linhagens industriais e laboratoriais bem conhecidas. Descrevemos o cultivo dessas linhagens em condições padrão e condições de estresse relevantes para a indústria do etanol, como pH ácido, temperatura elevada e alta concentração de inibidores, como etanol e ácido acético. Nossa análise comparativa dos principais parâmetros de processo que afetam a produção de etanol combustível revelaram linhagens selvagens com características promissoras para aplicações biotecnológicas. Em geral, em pH ácido, as linhagens industriais apresentaram maior tolerância que as linhagens selvagens. No entanto, quando 40 °C foi aplicado como fator de stress, uma linhagem selvagem (UFMG-CM-Y256) apresentou a maior termotolerância, e demonstrou velocidade de crescimento relativa duas vezes maior que a linhagem industrial

PE-2 (79% vs. 39%) e que a linhagem laboratorial CEN.PK113-7D (79% vs. 37%). Com os avanços na indústria de etanol combustível de segunda geração (2G), tolerância ao ácido acético também se tornou relevante, e esse estudo permitiu a identificação de algumas linhagens selvagens tolerantes. A linhagem UFMG-CM-Y267, por exemplo, apresentou a segunda maior velocidade de crescimento relativo (83%), quando 10 g L^{-1} de ácido acético foram adicionados ao meio de cultivo, semelhante às velocidades apresentadas por CEN.PK113-7D e PE-2 (75% e 76%, respectivamente), e 30% maior que a linhagem industrial JP-1 (64%). Finalmente, o cultivo mimetizando a produção industrial de etanol revelou linhagens selvagens com rendimentos de etanol maiores (5%) e produtividades semelhantes aos valores apresentados pelas linhagens industriais, como por exemplo, a linhagem UFMG-CM-Y260, que ainda apresentou tolerância levemente maior a pH ácido, em relação a outras linhagens selvagens, e à alta concentração de etanol (80 g L^{-1}), em relação às linhagens industriais.

Palavras-chave: Levedura. *Saccharomyces cerevisiae.* Etanol combustível. Tolerância a estresse. Biodiversidade. Linhagens selvagens. Biotecnologia industrial.

Abstract

BEATO, F. B. **Physiology of *Saccharomyces cerevisiae* strains isolated from Brazilian biomes: new insights into biodiversity and industrial applications**. 2016. 142 p. Master thesis (Biotechnology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2016

The yeast *Saccharomyces cerevisiae* has served human society for thousands of years and today it is a common model organism in biology and one of the preferred organisms for production of ethanol and other chemicals of commercial value. In Brazil, *S. cerevisiae* is widely used for 1st generation (1G) fuel ethanol production. However, the conditions of the non-aseptic fermentation process can be quite challenging for yeasts and usually indigenous strains naturally replace the commercial ones (used annually to start the industrial process) during the sugarcane harvesting season. Besides the fact that the origin of these yeasts is not known, the natural habitat of *S. cerevisiae* has not yet been fully understood and characterized, and the few studies aiming at isolating individuals from natural environments were conducted mainly in Europe and North America. The isolation and study of indigenous strains from selected natural habitats has the potential to shed light on these topics and also to reveal strains that present interesting industrial traits. In this study, we analyzed genetic and physiologic properties of 14 indigenous *S. cerevisiae* strains isolated from the Brazilian Atlantic Rain Forest and Cerrado biomes in comparison to industrial and laboratory strains. We describe the cultivation of these strains under standard and industrially relevant stressful conditions, *e.g.* high ethanol concentration, high acetic acid concentration, low pH and high temperature. Our comparative analysis of the main process parameters influencing fuel ethanol production revealed indigenous strains with potential traits for biotechnological applications. In general, under low pH, for example, industrial strains presented higher tolerance than indigenous strains. However, when 40 °C was applied as stress factor, an indigenous strain (UFMG-CM-Y256) showed the highest thermotolerance, and presented two times higher relative growth rate than PE-2 (79% vs. 39%) and CEN.PK113-7D (79% vs. 37%). With the advances in the 2G fuel ethanol industry, acetic acid tolerance also became relevant, and this study allowed some tolerant indigenous strains to be identified.

Indigenous strain UFMG-CM-Y267, for example, displayed the second highest relative growth rate (83%), when 10 g L^{-1} acetic acid was applied, similar to the values displayed by CEN.PK113-7D and PE-2 (75% and 76%, respectively), and 30% higher than JP-1 (64%). Finally, the cultivation mimicking industrial fuel ethanol production revealed indigenous strains with higher ethanol yields (5%) and similar productivities to industrial strains, *e.g.* UFMG-CM-Y260, which also presented slightly higher tolerance towards acidic pH, in comparison to other indigenous strains, and higher tolerance to ethanol (80 g L^{-1}), in comparison to the industrial strains tested.

Keywords: Yeast. *Saccharomyces cerevisiae.* Fuel ethanol. Stress tolerance. Biodiversity. Indigenous strains. Industrial biotechnology.

List of Figures

Figure 1 - Polypropylene square 24-deepwell microplates used for obtaining inocula. 51
Figure 2 - Polystyrene 24-roundwell microplates used for screening on Growth Profiler 1152. 51
Figure 3 - Example of a calibration curve correlating Green-Value and Abs_{600nm}..... 52
Figure 4 - Setup used for the cultivation mimicking the 1G fuel ethanol industrial process. The figure shows a schematic illustration of the serum flask containing medium (left) connected by a Norprene tube and a needle to another flask containing water (right). The air-lock created using the Norprene tube allows the gas produced during the fermentation to be released and avoids return of air into the cultivation flask (left)..... 53
Figure 5 - Electrophoresis agarose gel of *S. cerevisiae* total DNA digested or not with *Hinf* I. Well 1 shows the DNA without restriction enzyme treatment. Well 2 shows the DNA treated with *Hinf* I. M represents the DNA marker, the black arrow indicates a band which is probably composed of total DNA and the white arrow indicates RNA. O'GeneRuler DNA Ladder Mix (Life Technologies) was used as DNA ladder. 57
Figure 6 - Electrophoresis gels of mtDNA extraction using Zymolyase or Lyticase as lytic enzymes. Panel A shows non-digested DNA and Panel B shows DNA digested with *Hinf* I. O'GeneRuler DNA Ladder Mix (Life Technologies) was used as DNA ladder. 58
Figure 7 - Electrophoresis gel after DNA extraction using the protocol described by Querol et al. (1992). Well 1: non-digested DNA, without RNase treatment; well 2: non-digested DNA treated with RNase; well 3: DNA digested with *Hinf* I and treated with RNase; well 4: DNA digested with Hinf-1 without RNase treatment; well 5 shows DNA digested with *Eco*RI and treated with RNase; well 6: DNA digested with *Eco*RI without RNase treatment. Letter P means present. M represents the molecular marker (O'GeneRuler DNA Ladder Mix, Life Technologies). The black arrow indicates the 10 kb band and the white arrow indicates the 1 kb band. 59
Figure 8 - Result of *S. cerevisiae* total DNA digestion testing different RNases. Well 1: Non-digested DNA without RNase. Well 2: DNA digested with *Hinf* I without RNase. Well 3: DNA digested with *Hinf* I and treated with 1 x RNase from our lab. Well 4: DNA digested with *Hinf* I and treated with 0.05x RNase from our lab. Well 5: DNA digested with *Hinf* I and treated with 1x RNase from a different lab. Well 6: DNA digested with *Hinf* I and treated with 0.05x RNase from a different lab. M represents the molecular marker (O'GeneRuler DNA Ladder Mix, Life Technologies). The black arrow indicates the 10 kb band and the white arrow points the 1 kb band. 60
Figure 9 - Digestion of *S. cerevisiae* total DNA with *Hinf* I after DNA concentration using SpeedVac. 1 x: DNA not concentrated. 2 x: DNA concentrated 2 x. 3 x: DNA concentrated 3 x. M shows the molecular marker (O'GeneRuler DNA Ladder Mix, Life Technologies). The black arrow indicates the 10 kb band and the white arrow points the 1 kb band. 61

Figure 10 - Electrophoresis gels of the DNA fingerprinting based on minisatellite-PCR (CARVALHO-NETTO et al., 2013). DNA fingerprinting of twenty *S. cerevisiae* strains was evaluated by nine PCR preparations using different pairs of oligonucleotide primers. The names of the strains are shown above, and the ORF identifiers are shown on the left side of the figure. M indicates the molecular mass marker. The black arrow indicates the 500 bp marker band and the grey arrow indicates the 800 bp marker band. 63
Figure 11 - PCR fingerprinting using oligonucleotide primer EI-1 (DE BARROS LOPES et al., 1998) performed for 20 strains. M indicates the molecular marker. The symbols (■, □ and ●) represent the three distinct banding patterns identified. 66
Figure 12 - Result of amplification and purification of PCR products of 15 *S. cerevisiae* strains using oligonucleotide primers NL-1 and NL-4. Wells 1-14: *S. cerevisiae* indigenous strains. Well 15: CEN.PK113-7D (*S. cerevisiae* laboratorial strain). Data is not shown for *Y. lipolytica*. M indicates the marker (O'GeneRuler DNA Ladder Mix, Life Technologies) and the black arrow points to the 600 bp marker band. 67
Figure 13 - Maximum specific growth rates of 14 indigenous *S. cerevisiae* strains (UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y259, UFMG-CM-Y260, UFMG-CM-Y262, UFMG-CM-Y263, UFMG-CM-Y264, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y455, UFMG-CM-Y636 and UFMG-CM-Y643), 3 industrial *S. cerevisiae* strains (CAT-1, PE-2, JP-1) and one laboratory strain (CEN.PK113-7D). Cells were cultivated as described in the Methods section in a defined medium containing initially either 20 g L^{-1} sucrose (dark bars) or 20 g L^{-1} glucose (light bars). The means containing "*"are significantly different from their pairs within a 90% confidence level (Tukey – HSD / Multiple Sample Comparison). The error bars represent the pooled standard error using three culture replicates. 70
Figure 14 - Growth of 14 indigenous, 3 industrial (CAT-1, PE-2, JP-1) and one laboratory (CEN.PK113-7D) *S. cerevisiae* strains under stress conditions relative to their growth under standard conditions (pH 5.5 and 30 °C). Strains were cultivated as described in the Methods section in a defined medium containing initially 20 g L^{-1} glucose under four different stress conditions: pH 3.0 (A), 40 °C (B), 10 g L^{-1} acetic acid (C) and 80 g L^{-1} ethanol (D). Performance was evaluated using relative growth (%), which was calculated dividing specific growth rate under stressful condition by specific growth rate under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol). Strains were sorted in ascending order of average values (bars) calculated using two independent culture replicates (dots). 74
Figure 15 - Relative biomass increase of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions. The relative biomass increase (ΔX%) was taken as the difference in

the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. The strains were sorted in ascending order of mean values. The error bars represent the mean deviation between two independent culture replicates. 78
Figure 16 - Ethanol yield on sugars of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions. Ethanol yield (in terms of moles of ethanol produced divided by moles of hexose equivalents consumed) was obtained as the slope of the straight line from the plot of ethanol concentration against hexose-equivalent concentration. The strains were sorted in ascending order of mean values. The error bars represent the mean deviation between two independent culture replicates. *ETH* ethanol, *Eq. Hexose* equivalent hexose. 80
Figure 17 - Ethanol productivity of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions Volumetric productivity of ethanol was calculated using the formula: $\Delta P_{Eth} / \Delta t$, where ΔP_{Eth} corresponds to the final ethanol concentration minus the initial ethanol concentration, in mol L^{-1}, and Δt is 10 h. The strains were sorted in ascending order of mean values. The error bars represent the mean deviation between two independent culture replicates. 81
Figure 18 - Residual sugar at 10 h cultivation of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions. The strains were sorted in ascending order of fructose values. The error bars represent the mean deviation between two independent culture replicates. 83
Figure 19 - Spotting on low pH YPD plates of 14 indigenous *S. cerevisiae* strains, 3 industrial (CAT-1, JP-1 and PE-2) and one laboratory strain (CEN.PK113-7D). Cells were plated at five dilutions, ranging from Abs_{600nm} 10^{-1} to 10^{-5}, on YPD plates with pH adjusted to 2.5 using H_2SO_4. Pictures were taken after 72 hours of incubation at 30 °C. 85
Figure 20 – Growth of *S. cerevisiae* strains Y254, Y255, Y256, Y257 and Y259 under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 98

Figure 21 - Growth of *S. cerevisiae* strains Y260, Y262, Y263, Y264 and Y266 under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 99
Figure 22 - Growth of *S. cerevisiae* strains Y267, Y455, Y636, Y643 and CAT-1 under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 100
Figure 23 - Growth of *S. cerevisiae* strains JP-1, PE-2 and CEN.PK113-7D under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 101
Figure 24 - Growth of *S. cerevisiae* strains Y254, Y255, Y256, Y257 and Y259 under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 102
Figure 25 - Growth of *S. cerevisiae* strains Y260, Y262, Y263, Y264 and Y266 under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 103
Figure 26 - Growth of *S. cerevisiae* strains Y267, Y455, Y636, Y643 and CAT-1 under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 104
Figure 27 - Growth of *S. cerevisiae* strains JP-1, PE-2 and CEN.PK113-7D under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line

corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 105
Figure 28 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 107
Figure 29 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 108
Figure 30 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 109
Figure 31 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 110
Figure 32 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 111
Figure 33 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates

(μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 112
Figure 34 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 113
Figure 35 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 114
Figure 36 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). (○). 115
Figure 37 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 116
Figure 38 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 117
Figure 39 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the

slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 118
Figure 40 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 119
Figure 41 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 120
Figure 42 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 121
Figure 43 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 122
Figure 44 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 123
Figure 45 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 124

Figure 46 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 125
Figure 47 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 126
Figure 48 Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under ethanol stress condition (80 g L^{-1} ethanol, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 127
Figure 49 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under ethanol stress condition (80 g L^{-1} ethanol, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). 128
Figure 50 - Calculation of ethanol yield for strains UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y262 and UFMG-CM-Y263, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). 129
Figure 51 - Calculation of ethanol yield for strains UFMG-CM-Y264, UFMG-CM-Y455, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y643 and UFMG-CM-Y259, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). 130

Figure 52 - Calculation of ethanol yield for strains UFMG-CM-Y260, UFMG-CM-Y636, CAT-1, JP-1, PE-2 and CEN.PK113-7D, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). 131
Figure 53 - Calculation of ethanol yield for strains CEN.PK122 and S288C, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). 132
Figure 54 - Growth of strains UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y262 and UFMG-CM-Y263 under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■). 134
Figure 55 - Growth of strains UFMG-CM-Y264, UFMG-CM-Y455, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y643 and UFMG-CM-Y259 under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■). 135
Figure 56 - Growth of strains UFMG-CM-Y260, UFMG-CM-Y636, CAT-1, JP-1, PE-2 and CEN.PK113-7D under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■). 136
Figure 57 - Growth of strains CEN.PK122 and S288C under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown

against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■). 137

List of Tables

Table 1 - Location and name of the plant species where the indigenous strains used in this work were collected from (adapted from SAFAR, 2013). 39
Table 2 - Reference *S. cerevisiae* strains used in the study..... 42
Table 3 - Oligonucleotide primers utilized in the minisatellite PCR-based DNA fingerprinting method (adapted from CARVALHO-NETTO et al., 2013). 46
Table 4 - Conditions tested during the Lyticase evaluation test. DNA quantification was performed using NanoDrop (ThermoScientific) before enzymatic digestion (10x diluted). ... 57
Table 5 - Banding patterns obtained by the minisatellite PCR-based DNA fingerprinting method. 65
Table 6 - Result of alignment of consensus sequences of 15 *S. cerevisiae* strains and one *Y. lipolytica* strain with the reference sequence of *S. cerevisiae* NRRL Y-12632 (KURTZMAN; ROBNETT, 1998) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). 68

List of Abbreviations

1G	First generation (ethanol production process)
2G	Second generation (ethanol production process)
bp	Base pair
Eq. Hex.	Equivalent Hexose
ETH	Ethanol
FAPEMIG	Fundação de Amparo à Pesquisa do Estado de Minas Gerais
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
HPLC	High Performance Liquid Chromatography
kb	Kilobase
Lab	Laboratorial
mtDNA	Mitochondrial DNA
nDNA	Nuclear DNA
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
rDNA	Ribosomal DNA
rRNA	Ribossomal RNA
SDS	Sodium dodecyl sulfate
SE	Standard Error
TAE	Tris-EDTA acetate buffer
Tm	Melting temperature
UFMG	Federal University of Minas Gerais
UnB	University of Brasilia
UNICAMP	University of Campinas
USP	University of Sao Paulo
v/v	Volume/Volume
YP	Yeast extract-peptone culture medium
YPD	Yeast extract-peptone-dextrose medium

List of Symbols

Abs_{600}	Absorbance at 600 nm
μ_{max}	Maximum specific growth rate
P_{Eth}	Ethanol concentration in the medium
t	Time
ΔX	Biomass increase (%)

Summary

1 INTRODUCTION 27
2 OBJECTIVES 30
3 LITERATURE REVIEW 30
3.1 Biotechnological applications of *Saccharomyces cerevisiae* 30
3.2 Natural habitat and ecology of *S. cerevisiae* 32
3.3 The concept of species in microbiology 33
3.4 Mini and microscale cultivation systems 34
3.5 Strain selection for industrial fuel ethanol production 35
4 MATERIAL & METHODS 38
4.1 Microorganisms and preservation 38
4.2 Strain Identification 43
4.2.1 Restriction analysis of mtDNA 43
4.2.2 DNA fingerprinting: Minisatellite-PCR 44
4.2.3 PCR-Fingerprinting using El-1 primer 47
4.3 Species confirmation 48
4.4 Media and growth conditions 49
4.4.1 Cultivation under standard conditions 49
4.4.2 Cultivation under stressful conditions 50
4.4.3 Cultivation under conditions mimicking the 1G fuel ethanol industrial process 52
4.5 Calculation of physiological and process parameters 54
4.6 Serial dilution spot assays at low pH 55
5 RESULTS AND DISCUSSION 56
5.1 Strain identification 56
5.1.1 Restriction analysis of mtDNA provided unsatisfactory results 56
5.1.2 *Minisatellite PCR-based DNA fingerprinting resolved five out of fourteen indigenous strains* .. 62
5.1.3 PCR-Fingerprinting using EI-1 primer could work as accessory method to complement other strategies 65
5.2 Species identity was confirmed using a molecular method 67
5.3 None of the *S. cerevisiae* strains grew faster on glucose than on sucrose 69
5.4 Indigenous strains present higher tolerance to different stressors in comparison to industrial strains 71
5.5 Some indigenous strains displayed slightly higher ethanol yields than industrial strains 76

5.6 Spotting on low pH YPD plates evidences the higher tolerance of industrial strains to acidic conditions 83

6 CONCLUSIONS AND OUTLOOK 86

References* 88

APPENDIX A – SUPPLEMENTARY RESULTS 97

A.1 Growth of *S. cerevisiae* strains under standard conditions using the automated plate reader Infinity M200 97

A.2 Growth of *S. cerevisiae* strains under standard and stressful conditions using Growth Profiler 1152 (Enzyscreen) 105

A.3 Ethanol yield calculation 128

A.4 Relative biomass increase 133

A.5 Sequences of the D1/D2 domain of the 26S ribosomal DNA 138

1 INTRODUCTION

The yeast *Saccharomyces cerevisiae* is one of the most important microorganisms utilized in fermentation processes and it has been associated to human activities for thousands of years. Besides its application in the food industry, in the last few decades this yeast has also been utilized for the production of heterologous proteins, biofuels and other chemicals within the "white biotechnology" concept. Furthermore, *S. cerevisiae* was the first eukaryotic organism to have its genome sequence completely determined (MEWES et al., 1997), and it has been extensively used as an eukaryotic model organism for research in various areas, including human disease (SIMON et al., 2000). In spite of its undoubtable importance to human life, the natural habitat of *S. cerevisiae* has not yet been fully understood and characterized.

In Brazil and in North America, *S. cerevisiae* is widely used for 1st generation (1G) fuel ethanol production (GOMBERT; VAN MARIS, 2015). Particularly in Brazil, the conditions of the non-aseptic fermentation system can be quite challenging for yeast. Typically, a number of stress factors is present, including: the process temperature usually varies between 32 °C and 35 °C, and can reach up to 40 °C in some warmer regions of Brazil (BASSO; BASSO; ROCHA, 2011); the ethanol concentration reaches 8-11% (v/v) towards the end of each fermentation cycle; bacterial contamination is a common issue (BASSO; BASSO; ROCHA, 2011); high sugar concentration is present in the beginning of the fermentation (DELLA-BIANCA et al., 2013); osmotic stress caused by salts in molasses is also a concern (BASSO; BASSO; ROCHA, 2011; DELLA-BIANCA et al., 2013) and the cells are intensively recycled, passing through a severe acid treatment in order to decrease bacterial contamination (incubation under pH 1.8 to 2.5 is used for 1 to 2 h) (WHEALS et al., 1999). Currently, several initiatives aim at introducing large-scale commercial 2nd generation (2G) fuel ethanol production. In this type of process, differently from the 1G technology, pre-treatment and hydrolysis of lignocellulosic materials are required, which lead to the release of compounds that act as inhibitors of yeast performance, acetic acid being one of the strongest inhibitors,

particularly in the case of sugarcane bagasse (ALMEIDA et al., 2007). Because of these harsh conditions, very few yeast strains that present favorable technological properties (such as absence of flocculation and foam formation), are able to survive and persist in the fermentation vats during the entire sugar cane crushing season (which encompasses an 8-month period from April to November in the Southeast of Brazil). In fact, there are currently only five strains commercially available for the almost 400 industrial sugar-cane mills in Brazil: PE-2, CAT-1, BG-1 and VR-1 in the Southeast (BASSO et al., 2008) and JP-1 in the Northeast (DA SILVA-FILHO et al., 2005a). Because these mills are located in many geographical regions, the process environments can be remarkably different due to varying microclimates and type of sugarcane used as raw material. The composition of molasses, for instance, is extremely variable. Since this substrate is a byproduct of the sugar production, its composition may vary not only with the sugar cane variety but also with the operational conditions of the sugar factory, which is dependent on the world's sugar market (ANDRIETTA; ANDRIETTA; STUPIELLO, 2011). Besides the variable sugar concentration, a series of secondary compounds may be present in molasses and influence the fermentation process, such as hydroxymethylfurfural (HMF), organic acids (lactic, acetic, formic), melanoidines and others (ANDRIETTA; ANDRIETTA; STUPIELLO, 2011; BASSO, 2015). In addition to the feedstock variability, the operation method may vary among different mills. Although the fermentation processes present the same stages (fermentation per se, cell separation unit and recycled yeast unit), the operation method can vary between fed-batch and continuous processes (ANDRIETTA; ANDRIETTA; STUPIELLO, 2011).

Presently, baker's yeast is normally used as the major starter strain in the Brazilian fuel ethanol industry. However, already some weeks after the start of the crushing season and operation, the starter strain is naturally replaced by indigenous *S. cerevisiae* strains (such as PE-2, CAT-1 and others) (BASSO et al., 2008). One important question to be answered is therefore: Where do these strains, which can persist and dominate in the fermentation environment, come from? Do they originate from the sugar-cane field or are they brought to the industrial environment by birds or insects?

Brazil is known to have a wide biodiversity. The country's Atlantic Forest, for example, contains more than 8,000 endemic plant species and is classified as one of the

richest areas in the world in terms of plants and vertebrates species (MYERS et al., 2000). Therefore, it is expected that the microbiological diversity is also high, regarding both the number of species and the variety of strains within a given species.

Within this context, research on indigenous strains represents an opportunity to give us a deeper understanding on the natural habitat of *S. cerevisiae*. Such studies may also uncover strains with relevant features for industrial application, such as stress tolerance, high specific growth rates or high product yields, as already exemplified by Ramos et al., (2013). In this work, we performed genetic and physiological comparisons among indigenous, industrial and laboratory strains of the yeast *Saccharomyces cerevisiae*, using cultivations under standard and stressful conditions. Our comparative analysis of the main process parameters influencing fuel ethanol production revealed indigenous strains with potential traits for biotechnological applications, such as high specific growth rates and tolerance towards low pH, acetic acid and high temperatures.

The indigenous strains studied in this work were provided by Prof. Carlos A. Rosa (Federal University of Minas Gerais, UFMG), as part of our research network within a FAPESP/FAPEMIG/VALE initiative. Part of the work was conducted at the Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark (DTU Biosustain), during a six months internship performed by F.B.B., as part of a research collaboration with Prof. Jochen Forster. Cultivations under standard conditions using the microplate reader Tecan Infinity® were performed at Prof. Luiz Carlos Basso's laboratory at ESALQ/USP. The stress spot assays and the restriction analysis of mtDNA were performed at the Bioprocess & Metabolic Engineering Lab (LEMeB) at Unicamp, where the project was initially designed.

2 OBJECTIVES

The objective of this work was to analyze fourteen indigenous *S. cerevisiae* strains, isolated from Brazilian biomes, by comparing their genetic and physiological properties, under standard and industrially relevant stressful conditions, with traditional laboratory and industrial strains used for fuel ethanol production in Brazil.

3 LITERATURE REVIEW

3.1 Biotechnological applications of *Saccharomyces cerevisiae*

Since ancient times, human societies have learned how to produce fermented products. Evidence for the production of a fermented beverage, in China, dates from 7,000 BC (MCGOVERN, 2003). There are also evidences that wine was produced in Iran at around 6,000 BC and in Egypt at around 3,000 BC (CAVALIERI et al., 2003; MCGOVERN et al., 1997). In this context, *S. cerevisiae* has been one of the most important microorganisms utilized in fermentation processes and it has been associated to different human activities (LEGRAS et al., 2007). Analysis of DNA samples isolated from a presumed wine jug at Abydos (3150 BC) has shown high similarities to the sequence of modern *S. cerevisiae* (CAVALIERI et al., 2003), indicating that this yeast has been used for millennia. Today, it is still considered the main agent in wine fermentation, one of the most expressive alcoholic beverage markets in the world. In 2012, global wine production reached 25 billion liters (Wine Institute, World Wine Production by Country, 2013), resulting in revenues of 25.3 billion Euros (International

Organization of Vine and Wine, The wine market: evolution and trends. 2014). Although beer technology is supposed to be almost as ancient as wine with origins around 10,000 BC (AXCELL, 2007), archeological evidence of brewing activity dates from 1,800 BC (KATZ; MAYTAG, 1991). Since then, many changes have been made to beer technology; however, *S. cerevisiae* still plays an important role in beer brewing by its utilization in ale and weiss beer production (LODOLO et al., 2008). Another activity strictly related to *S. cerevisiae* and of great importance to human society is bread making. Evidence shows that use of leavening in baking appeared around 4,000 BC (MCGEE, 1984), resulting in a lighter and more pleasant bread since then.

Besides its application in the food industry, *S. cerevisiae* has been utilized to produce fuel ethanol, the main biofuel used for transportation. In 2014, Brazil produced approximately 27.5 billion liters of this fuel using sugar-cane as a raw material (Brazilian Sugar Cane Industry Association 2015, 2013/2014 season) and is considered the greatest exporter and the second largest producer in the world (BASSO; BASSO; ROCHA, 2011). Together, the sugar and energy market in Brazil employs more than 4.3 million people, and the ethanol market alone generates more than R$ 35.9 billion per year as revenue (BATISTA, 2013).

Biosynthesis of heterologous proteins can also be performed industrially with yeasts. Unlike prokaryotic cells, yeasts present many of the post–translational folding, processing and modification events required to produce bioactive mammalian proteins (BUCKHOLZ; GLEESON, 1991). In addition, as unicellular microorganisms, they exhibit fast growth, simple cultivation technology and easiness for genetic engineering. Because of the vast amount of knowledge generated by the scientific community on *S. cerevisiae*, this species has been the most common host for heterologous protein production.

It is important to recall that *S. cerevisiae* presents properties that lower the risk of contamination in industrial fermentations. It is relatively tolerant to low pH, high ethanol and high sugar concentrations. For this reason, its use in the industrial “white biotechnology” context, focusing on the fermentative production of industrially relevant biochemicals, *e.g.*, propanediol, glycerol, organic acids, sugar alcohols, steroids, and isoprenoids, has been increasing (NEVOIGT, 2008). Lately, with advances in metabolic

engineering, *S. cerevisiae* has also been used for the production of higher value compounds, *e.g.*, the artemisinin drug used for malaria treatment, a disease that affected more than 200 million people in 2010 (PADDON et al., 2013).

3.2 Natural habitat and ecology of *S. cerevisiae*

Since *S. cerevisiae* has been largely used by humans for thousands of years in baking, brewing and wine making, it has been for a long time considered a domesticated organism, distinct from any of its counterparts in natural habitats (CIANI et al., 2004). However, recent studies have shown that *S. cerevisiae* existed as a distinct species in natural environments before it was utilized by humankind (FAY; BENAVIDES, 2005; LITI et al*.,* 2009; WANG et al*.,* 2012). By analyzing the nucleotide variation of more than 80 *S. cerevisiae* isolates collected worldwide, Fay and Benavides (2005) obtained indications that support the hypothesis that domesticated strains derive from natural populations.

In wine making, when grapes are gathered and crushed, spontaneous fermentation occurs (GODDARD, 2008), and usually *S. cerevisiae* is the primary microbe associated with this process. Thus, it would seem logical to conclude that its natural habitat includes grapes and the other fruits used to make alcoholic beverages (GODDARD; GREIG, 2015). However, the artificial fermentation environment differs dramatically from the natural fruit habitat. Indeed, while *S. cerevisiae* can be easily found in grapes damaged by birds and/or insects (24% of the damaged grapes contain *S. cerevisiae*), it is relatively unusual to encounter this species on the surface of undamaged grapes (< 0.1% of undamaged grapes contain *S. cerevisiae*) (LANDRY et al., 2006; MORTIMER; POLSINELLI, 1999). Several studies have aimed at characterizing indigenous strains, and *S. cerevisiae* isolates have been obtained not only from grapes damaged by birds or insects (MORTIMER; POLSINELLI, 1999), but also from different locations such as: exudates of trees or various fruits (NAUMOV et al., 2003); flower nectars (LITI et al*.,* 2009); the intestines of insects (STEFANINI et al., 2012); and soil or oak

trees (SAMPAIO; GONÇALVES, 2008; SNIEGOWSKI; DOMBROWSKI; FINGERMAN, 2002). In Brazil, isolates have been obtained occasionally in lakes or associated to flies, but SAFAR (2013) and BARBOSA et al., (2016) describe the isolation of several strains from tree barks of *Quercus rubra* (Red oak or *Carvalho*) and *Tapirira guianensis (Pau-pombo* or *Tapiriri)* species in different Brazilian biomes. In spite of its undoubtable importance to human life, the natural habitat of *S. cerevisiae* has not yet been fully understood and characterized. Perhaps, as proposed by Goddard and Greig (2015) and addressed by Marques et al., (2015), this vast number of habitats indicates that *S. cerevisiae* might be a nomad yeast, able to survive as a generalist in a wide range of environments.

3.3 The concept of species in microbiology

The concept of species in microbiology is still very controversial. There is not such a theory for the microbial world as the biological species concept proposed by Mayr (1942): “Species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups.”. As described by Ogunseitan (2004): “The concept of biological diversity implies consensus on the discrete nature of independent species and on the mechanisms that generate speciation. The recognition of differences and similarities among the discrete features of microorganisms is more challenging and less well understood than for large multicellular organisms.”

Various concepts have been proposed for microbial species, but none has been widely accepted, very likely because all of these concepts make use of methodological considerations (ACHTMAN; WAGNER, 2008). Currently, microbial species are defined by a pragmatic approach based on rules for both phenotypic and genotypic properties discussed in 1996 by Vandamme et al., (1996). Until 2008, this definition provided the differentiation of more than 7,031 accepted microbial species (ACHTMAN; WAGNER, 2008), but it relies on a certain cut-off level for pairwise genomic DNA-DNA hybridization. The cut-off number was

chosen 20 years ago to match previously identified species and it is not based on any theoretical justification. According to Wayne et al*.,* 1987, microorganisms are assigned to a common species if their pairwise DNA re-association levels are ≥ 70% in DNA-DNA hybridization experiments under standardized conditions and if their ΔTm (melting temperature) is ≤ 5 °C. Additionally, species description may be based on more than one strain and all strains within a species have to present a certain degree of phenotypic consistency. Similarity between 16S rDNA sequences of prokaryotes or 26S rDNA of eukaryotes present strong correlation with the DNA-DNA hybridization and are also used to differentiate species. Microorganisms that are ≤ 98.7% identical on the 16S rDNA sequence present < 70% DNA-DNA similarity and are considered members of different species (STACKEBRANDT; EBERS, 2006). Still, the opposite is not certainly true, and distinct species have been occasionally described with identity of 16S rDNA > 98.7*%* (ACHTMAN; WAGNER, 2008).

3.4 Mini and microscale cultivation systems

Mini- (< 100 mL) and microbioreactors (< 1 mL) play an important role in small-scale cultivation and high-throughput screening experiments in modern biotechnological process development. Recently, new monitoring technologies and control strategies for small-scale bioreactors have been developed, increasing the utilization of smaller working volumes and automated equipment for microorganism cultivations (LATTERMANN; BÜCHS, 2015). Two pertinent examples of stirred minibioreactor systems that are commercially available are: bioREACTOR 48 system, 8–15 mL (2mag, Germany) and the automated ambr system for cell line cultivation, 10–15 mL (TAP Biosystems, United Kingdom). For microscale applications, there are, for example: the BioLector system, 0.8–2.4 mL (m2p-labs GmbH, Germany) and the MicroFlask system, 0.1–4 mL (Applikon, The Netherlands). For shake flasks, there are systems available that allow for oxygen transfer rate (OTR) measurements in 250 mL working

volumes: HiTec Zang GmbH, Germany and Kühner AG, Switzerland. Among other examples, there is also the Growth Profiler 1152 (Enzyscreen, The Netherlands), which provides high oxygen transfer rates and automated cell density measurement, enabling aerobic cultivations in microwell-plates (0.1 mL to 25 mL).

Employment of these technologies brings many benefits to research, as it accelerates process development and reduces the large costs related to the utilization of pilot- and industrial-scale bioreactors (NEUBAUER et al*.,* 2013). For that reason, intensive research has been made in this area, trending to a decrease in working volumes and increase in the utilization of automated cultivation platforms to determine quantitative physiological data of microorganisms. Still, larger fermenters and pilot- or industrial-scale bioreactors are highly needed in the scaling-up of bioprocesses, since many transport phenomena can only be observed and investigated in larger volumes.

3.5 Strain selection for industrial fuel ethanol production

Fuel ethanol mills in Brazil traditionally use baker's yeast as the starter strain for fermentation, mainly because of the low costs and high availability. However, in each sugarcane crushing season, already a few weeks after operation has started, these baker's strains are naturally substituted by contaminant *S. cerevisiae* strains, which tend to persist and dominate in the fermentation environment (DA SILVA-FILHO et al., 2005b). Basso et al., (1993) observed that baker's strains were unable to compete with indigenous strains, and Da Silva-Filho et al., (2005b) confirmed that only strains that had previously been isolated from ethanol mills and used as starter strains could be detected during almost the whole crop season. Irrespective of their dominance and persistence, most of these strains present undesirable traits for industrial ethanol production, such as excessive foam formation, flocculation, long fermentation time and high residual sugar after fermentation.

Basso et al., (2008) performed a study in which 340 indigenous strains were isolated from 50 different distilleries during the 1993-2005 crop seasons and had their fermentation performances evaluated under industrial-like conditions. Among the isolated strains, 67% produced excessive foam, 33% were flocculant and 53% were not able to consume all the sugar present in the medium; only 20% of all strains tested did not present such unfavorable traits. After a round of screening in which ethanol yield, glycerol formation, cell viability and intracellular levels of trehalose and glycogen were evaluated, 14 selected strains were reintroduced in different distilleries to assess their implantation capability. Due to their persistence and dominance after several fermentations, strains such as PE-2 and CAT-1 were identified and are still used today as starter strains (usually combined with a higher proportion of baker's yeast), in order to improve ethanol yields and contribute to good fermentation performance and stability in many distilleries in Brazil.

In another strain selection program, Da Silva-Filho et al., (2005a) isolated and selected the JP-1 strain, an indigenous *S. cerevisiae* strain that exhibited similar fermentative capacity and stress tolerance compared to the most commonly used commercial strains and which later showed to be highly adapted to different industrial units in the Northeast of Brazil.

In a different study, Stambuk et al., (2009) evaluated five different yeast strains (BG-1; CAT-1 PE-2, SA-1 and VR-1) isolated from Brazilian fuel ethanol plants using microarray-based comparative genome hybridization (aCGH). The authors observed increased copy numbers of the genes involved in the biosynthesis of vitamins B6 (pyridoxine) and B1 (thiamin) for all five strains, in comparison to the laboratory S288c strain. They also showed that these gene copy number variations (CNVs) confer the ability to laboratory strains to grow more efficiently under the repressive effects of thiamin, especially in the absence of pyridoxine and with high sugar concentrations, suggesting that these genetic changes provide adaptive advantages to the industrial fermentation conditions.

In another work, Steckelberg (2001) conducted a study in which strains were isolated from 19 Brazilian ethanol fermentation processes and characterized under standardized conditions. The author observed a high variability among the strains studied, which,

according to Andrietta, Andrietta and Stupiello (2011), points to a high diversity of yeast strains inhabiting the Brazilian industrial processes.

4 MATERIAL & METHODS

4.1 Microorganisms and preservation

The fourteen indigenous *S. cerevisiae* strains studied in this work were isolated by the group of Dr. Carlos A. Rosa (Federal University of Minas Gerais, Belo Horizonte, Brazil) from tree barks of three different plant species (SAFAR, 2013; BARBOSA et al., 2016). The strains UFMG-CM-Y254, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y643 and UFMG-CM-Y259 were obtained from barks of *Quercus rubra* (Red Oak) located in Santuário do Caraça (Minas Gerais) within the Atlantic Forest biome; UFMG-CM-Y255 was isolated from barks of a non-identified tree located in the same region; and UFMG-CM-Y262, UFMG-CM-Y263, UFMG-CM-Y264, UFMG-CM-Y455, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y260 and UFMG-CM-Y636 strains were isolated from barks of *Tapirira guianensis* (Tapirirá) located in Taquaruçu (Tocantins) within the Cerrado biome (**Table 1**). It is noteworthy that *Q. rubra* trees are not native to Brazil and were imported from North America in the first half of the 1900's (BARBOSA et al., 2016). The methods for yeast isolation are described in Barbosa et al., (2016). These strains are deposited in the Collection of Microorganisms and Cells of the Federal University of Minas Gerais, Brazil (AUTORIZAÇÃO DE ACESSO AO PATRIMÔNIO GENÉTICO PARA FINS DE PEQUISAS CIENTÍFICA No 010625/2015-1).

Table 1 - Location and name of the plant species where the indigenous strains used in this work were collected from (adapted from SAFAR, 2013).

Collection site	Strain*	Plant species (sample)
Santuário do Caraça (Jardins), MG	UFMG-CM-Y254 [a]	*Quercus rubra* (Red oak or Carvalho)
	UFMG-CM-Y256 [a]	*Quercus rubra* (Red oak or Carvalho)
	UFMG-CM-Y257 [a]	*Quercus rubra* (Red oak or Carvalho)
	UFMG-CM-Y643 [d]	*Quercus rubra* (Red oak or Carvalho)
	UFMG-CM-Y259 [d]	*Quercus rubra* (Red oak or Carvalho)
Santuário do Caraça (Trilha da Cascatona), MG	UFMG-CM-Y255 [f]	Non-identified tree
Taquaruçu (Mata ripária do Córrego Buritizal), TO	UFMG-CM-Y262 [g]	*Tapirira guianensis* (Tapirirá)
	UFMG-CM-Y260 [i]	*Tapirira guianensis* (Tapirirá)
Taquaruçu (Mata ripária do Córrego Bela Vista), TO	UFMG-CM-Y636 [j]	*Tapirira guianensis* (Tapirirá)
	UFMG-CM-Y263 [h]	*Tapirira guianensis* (Tapirirá)
	UFMG-CM-Y264 [b]	*Tapirira guianensis* (Tapirirá)
	UFMG-CM-Y266 [b]	*Tapirira guianensis* (Tapirirá)
	UFMG-CM-Y455 [c]	*Tapirira guianensis* (Tapirirá)
	UFMG-CM-Y267 [c]	*Tapirira guianensis* (Tapirirá)

*a, b, c, d, f, g, h, i and j represent the groups in which indigenous strains were clustered using the minisatellite-PCR described in the Results & Discussion section. Strains assigned to the same group presented the same banding pattern when the minisatellite-PCR was performed.

The cells of indigenous strains were received in dry filter membranes and were inoculated onto YPD medium plates (10 g L^{-1} yeast extract, 20 g L^{-1} peptone, 20 g L^{-1} glucose and 15 g L^{-1} agar) and cultivated at 30 °C until growth was observed. Using an inoculation loop, cells (approximately one colony) were transferred to conical tubes containing 5 mL of YPD liquid medium (10 g L^{-1} yeast extract, 20 g L^{-1} peptone and 20 g L^{-1} glucose) and incubated at 30 °C and 250 rpm. On the next day, 30% (v/v) glycerol was added and 1.5 mL aliquots were stored at -80 °C. All experiments were started by transferring an aliquot directly from the frozen stocks onto the surface of solid YPD plates. The plates were streaked using an inoculation loop in order to obtain isolated colonies and incubated at 30 °C for 48 h.

A list of all the reference *S. cerevisiae* strains used in the study is shown in **Table 2**. The laboratory reference strains of the CEN.PK family (CEN.PK113-7D haploid and CEN.PK122 diploid) were chosen as control/reference strains, since they are widely used by the scientific community (VAN DIJKEN et al., 2000). In addition, the genome sequence of the CEN.PK113-7D strain is publicly available (NIJKAMP et al., 2012). The reference strain S288c was also included in this study, due to the detailed genome sequencing carried out in 1996 (MEWES et al., 1997) and to the *Saccharomyces* Genome Database project (http://www.yeastgenome.org). The industrial strains PE-2 and CAT-1 were chosen because they are currently widely used in industrial fuel ethanol plants in Brazil and present remarkable capacity for surviving and dominating in non-aseptic industrial fermentations (BASSO et al., 2008). For the CAT-1 strain a full diploid genome sequence is available publicly, whereas for the PE-2 strain the sequence of a haploid spore was published (ARGUESO et al., 2009; BABRZADEH et al., 20121 x). The strain JP-1 was also included in this study because it was isolated from an ethanol plant in Northeastern Brazil, which presents different process conditions when compared to the Southeastern region (DA SILVA-FILHO et al., 2005a) and because its genome sequence will soon be made available (Fernando Araripe Torres, University of Brasília, UnB, personal communication). For stock preparation purposes, all reference strains were re-streaked onto new YPD plates from cryotubes or from YPD plates (**Table 2**) and incubated at 30 °C. On the next 24 - 48 h, cells (approximately one colony) were transferred to conical tubes containing 5 mL of YPD liquid medium and incubated at 30 °C and 250 rpm. On the next day, 30% (v/v) glycerol was added and 1.5 mL

aliquots were stored at -80 °C. The experiments were started following the same procedure described for the indigenous strains.

Table 2 - Reference *S. cerevisiae* strains used in the study

Strain	Group	Origin	Receipt	Laboratory	References
CEN.PK113-7D	Laboratory	Dr. P. Kötter (Euroscarf, Germany)	Cryotube stock	UNICAMP/DTU	Nijkamp et al., (2012);van Dijken et al., (2000)
CEN.PK122	Laboratory	Dr. P. Kötter (Euroscarf, Germany)	YPD plates	DTU	van Dijken et al., (2000)
S288C	Laboratory	Mewes et al., (1997)	YPD plates	DTU	Mewes et al., (1997)
CAT-1	Industrial	Dr. L. C. Basso (USP, Brazil)	YPD plates	UNICAMP/DTU	Basso et al., (2008); Babrzadeh et al., (2012)
PE-2	Industrial	Dr. L. C. Basso (USP, Brazil)	Cryotube stock	UNICAMP/DTU	Basso et al., (2008); Argueso et al., (2009)
JP-1	Industrial	Dr. M. Morais Jr. (UFPE, Brazil)	YPD plates	UNICAMP/DTU	da Silva-Filho et al., (2005a)

4.2 Strain Identification

Three different methodologies were performed in an attempt to differentiate and verify the strains used in this work: restriction analysis of mtDNA (QUEROL et al., 1992), minisattelite-PCR (CARVALHO-NETTO et al., 2013) and PCR-fingerprinting using oligonucleotide primer EI-1 (DE BARROS LOPES et al., 1998).

4.2.1 Restriction analysis of mtDNA

The technique of restriction analysis of mtDNA described by Querol et al., 1992 was the first method performed in this work for strain differentiation. The procedure consists of total DNA extraction (mitochondrial and genomic), digestion with specific restriction enzymes and analysis of the banding pattern produced by electrophoresis. Since mtDNA contains 75% of A/T base pairs (GRAY, 1989) with only 200 regions rich in G/C base pairs, the occurrence of GCAT sites on this nucleic acid is low. Therefore, enzymatic digestion of total DNA using restriction enzymes specific for GCAT sites promotes many cuts on the nuclear DNA, but only a few cuts on mtDNA. Thus, the nDNA is cut in very small molecules, and using electrophoresis it is then possible to differentiate yeast strains by comparing the banding pattern produced by the mtDNA digestion (FRUTOS, 2007).

The DNA extraction was performed according to Querol et al., 1992. First, a colony was transferred from YPD plates to 1.5 mL Eppendorf tubes containing 800 µL of YPD medium (10 g L^{-1} yeast extract, 20 g L^{-1} peptone and 20 g L^{-1} glucose, pH 6.5). The cells were incubated at 30 °C and 200 rpm in an Innova shaker (New Brunswick, USA) overnight. The tubes were then centrifuged at 13,500 g for 3 min. The supernatant was discarded, cells were ressuspended in 1 mL of milli-Q water and centrifuged again at 13,500 g for 3 min. The

pellet was ressuspended in 500 µL of a solution containing 0.9 M Sorbitol and 0.1 M EDTA. 30 µL of Zymolyase (0.02 U/µL) (from Yeast Plasmid Miniprep kit - Zymoprep) or Lyticase (0.3 U/µL) (Sigma-Aldrich) were added, and the tubes were vortexed. After 20 min at 37 °C, the cells were harvested at 13,500 g and 3 min, and the pellet was ressuspended in 500 µL of a solution containing 50 mM Tris-HCl and 20 mM EDTA (pH 7.4). Thirteen microliters (13 µL) of 10% SDS were added and tubes were vortexed. After 5 min at 65 °C, 200 µL of 5 M Potassium Acetate (4 °C) were added, and tubes were kept on ice for 5 min. The suspension was centrifuged at 18,400 g and 4 °C for 15 min, and the supernatant (700 µL) was transferred to another tube containing 700 µL of isopropanol. The tubes were homogenized by inversion and kept at room temperature for 10 min. After centrifugation at 13,500 g for 10 min, the supernatant was discarded and the pellet was washed twice with 500 µL ethanol. The DNA pellet was kept at 37 °C during 1 h for drying. Finally, DNA was ressuspended in 20 µL of water and stored at -20 °C. The quantification was carried out using NanoDrop 2000 UV-Vis Spectrophotometer (ThermoScientific).

Eighteen microliters (18 µL) of DNA were then treated with restriction enzyme *Hinf* I by adding 2.5 µL of 10x Buffer Red (Sigma-Aldrich), 1 µL of Ribonuclease A, 1 µL of *Hinf* 1 (Sigma-Aldrich) and 2.5 µL of water. After incubation at 37 °C overnight for 25 min, the digestion product was resolved in 1.5% agarose gel containing SYBR safe Gel Stain (Life Technologies).

4.2.2 DNA fingerprinting: Minisatellite-PCR

Still for strain identification, a minisatellite-PCR methodology based on polymorphic sites of the *S. cerevisiae* genome was applied **(Table 3)** (CARVALHO-NETTO et al., 2013). Colony PCR was performed by transferring cells from a colony on YPD plates to 300 µL PCR tubes containing 10 µL Milli-Q water. For cell lysis and DNA release, tubes were submitted to temperature shocks as follows: 65 °C for 0.5 min, 8 °C for 0.5 min, 65 °C for 1.5 min, 97 °C for

3 min, 8 °C for 1 min, 65 °C for 3 min, 97 °C for 1 min, 65 °C for 1 min and 80 °C for 5 min. The samples were then spun down for 1 min and the supernatant was used as DNA template. Nine different PCR reactions were performed individually using 12.5 µL OneTaq Quick-Load 2 x Master Mix (New England Biolabs, Ipswich, USA), 1 µL of forward primer (20 µM), 1 µL of reverse primer (20 µM) **(Table 3)**, 8 µL of DNA template solution and 2.5 µL of water (25 µL total volume). The reaction was run for 8 cycles as follows: 94 °C for 35 sec, 59-52 °C (starting at 59 and decreasing 1 °C every cycle) for 45 sec, 68 °C for 2.5 min, and for 32 cycles as follows: 94 °C for 35 sec, 52 °C for 45 sec and 68 °C for 3.5 min. An initial 6 min denaturation at 94 °C and a final 5 min extension at 68 °C were used. PCR products were resolved in 2% agarose gel containing Gel Red (ThermoScientific, Waltham, USA) in TAE buffer.

The amplification products of each pair of primers were analyzed separately. The number one was attributed to the bands presented by the S288c strain, while random numbers were attributed to the other bands encountered according to their size. Strains presenting different band sizes for a particular ORF were identified with different numbers, while strains displaying the same band size for a particular ORF were attributed the same number for this specific ORF. The numbers were used to simplify the analysis and label the DNA bands, so they do not correlate to band size proportionally.

Table 3 - Oligonucleotide primers utilized in the minisatellite PCR-based DNA fingerprinting method (adapted from CARVALHO-NETTO et al., 2013).

ORF identifier	Gene	Expected PCR product for the S288C strain (bp)	5'-3' primer sequence		Melting temp. (°C)
			Forward	Reverse	
YBL046W	*PSY4*	514	GCAGTTGGCAGGTTTTTGAT	GCCTGCCTATCTTCGTCATC	55
YDR299W	*BFR2*	433	AGCGGATCAAATTTCCGATA	TGCATGCCTTTTCTGTTGAG	52
YDR504C	*SPG3*	564	TCGACAACTTTCAGATTGCTG	AATGGAAGCGGCTGTTAGAA	55
YDR505C	*PSP1*	425	ACAGTTTCGCCGGAATACAC	CCATGTTCTTCGCGTCATAA	52
YFL024C	*EPL1*	385	ACGATTCCAAATACGACGAA	TTCTGTTTCGCTTCTGAATTG	57
YKL163W	*PIR3c*	391	TGTCGCCTCATCTAAAGCAA	TGTAATTTGGGATGCAGCAG	54
YKL201C	*MNN4c*	699	TAGACCTTTTTGCGCCAACT	ATTACCACGATTCCGTCGAA	57
YLL021W	*SPA2c*	655	GAAAATGACGATGCAGACGA	AGGACTCGCTTTCCCTTACC	55
YPR143W	*RRP15*	562	GGGCTCCAAGCACAGAGTAG	CTTTCTGAGCGCTGTCTCCT	57

4.2.3 PCR-Fingerprinting using El-1 primer

The PCR Fingerprinting method described by DE BARROS LOPES et al., (1998) was also performed to complement strain differentiation. This methodology is based on the analysis of banding patterns generated by PCR using the oligonucleotide primer El-1, complementary to intron splice sites. Due to the presence of various intron splice sites along the genome, primer El-1 can anneal at different positions to both DNA strands and promote the amplification of different bands sizes. The existing polymorphisms on these sites permit both strain differentiation and species identification (DE BARROS LOPES et al., 1998).

After over-night growth on YPD medium, 200 µL of each cell suspension were harvested and ressuspended in 100 µL of 0.2 M lithium acetate + 1% SDS solution. The tubes were then vortexed and incubated at 70 °C for 5 min for cell lysis. The suspension was centrifuged and ressuspended in 300 µL of ethanol 100% for DNA precipitation. After centrifugation, the DNA pellet was washed with 400 µL of ethanol 70% and centrifuged again. The pellet was ressuspended in 100 µL of milli-Q water and centrifuged for the last time. The supernatant was then used as DNA template. The PCR reaction was performed using 12.5 µL OneTaq® Quick-Load® 2 x Master Mix (New England Biolabs), 6.0 µL of primer El-1 (5'-CTGGCTTGGTGTATG) at 20 µM, 1 µL of DNA template, and 5.5 µL of water (25 µL total volume). The reaction was run for 33 cycles as follows: 94 °C for 1 min, 45 °C for 2 min and 74 °C for 1.5 min. An initial 3 min denaturation at 94 °C and a final 5 min extension at 74 °C were used. PCR products were resolved in 2% agarose gel containing Gel Red (ThermoScientific) in TAE buffer.

4.3 Species confirmation

In order to establish a reliable method for identifying different yeast species and confirm the identity of the strains used in this work, the sequencing of the region D1/D2 of rDNA 26S was performed according to Lachance et al., (1999) and Kurtzman and Robnett (1998). The DNA of all 14 indigenous *S. cerevisiae* strains, the laboratory CEN.PK113-7D *S. cerevisiae* strain and the commercial Po1g *Yarrowia lipolytica* strain (used as an outlier) were extracted and submitted to PCR using the oligonucleotide primers NL-1 (5'-GCATATCAATAAGCGGAGGAAAAG) and NL-4 (5'-GGTCCGTGTTTCAAGACGG) following the exact conditions described by Kurtzman and Robnett (1998). After amplification of the desired bands, DNA was purified using the GenElute™ PCR Clean-Up kit (Sigma-Aldrich, St. Louis, MO) and submitted for sequencing at Centro de Estudos do Genoma Humano, Universidade de São Paulo (*S. cerevisiae* strains) or at LaCTAD, Universidade Estadual de Campinas (*Y. lipolytica*).

The sequencing of *S. cerevisiae* strains was performed using oligonucleotide primer NL-1 in duplicate, while sequencing of the *Y. lipolytica* strain was performed using 4 oligonucleotide primers: NL-1, NL-2A (5'-CTTGTTCGCTATCGGTCTC), NL-3A (5'-GAGACCGATAGCGAACAAG) and NL-4. After consensus sequence was assembled and low quality detected peaks were eliminated, sequences of different lengths were generated (439-542 bp). They were then aligned to the reference sequence of the rRNA 26S gene of *S. cerevisiae* strain NRRL Y-12632 (KURTZMAN; ROBNETT, 1998) using the Align tool of the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

4.4 Media and growth conditions

4.4.1 Cultivation under standard conditions

The growth profiles of all indigenous and reference strains were evaluated under standard conditions using 96 micro-well plates (polystyrene flat bottom plates -TPP Techno Plastic Products AG, Trasadingen, Switzerland) and the automated plate reader Infinity M200 (Tecan, Männedorf, Switzerland) at Prof. Luiz Carlos Basso's laboratory, ESALQ/USP. Pre-cultures were prepared by transferring cells from one colony on YPD plates to Eppendorf tubes (1.5 mL) containing 200 µL of a defined medium according to Verduyn et al., (1992) with 2% glucose and vitamins.

After 48 h incubation at 30 °C, a 10 times dilution was performed and 10 µL of this suspension was used to inoculate one well containing 90 µL of medium. The 96-well plates were then sealed with PCR sealing films (UC-500 ultra-clear pressure sensitive sealing film - E&K Scientific, Santa Clara, CA) and incubated at 30 °C and 44 rpm (3.5 mm of amplitude) for 24 h in the automated plate reader Infinity M200 (Tecan, Männedorf, Switzerland). Cultivations were run in triplicates and cell growth was monitored by measuring the absorbance at 570 nm every 15 min. Specific growth rates were calculated and results were submitted to the multiple comparison test Tukey's HSD (Honestly Significant Difference) with a significance level of 0.10 using Statgraphics Centurion XVI (Statpoint Technologies, Warrenton, VA). The defined medium contained vitamins, trace metals and salts (VERDUYN et al., 1992), including 5.0 g L^{-1} $(NH_4)_2SO_4$, 3.0 g L^{-1} KH_2PO_4, and 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, with 20 g L^{-1} glucose or 20 g L^{-1} sucrose and initial pH adjusted to 6.0 using KOH. Vitamins and glucose solutions were filter-sterilized and added to the medium after autoclavation (121 °C for 20 min).

4.4.2 Cultivation under stressful conditions

Performance of strains under stressful conditions was evaluated at DTU Biosustain, The Novo Nordisk Foundation Center for Biosustainability (Hørsholm, Denmark). Cells were grown in a defined medium containing vitamins, trace metals and salts according to Verduyn et al., (1992) with the following modifications: 6.6 g L^{-1} $(NH_4)_2SO_4$, 6.0 g L^{-1} NaH_2PO_4, and 0.5 g L^{-1} $MgSO_4.7H_2O$ with 20 g L^{-1} glucose. The medium was buffered with 0.01 g L^{-1} potassium phthalate monobasic and the initial pH was adjusted to 5.5 using 2 M NaOH. Vitamins and glucose solutions were filter-sterilized and added to the medium after autoclavation. For ethanolic stress, 80 g L^{-1} ethanol were added to the medium, whereas for acetic acid stress, 10 g L^{-1} acetic acid were added (still pH was adjusted to 5.5 using 2 M NaOH). For acid stress, potassium phthalate buffer was replaced by tartaric acid buffer (pKa = 2.98) and initial pH was adjusted to 3.0. Temperature stress was induced by incubation at 40 °C, instead of 30 °C. Pre-cultures were prepared by transferring cells from one colony on YPD plates to square 24-deepwell microplates (CR1424 Enzyscreen, Haarlen, The Netherlands) (**Figure 1**) containing 1 mL of the standard medium at pH 6.0. After overnight incubation at 30 °C and 300 rpm, cells were harvested, resuspended in water and inoculated at initial Abs_{600nm} = 0.3 in 24-roundwell microplates (CR1424f Enzyscreen, Haarlen, The Netherlands) (**Figure 2**) containing 700 µL of the medium described above at pH 5.5. The plates were incubated at 30 °C or 40 °C and 225 rpm in the automated plate reader Growth Profiler 1152 (Enzyscreen, Haarlen, The Netherlands) and growth was measured every 15 min until stationary phase was reached. Growth experiments were performed in duplicates.

Figure 1 - Polypropylene square 24-deepwell microplates used for obtaining inocula.

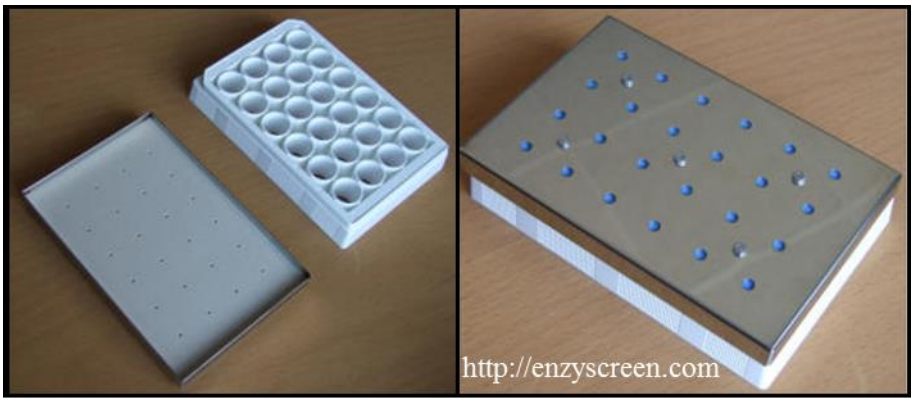

Figure 2 - Polystyrene 24-roundwell microplates used for screening on Growth Profiler 1152.

Growth Profiler 1152 is a high-performance piece of equipment that permits an automated recording of the progression of growth by measurement of cell density at regular time intervals. Using a flatbed scanner, the machine is able to count the number of green pixels in a circular area of 2 mm diameter at the bottom of each well. The values obtained (Green-value) can then be converted to Abs_{600nm} using a calibration curve: after overnight growth in YPD at 30 °C, cells were harvested, ressuspended in standard medium at pH 6.0 and diluted in order to obtain a range from Abs_{600}=0 to Abs_{600}=115. The content of each dilution was then distributed in three wells (for triplicates) of a 24 roundwell plate **(Figure 2)** and Green-value was obtained utilizing Growth Profiler. Using a non-linear regression model,

the Abs_{600nm} values obtained using a spectrophotometer were correlated to Green-values (see **Figure 3** as an example).

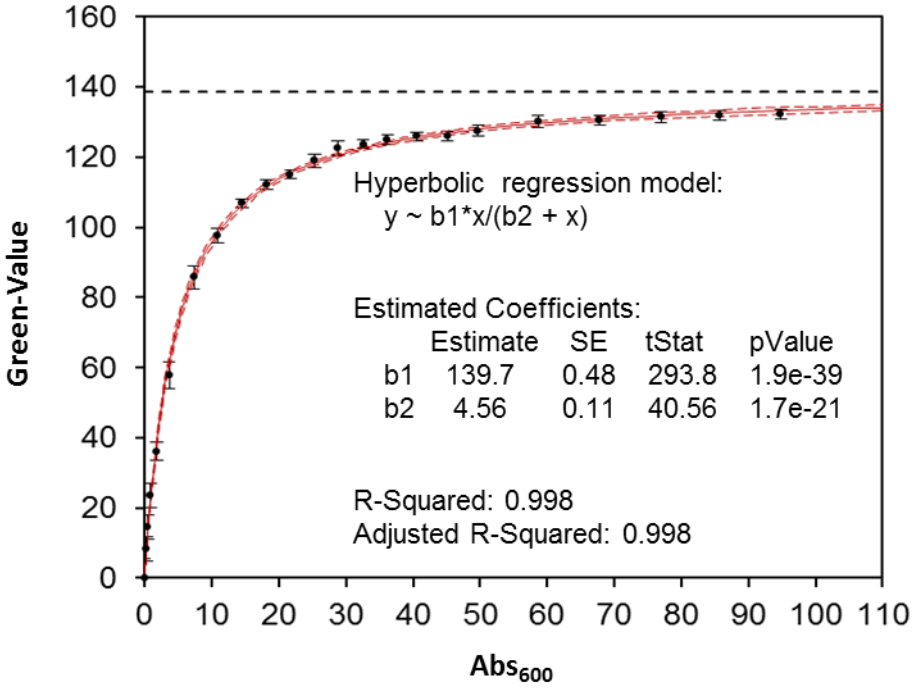

Figure 3 - Example of a calibration curve correlating Green-Value and Abs_{600nm}.

4.4.3 Cultivation under conditions mimicking the 1G fuel ethanol industrial process

The cultivations mimicking 1G fuel ethanol production were performed at DTU Biosustain, The Novo Nordisk Foundation Center for Biosustainability (Hørsholm, Denmark). Pre-cultures were prepared by transferring cells from one colony on YPD plates to 500 mL baffled shake flasks containing 100 mL concentrated YPD medium (200 g L^{-1} glucose, 12 g L^{-1} yeast extract, 24 g L^{-1} peptone). The flasks were incubated at 30 °C and 175 rpm for 96 h. At 24 h and 72 h, additional 100 mL of the same medium were added to the flasks. Afterwards, cells were harvested by centrifugation, resuspended in water and inoculated at initial cell concentration of 9% (wet mass/v) in 200 mL serum bottles containing 50 mL of YP medium

(10 g L^{-1} yeast extract and 20 g L^{-1} peptone, pH adjusted to 4.5 using 2 M H_2SO_4 prior to autoclavation) with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose, 40 g L^{-1} fructose). To approach anaerobic conditions as far as possible, serum bottles were closed with butyl stoppers, sealed with aluminum caps, and flushed with N_2 (99.999% pure) at 1 bar for 20 min prior to inoculation. In order to allow gas release and at the same time avoid air entrance during cultivation, an air-lock was created using a Norprene tube and a needle to connect the headspace of the flasks to a bottle filled with water (**Figure 4**). Cultivations were performed in duplicates at 35 °C and 200 rpm for 27 hours. Samples were taken at selected time intervals by piercing the stopper with a syringe and a needle. Cell growth was monitored by following the Abs_{600nm}. For metabolite concentration analysis, samples were centrifuged and supernatants were analyzed by HPLC (UltiMate 3000, Dionex) with an Aminex HPX-87H column (BioRad, Richmond, USA) at 30 °C eluted with 5 mM H_2SO_4 at 0.6 mL min^{-1}, coupled to a RI-101 Refractive Index Detector (Dionex) (for detection of sucrose, fructose, glucose, ethanol and glycerol).

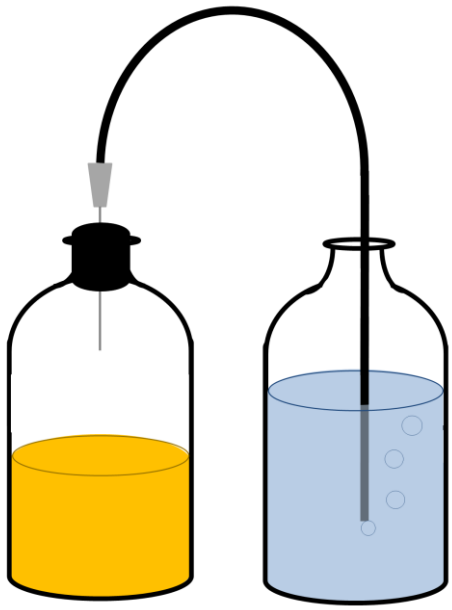

Figure 4 - Setup used for the cultivation mimicking the 1G fuel ethanol industrial process. The figure shows a schematic illustration of the serum flask containing medium (left) connected by a Norprene tube and a needle to another flask containing water (right). The air-lock created using the Norprene tube allows the gas produced during the fermentation to be released and avoids return of air into the cultivation flask (left).

4.5 Calculation of physiological and process parameters

For the cultivations under standard and stressful conditions, maximum specific growth rates (μ_{max}; in h^{-1}) were obtained by plotting the natural logarithm of absorbance values against time and calculating the slope of the straight line corresponding to the exponential growth phase. Performance of the strains under stressful conditions was evaluated using relative growth rate (%), which was calculated dividing the maximum specific growth rate (μ_{max}) under a particular stressful condition by the maximum specific growth rate under standard conditions.

For the cultivations mimicking the industrial 1G fuel ethanol production in Brazil, process parameters were calculated based on a 10 h fermentation time. **Volumetric productivity of ethanol** was calculated using the formula: $\Delta P_{Eth} / \Delta t$, where ΔP_{Eth} corresponds to the final ethanol concentration minus the initial ethanol concentration, in mol L^{-1}, and Δt is 10 h. The **ethanol yield** (in terms of *moles of ethanol produced divided by moles of hexose equivalents consumed*) was obtained as the slope of the straight line from the plot of ethanol concentration against hexose-equivalent concentration. For this purpose, it was considered that 1 mol of sucrose corresponds to 2 moles of hexose-equivalent. In this way, the hexose-equivalent concentration (in mol L^{-1}) was calculated by summing the molar concentrations of glucose, fructose and sucrose, the latter multiplied by two. The **relative biomass increase** ($\Delta X\%$) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. For some datasets, different data points were used when data corresponding to 10 h or 0 h were classified as outliers. Errors between duplicates were expressed as the deviations of the mean.

Carbon/nitrogen (C/N) ratio was calculated by dividing the mass of carbon added initially in the medium (within yeast extract, peptone and sugars) by the mass of nitrogen (added with yeast extract and peptone). It was considered that yeast extract contains approximately 10.5% nitrogen and 43% carbon, peptone contains approximately 11.5% nitrogen and 39% carbon, and sucrose, glucose and fructose contain 42, 40 and 40% of

carbon, respectively. The N content in yeast extract and peptone were those informed by the manufacturer (Sigma-Aldrich). The C content in yeast extract and peptone were determined using a CHNS-O Flash 2000 elemental analyzer (ThermoFisher Scientific, Delft, The Netherlands). Conditions of analyses were: reaction column temperature, 950 °C; Helium gas flow, 140 mL/min; cycle time, 720 s (with oxygen injection at 250 mL/min during the initial 5 s); separation column temperature, 65 °C; thermal conductivity detector. The calibration curve was obtained with Cystine, BBOT (2,5-Bis(5-tert-butyl-2benzo-oxazol-2yl Thiophene) and SUL (4-aminobenzenesulphonamide).

4.6 Serial dilution spot assays at low pH

Cells from a colony were transferred from YPD plates to 50 mL conical tubes containing 10 mL of liquid YPD medium and grown overnight at 30 °C and 200 rpm. Biomass was harvested by centrifugation, resuspended in fresh YPD medium and inoculated at an initial Abs_{600nm} = 0.1 in new 50 mL conical tubes containing 10 mL of YPD medium. Cells were then incubated at 30 °C for 3 h (until early exponential phase), harvested, washed with demineralized water and diluted to Abs_{600nm} = 0.1. Four successive dilutions were performed in order to obtain tubes with Abs_{600nm} of 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5}. The spot assay was performed as described by Della-Bianca and Gombert (2013) (based on NETTO, 2006), by applying 3 µL of each suspension onto YPD plates with pH adjusted to 2.5 using 2 M H_2SO_4. Plates were incubated at 30 °C, and pictures were taken after 72 h of incubation.

5 RESULTS AND DISCUSSION

5.1 Strain identification

5.1.1 Restriction analysis of mtDNA provided unsatisfactory results

We performed different tests using the restriction analysis of mtDNA in an attempt to differentiate the indigenous *S. cerevisiae* strains. The enzymatic digestion of total DNA using restriction enzymes specific for GCAT sites promotes many cuts on the nuclear DNA, but only a few cuts on mtDNA. Thus, the nDNA is cut in very small molecules, while mtDNA is cut in larger fragments and can be visualized by electrophoresis. When we tested the protocol provided by Dr. Carlos A. Rosa (Federal University of Minas Gerais, Belo Horizonte, Brazil), the total DNA extraction step did not present the expected results, since we could not visualize the DNA bands on the gel after digestion with the restriction enzyme *Hinf* I. We then started to use a protocol provided by Amparo Querol's group (IATA-CSIC, Valencia – Spain) (QUEROL, 1992). With this protocol we obtained an intense band on the gel after the DNA extraction step (black arrow on well 1 - **Figure 5**), indicating a higher yield of DNA compared to what we were obtaining with the first protocol (data not shown). Nevertheless, as shown on well 2, we were not able to visualize the mtDNA bands after digestion with the restriction enzyme, but only a low intensity smear (well 2 - **Figure 5**).

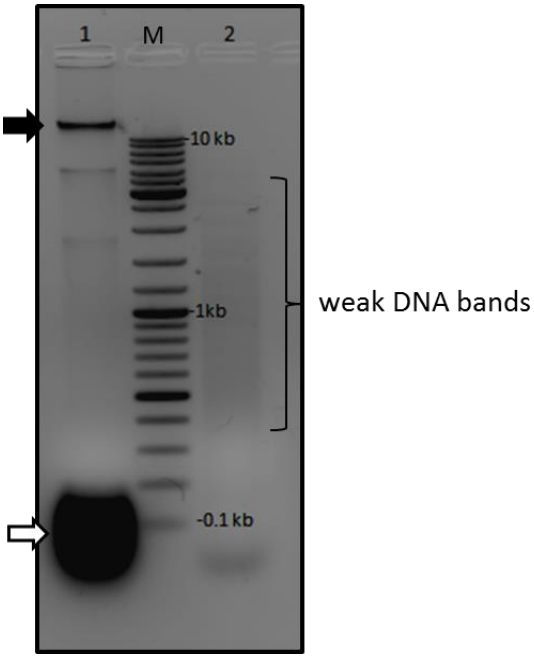

Figure 5 - Electrophoresis agarose gel of *S. cerevisiae* total DNA digested or not with *Hinf* I. Well 1 shows the DNA without restriction enzyme treatment. Well 2 shows the DNA treated with *Hinf* I. M represents the DNA marker, the black arrow indicates a band which is probably composed of total DNA and the white arrow indicates RNA. O'GeneRuler DNA Ladder Mix (Life Technologies) was used as DNA ladder.

Since we had a limited stock of Zymolyase available, before trying to optimize the protocol, we tested if the enzyme Lyticase could be used to replace Zymolyase. On this assay, we evaluated two different enzyme concentrations (0.1 U μL^{-1} and 0.3 U μL^{-1}) and two different incubation times (25 or 60 minutes), as shown in **Table 4**.

Table 4 - Conditions tested during the Lyticase evaluation test. DNA quantification was performed using NanoDrop (ThermoScientific) before enzymatic digestion (10x diluted).

Enzyme	Zymolyase	Lyticase	Lyticase	Lyticase	Lyticase
Concentration (U μL^{-1})	0.02	0.1	0.1	0.3	0.3
Incubation time (min)	25	25	60	25	60
DNA obtained (ng μL^{-1})	718.1	524.8	559.7	755.8	690.7

According to the intensities of the non-digested DNA bands, Lyticase 0.1 U μL^{-1} provided a lower yield of DNA compared to the other conditions tested (**Figure 6A**). On the other hand, Lyticase 0.3 U μL^{-1} provided similar results to what we obtained with Zymolyase.

Moreover, the longer incubation time seemed not to affect the DNA yield. Again, only smears could be observed when *Hinf* I treatment was applied (**Figure 6B**).

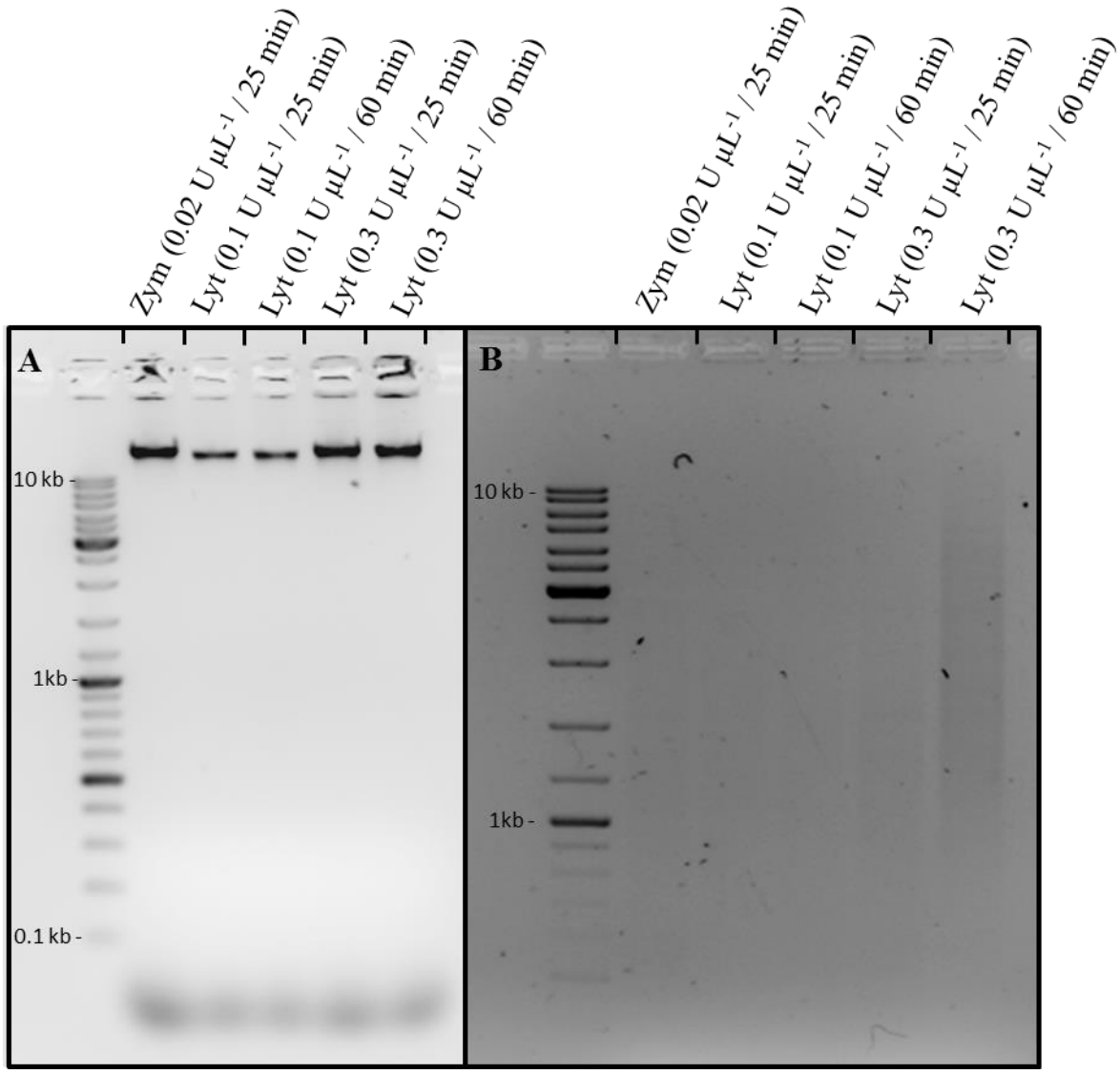

Figure 6 - Electrophoresis gels of mtDNA extraction using Zymolyase or Lyticase as lytic enzymes. Panel A shows non-digested DNA and Panel B shows DNA digested with *Hinf* I. O'GeneRuler DNA Ladder Mix (Life Technologies) was used as DNA ladder.

The similar DNA yields obtained when Zymolyase (0.02 U μL^{-1}) and Lyticase (0.3 U μL^{-1}) were used as lytic enzymes were confirmed by quantification using NanoDrop **(Table 4)**. After this assay, Lyticase (0.3 U μL^{-1}) was adopted as the lytic enzyme for DNA extraction.

Despite the high DNA concentrations obtained after the extraction step, the digestion with endonuclease was not providing the expected results. In order to understand the reasons for this issue, we performed the DNA digestion testing a different endonuclease (*Eco*RI) and the non-utilization of the RNase treatment. As shown in **Figure 7**, when *Hinf* I was used and RNase treatment was not applied (well 4), the DNA bands after the digestion could finally be visualized. When *Eco*RI was used without RNase we also observed clearer DNA bands than when RNase was added. With this test, we concluded that *Hinf* I was working properly, and it seemed possible that RNase was somehow promoting DNA degradation.

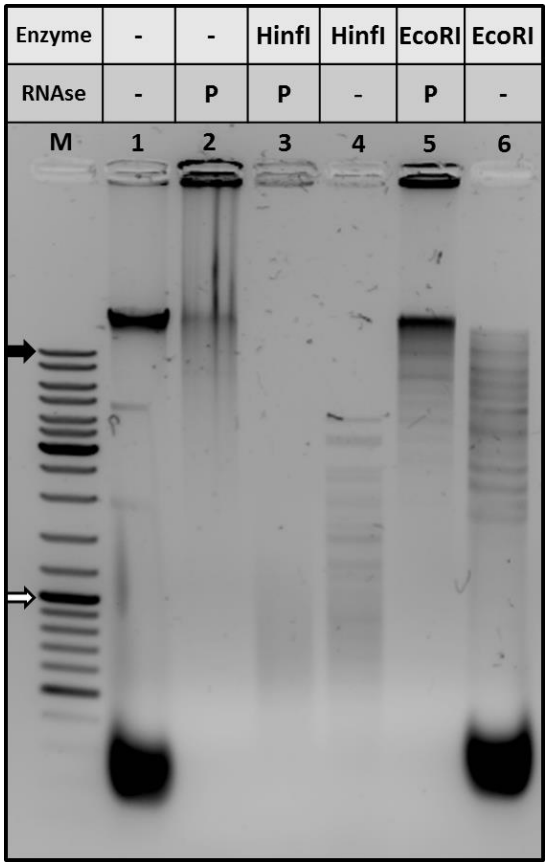

Figure 7 - Electrophoresis gel after DNA extraction using the protocol described by Querol et al., (1992). Well 1: non-digested DNA, without RNase treatment; well 2: non-digested DNA treated with RNase; well 3: DNA digested with *Hinf* I and treated with RNase; well 4: DNA digested with Hinf-1 without RNase treatment; well 5 shows DNA digested with *Eco*RI and treated with RNase; well 6: DNA digested with *Eco*RI without RNase treatment. Letter P means present. M represents the molecular marker (O'GeneRuler DNA Ladder Mix, Life Technologies). The black arrow indicates the 10 kb band and the white arrow indicates the 1 kb band.

In order to investigate whether DNA degradation was indeed being caused by the RNase we had available in our laboratory or eventually by the concentration we were using, another test was performed using our RNase diluted (0.05x) and another RNase provided by a different laboratory (Qiagem, Hilden, Germany). As shown in **Figure 8**, when diluted RNase (0.05x Lab) or a different RNase (≠Lab) was used, the DNA bands could be observed on the gel. These results corroborate the hypothesis that the concentration of RNase used was causing DNA degradation. We believe that some contaminant DNase can be present in our RNase stock, and depending on the concentration used, it was promoting DNA degradation. We also observed that the *Hinf* I stock we used may contain some RNase since it promoted RNA degradation alone, independently of RNase addition (**Figure 7** - well 4 and **Figure 8** - well 2).

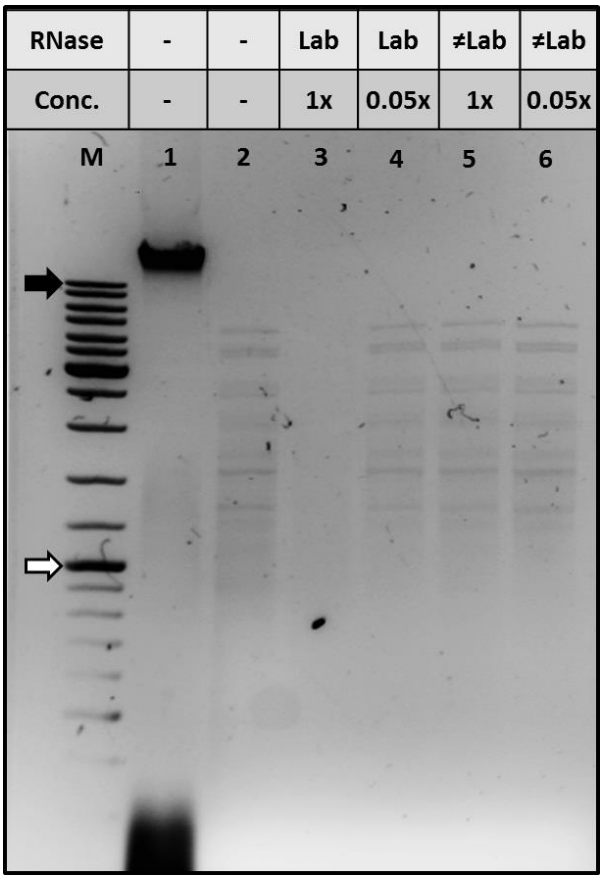

Figure 8 - Result of *S. cerevisiae* total DNA digestion testing different RNases. Well 1: Non-digested DNA without RNase. Well 2: DNA digested with *Hinf* I without RNase. Well 3: DNA digested with *Hinf* I and treated with 1 x RNase from our lab. Well 4: DNA digested with *Hinf* I and treated with 0.05x RNase from our lab. Well 5: DNA digested with *Hinf* I and treated with 1x RNase from a different lab. Well 6: DNA digested with *Hinf* I and treated with 0.05x RNase from a different lab. M represents the molecular marker (O'GeneRuler DNA Ladder Mix, Life Technologies). The black arrow indicates the 10 kb band and the white arrow points the 1 kb band.

In an attempt to optimize the protocol and obtain more intense DNA bands, the DNA was concentrated by the end of extraction using the vacuum concentrator SpeedVac Plus (Eppendorf, Germany). **Figure 9** shows the result of DNA digestion with *Hinf* I after concentration of DNA: 1 x (20 µL of DNA, not concentrated), 2 x (40 µL of DNA concentrated in 20 µL) and 3 x (60 µL of DNA concentrated in 20 µL).

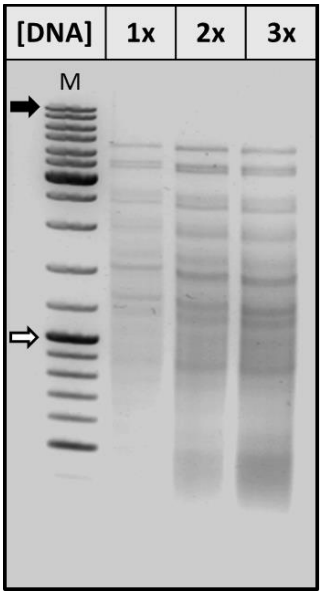

Figure 9 - Digestion of *S. cerevisiae* total DNA with *Hinf* I after DNA concentration using SpeedVac. 1 x: DNA not concentrated. 2 x: DNA concentrated 2 x. 3 x: DNA concentrated 3 x. M shows the molecular marker (O'GeneRuler DNA Ladder Mix, Life Technologies). The black arrow indicates the 10 kb band and the white arrow points the 1 kb band.

The results obtained with DNA concentrated 2 x or 3 x were similar, and both of them provided more intense bands than those obtained when DNA was not concentrated (1 x). Eventhough the protocol of restriction analysis of mtDNA was optimized and the bands could now be visualized on agarose gel, we considered the results were still not adequate to allow strain differentiation. For that reason, we decided to try a different method for fingerprinting our *S. cerevisiae* strains, which could provide more reliable results and be performed in a more straightforward manner, when compared to the restriction analysis of mtDNA.

5.1.2 *Minisatellite PCR-based DNA fingerprinting resolved five out of fourteen indigenous strains*

We chose the minisatellite-PCR method described by Carvalho-Netto et al., (2013) based on polymorphic insertion and deletion sites (*indels*) to differentiate the twenty *S. cerevisiae* strains utilized in this work. Using nine oligonucleotide primer pairs (**Table 3**) and analysis of the banding patterns generated by PCR using those primer pairs, we were able to differentiate nine of the strains tested. The identity of four industrial and laboratory strains could also be confirmed using the patterns reported in the referred work.

The bands obtained for strains PE-2, CAT-1 and S288c were nearly identical to the results reported by Carvalho-Netto et al., (2013), with the exception of one band representing the site YFL024C in PE-2 (**Figure 10**). In this case, instead of two bands, only one of the bands was amplified in our work. This divergent result could be due to an intrinsic heterogeneity in commercially available stocks of PE-2, which have been shown to contain a mixture of cells with slightly different karyotypes (ARGUESO et al*.,* 2009). Hence, because all the other bands from PE-2 matched the bands observed by Carvalho-Netto et al., (2013), we considered the results sufficiently reliable for confirming the identity of this strain.

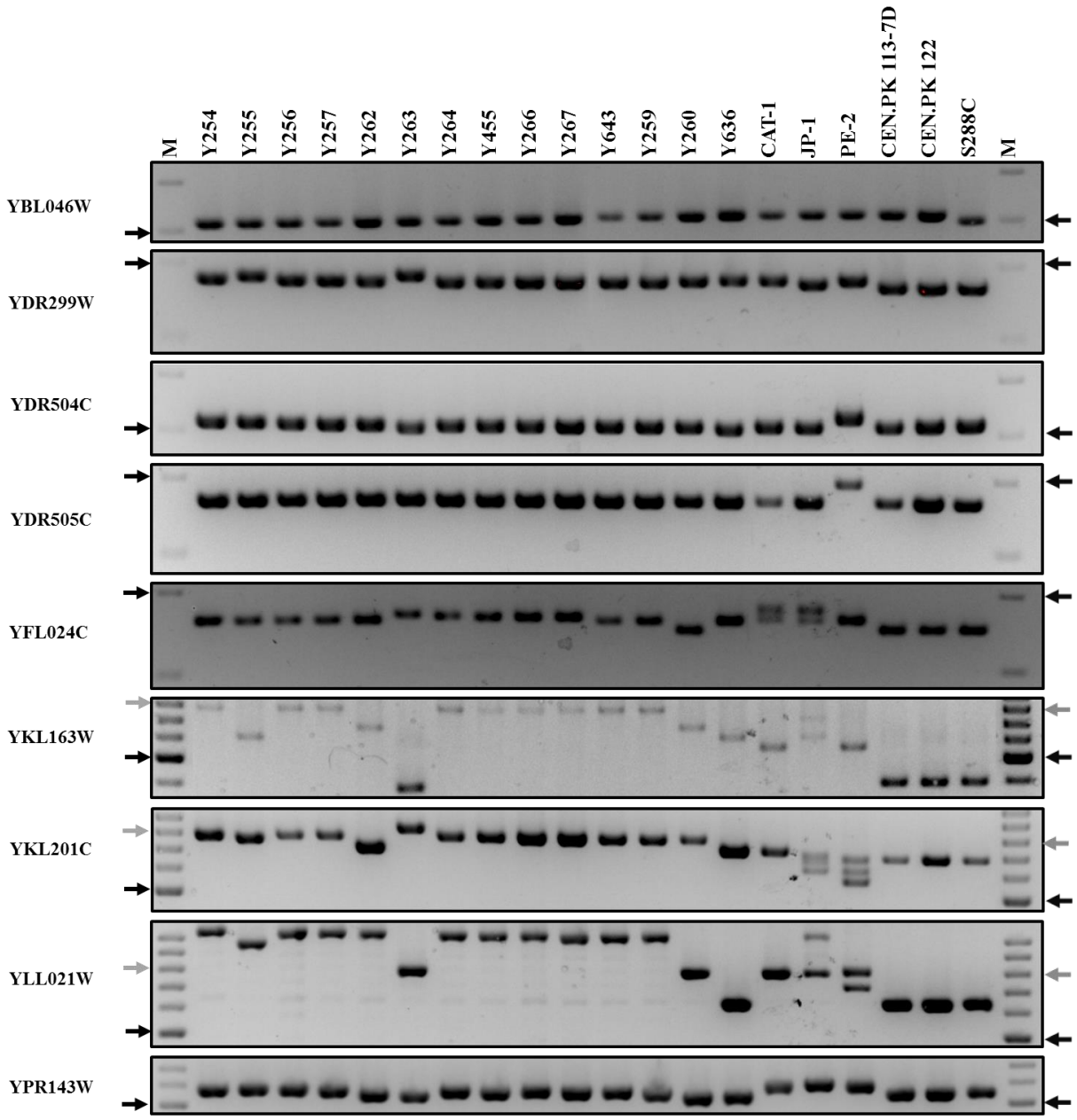

Figure 10 - Electrophoresis gels of the DNA fingerprinting based on minisatellite-PCR (CARVALHO-NETTO et al., 2013). DNA fingerprinting of twenty *S. cerevisiae* strains was evaluated by nine PCR preparations using different pairs of oligonucleotide primers. The names of the strains are shown above, and the ORF identifiers are shown on the left side of the figure. M indicates the molecular mass marker. The black arrow indicates the 500 bp marker band and the grey arrow indicates the 800 bp marker band.

To facilitate the analysis, the strains were grouped according to the banding pattern **(Table 5)**. The strains UFMG-CM-Y255, UFMG-CM-Y262, UFMG-CM-Y263, UFMG-CM-Y260, UFMG-CM-Y636, CAT-1, JP-1, PE-2 and S288c presented unique amplification patterns and hence formed nine distinct groups **(Table 5).** The remaining strains formed five small clusters

in which the differentiation among its members was not possible: (a) UFMG-CM-Y254, UFMG-CM-Y256 and UFMG-CM-Y257, (b) UFMG-CM-Y264 and UFMG-CM-Y266, (c) UFMG-CM-Y455 and UFMG-CM-Y267, (d) UFMG-CM-Y643 and UFMG-CM-Y259, and (e) CEN.PK113-7D and CEN.PK122. As expected, the strains of the CEN.PK family clustered together, since the latter is the diploid of the former (VAN DIJKEN et al., 2000). The amplification patterns of the indigenous strains correlated with the location from which these strains were isolated. The strains that were classified in a common group were isolated from the same tree species located in the same region in Brazil **(Table 1).**

Despite the fact that some of the strains could not be differentiated by this method, whole genome sequencing analysis showed significant differences among these strains, which strengthen their identity as distinct strains (Prof. C. Rosa, Federal University of Minas Gerais, Brazil, personal communication). These findings imply that more than one method (or more ORF's) should be used in order to completely differentiate natural strains isolated from the same habitat.

Recently, seven of the indigenous strains used in this work had their phylogeny compared to representatives of all the known *S. cerevisiae* populations for which genomic data is available (BARBOSA et al., 2016). Most of the indigenous strains clustered in new clades: five of them were grouped in the Brazil 1 clade (UFMG-CM-Y257, UFMG-CM-Y262, UFMG-CM-Y264 and UFMG-CM-Y643) and other two strains were grouped in the Brazil 2 clade (UFMG-CM-Y636 and UFMG-CM-Y260). Both groups do not overlap with previously known populations of *S. cerevisiae* that were also used in the analysis. For this reason, according to the authors, these isolates appear to represent new lineages of *S. cerevisiae*. On the other hand, UFMG-CM-Y255 strain was placed in the Japan - North America clade, and the authors believe that this strain might represent a North American migrant since it had a pure North American genome ancestry. They have also verified a lower genetic divergence between the Brazilian B1 population and the Japanese and North American populations in comparison to the calculated divergence from other groups, and together with the fact that the genetic diversity of *S. cerevisiae* is the highest in oak environments of China and Japan, they propose that South American lineages might have their origin related to a migration of animal species, including humans, from Asia to the New World, with South American lineages deriving from their North American counterparts.

Table 5 - Banding patterns obtained by the minisatellite PCR-based DNA fingerprinting method.

Strain	YBL 046W	YDR 299W	YDR 504C	YDR 505C	YFL 024C	YKL 163W	YKL 201C	YLL 021W	YPR 143W	Groups *
UFMG-CM-Y254	2	2	1	1	2	6	5	5	2	**a**
UFMG-CM-Y256	2	2	1	1	2	6	5	5	2	**a**
UFMG-CM-Y257	2	2	1	1	2	6	5	5	2	**a**
UFMG-CM-Y264	2	2	1	1	3	6	5	5	4	**b**
UFMG-CM-Y266	2	2	1	1	3	6	5	5	4	**b**
UFMG-CM-Y267	2	2	1	1	3	6	5	5	1	**c**
UFMG-CM-Y455	2	2	1	1	3	6	5	5	1	**c**
UFMG-CM-Y643	2	2	1	1	2	6	5	5	1	**d**
UFMG-CM-Y259	2	2	1	1	2	6	5	5	1	**d**
CEN.PK113-7D	2	1	1	1	1	1	1	1	1	**e**
CENPK122	2	1	1	1	1	1	1	1	1	**e**
UFMG-CM-Y255	2	3	1	1	2	3	6	4	2	**f**
UFMG-CM-Y262	2	2	1	1	2	4	4	5	1	**g**
UFMG-CM-Y263	2	3	1	1	3	1	7	2	1	**h**
UFMG-CM-Y260	2	2	1	1	1	4	5	2	3	**i**
UFMG-CM-Y636	2	2	1	1	2	3	4	1	3	**j**
CAT-1	2	2	1	1	2/4	2	4	2	2	**k**
JP-1	2	2	1	1	2/4	3/5	1/2	2/5	2	**l**
PE-2	2	2	2	2	2	2	1/2/3	2/3	2	**m**
S288C	1	1	1	1	1	1	1	1	1	**n**

* Strains were assigned to a common group when the same banding pattern was obtained.

5.1.3 PCR-Fingerprinting using EI-1 primer could work as accessory method to complement other strategies

In order to differentiate the strains that presented similar banding patterns in the previous assay, we performed another PCR fingerprinting method using the oligonucleotide primer EI-1 (DE BARROS LOPES et al., 1998). As shown in **Figure 11**, we were able to group the strains into three distinct clusters using this method, according to their banding pattern:

(■) UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y643 and UFMG-CM-Y259; (□) UFMG-CM-Y262, UFMG-CM-Y263, UFMG-CM-Y264, UFMG-CM-Y455, UFMG-CM-Y266, UFMG-CM-Y267, CAT-1 and JP-1; and (●) UFMG-CM-Y260, UFMG-CM-Y636, PE-2, CEN.PK113-7D, CEN.PK122 and S288C. In conclusion, none of the strains could be uniquely identified when the EI-1 primer-based method was employed.

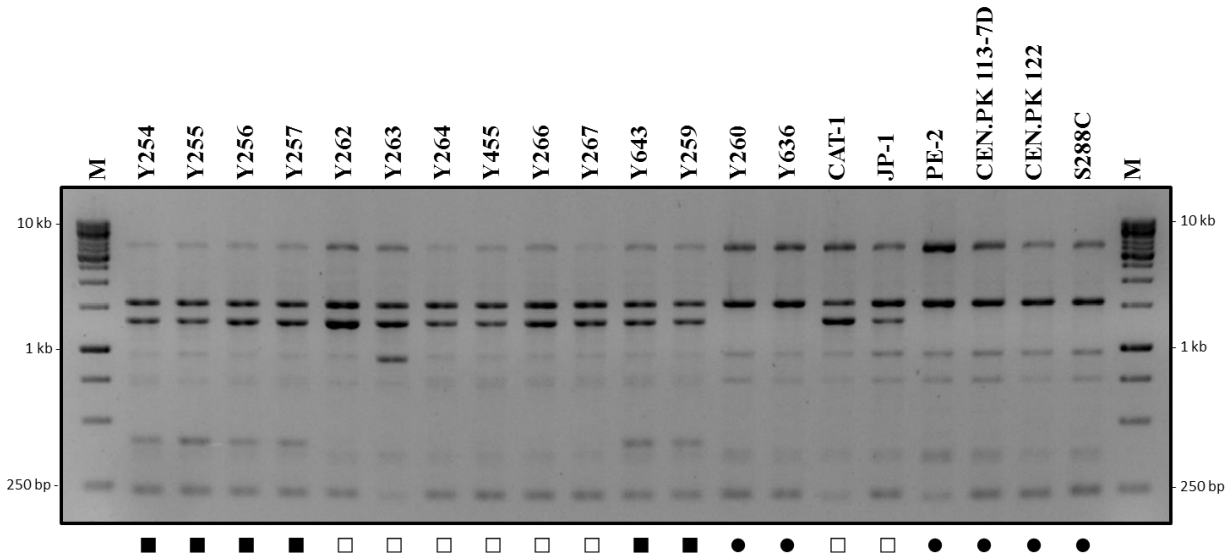

Figure 11 - PCR fingerprinting using oligonucleotide primer EI-1 (DE BARROS LOPES et al*.,* 1998**)** performed for 20 strains. M indicates the molecular marker. The symbols (■, □ and ●) represent the three distinct banding patterns identified.

In comparison to the results reported by De Barros Lopes et al., (1998), we obtained fewer and lower-resolution bands, which compromises the discriminatory potential of the method. Therefore, in this case, optimization of the PCR is required in order to evaluate the possibility of using this method to differentiate such similar strains. However, the method could be used as an accessory method to complement other differentiation strategies.

5.2 Species identity was confirmed using a molecular method

In order to establish a reliable method for identifying different yeast species and to confirm the identity of the strains used in this work, sequencing of the DNA corresponding to the D1/D2 region of 26S rRNA was performed as described by Lachance et al., (1999) and Kurtzman and Robnett (1998). According to the authors, the nucleotide substitutions in a 600 bp region of the D1/D2 domain do not generally exceed 1%, when strains belonging to the same species are compared. After DNA extraction and PCR amplification, samples were purified and submitted to electrophoresis in order to confirm the size of the PCR products obtained. The results presented in **Figure 12** show the amplification of 600 bp DNA fragments that were subsequently used for sequencing.

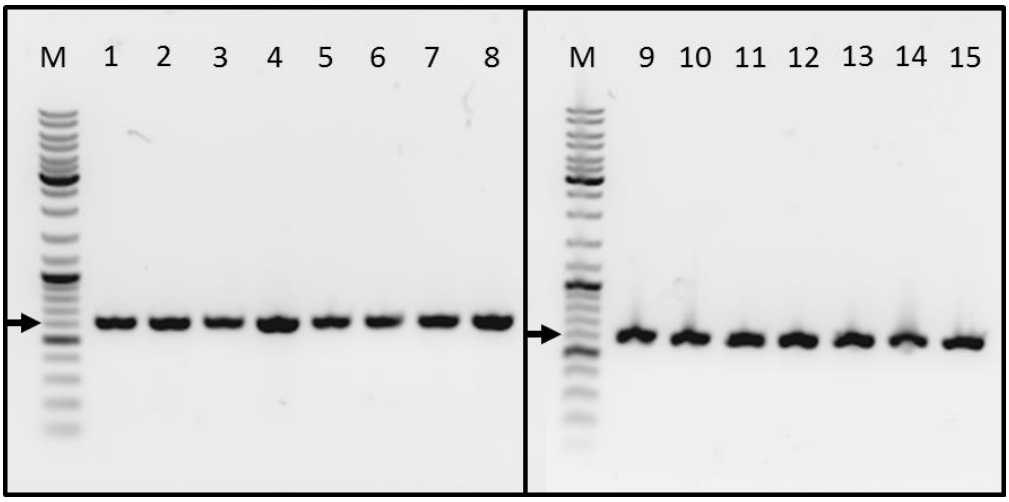

Figure 12 - Result of amplification and purification of PCR products of 15 *S. cerevisiae* strains using oligonucleotide primers NL-1 and NL-4. Wells 1-14: *S. cerevisiae* indigenous strains. Well 15: CEN.PK113-7D (*S. cerevisiae* laboratorial strain). Data is not shown for *Y. lipolytica*. M indicates the marker (O'GeneRuler DNA Ladder Mix, Life Technologies) and the black arrow points to the 600 bp marker band.

The sequencing of *S. cerevisiae* strains was performed using oligonucleotide primer NL-1 in duplicate, while sequencing of the *Y. lipolytica* strain (used as an external control) was performed using four oligonucleotide primers: NL-1, NL-2A, NL-3A and NL-4. After low

quality detected peaks were eliminated and consensus sequence was assembled, sequences of different lengths were obtained (439-542 bp). They were then aligned to the reference sequence of the rRNA 26S region (3,301 bp) of *S. cerevisiae* strain NRRL Y-12632 (KURTZMAN; ROBNETT, 1998) using the Align tool of the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi). All 15 *S. cerevisiae* sequences tested presented 100% identity to the reference sequence (query) (**Table 6**), confirming that the strains used belong to the *Saccharomyces cerevisiae* species.

Table 6 - Result of alignment of consensus sequences of 15 *S. cerevisiae* strains and one *Y. lipolytica* strain with the reference sequence of *S. cerevisiae* NRRL Y-12632 (KURTZMAN; ROBNETT, 1998) (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Strain	Max score[a]	Total score[b]	Query cover[c] (%)	E value[d]	Identity (%)
UFMG-CM-Y254	1002	1002	16	0	100
UFMG-CM-Y255	979	979	16	0	100
UFMG-CM-Y256	976	976	15	0	100
UFMG-CM-Y257	1002	1002	16	0	100
UFMG-CM-Y259	990	990	16	0	100
UFMG-CM-Y260	981	981	16	0	100
UFMG-CM-Y262	981	981	16	0	100
UFMG-CM-Y263	972	972	15	0	100
UFMG-CM-Y264	970	970	15	0	100
UFMG-CM-Y266	941	941	15	0	100
UFMG-CM-Y267	955	955	15	0	100
UFMG-CM-Y455	972	972	15	0	100
UFMG-CM-Y636	811	811	13	0	100
UFMG-CM-Y643	933	933	15	0	100
CEN.PK113-7D	990	990	16	0	100
Po1g (*Y. lipolytica*)	283	283	10	6e-80	82

[a] Highest alignment score (bit-score) between the query sequence and the database sequence segment . In the context of sequence alignments, a score is a numerical value that describes the overall quality of an alignment. Higher numbers correspond to higher similarity.

[b] Sum of alignment scores of all segments from the same database sequence that match the query sequence (calculated over all segments). This score is different from the max score if several parts of the database sequence match different parts of the query sequence.

[c] Percent of the query length that is included in the aligned segments.

[d] Number of alignments expected by chance with a particular score or better.

The sequence of D1/D2 of rRNA 26S gene of the *Yarrowia lipolytica* Po1g strain was also determined and used as outlier. Differently from the results obtained with *S. cerevisiae* strains, only 82% identity was obtained when the consensus sequence of *Y. lipolytica* Po1g strain was aligned to the rRNA 26S region of *S. cerevisiae* strain NRRL Y-12632 (KURTZMAN; ROBNETT, 1998) **(Table 6)**, confirming the reliability of the analysis.

5.3 None of the *S. cerevisiae* strains grew faster on glucose than on sucrose

Although many studies describe physiological traits of *S. cerevisiae* laboratory strains, few works address quantitative physiological data on the Brazilian fuel ethanol strains (DELLA-BIANCA et al., 2013, 2014). In this study, we investigated the growth performance of all strains on two different carbon sources (separately): glucose or sucrose. We observed a significant difference in the specific growth rate when the indigenous and the reference strains are compared, which highlights the existing physiological biodiversity and the potential of this type of study in seeking for traits of industrial interest.

When the two carbon sources (glucose and sucrose) are compared, interestingly, none of the strains investigated here presented a higher specific growth rate (μ_{max}) on glucose than on sucrose (**Figure 13**). Two of the strains (UFMG-CM-Y257 and UFMG-CM-Y259) presented higher specific growth rates when sucrose was used as carbon source, when compared to glucose. The reasons underlying faster growth on sucrose remain uncertain, however this may be related to the higher affinity for sucrose of the G-Protein Coupled Receptor (Gpr1), which can increase the glycolytic flux through activation of the cAMP signaling pathway (BADOTTI et al., 2008). Nonetheless, a deeper analysis on this topic is required, and the indigenous strains described here represent an excellent material for investigating this behavior.

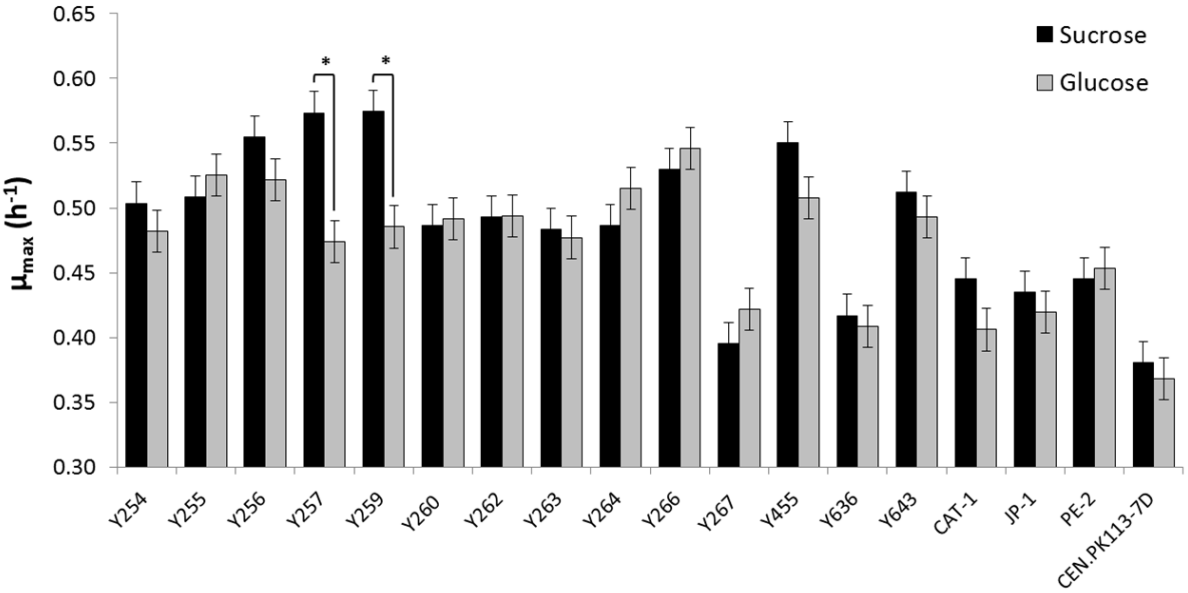

Figure 13 - Maximum specific growth rates of 14 indigenous *S. cerevisiae* strains (UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y259, UFMG-CM-Y260, UFMG-CM-Y262, UFMG-CM-Y263, UFMG-CM-Y264, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y455, UFMG-CM-Y636 and UFMG-CM-Y643), 3 industrial *S. cerevisiae* strains (CAT-1, PE-2, JP-1) and one laboratory strain (CEN.PK113-7D). Cells were cultivated as described in the Methods section in a defined medium containing initially either 20 g L^{-1} sucrose (dark bars) or 20 g L^{-1} glucose (light bars). The means containing "*"are significantly different from their pairs within a 90% confidence level (Tukey – HSD / Multiple Sample Comparison). The error bars represent the pooled standard error using three culture replicates.

In the Brazilian 1G fuel ethanol industry, a high growth rate on sucrose represents a competitive advantage, since most of the sugar present in sugar-cane juice and molasses is available in the form of this disaccharide (BASSO; BASSO; ROCHA, 2011). When sucrose was used as carbon source, PE-2 displayed a 17% higher specific growth rate than the laboratory strain CEN.PK113-7D (0.45 ± 0.02 vs. 0.38 ± 0.02 h^{-1}), the slowest strain among all tested in our setup (**Figure 13**). This trait may, in part, contribute to the capacity of PE-2 to dominate the industrial fermentation vats in many Brazilian mills only a few weeks after start of operation (BASSO et al., 2008). Interestingly, even higher values were observed for all indigenous strains, except for UFMG-CM-Y636 and UFMG-CM-Y267. Strains UFMG-CM-Y257 and UFMG-CM-Y259, for instance, displayed a 29% higher specific growth rate than PE-2 (0.57 ± 0.02 vs. 0.45 ± 0.02 h^{-1}). In accordance with our results, van Dijken et al., (2000) reported similar growth rates for CEN.PK122 (diploid strain of CEN.PK113-7D) using sucrose as sole carbon source (0.38 h^{-1}).

When glucose was used as the sole carbon and energy source, CEN.PK113-7D also exhibited the lowest specific growth rate among the strains studied in this work (**Figure 13**) and PE-2 displayed a 23% higher value than this laboratory strain (0.45 ± 0.02 vs. 0.37 ± 0.02 h^{-1}). Under similar conditions (same defined medium), Da Silva Filho et al., (2005) reported the same values of μ_{max} for PE-2, but higher values for the strain JP-1, when compared to our results (0.49 vs. 0.42 ± 0.02 h^{-1}). The three highest specific growth rate values on glucose were displayed by the indigenous strains UFMG-CM-Y266 (0.55 ± 0.02 h^{-1}), UFMG-CM-Y255 (0.53 ± 0.02 h^{-1}) and UFMG-CM-Y256 (0.52 ± 0.02 h^{-1}), which were between 15 and 20% higher than the rate exhibited by PE-2.

Strains UFMG-CM-Y256 and UFMG-CM-Y266 presented high growth rates (> 0.52 ± 0.02 h^{-1}) either when sucrose or glucose were used as the sole carbon and energy source. Thus, it would be interesting to investigate which mechanisms underlie this trait. These specific growth rates are, to our knowledge, the highest values ever reported for *S. cerevisiae* on a defined medium with a single carbon source.

5.4 Indigenous strains present higher tolerance to different stressors in comparison to industrial strains

To evaluate tolerance of indigenous strains towards stressors that are relevant in the context of Brazilian industrial fuel ethanol production, experiments were performed under four different conditions: low pH (3.0), high ethanol concentration (80 g L^{-1}), high temperature (40 °C) and high acetic acid concentration (10 g L^{-1}), which were investigated individually. The three former stresses can typically be found in 1G ethanol production in Brazil (BASSO; BASSO; ROCHA, 2011) and are considered some of the major selective pressures regarding strain persistence and survival inside the fermenters, especially low pH (DELLA-BIANCA et al., 2014). The latter is a well-known inhibitory compound released during hydrolysis of lignocellulosic biomass that became relevant when 2G technologies started to be applied (ALMEIDA et al., 2007). Besides the fourteen indigenous strains, three industrial

strains (PE-2, CAT-1 and JP-1) and one laboratory strain (CEN.PK113-7D) were included in the study. Performance of the strains was evaluated using relative growth rate (%), which was calculated dividing the maximum specific growth rate (μ_{max}) under a particular stressful condition by the maximum specific growth rate under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol). When pH 3.0 was applied as a stress factor, strains PE-2, JP-1 and CEN.PK113-7D displayed similar results in terms of the relative growth rate (88% on average), the highest values among all strains tested (**Figure 14A**). The higher relative growth rates of these strains compared to indigenous strains demonstrate the lower tolerance of the latter ones towards acidic pH. On average, indigenous strains presented a relative growth rate of 67%. The higher tolerance to pH 3.0 of industrial strains PE-2 and JP-1 could suggest that low pH tolerance is a common feature among industrial strains used in Brazilian fuel ethanol industries. However, under the conditions tested, industrial strain CAT-1 did not present the same tolerance as PE-2 and JP-1. Previous data from our group showed that PE-2 only presents better growth and fermentation performance than CEN.PK113-7D when grown in a complex medium at low pH (2.7) (DELLA-BIANCA et al., 2014). Use of pH 5.0 or of a synthetic medium results in similar performances of the two strains. Since we here used a synthetic medium at pH 3.0, it is possible to conclude that the present results corroborate our previous ones, but the reasons for these observations remain unclear.

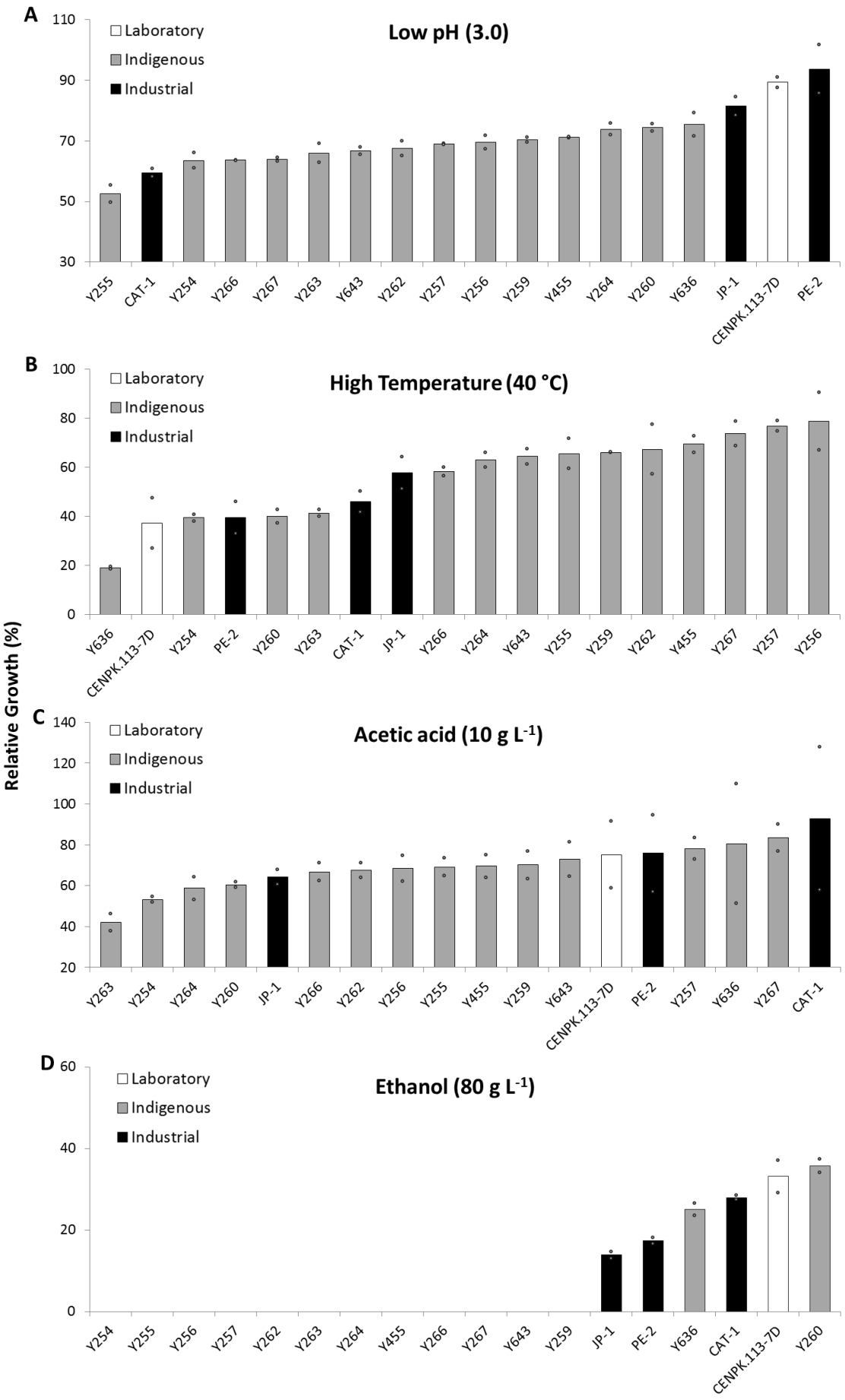
A
Low pH (3.0)
Laboratory
Indigenous
Industrial
110
90
70
50
30
Y255
CAT-1
Y254
Y266
Y267
Y263
Y643
Y262
Y257
Y256
Y259
Y455
Y264
Y260
Y636
JP-1
CENPK.113-7D
PE-2
B
High Temperature (40 °C)
Laboratory
Indigenous
Industrial
100
80
60
40
20
0
Y636
CENPK.113-7D
Y254
PE-2
Y260
Y263
CAT-1
JP-1
Y266
Y264
Y643
Y255
Y259
Y262
Y455
Y267
Y257
Y256
Relative Growth (%)
C
Acetic acid (10 g L-1)
Laboratory
Indigenous
Industrial
140
120
100
80
60
40
20
Y263
Y254
Y264
Y260
JP-1
Y266
Y262
Y256
Y255
Y455
Y259
Y643
CENPK.113-7D
PE-2
Y257
Y636
Y267
CAT-1
D
Ethanol (80 g L-1)
Laboratory
Indigenous
Industrial
60
40
20
0
Y254
Y255
Y256
Y257
Y262
Y263
Y264
Y455
Y266
Y267
Y643
Y259
JP-1
PE-2
Y636
CAT-1
CENPK.113-7D
Y260

Figure 14 - Growth of 14 indigenous, 3 industrial (CAT-1, PE-2, JP-1) and one laboratory (CEN.PK113-7D) *S. cerevisiae* strains under stress conditions relative to their growth under standard conditions (pH 5.5 and 30 °C). Strains were cultivated as described in the Methods section in a defined medium containing initially 20 g L^{-1} glucose under four different stress conditions: pH 3.0 (A), 40 °C (B), 10 g L^{-1} acetic acid (C) and 80 g L^{-1} ethanol (D). Performance was evaluated using relative growth (%), which was calculated dividing specific growth rate under stressful condition by specific growth rate under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol). Strains were sorted in ascending order of average values (bars) calculated using two independent culture replicates (dots).

Under high temperature stress (40 °C), most of the indigenous strains had similar performances to one another, but the relative growth rates were generally higher when compared to the industrial strains (**Figure 14B**). For example, with exceptions of UFMG-CM-Y636, UFMG-CM-Y254, UFMG-CM-Y260 and UFMG-CM-Y263, all remaining strains presented higher relative growth rates than PE-2. The capability of these strains to grow in elevated temperatures could be explained by the larger temperature fluctuations faced by the cells in their natural habitat (usually bark of trees), when compared to the industrial environment, where the variation is narrower. Della-Bianca and Gombert (2013) described a higher tolerance of industrial strains JP-1 and CAT-1 in comparison to PE-2 or to the lab strain CEN.PK113-7D, when evaluated by spot assays on YPD plates at 40 ^{o}C. In accordance, we observed similar results in our assay (**Figure 14B**). Indigenous strain UFMG-CM-Y256 showed the highest thermotolerance, and presented 100% higher relative growth rate than PE-2 (79% vs. 39%) and CEN.PK113-7D (79% vs. 37%).

The industrial yeasts represent an interesting background for strain development since they present features that allow them to survive and persist through the stressful conditions encountered in the production environment. The renewable products company Amyris based in Emeryville, California, for example, is using the industrial yeast strain *S. cerevisiae* PE-2 as background for farnesene production from sugarcane juice in Brotas-SP (PERALTA-YAHYA et al., 2012; UBERSAX; PLATT, 2010). The same industrial yeast was used by Granbio, a Brazilian biotech company based in Sao Paulo, as the main background to develop its commercially approved strain for production of fuel ethanol from sugarcane straw in Alagoas, Brazil (Prof. Gonçalo Pereira, University of Campinas, personal communication). Thus, our current fuel ethanol strains (PE-2, CAT-1, JP-1 and others) do present potential traits to be used in the new 2G processes.

In this context, the strains were evaluated in the presence of acetic acid, one of the most common weak acids present in lignocellulosic hydrolysates (ALMEIDA et al., 2007). When 10 g L^{-1} acetic acid were applied as a stressor, PE-2 and CEN.PK113-7D presented similar performances again (**Figure 14C**). Industrial strain CAT-1 displayed the highest average relative growth rate (92%), but very high variation among the replicates was observed (± 35%). Indigenous strain UFMG-CM-Y267 displayed the second highest relative growth rate (83%), similar to the values displayed by CEN.PK113-7D and PE-2 (75% and 76%, respectively), and 30% higher than JP-1 (64%).

To evaluate tolerance of the strains towards high ethanol concentrations, the medium was supplemented with 80 g L^{-1} ethanol, and the performance of strains was again evaluated in terms of growth relative to standard conditions. Most of the strains could not be analyzed using the applied growth profiling method due to abnormal growth and cell agglomeration. We attribute this behavior to the high level of stress imposed, which is not naturally or industrially faced by the cells. In the 1G fuel ethanol production process, titers as high as 120 g L^{-1} can be achieved (BASSO; BASSO; ROCHA, 2011), but, differently from the conditions applied here (where ethanol was added to the medium), this concentration is gradually reached during the fermentation and the ethanol is produced and excreted by the cells. Many studies have been published describing the responses of *S. cerevisiae* to ethanol stress. To mention some of the main observations: membrane fluidity and integrity are compromised (MISHRA; PRASAD, 1989), cell viability and growth rates are reduced (STANLEY; HOBLEY; PAMMENT, 1997), vacuole morphology is altered (MEADEN et al., 1999), trehalose accumulation is increased (LUCERO et al., 2000) and production of heat shock proteins is induced (PLESSET et al., 1982). But few works address induction of cell clumping or agglomeration, as we observed. Mill (1964) observed flocculation of non-flocculent yeast after addition of organic solvents, such as ethanol, to cell suspensions containing traces of calcium-chloride, but much higher concentrations were applied in that case (29-44% v/v). Nevertheless, the three industrial, one laboratory and two indigenous strains could be analyzed using our set-up, as they did not form cell clumps. The non-flocculent behavior of industrial yeasts is expected since they were obtained from ethanol fermentation processes and were chosen through selection programs that pondered such constraints (BASSO et al., 2008; DA SILVA-FILHO et al., 2005a). Indigenous strain UFMG-CM-Y260 presented the

highest relative growth rate (36%), similar to the value displayed by CEN.PK113-7D, CAT-1 and UFMG-CM-Y636 (29% in average), but 100% higher than did PE-2 (17%) and 150% higher than did JP-1 (14%) **(Figure 14D)**.

5.5 Some indigenous strains displayed slightly higher ethanol yields than industrial strains

We evaluated some relevant physiological and process parameters regarding industrial 1G ethanol production by mimicking the Brazilian fermentation process: anaerobic conditions (as far as possible), high cell density (9% wet mass/v), high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose, 40 g L^{-1} fructose), pH 4.5 and 35 °C. The cultivations were carried out in duplicates as described in the Methods section. Because the industrial ethanol fermentation cycle is short (6 to 10 h) (BASSO; BASSO; ROCHA, 2011), all the parameters were calculated for the first 10 h of cultivation. Besides the fourteen indigenous strains, three industrial strains (PE-2, CAT-1 and JP-1) and three laboratory strains (CEN.PK113-7D, CEN.PK122 and S288c) were included in the study.

Due to the very high inoculum fraction and high carbon/nitrogen (C/N) ratio (approximately 28) employed in this set-up, a lower biomass formation was observed in comparison to classical laboratory experiments (C/N ratio of medium described by VERDUYN et al., 1992 is 3.8). When low inoculum fractions and minimal media are used, a *S. cerevisiae* strain growing at constant $\mu = 0.3\ h^{-1}$ would result in a 30% biomass increase with respect to the initial biomass every hour and a 1,900% increase in biomass after 10 h of cultivation, as long as no nutrient limitations occur. In this assay, the highest biomass increase after 10 h of cultivation was displayed by indigenous strain UFMG-CM-Y262 (54 ± 3%) **(Figure 15)**. The other indigenous strains presented an increase in biomass ranging from 15 ± 2% (strain UFMG-CM-Y267) to 39 ± 2% (strain UFMG-CM-Y636). JP-1 presented the highest growth in terms of biomass formation (38 ± 6%) among the industrial strains, whereas strain PE-2 had an increment of only 24 ± 0% in relation to its initial cell concentration. Laboratory strains

CEN.PK113-7D, CEN.PK122 and S288c presented similar performance in terms of biomass formation with an average increase of 23 ± 1% after 10 h of cultivation. The relatively low values obtained for biomass production (15-54%) resemble the situation in industrial fermenters as we had aimed for in these experiments. However, the values reported by 1G ethanol mills are even lower than the numbers we observed. Basso, Basso and Rocha (2011) describe an estimated yeast biomass increase of only 5 to 10% during each fermentation cycle. The C/N ratio we used in our medium (28) is lower in comparison to the proportion required for biomass production during *S. cerevisiae* anaerobic growth reported by VERDUYN (1992) (C/N = 44), indicating that the limiting nutrient in our medium is still carbon. On the other hand, the numbers reported for industrial fermentations are much higher. Considering that sugar cane based substrates contain 18-22% of TRS (Total Reducing Sugars) (BASSO; BASSO; ROCHA, 2011) and 70-350 mg L^{-1} of nitrogen (AMORIM; LEÃO, 2005; BASSO; BASSO; ROCHA, 2011), a C/N ratio in the order of 200-1,000 is expected to be found in the industrial fermentation broth. These numbers are much higher than the C/N ratio of 44 described by VERDUYN (1992) and demonstrate the utilization of a severely nitrogen-limited medium. This could be an explanation for the differences observed between our results and the values reported for the ethanol industry.

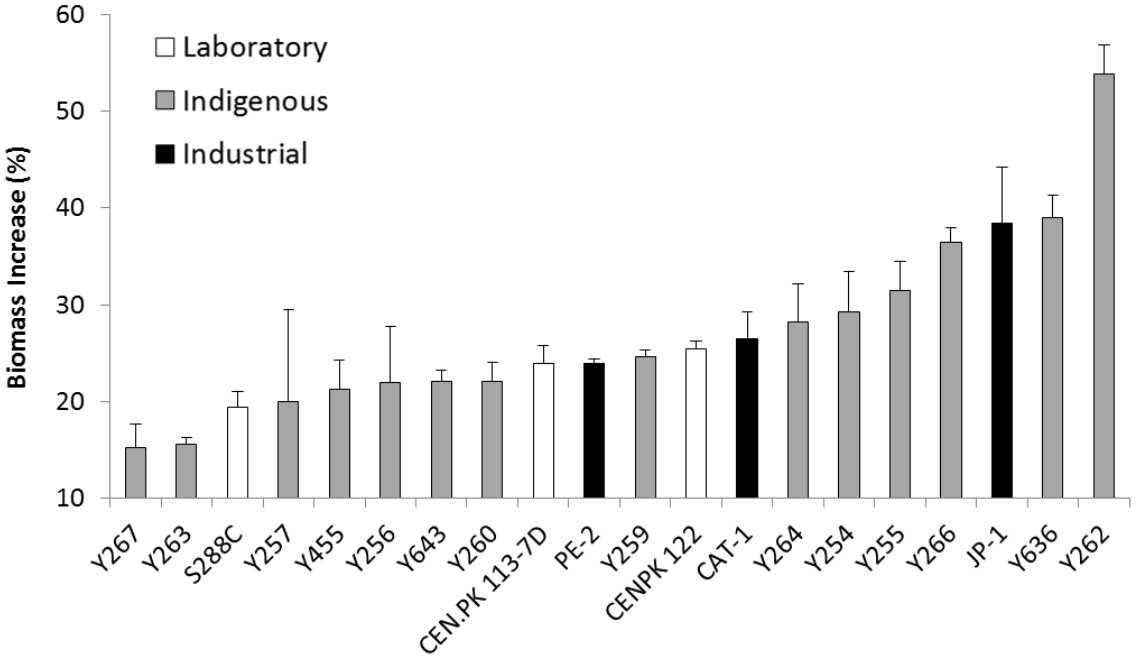

Figure 15 - Relative biomass increase of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions. The relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. The strains were sorted in ascending order of mean values. The error bars represent the mean deviation between two independent culture replicates.

Since more than 60% of the final costs of fuel ethanol production are due to the cost of sugarcane, the ethanol yield on sugars represents one of the most important parameters regarding improvement of fuel ethanol production. Considering that Brazilian ethanol mills have achieved yields between 90 and 92% of the stoichiometric conversion (2 mol_{Eth} $mol_{Equivalent\ Hexose}^{-1}$) during the past decade (DELLA-BIANCA et al., 2013), and that 27.5 billion liters were produced in the 2013/14 season (Brazilian Sugarcane Industry Association, 2015), any improvement of this parameter can lead to substantial economic gains to the whole production chain. In our batch cultivations, the indigenous strains UFMG-CM-Y260, UFMG-CM-Y257, UFMG-CM-Y259, UFMG-CM-Y256 and UFMG-CM-Y255 and the lab strain S288c displayed on average a 5% higher ethanol yield than the industrial strain JP-1, the best performing industrial strain under these conditions in terms of ethanol yield (1.64 ± 0.07 vs. 1.57 ± 0.03 mol_{Eth} $mol_{Eq.\ Hex.}^{-1}$) (**Figure 16**). The industrial strain PE-2 presented an ethanol

yield similar to JP-1 (1.55 ± 0.01 mol_{Eth} $mol_{Eq.\ Hex.}^{-1}$) as well as the lab strains CEN.PK113-7D and CEN.PK122 (1.56 ± 0.03 and 1.53 ± 0.00 mol_{Eth} $mol_{Eq.\ Hex.}^{-1}$, respectively). CAT-1 displayed a slightly lower ethanol yield, when compared to the other industrial strains (1.50 ± 0.03 mol_{Eth} $mol_{Eq.\ Hex.}^{-1}$). In comparison to our results, Guo, Olsson (2014) reported a similar ethanol yield on glucose for CEN.PK113-7D (1.48 mol_{Eth} $mol_{Eq.\ Hex.}^{-1}$) in anaerobic batch cultures using a defined medium, which is very different from the complex medium we used. On the other hand, the results we obtained with the industrial strains were much lower than those reported by Brazilian ethanol mills (the 90% stoichiometric conversion corresponds to 1.80 mol_{Eth} $mol_{Eq.\ Hex.}^{-1}$). Again, these facts indicate that the conditions applied here do not accurately resemble the real industrial set-up. However, as already discussed by Della-Bianca et al., (2014), it is not a simple task to mimic an industrial environment in the laboratory. It should be emphasized that, although some indigenous strains presented higher ethanol yields than industrial fuel ethanol strains, this trait *per se* is probably not the one that defines persistence and dominance in non-aseptic industrial fermentors, which are more related to the specific growth rate and tolerance to the stress factors imposed.

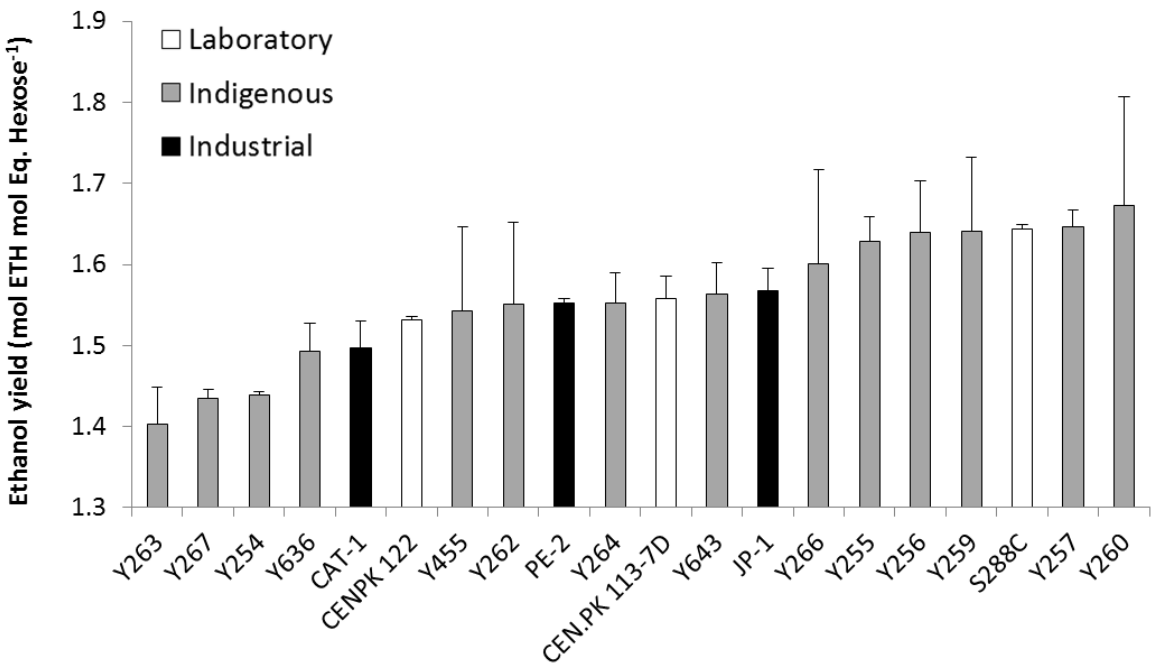

Figure 16 - Ethanol yield on sugars of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions. Ethanol yield (in terms of moles of ethanol produced divided by moles of hexose equivalents consumed) was obtained as the slope of the straight line from the plot of ethanol concentration against hexose-equivalent concentration. The strains were sorted in ascending order of mean values. The error bars represent the mean deviation between two independent culture replicates. *ETH* ethanol, *Eq. Hexose* equivalent hexose.

Productivity of ethanol is another relevant parameter regarding fuel ethanol production and was also evaluated in this assay for the first 10 h of cultivation. The lowest values of ethanol productivities were obtained for indigenous strains UFMG-CM-Y263 and UFMG-CM-Y636 (0.12 ± 0.00 mol L^{-1} h^{-1}, on average) (**Figure 17**) and the highest values were obtained by indigenous strains UFMG-CM-Y255, UFMG-CM-Y256 and UFMG-CM-Y257 (0.19 ± 0.01 mol L^{-1} h^{-1}, on average), which also displayed higher ethanol yields than industrial strains. Industrial strain JP-1 also presented a high ethanol productivity (0.19 ± 0.00 mol L^{-1} h^{-1}) even though the yield was slightly lower than those presented by the indigenous strains. The ethanol productivities obtained by the other industrial strains PE-2 and CAT-1 (0.16 and 0.17 mol L^{-1} h^{-1}, respectively) were similar but in the lower range of the values reported by the fuel ethanol industry. The lowest ethanol productivity was observed for the laboratory strains S288c and CEN.PK122 (0.15 ± 0.01 mol L^{-1} h^{-1} for both) and the productivity displayed

by CEN.PK113-7D was only slightly higher (0.16 ± 0.00 mol L^{-1} h^{-1}). Considering that during industrial production ethanol titers of 8 to 12% (v/v) are achieved in 6 to 10 hours fermentation (BASSO; BASSO; ROCHA, 2011), the industrial productivities vary between 8 g L^{-1} h^{-1} (0.17 mol L^{-1} h^{-1}) and 15 g L^{-1} h^{-1} (0.33 mol L^{-1} h^{-1}). Therefore, the lower values of productivities and also ethanol yields observed in our assay may be related to the differences in the medium composition, particularly in the C/N ratio applied, which is much higher in the sugar cane based substrates than in the YP sugar concentrated medium we used. Hence, the nitrogen-limited medium used in the industrial set-up probably impairs biomass formation and instead favors ethanol production.

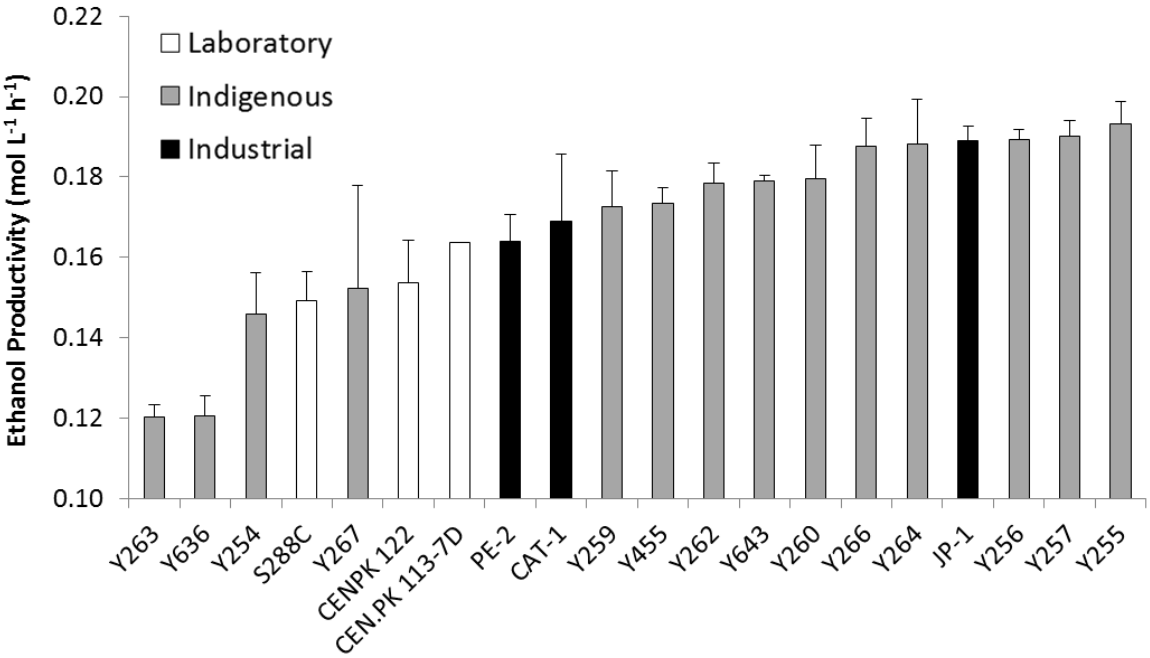

Figure 17 - Ethanol productivity of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions Volumetric productivity of ethanol was calculated using the formula: ΔP_{Eth} / Δt, where ΔP_{Eth} corresponds to the final ethanol concentration minus the initial ethanol concentration, in mol L^{-1}, and Δt is 10 h. The strains were sorted in ascending order of mean values. The error bars represent the mean deviation between two independent culture replicates.

In general, fructose was the most abundant residual sugar after 10 h of cultivation (**Figure 18**). There was a large variation among the strains tested and between the replicates. However, on average the residual fructose concentration was 0.11 ± 0.03 mol L^{-1} for all strains tested. The faster consumption of glucose in comparison to fructose is foreseen since *S. cerevisiae* is known to be a glucophilic yeast (GUILLAUME et al., 2007), leading to an increase of the fructose-to-glucose ratio along the fermentation (BERTHELS et al., 2004; BISSON, 1999). As expected, we observed an inversely proportional relation between ethanol productivity and final residual sugar concentration. In other words, the strains that presented lower ethanol productivities had, in general, higher concentrations of residual sugars at 10 h of cultivation. For example, the same two strains that presented the lowest ethanol productivities (UFMG-CM-Y636 and UFMG-CM-Y263) displayed the highest concentrations of residual fructose (0.26 ± 0.03 and 0.27 ± 0.03 mol L^{-1}, respectively) and glucose (0.11 ± 0.01 and 0.09 ± 0.02 mol L^{-1}, respectively) at 10 h of fermentation, illustrating their lower speed in converting these sugars into ethanol compared to the other strains. Accordingly, the strains that displayed some of the highest ethanol productivities, such as UFMG-CM-Y255, UFMG-CM-Y264, JP-1, UFMG-CM-Y256 and UFMG-CM-Y266, ended the 10 h cultivation with the lowest amounts of residual sugars. Most of the strains were able to consume most of the sucrose added, however some of the replicates of indigenous strains UFMG-CM-Y267 and UFMG-CM-Y636, laboratory strain S288c and industrial strain CAT-1 had high residual sucrose hence the average residual sucrose concentration was higher for these strains compared to the other strains.

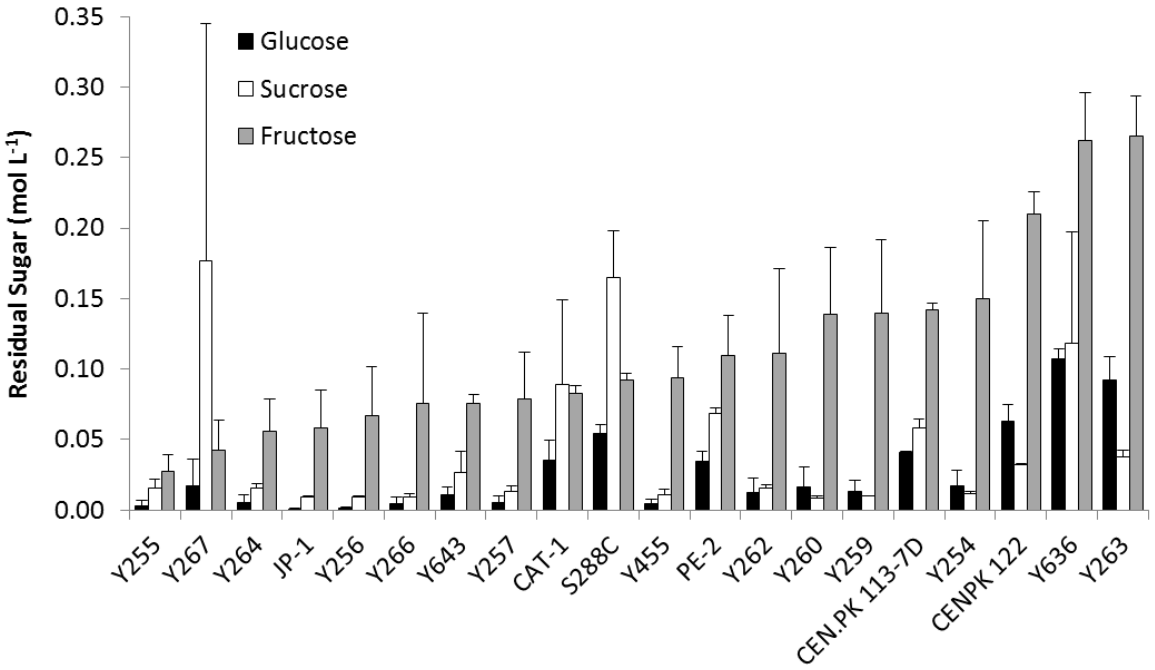

Figure 18 - Residual sugar at 10 h cultivation of 14 *S. cerevisiae* indigenous strains, 3 industrial strains (CAT-1, PE-2, JP-1) and 3 laboratory strains (CEN.PK113-7D, CEN.PK122 and S288C) during cultivations mimicking the 1G fuel ethanol industrial process. Strains were cultivated as described in the Methods section at pH 4.5, 35 °C, using YP medium with high sugar content (120 g L^{-1} sucrose, 40 g L^{-1} glucose and 40 g L^{-1} fructose) under anaerobic conditions**.** The strains were sorted in ascending order of fructose values. The error bars represent the mean deviation between two independent culture replicates.

5.6 Spotting on low pH YPD plates evidences the higher tolerance of industrial strains to acidic conditions

In the Brazilian 1G fuel ethanol industry, cells are recycled and repeatedly washed for 1-2 hours using sulphuric acid to reduce bacterial contamination, between two consecutive fermentation rounds. During this procedure, yeasts need to maintain high cell viability at very low pH (1.8-2.5) (BASSO; BASSO; ROCHA, 2011). Surviving these conditions represents a vital trait for strains to persist in the fermentation vats during the sugar-cane crushing season.

To further investigate the impact of the stress imposed to yeasts cells during the acid washing step and evaluate the performance of indigenous strains under even lower pH than that applied in the screening, cells were submitted to a non-proliferating environment containing water and sulphuric acid (pH adjusted to 2.0) after growth on liquid YPD. Unfortunately, we were not able to get reproducible results of cell viability using this setup (data not shown), probably due to differences in the metabolic state in which cells were transferred from YPD cultures to the sulphuric acid solution, as stress tolerance is known to be growth phase-dependent (LEWIS et al., 1993).

In order to have more consistent results, cells were submitted to a spotting test on YPD plates with pH adjusted to 2.5 using H_2SO_4. Five different concentrations (Abs_{600nm}= 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4} and 10^{-5}) were used and strains were incubated at 30 °C for 72 h to identify the most tolerant strains towards low pH stress. According to some previous results obtained (DELLA-BIANCA; GOMBERT, 2013), industrial strain PE-2 was able to grow at pH 2.5, while laboratory strain CEN.PK113-7D was not (**Figure 19**). These results show a different scenario from what was obtained using defined liquid medium at pH 3.0, in which similar growth was observed for the two strains. The performance of industrial and laboratory strains during acid stress has previously been studied by Della-Bianca et al., (2014), and when PE-2 and CEN.PK113-7D were cultivated in synthetic medium, no differences could be observed regarding acid tolerance. However, when cultivated in anaerobic batch cultures in complex medium and pH 2.7, strain PE-2 displayed a 36% higher maximum specific growth rate than CEN.PK113-7D. This work highlights how changes in medium composition and other cultivation parameters can influence the tolerance towards stress conditions in yeast strains, which was also confirmed in this work. Because of the different tolerances displayed by PE-2 and CEN.PK113-7D, these strains were used as positive and negative controls, respectively, and inoculated in all plates used in this assay. Industrial strain JP-1 also displayed the same results obtained by Della-Bianca and Gombert (2013) and was able to grow under pH 2.5. On the other hand, differently from what the authors described for industrial strain CAT-1, this strain was able to grow in our experiments, although to a lower extent compared to PE-2 and JP-1 (**Figure 19**). The higher tolerance of PE-2 and JP-1 towards low pH in comparison to CAT-1 and indigenous strains was also observed in the experiment using defined liquid medium (**Figure 14A**). Among the

indigenous strains, only three out of the fourteen strains tested displayed growth under the spot assay conditions: UFMG-CM-Y254, UFMG-CM-Y255 and UFMG-CM-Y260 (**Figure 19**). No correlation to the natural habitat could be established since the indigenous strains were isolated from three different plant species (**Table 1**). Surprisingly, UFMG-CM-Y254 and UFMG-CM-Y255, which displayed lower tolerance towards acidic stress when defined medium was used (**Figure 14A**), presented the highest tolerance among indigenous strains on the YPD spot stress test. The colony size of all strains that were able to grow under these conditions was smaller compared to results obtained on YPD plates with no pH adjustment (data not shown). The reduced colony size of stress-sensitized cells might be associated with a decrease in the total number of cells per colony as a result of lower growth rates (YOUSEF; JUNEJA, 2003).

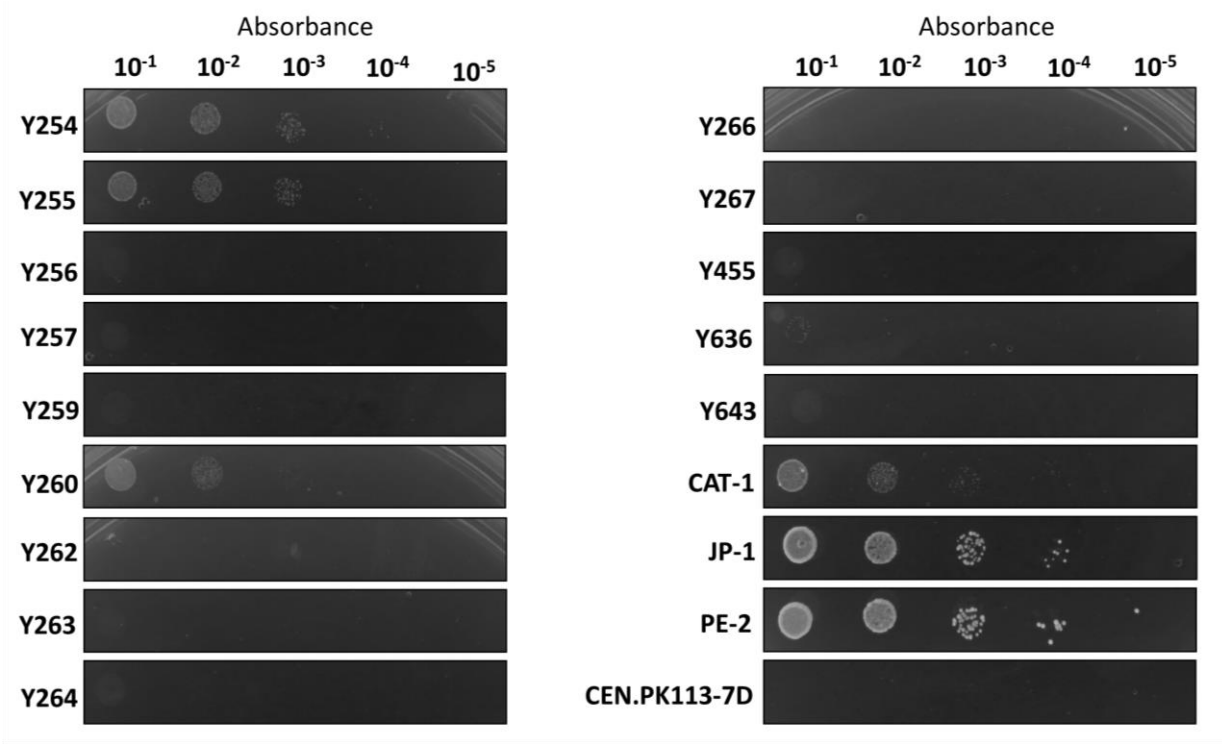

Figure 19 - Spotting on low pH YPD plates of 14 indigenous *S. cerevisiae* strains, 3 industrial (CAT-1, JP-1 and PE-2) and one laboratory strain (CEN.PK113-7D). Cells were plated at five dilutions, ranging from Abs_{600nm} 10^{-1} to 10^{-5}, on YPD plates with pH adjusted to 2.5 using H_2SO_4. Pictures were taken after 72 hours of incubation at 30 °C.

6 CONCLUSIONS AND OUTLOOK

In summary, we report a physiological description of 14 indigenous strains isolated from Brazilian biomes in comparison to laboratory and fuel ethanol strains, both under standard laboratory cultivation conditions and under conditions related to 1G and 2G fuel ethanol industrial processes.

- After trying to implement a molecular method for strain differentiation, the minisatellite-PCR method applied here proved to be the most useful one in comparison to the restriction analysis of mtDNA and the EI-1 PCR-fingerprinting. Using this technique, we were able to differentiate 9 out of the 20 strains tested and also to confirm the identity of strains PE-2, CAT-1 and S288C, by comparing the banding patterns obtained with the results described by Carvalho-Netto et al., (2013). The minisatellite-PCR method applied also revealed a molecular congruency among strains isolated from both the same geographical locations and plant species, suggesting a genotypic similarity among yeast strains isolated from the same habitat.

- We also confirmed that all strains used in this work belong to the species *Saccharomyces cerevisiae*, for which we applied the sequencing of the D1/D2 region of the 26S rDNA (KURTZMAN; ROBNETT, 1998; LACHANCE et al., 1999).

- Under standard conditions, none of the strains studied presented faster growth on glucose than on sucrose, which represents an interesting topic for further investigation. When grown on sucrose, industrial strain PE-2 displayed 17% higher specific growth rate than did laboratory strain CEN.PK-113-7D, which we believe may in part contribute to PE-2's ability to overcome other starter strains during industrial fermentation. Moreover, some indigenous strains, such as UFMG-CM-Y259 (μ_{max} = 0.57 ± 0.02 h^{-1}), presented even faster growth rates than PE-2 (μ_{max} = 0.45 ± 0.02 h^{-1}) when sucrose was used, giving evidence on the existing biodiversity among indigenous strains and their adequacy in searching for potential traits for industrial applications.

- Industrial strains presented in general higher tolerance towards low pH than indigenous strains. However, when 40 °C was applied as stress factor, some indigenous strains, such as UFMG-CM-Y256, presented higher relative growth rates than industrial or laboratorial strains (100% and 110% higher relative growth rate than PE-2 and CEN.PK113-7D, respectively). With the advances in the 2G fuel ethanol industry, acetic acid tolerance also became relevant, and this study allowed some tolerant indigenous strains to be identified (*e.g.* UFMG-CM-Y267). Finally, the cultivation mimicking industrial fuel ethanol production revealed indigenous strains with slightly higher ethanol yields (5% higher yield) and similar productivities to industrial strains (*e.g.* UFMG-CM-Y260).

- In conclusion, wild *S. cerevisiae* strains with useful traits for industrial applications were identified. These strains can serve in fuel ethanol production, but can also be used as platform strains and engineered for the production of other chemicals (STOVICEK et al., 2015). Strain UFMG-CM-Y257 presented high specific growth rate on sucrose (0.57 ± 0.02 h^{-1}), high ethanol yield (1.65 ± 0.02 mol_{Eth} $mol_{Eq.\ Hex}.^{-1}$), high ethanol productivity (0.19 ± 0.00 mol L^{-1} h^{-1}), and high tolerance to acetic acid and high temperature (40 °C). Strain UFMG-CM-Y260 displayed high ethanol yield (1.67 ± 0.13 mol_{Eth} $mol_{Eq.\ Hex}.^{-1}$), high tolerance to ethanol and to low pH, both in liquid and solid medium, compared to other indigenous strains. Strain UFMG-CM-Y267 displayed high tolerance to acetic acid and to high temperature (40 °C), which is of particular interest to simultaneous saccharification and fermentation processes (SSF), in which acetic acid is present in large quantities and higher temperatures are used in order to have cellulolytic enzymes working closer to their optimal temperatures.

References*

ACHTMAN, M.; WAGNER, M. Microbial diversity and the genetic nature of microbial species. **Nature Reviews** v. 6, n.6, p. 431-440, 2008.

ALMEIDA, J. R. M.; MODIG, T.; PETERSSON, A. et al. Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. **Journal of Chemical Technology and Biotechnology,** v. 82, n. 4, p. 340–349, 2007.

AMORIM, H. V.; LEÃO, R. M. **Fermentação Alcoólica – ciência e tecnologia**. Piracicaba: Fermentec, 2005. 448p.

Andrietta, M. G. S.; Andrietta, S. R. and Stupiello, E. N. A. Bioethanol – What Has Brazil Learned About Yeasts Inhabiting the Ethanol Production Processes from Sugar Cane?. In Bernardes, M. A. S. (ed) **Biofuel production**. InTech, Rijeka, 2011. p. 67-84.

ARGUESO, J. L.; CARAZZOLLE, M. F.; MIECZKOWSKI, P. A. et al. Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. **Genome Research,** v. 19, n. 12, p. 2258-2270, 2009.

AXCELL, B.C. R&D in brewing: the importance of underpinning knowledge**. Proc. European Brewery Convention,** v. 31, p. 173, 2007.

BABRZADEH, F.; JALILI, R.; WANG, C. et al. Whole-genome sequencing of the efficient industrial fuel-ethanol fermentative *Saccharomyces cerevisiae* strain CAT-1. **Molecular Genetics and Genomics,** v. 287, n. 6, p. 485-494, 2012.

BADOTTI, F.; DÁRIO, M. G.; ALVES JUNIOR, S. I. et al. Switching the mode of sucrose utilization by *Saccharomyces cerevisiae*. **Microbiology Cell Factories**, v. 7, n.4, p. 1–11, 2008.

*According to:
ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação: referências: elaboração, Rio de Janeiro, 2002.

BARBOSA, B.; ALMEIDA, P.; SAFAR, S. V. B. et al. Evidence of natural hybridization in Brazilian wild lineages of *Saccharomyces cerevisiae*. **Genome Biology and Evolution**, v. 8, n. 2, p. 317–329, 2016.

BASSO, L. C.; AMORIM, H. V.; OLIVEIRA, A. J. et al. Yeast selection for fuel ethanol production in Brazil. **FEMS Yeast Research**, v. 8, n.7, p. 1155–1163, 2008.

BASSO, L. C.; BASSO, T.O.; ROCHA, S.N. Ethanol Production in Brazil: The Industrial Process and Its Impact on Yeast Fermentation. Bernardes, M. A. S. (ed) **Biofuel production**. InTech, Rijeka, 2011. p. 85-100

BASSO, L. C.; OLIVEIRA, A. J.; ORELLI, V. F. D. M. et al. **Dominância das leveduras contaminantes sobre as linhagens industriais avaliada pela técnica da cariotipagem**. In: V Congresso Nacional da STAB. Águas de São Pedro: STAB, 1993. p. 246-250 v. 5.

BASSO, T. P. **Improvement of *Saccharomyces cerevisiae* hybridization for increased tolerance towards inhibitors from second-generation ethanol substrate**. 107. PhD Thesis (Agricultural Microbiology). Faculdade de Agricultura Luiz Queiroz, Universidade de São Paulo, Piracicaba. 2015.

BATISTA, F. Receita do etanol supera a do açúcar nesta safra. Available at: http://www.novacana.com/n/etanol/mercado/precos/receita-etanol-supera-acucar-safra-280813/. Accessed: March 2015. 2013.

BERTHELS, N. J.; CORDERO OTERO, R. R.; BAUER, F. F. et al. Discrepancy in glucose and fructose utilization during fermentation by *Saccharomyces cerevisiae* wine yeast strains. **FEMS Yeast Research**, v. 4, n.7, p. 683–689, 2004.

BISSON, L. F. Stuck and Sluggish Fermentations. **American Journal of Enology and Viticulture**, v. 50, n.1, p. 107-119, 1999.

Brazilian Sugarcane Industry Association. Moagem de cana-de-açúcar e produção de açúcar e etanol - safra 2013/2014. Available at: http://www.unicadata.com.br/. Accessed at 03/14/2015, 2015.

BUCKHOLZ, R.G.; GLEESON, M. A. G. Yeast Systems for the Commercial Production of Heterologous Proteins. **Nature Biotechnology**, v. 9, n. 3, p. 1067 – 1072, 1991.

CARVALHO-NETTO, O. V.; CARAZZOLLE, M. F.; RODRIGUES, A. et al. A simple and effective set of PCR-based molecular markers for the monitoring of the *Saccharomyces cerevisiae* cell population during bioethanol fermentation. **Journal of Biotechnology**, v. 168, n. 4, p. 701-709, 2013.

CAVALIERI, D.; MCGOVERN, P.E.; HARTL, D.L. et al. Evidence for *S. cerevisiae* fermentation in ancient wine. **Journal of Molecular Evolution***,* v. 57, s. 1, p. S226–S232, 2003.

CIANI, M.; MANNAZZU, I.; MARINANGELI P. et al. Contribution of winery-resident *Saccharomyces cerevisiae* strains to spontaneous grape mustfermentation. **Antonie Van Leeuwenhoek**, v. 85, p 159-164, 2004.

DA SILVA-FILHO, E. A.; DE MELO, H. F.; ANTUNES, D. F. et al. Isolation by genetic and physiological characteristics of a fuel-ethanol fermentative *Saccharomyces cerevisiae strain* with potential for genetic manipulation. **Journal of Industrial Microbiology and Biotechnology**, v. 32, n. 10, p. 481–486, 2005a.

DA SILVA-FILHO, E. A.; dos Santos, S. K. B.; Resende, A. M. et al. Yeast population dynamics of industrial fuel-ethanol fermentation process assessed by PCR fingerprinting. **Antonie van Leeuwenhoek**, v. 88, n. 1, p. 13–23, 2005b.

DE BARROS LOPES, M.; SODEN, A.; MARTENS, A. L. et al. Differentiation and species identification of yeasts using PCR. **International Journal of Systematic Bacteriology**. v. 48, p. 279-286, 1998.

DELLA-BIANCA, B. E.; BASSO, T. O.; STAMBUK, B. U. et al. What do we know about the yeast strains from the Brazilian fuel ethanol industry? **Applied Microbiology and Biotechnology**, v. 97, n. 3, p. 979–991, 2013.

DELLA-BIANCA, B. E.; DE HULSTER, E.; PRONK, J. T. et al. Physiology of the fuel ethanol strain *Saccharomyces cerevisiae* PE-2 at low pH indicates a context-dependent performance relevant for industrial applications. **FEMS Yeast Research**, v. 14, n. 18, p. 1196-1205, 2014.

DELLA-BIANCA, B. E.; GOMBERT, A. K. Stress tolerance and growth physiology of yeast strains from the Brazilian fuel ethanol industry. **Antonie van Leeuwenhoek**, v. 104, n. 6, p. 1083–1095, 2013.

FAY, J.C.; BENAVIDES, J.A. Evidence for domesticated and indigenous populations of *Saccharomyces cerevisiae*. **PLoS Genetics**, v. 1, n. 1, e5. 2005.

FRUTOS, R. de L. **Estudio comparative entre aislados clínicos y no clínicos de *Saccharomyces cerevisiae* y su papel como patógeno emergente.** 254. PhD Thesis (Biotecnologia). Univeristitat de València, Instituto de Agroquímica y Tecnología de Alimentos (CSIC), Departamento de Biotecnología, Valencia. 2007.

GODDARD, M. R. Quantifying the complexities of *Saccharomyces cerevisiae's* ecosystem engineering via fermentation. **Ecology**, v. 89, n. 8, p. 2077–82, 2008.

GODDARD, M. R.; GREIG, D. *Saccharomyces cerevisiae*: a nomadic yeast with no niche? **FEMS Yeast Research***,* v. 15, n. 3, p. 1-6, 2015.

GOMBERT, A. K.; VAN MARIS, A. J. A. Improving conversion yield of fermentable sugars into fuel ethanol in 1st generation yeast-based production processes. **Current Opinion in Biotechnology**, v. 33, p. 81-86, 2015.

GRAY, M. W. Origin and Evolution of Mitochondrial DNA. **Annual Review of Cell Biology**, v. 5, p. 25-50, 1989.

GUILLAUME, C.; DELOBEL, P.; SABLAYROLLES, J. M. et al. Molecular Basis of fructose utilization by the wine yeast *Saccharomyces cerevisiae*: a mutated *HXT3* allele enhances fructose fermentation. **Applied Environmental Microbiology**, v. 73, n. 8, p. 2432–2439, 2007.

GUO, Z.; OLSSON, L. Physiological response of *Saccharomyces cerevisiae* to weak acids present in lignocellulosic hydrolysate. **FEMS Yeast Research***,* v. 14, n. 8, p. 1234–1248, 2014.

International Organization of Vine and Wine, The wine market: evolution and trends. Available at: http://www.oiv.int/oiv/cms/index. Accessed: March 2015. 2014.

KATZ, S.H.; MAYTAG, F. Brewing an ancient beer. **Archaeology**. v. 44, n. 4, p. 24–27, 1991.

KURTZMAN, C. P.; ROBNETT, C. J. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. **Antonie van Leeuwenhoek***.* v. 73, n. 4, p. 331-371, 1998.

LACHANCE, M. A.; BOWLES, J. M.; STARMER, W. T. et al. *Kodamaea kakaduensis* and *Candida tolerans*, two new ascomycetous yeasts species from Australian Hibiscus flowers. **Canadian Journal of Microbiology**, v. 45, n. 2, p.172-177, 1999.

LANDRY, C.R.; TOWNSEND, J. P.; HARTL, D. L. et al. Ecological and evolutionary genomics of *Saccharomyces cerevisiae*. **Molecular Ecology**, v.15, n. 3, p. 575-91, 2006.

LATTERMANN, C.; BÜCHS, J. Microscale and miniscale fermentation and screening. **Current opinion in Biotechnology**. v. 35, p. 1-6, 2015.

LEGRAS, J. L.; DIDIER, M.; CORNUET, J. M. et al. Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. **Molecular Ecology**, v. 16, n. 10, p. 2091-2102, 2007.

LEWIS, J. G.; NORTHCOTT, C. J.; LEARMONTH, R. P. et al. The need for consistent nomenclature and assessment of growth phases in diauxic cultures of *Saccharomyces cerevisiae*. **Journal of General and Applied Microbiology***, v.* 139, n. 4, p. 835–839, 1993.

LITI, G.; CARTER, D. M.; MOSES, A. M. et al. Population genomics of domestic and wild yeasts. **Nature***, v.* 458, n. 7236, p. 337–341, 2009.

LODOLO, E. J.; KOCK, J. L. F.; AXCELL, B. C. et al. The yeast *Saccharomyces cerevisiae*- the main character in beer brewing. **FEMS Yeast Research**, v. 8, n. 7, p. 1018-1036, 2008.

LUCERO, P.; PEÑALVER, E.; MORENO, E. et al. Internal trehalose protects endocytosis from inhibition by ethanol in *Saccharomyces cerevisiae*. **Applied Environmental Microbiology***,* v. 66, n. 10, p. 4456–446, 2000.

MARQUES, W. L.; RAGHAVENDRAN, V.; STAMBUK, B. U. et al. Sucrose and *Saccharomyces cerevisiae*: a relationship most sweet. **FEMS Yeast Research**, v. 6, fov107. 2015.

MAYR, E. **Systematics and the Origin of Species — From the Viewpoint of a Zoologist.** Cambridge: Harvard University Press, 1942. 372 pp.

MCGEE, H. **On food and cooking**. Simon and Schuster, New York, NY. 1984. 704 pp.

MCGOVERN, P. E. **Ancient Wine: the Scientific Search for the Origins of Viniculture.** Princeton University Press, Princeton, New Jersey. 2003. 392 pp.

MCGOVERN, P. E.; HARTUNG, U.; BADLER V. R. et al. The beginnings of winemaking and viniculture in the ancient Near East and Egypt. **Expedition**. v. 39, n. 1, p. 3–21, 1997.

MEADEN, P. G.; ARNEBORG, N.; GULDFELDT, L. U. et al. Endocytosis and vacuolar morphology in *Saccharomyces cerevisiae* are altered in response to ethanol stress or heat shock. **Yeast***,* v. 15, n. 12, p. 1211–1222, 1999.

MEWES, H. W.; ALBERMANN, K.; BÄHR, M. et al. Overview of the yeast genome. **Nature**, v. 387, n. 737, p. 7-65, 1997.

MILL, P. J. The nature of interactions between flocculent cells in the flocculation of *Saccharomyces cerevisiae*. **Microbiology***,* v. 35, p. 61-68, 1964.

MISHRA, P.; PRASAD, R. Relationship between ethanol tolerance and fatty acyl composition of *Saccharomyces cerevisiae*. **Applied Microbiology and Biotechnology**, v. 30, n. 3, p. 294–298, 1989.

MORTIMER, R.; POLSINELLI, M. On the origins of wine yeast. **Research in Microbiology**, v. 150, n. 3, p. 199−204, 1999.

MYERS, N.; MITTERMEIR, R. A.; MITTERMEIR, C. G. et al. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, n. 6772, p. 853-858, 2000.

NAUMOV, G. I.; GAZDIEV, D. O.; NAUMOVA, E. S. Identification of the yeast species *Saccharomyces bayanus* in Far East Asia. **Mikrobiologiia***,* v. 72, n. 6, p. 834–839, 2003.

NETTO, L. E. S. A rapid screen for determination of the protective role of antioxidant proteins in yeast. In: Harris JR, Graham J, Rickwood D (eds) **Cell biology protocols**. John Wiley & Sons Ltd, Chichester, 2006. v. 6, p. 255–258.

NEUBAUER, P.; CRUZ, N.; GLAUCHE, F. et al. Consistent development of bioprocesses from microliter cultures to the industrial scale. **Engineering in Life Sciences**, v. 13, n. 3, p. 224-238, 2013.

NEVOIGT, E. Progress in Metabolic Engineering of *Saccharomyces cerevisiae*. **Microbiology and Molecular Biology Reviews**, v. 72, n. 3, p. 379–412, 2008.

NIJKAMP, J. F.; VAN DEN BROEK, M.; DATEMA, E. et al. De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. **Microbial Cell Factories**, v. 11, n. 36. 2012.

OGUNSEITAN, O. The Concept of Microbial Species, in Microbial Diversity: Form and Function in Prokaryotes. Blackwell Science Ltd, Malden, MA, USA. doi: 10.1002/9780470750490.ch1. 2004.

PADDON, C. J.; WESTFALL, P. J.; PITERA, D. J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. **Nature***.* v. 496, n. 25, p. 528-532, 2013.

PERALTA-YAHYA, P. P.; ZHANG, F.; CARDAYRE, S. B. et al. Microbial engineering for the production of advanced biofuels. **Nature**, v. 488, n. 7411, p. 320-328, 2012.

PLESSET, J.; PALM, C.; MCLAUGHLIN, C. S. Induction of heat shock proteins and thermotolerance by ethanol in *Saccharomyces cerevisiae*. **Biochemical and Biophysical Research Communications***,* v. 108, n. 3, p. 1340–1345, 1982.

QUEROL, A.; BARRIO, E.; HUERTA, T. et al. Molecular monitoring of wine fermentations conducted by active dry yeasts strains. **Applied Environmental Microbiology**. v. 58, n. 9, p. 2948-2953, 1992.

RAMOS, C. L.; DUARTE, W. F.; FREIRE, A. L. et al. Evaluation of stress tolerance and fermentative behavior of indigenous *Saccharomyces cerevisiae*. **Brazilian Journal of Microbiology**. v. 44, n. 3, p. 935-944, 2013.

SAFAR, S. V. B. **Diversidade molecular e fisiológica de isolados indígenas de *Saccharomyces cerevisiae* provenientes de diferentes ecossistemas brasileiros**. 115. MSc dissertation (Microbiologia). Departamento de Microbiologia, ICB/UFMG, Belo Horizonte. 2013.

SAMPAIO, J. P.; GONÇALVES, P. Natural Populations of *Saccharomyces kudriavzevii* in Portugal Are Associated with Oak Bark and Are Sympatric with *S. cerevisiae* and *S. paradoxus.* **Applied and Environmental Microbiology***,* v. 74, n. 7, p. 2144-2152, 2008.

SIMON, J. A.; SZANKASI, P.; NGUYEN, D. K. et al. Differential toxicities of anticancer agents among DNA repair and checkpoint mutants of *Saccharomyces cerevisiae*. **Cancer Research**, v. 60, n. 2, p. 328-333, 2000.

SNIEGOWSKI, P. D.; DOMBROWSKI, P. G.; FINGERMAN, E. *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics. **FEMS Yeast Research**, v. 1, n. 4, p. 299-306, 2002.

STACKEBRANDT, E.; EBERS, J. Taxonomic parameters revisited: tarnished gold standards. **Microbiology Today,** v. 33, p. 152–155, 2006.

STAMBUCK, B. U.; DUNN, B.; ALVES JR, S. L. et al. Industrial fuel ethanol yeasts contain adaptive copy number changes in genes involved in vitamin B1 and B6 biosynthesis. **Genome Research**, v. 19, n. 12, p. 2271-2278, 2009.

STANLEY, G. A.; HOBLEY, T. J.; PAMMENT, N. B. Effect of acetaldehyde on *Saccharomyces cerevisiae* and *Zymomonas mobilis* subjected to environmental shocks. **Biotechnology and Bioengineering**, v. 53, n. 1, p. 71-78, 1997.

STECKELBERG, C. **Caracterização de leveduras de processos de fermentação alcoólica utilizando atributos de composição celular e características cinéticas.** 202. PhD Thesis (Engenharia Química). Universidade Estadual de Campinas, Campinas. 2001.

STEFANINI, I.; DAPPORTO, L.; LEGRAS, J. L. et al. Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution. **PNAS**, n. 109, n. 33, p. 13398–13403, 2012.

STOVICEK, V.; BORIDINA, I.; FORSTER, J. CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. **Metabolic Engineering Communications**, n. 2, p. 13-22, 2015.

UBERSAX, J.; PLATT, D. Genetically modified microbes producing isoprenoids. Patent world 1,414,52 A1. Dec. 9, 2010.

VAN DIJKEN, J. P.; BAUERB, J.; BRAMBILLAC, L. et al. An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. **Enzyme and Microbial Technology**, v. 26, n. 9-10, p. 706-714, 2000.

VANDAMME, P.; POT, B.; GILLIS, M. et al. Polyphasic taxonomy, a consensus approach to bacterial systematics. **Microbiology Reviews***,* v. 60, n. 2, p. 407–438, 1996.

VERDUYN, C. Energetic aspects of metabolic fluxes in yeast (Doctoral dissertation). Retrieved from: http://repository.tudelft.nl/view/ir/uuid:872b54c8-6642-4e99-83b6-ee6c9d69cb04/. 1992.

VERDUYN, C.; Postma, E.; Scheffers, W. A. et al. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. **Yeast**, v. 8, n. 7, p. 501-517, 1992.

WANG, Q.-M.; LIU, W.-Q.; LITI, G. et al. Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. **Molecular Ecology,** v. 21, p. 5404–5417, 2012.

WAYNE, L. G.; BRENNER, D. J.; COLWELL, R. R. et al. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. **International Journal of Systematic Bacteriology**, v. 37, n. 4, p. 463–464, 1987.

WHEALS, A. E.; BASSO, L. C.; ALVES, D. M. G. et al. Fuel ethanol after 25 years. **Trends in Biotechnology**, v. 17, n. 12, p. 482-437, 1999.

Wine Institute, World Wine Production by Country. Avaiable at <http:// www.wineinstitute.org/resources/statistics>. Accessed: March 2015. 2013.

YOUSEF, A. E.; JUNEJA, V. K. Mechanism of Stress Adaptive Response. In: YOUSEF, A. E.; JUNEJA, V. K (ed) **Microbial Stress Adaptation and Food Safety**. Washington D. C., USA: CRC Press, 2003. p. 21.

APPENDIX A – SUPPLEMENTARY RESULTS

A.1 Growth of *S. cerevisiae* strains under standard conditions using the automated plate reader Infinity M200

Cells were cultivated as described in Materials and Methods using a defined medium, containing vitamins, trace metals and salts (VERDUYN et al., 1992), including 5.0 g L^{-1} $(NH_4)_2SO_4$, 3.0 g L^{-1} KH_2PO_4, and 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, with 20 g L^{-1} sucrose or 20 g L^{-1} glucose and initial pH adjusted to 6.0 using KOH. Strains were incubated for 24 h at 30 °C and 44 rpm in the automated plate reader Infinity M200 (Tecan, Männedorf, Switzerland). Cultivations were run in triplicates and cell growth was monitored by measuring the absorbance at 570 nm every 15 min. Maximum specific growth rates (μ_{max}; in h^{-1}) were obtained by plotting the natural logarithm of absorbance values against time and calculating the slope of the straight line corresponding to the exponential growth phase (**Figure 20**, **Figure 21**, **Figure 22, Figure 23, Figure 24**, **Figure 25**, **Figure 26** and **Figure 27**).

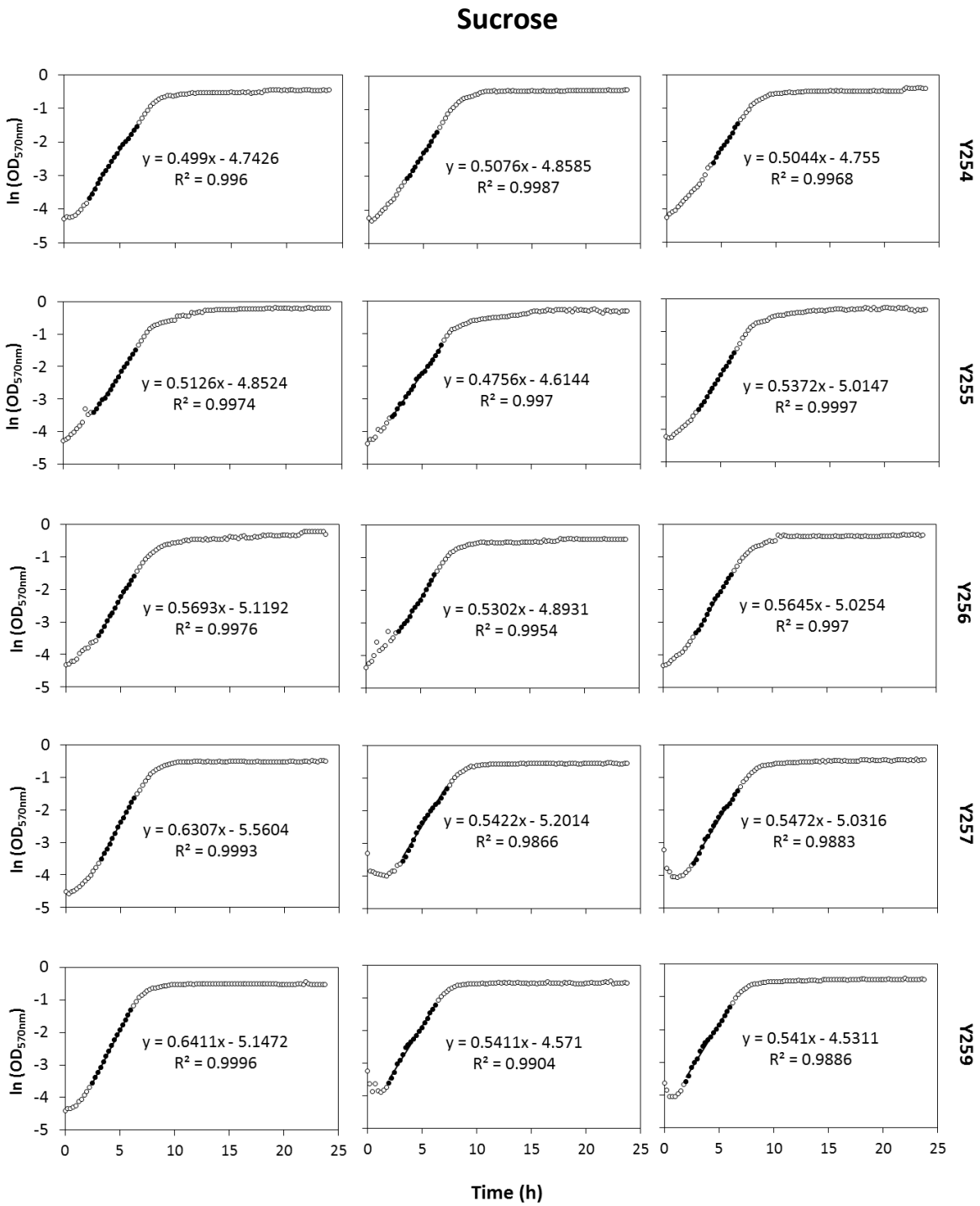

Figure 20 – Growth of *S. cerevisiae* strains Y254, Y255, Y256, Y257 and Y259 under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

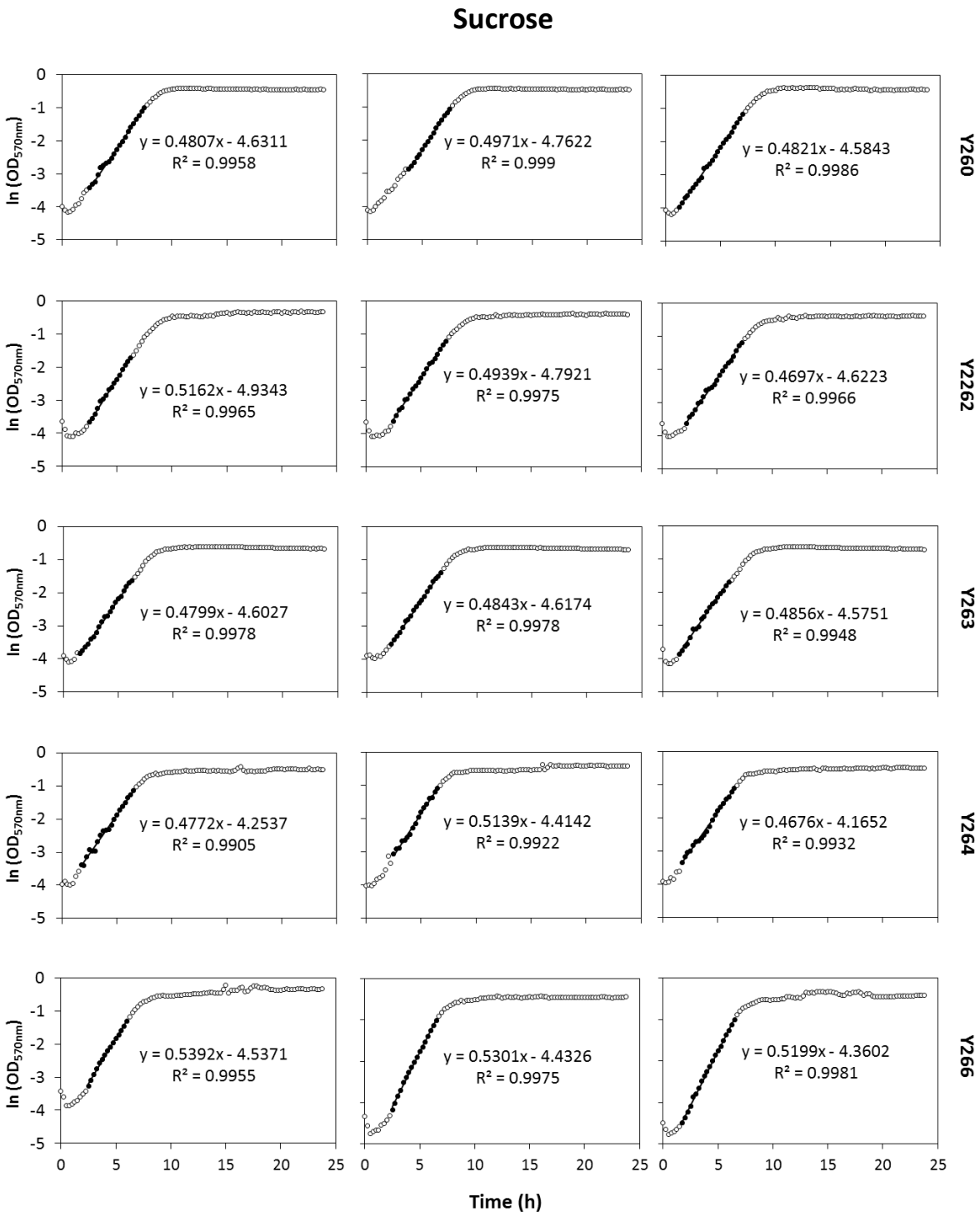

Figure 21 - Growth of *S. cerevisiae* strains Y260, Y262, Y263, Y264 and Y266 under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

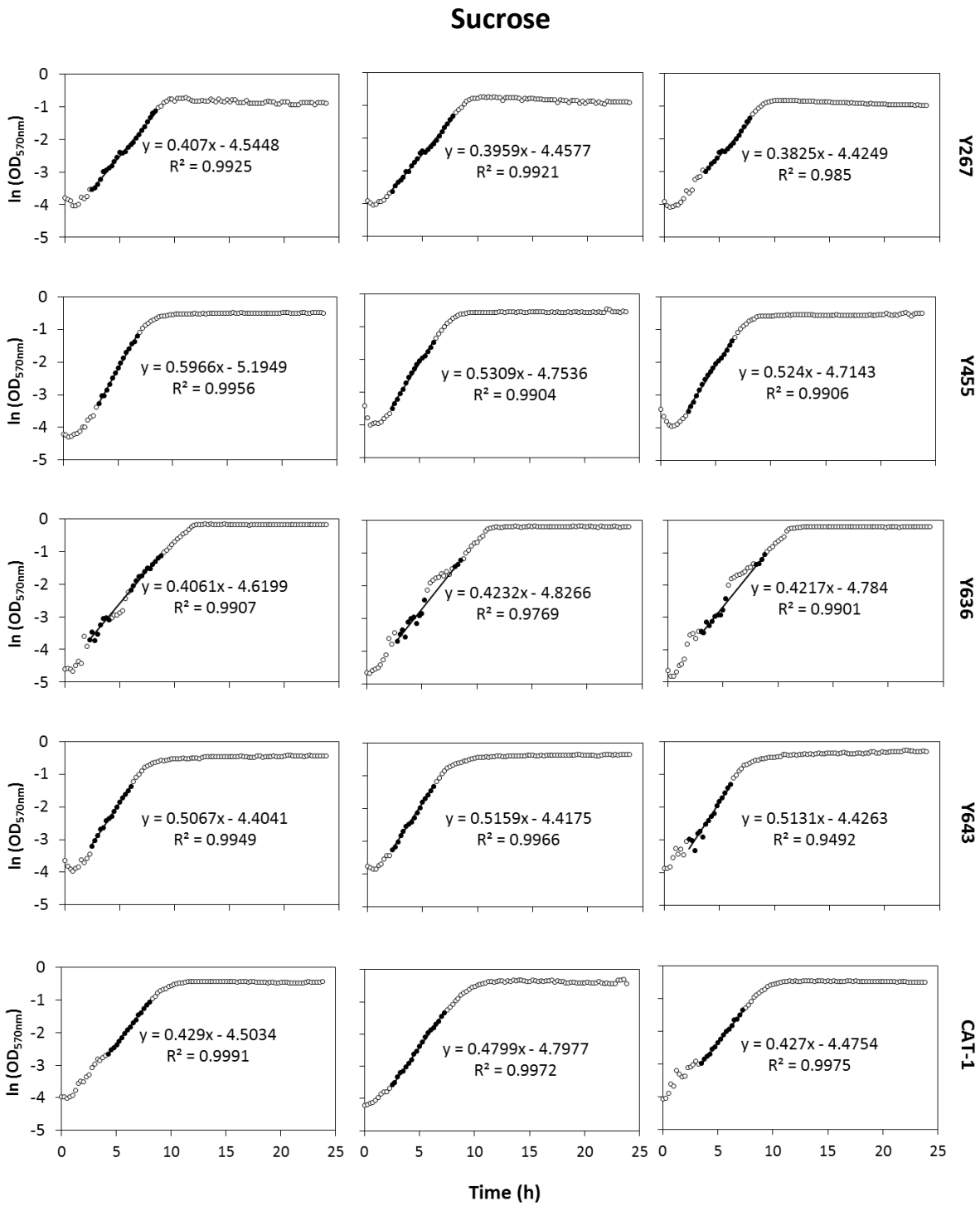

Figure 22 - Growth of *S. cerevisiae* strains Y267, Y455, Y636, Y643 and CAT-1 under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

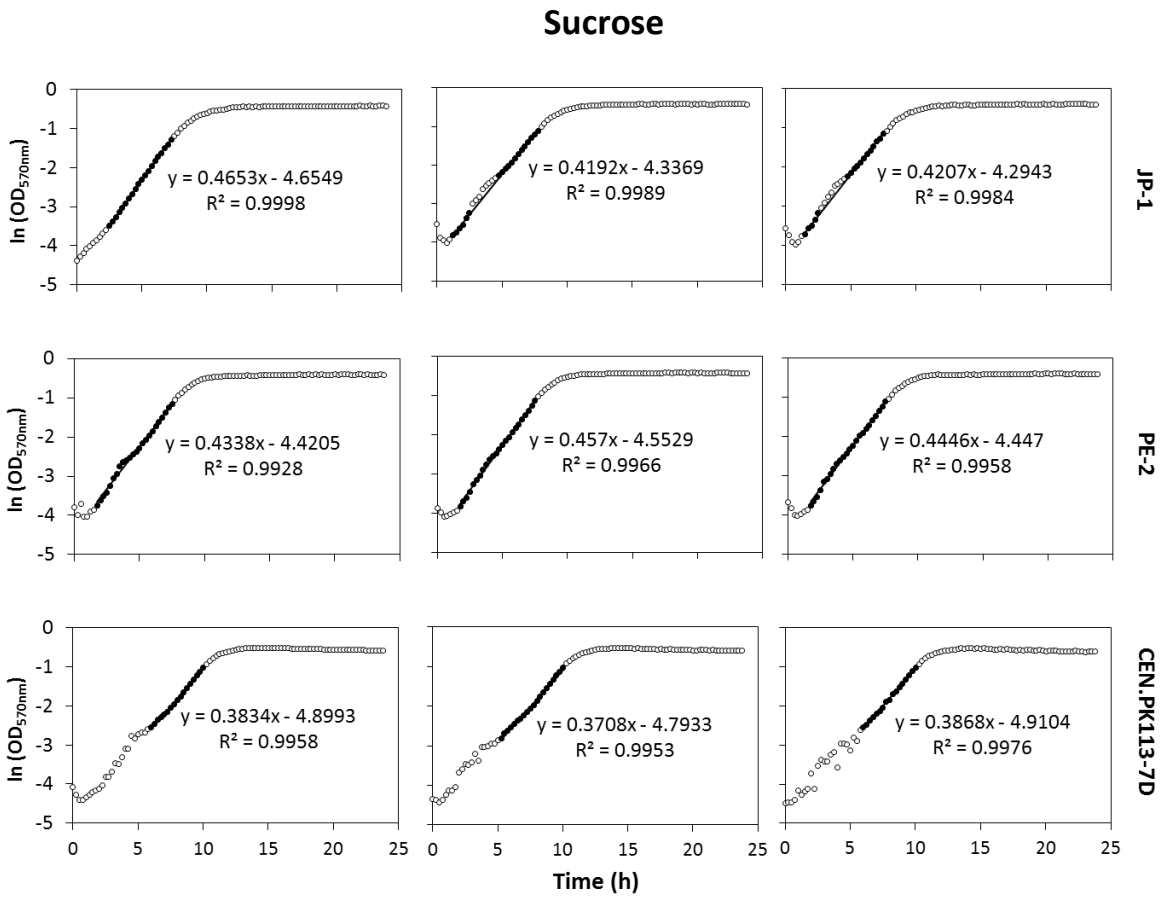

Figure 23 - Growth of *S. cerevisiae* strains JP-1, PE-2 and CEN.PK113-7D under standard conditions using 20 g L^{-1} sucrose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

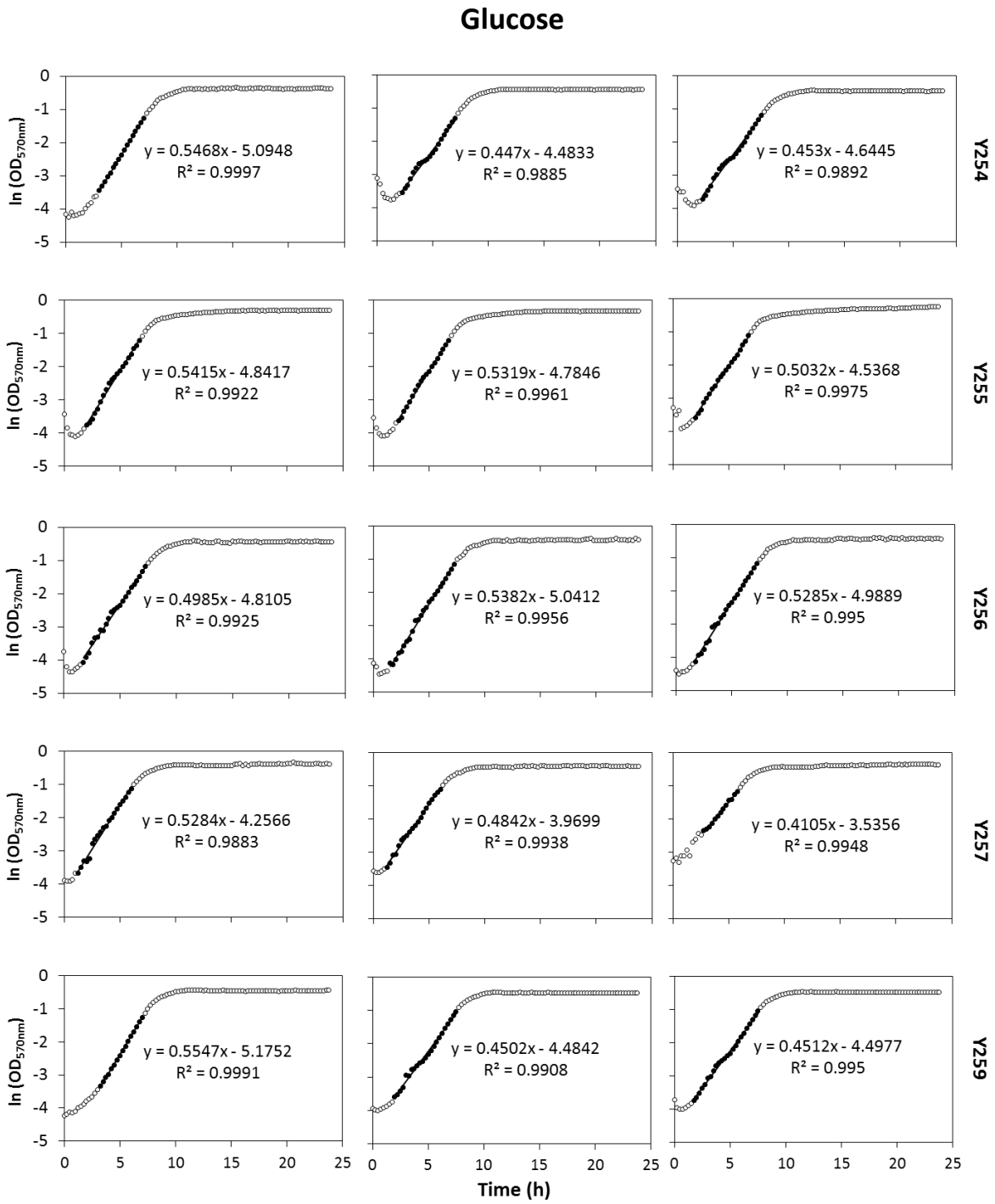

Figure 24 - Growth of *S. cerevisiae* strains Y254, Y255, Y256, Y257 and Y259 under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

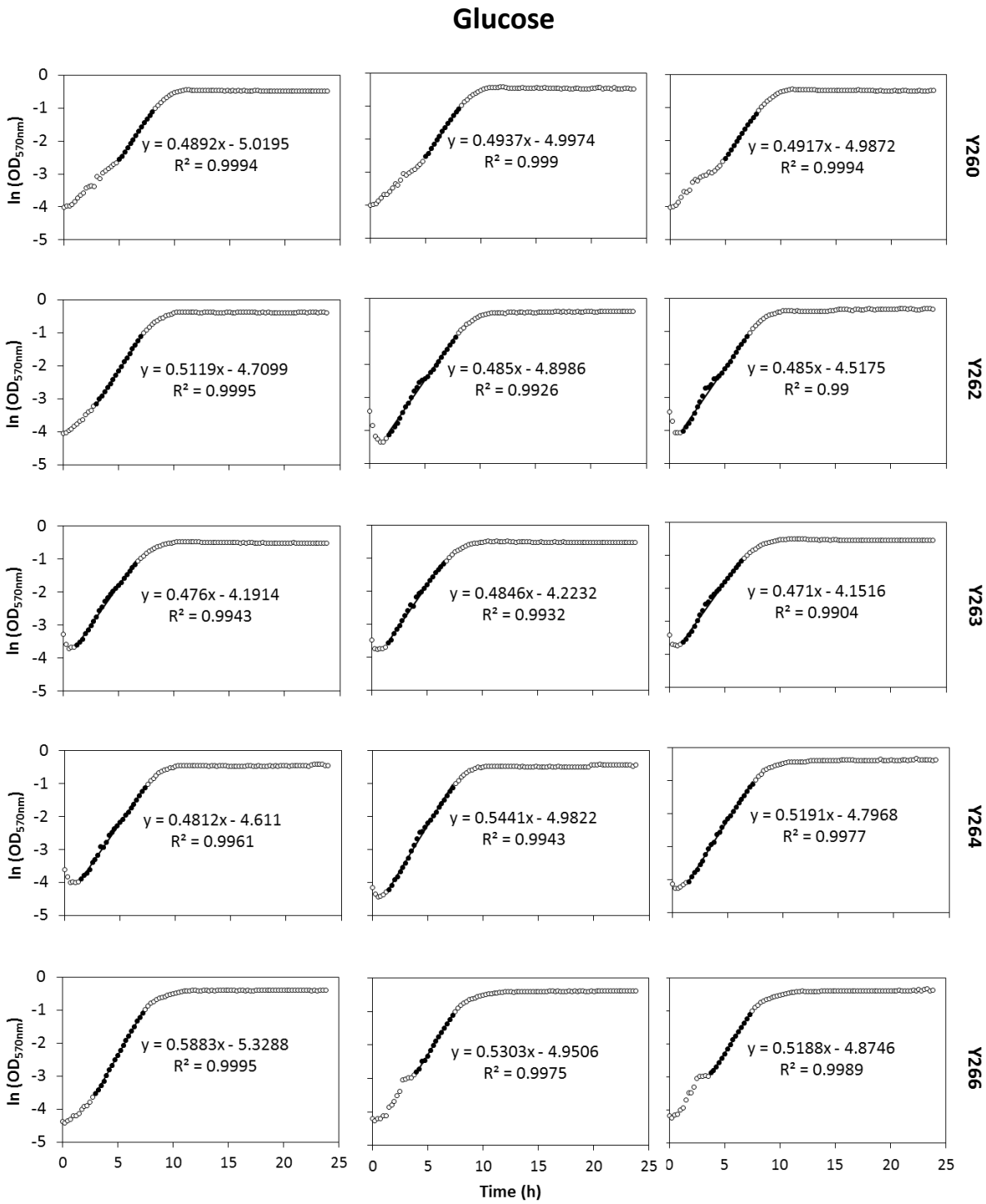

Figure 25 - Growth of *S. cerevisiae* strains Y260, Y262, Y263, Y264 and Y266 under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

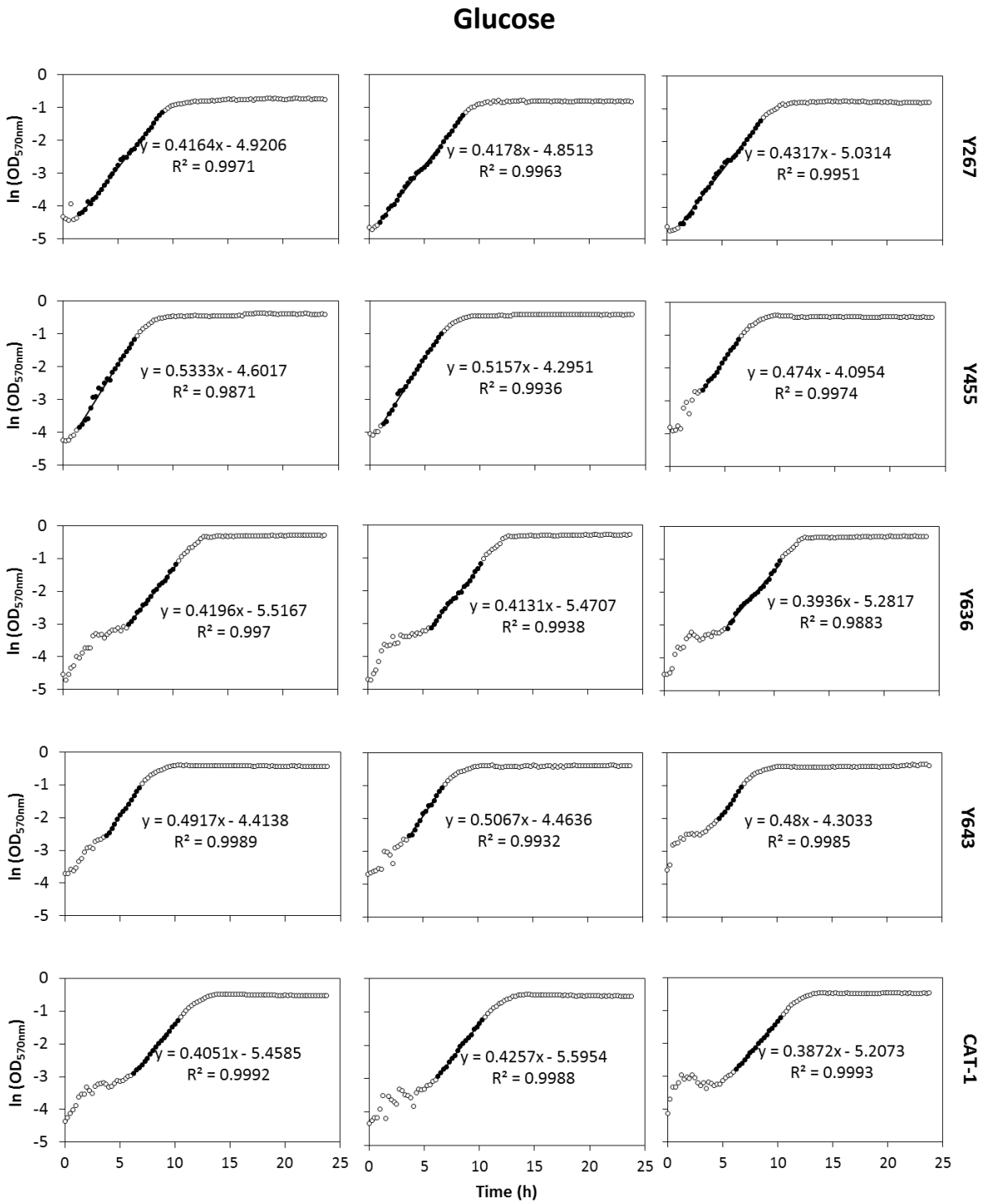

Figure 26 - Growth of *S. cerevisiae* strains Y267, Y455, Y636, Y643 and CAT-1 under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

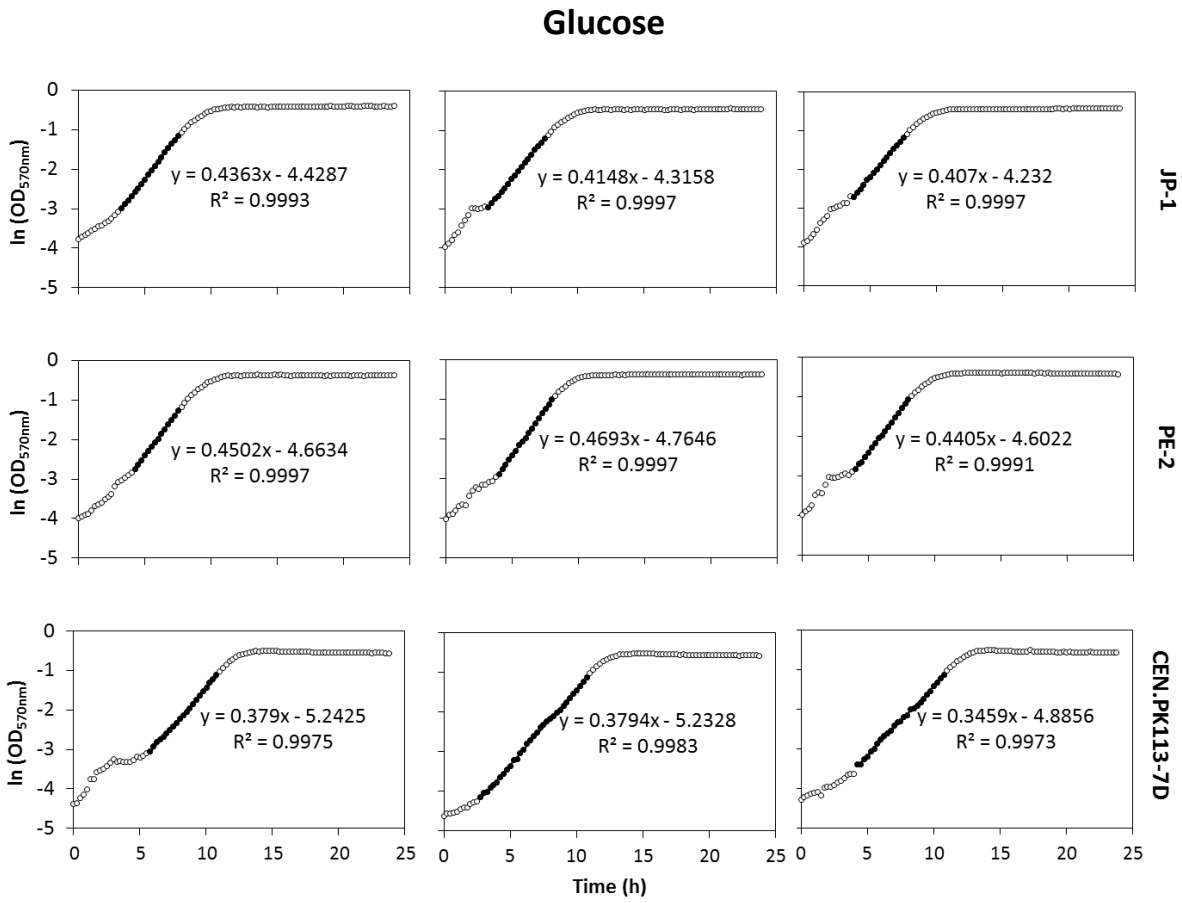

Figure 27 - Growth of *S. cerevisiae* strains JP-1, PE-2 and CEN.PK113-7D under standard conditions using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in triplicates using the automated plate reader TECAN Infinity M200. The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

A.2 Growth of *S. cerevisiae* strains under standard and stressful conditions using Growth Profiler 1152 (Enzyscreen)

Cells were cultivated as described in Materials and Methods using a defined medium containing vitamins, trace metals and salts according to Verduyn et al., (1992) with the following modifications: 6.6 g L^{-1} $(NH_4)_2SO_4$, 6.0 g L^{-1} NaH_2PO_4, and 0.5 g L^{-1} $MgSO_4.7H_2O$ with 20 g L^{-1} glucose. The medium was buffered with 0.01 g L^{-1} potassium phthalate monobasic and the initial pH was adjusted to 5.5 using 2 M NaOH. For ethanolic stress, 80 g L^{-1} ethanol were added to the medium, whereas for acetic acid stress, 10 g L^{-1} acetic acid were added

(still pH was adjusted to 5.5 using 2 M NaOH). For acid stress, potassium phthalate buffer was replaced by tartaric acid buffer (pKa = 2.98) and initial pH was adjusted to 3.0. Temperature stress was induced by incubation at 40 °C, instead of 30 °C. The plates were incubated at 225 rpm in the automated plate reader Growth Profiler 1152 (Enzyscreen, Haarlen, The Netherlands) and growth was measured every 15 min until stationary phase was reached. The values obtained (Green-value) were converted to Abs_{600nm} using a calibration curve. Maximum specific growth rates (μ_{max}; in h^{-1}) were obtained by plotting the natural logarithm of absorbance values against time and calculating the slope of the straight line corresponding to the exponential growth phase (**Figure 28, Figure 29, Figure 30, Figure 31, Figure 32, Figure 33, Figure 34, Figure 35, Figure 36, Figure 37, Figure 38, Figure 39, Figure 40, Figure 41, Figure 42, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47, Figure 48** and **Figure 49**). Growth experiments were performed in duplicates. Performance of the strains was evaluated using relative growth rate (%), which was calculated dividing the maximum specific growth rate (μ_{max}) under a particular stressful condition by the maximum specific growth rate under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol).

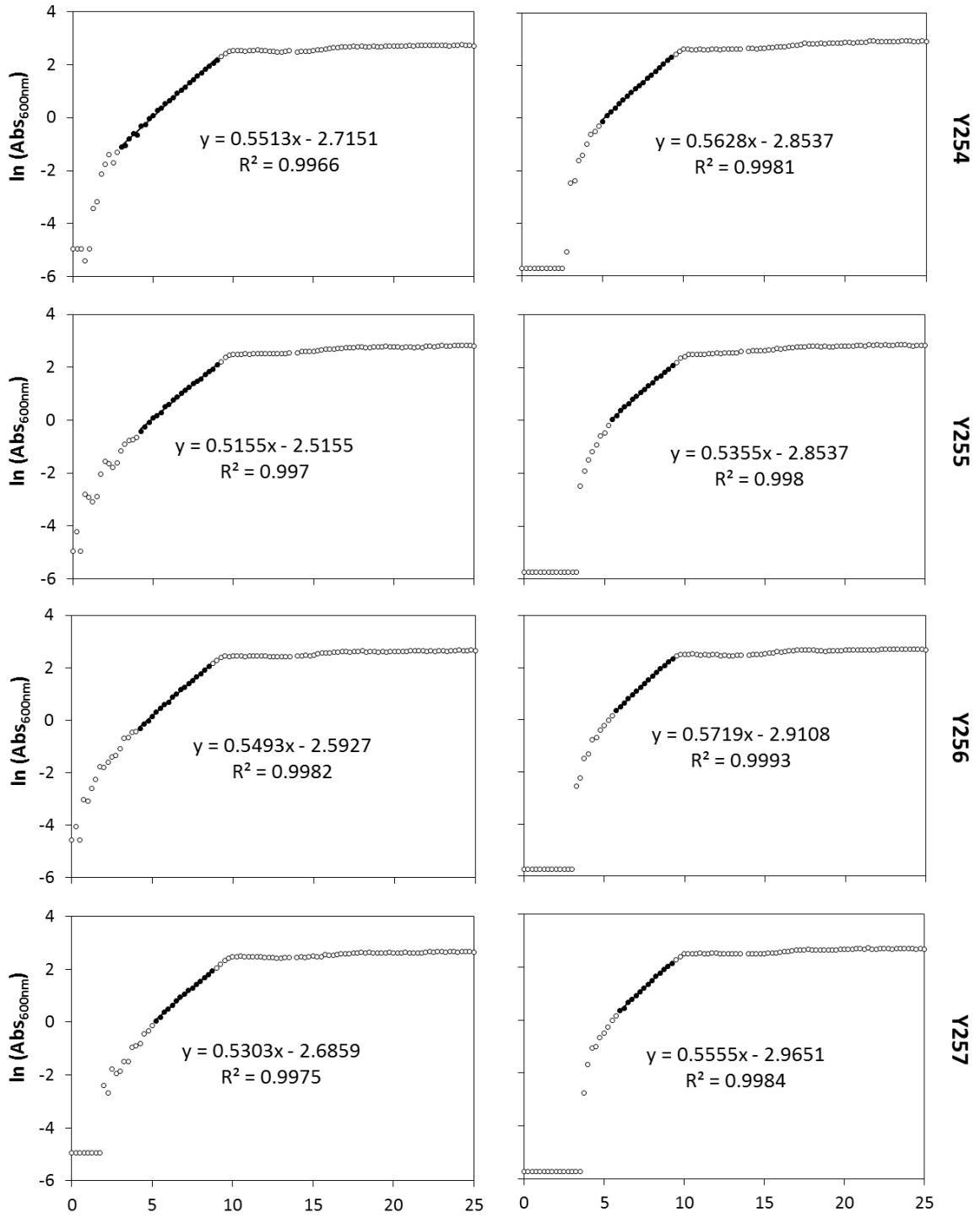

Figure 28 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

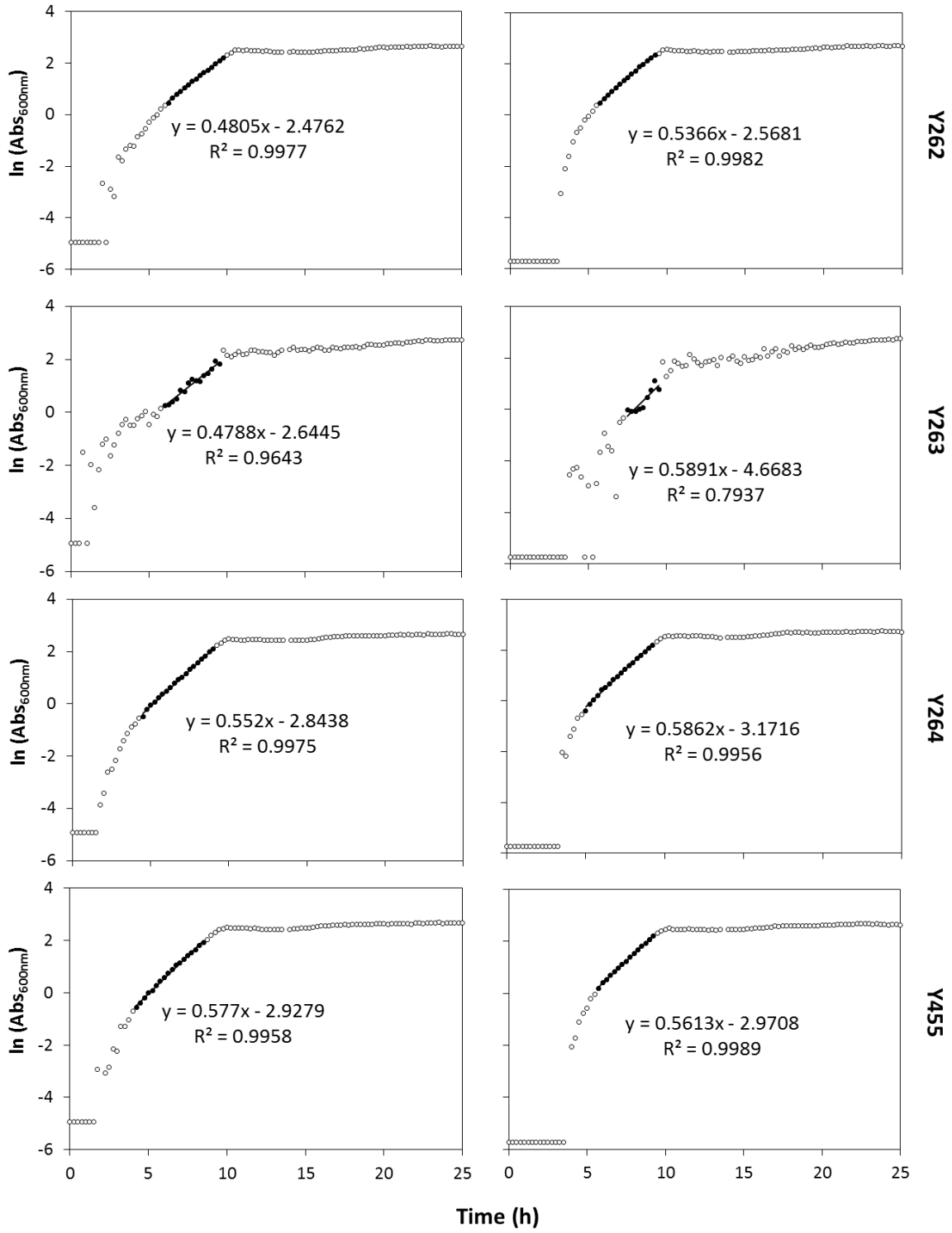

Figure 29 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

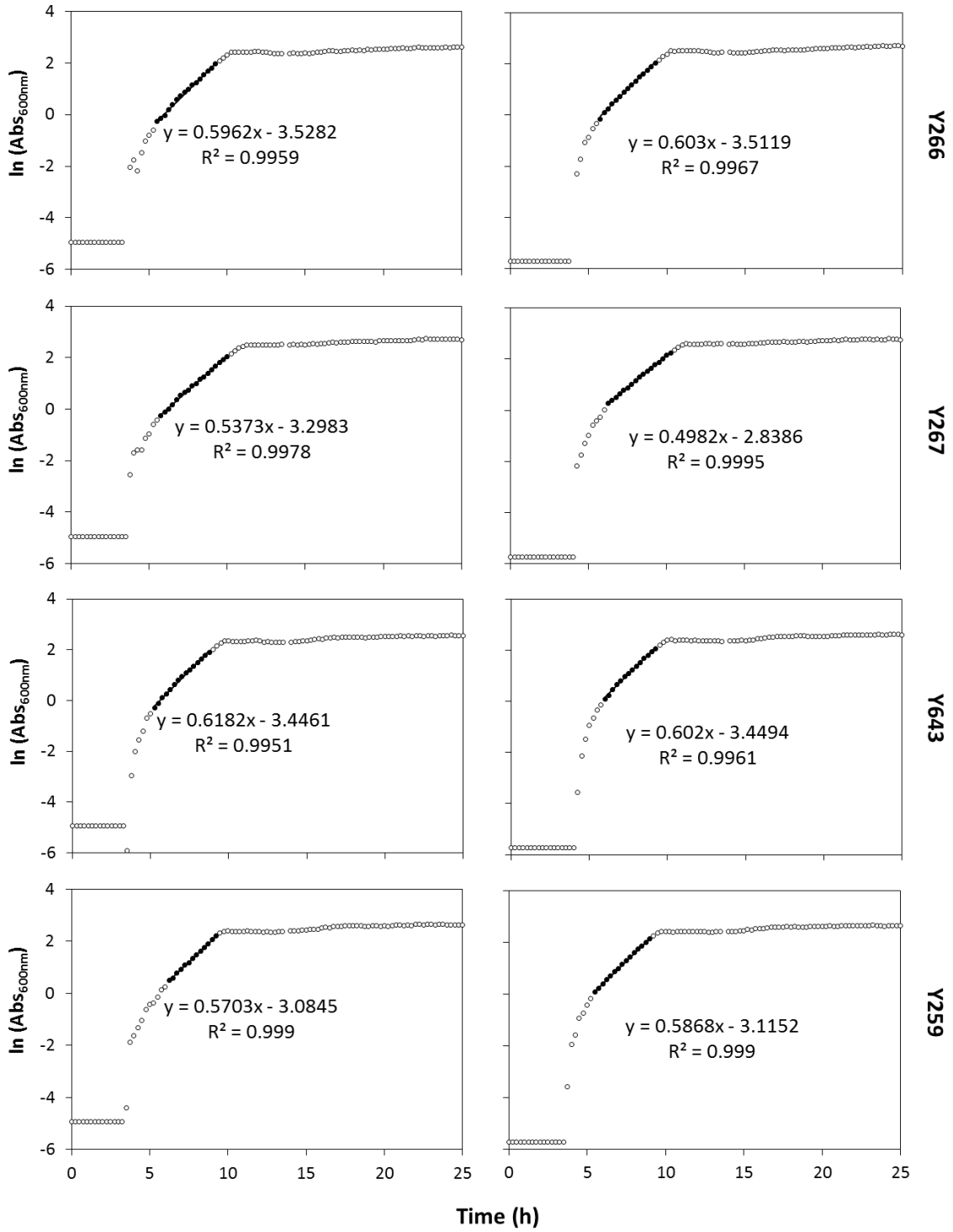

Figure 30 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

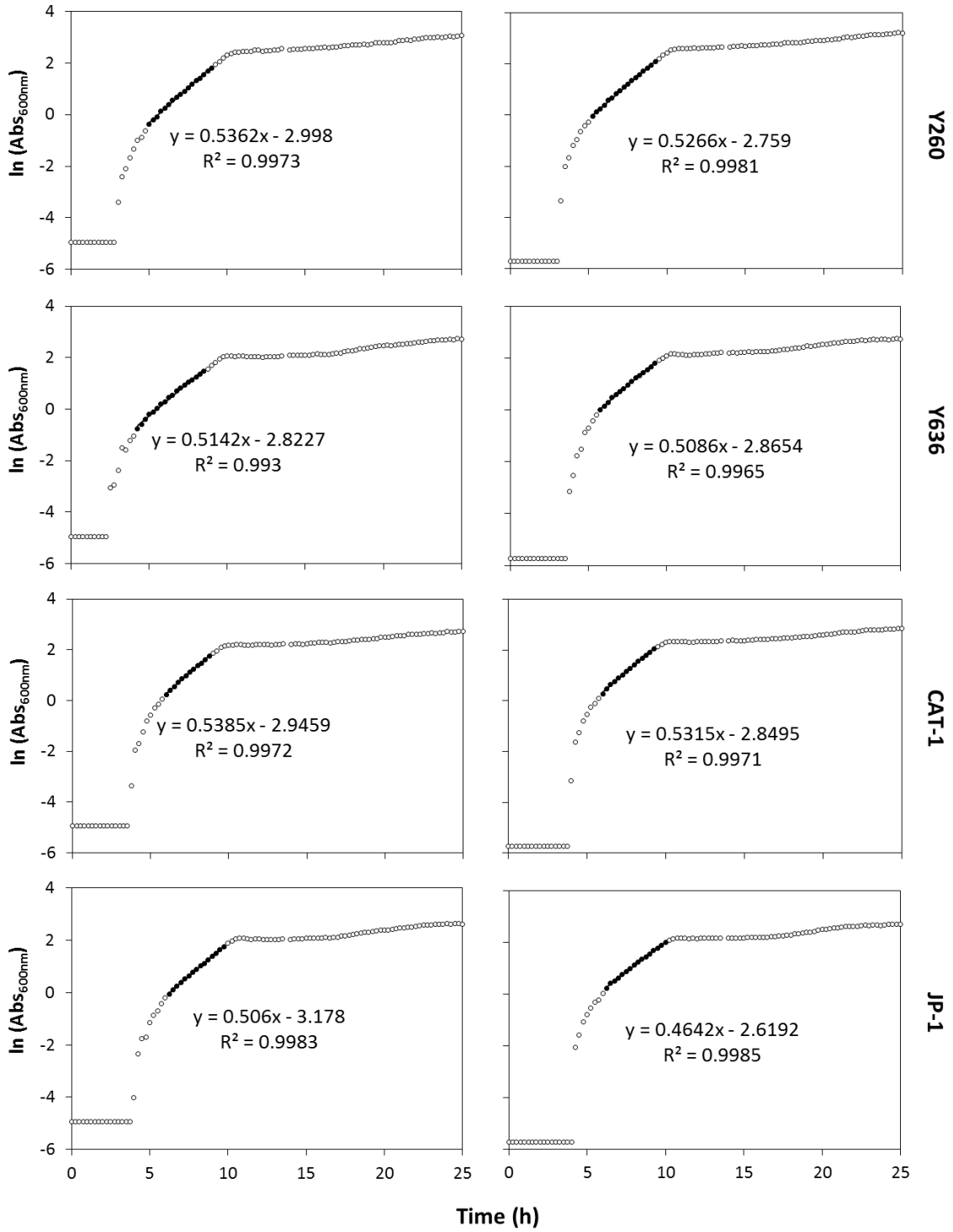

Figure 31 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

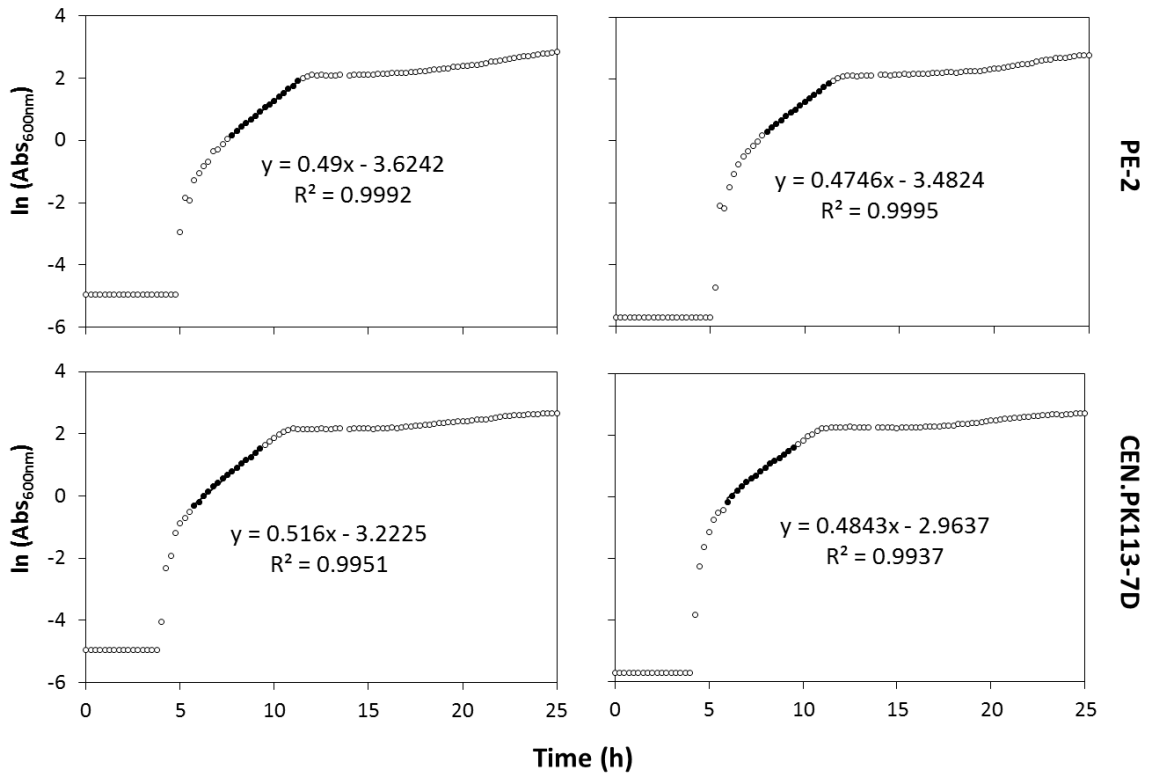

Figure 32 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under standard conditions (pH 5.5, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

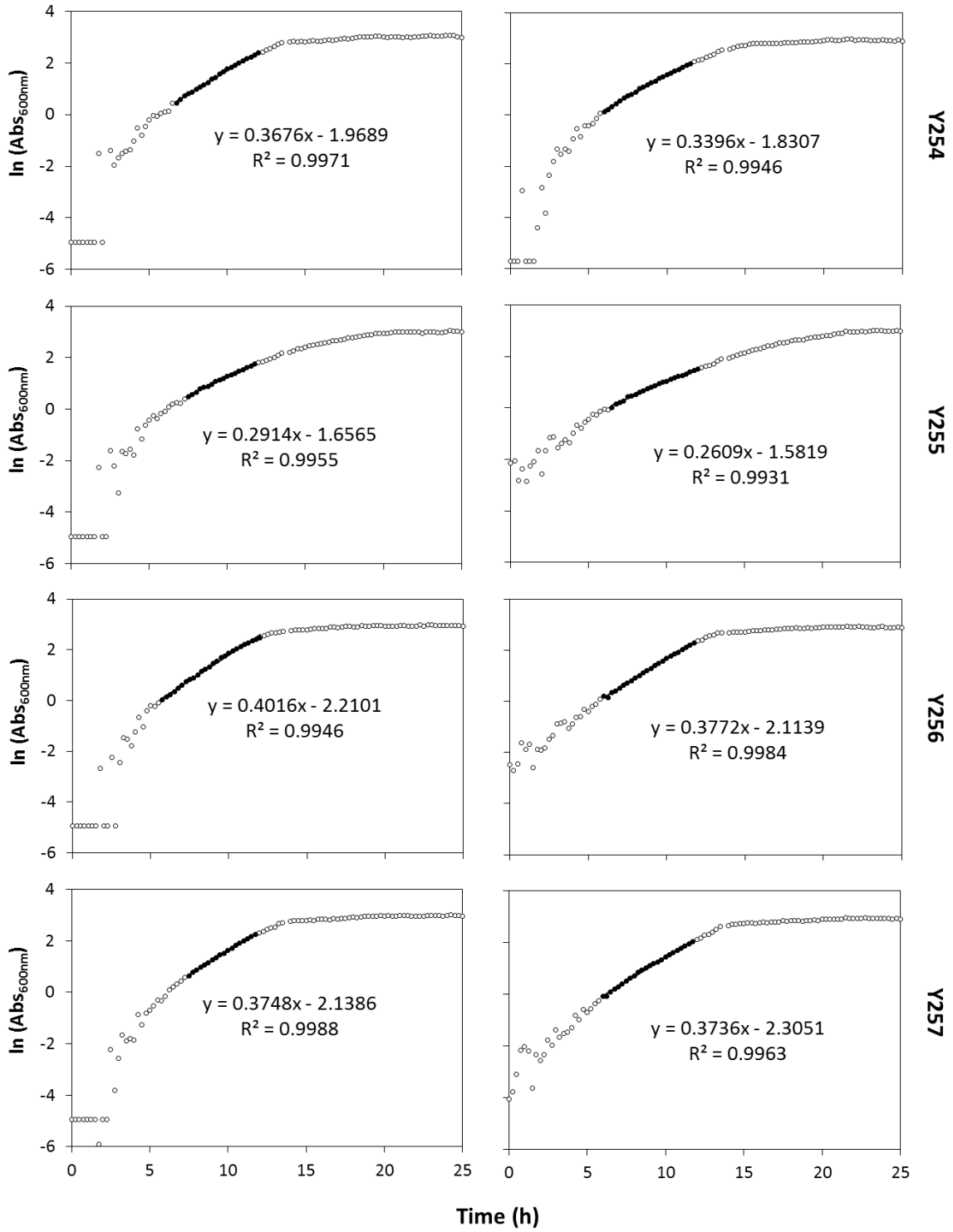

Figure 33 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

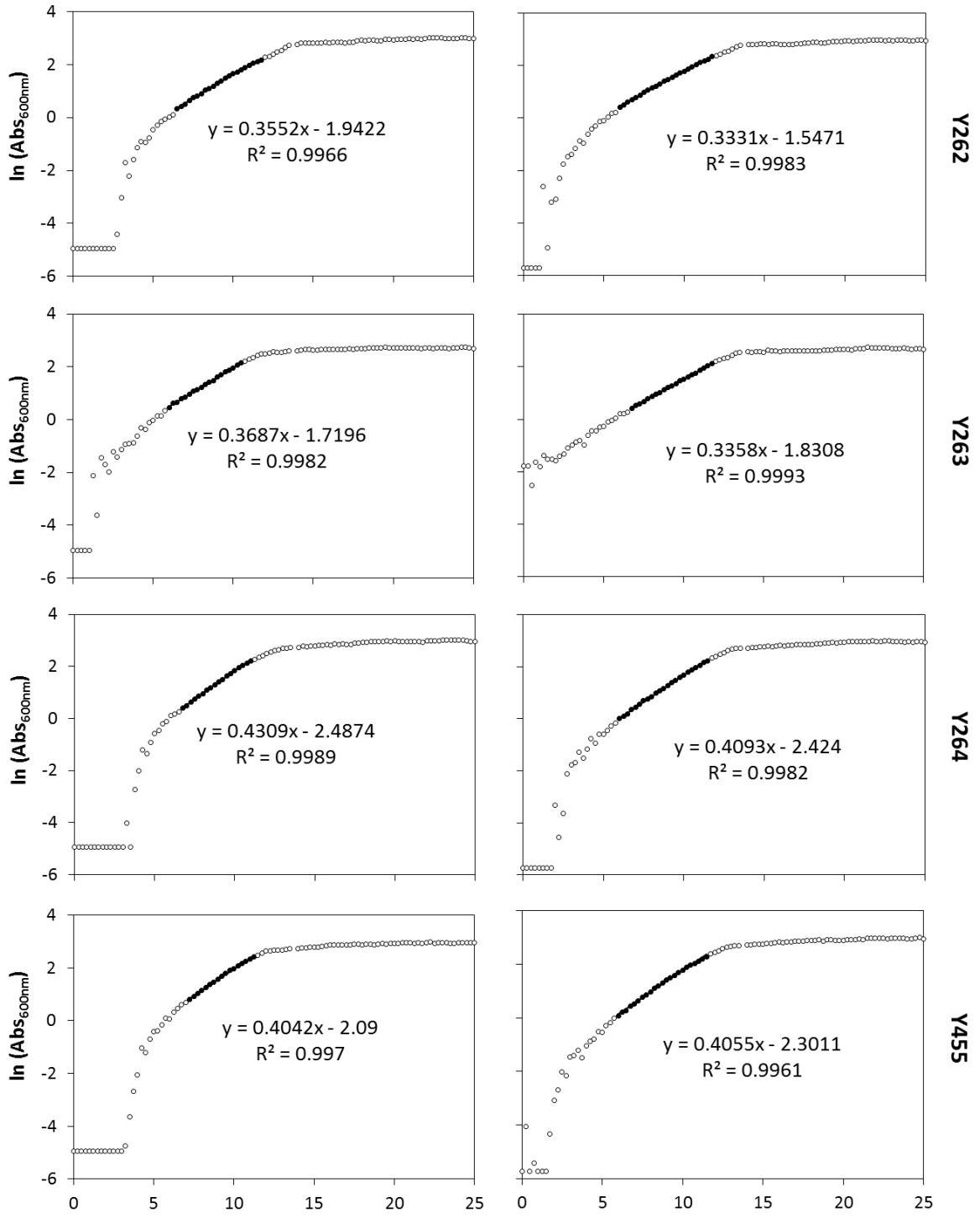

Figure 34 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

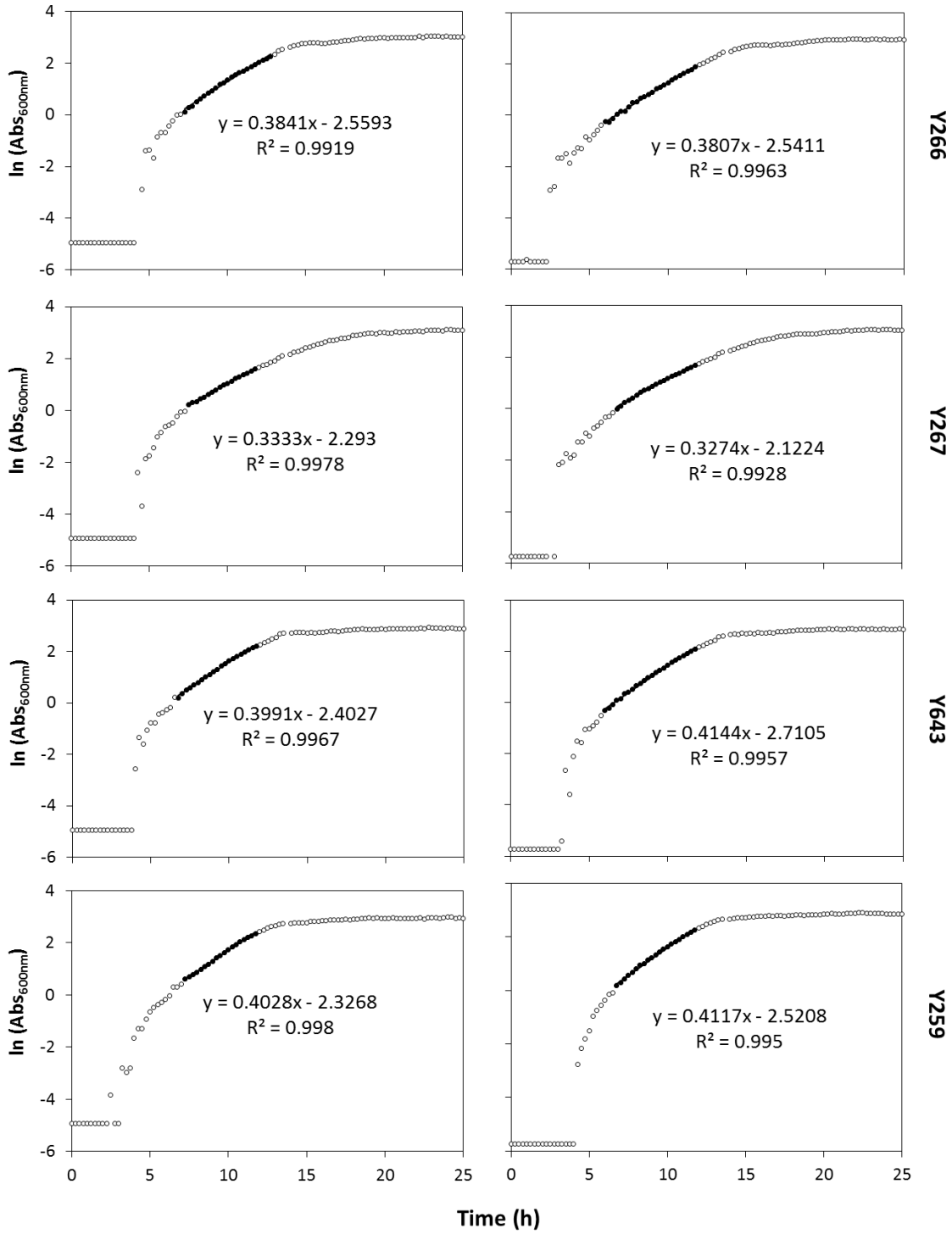

Figure 35 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

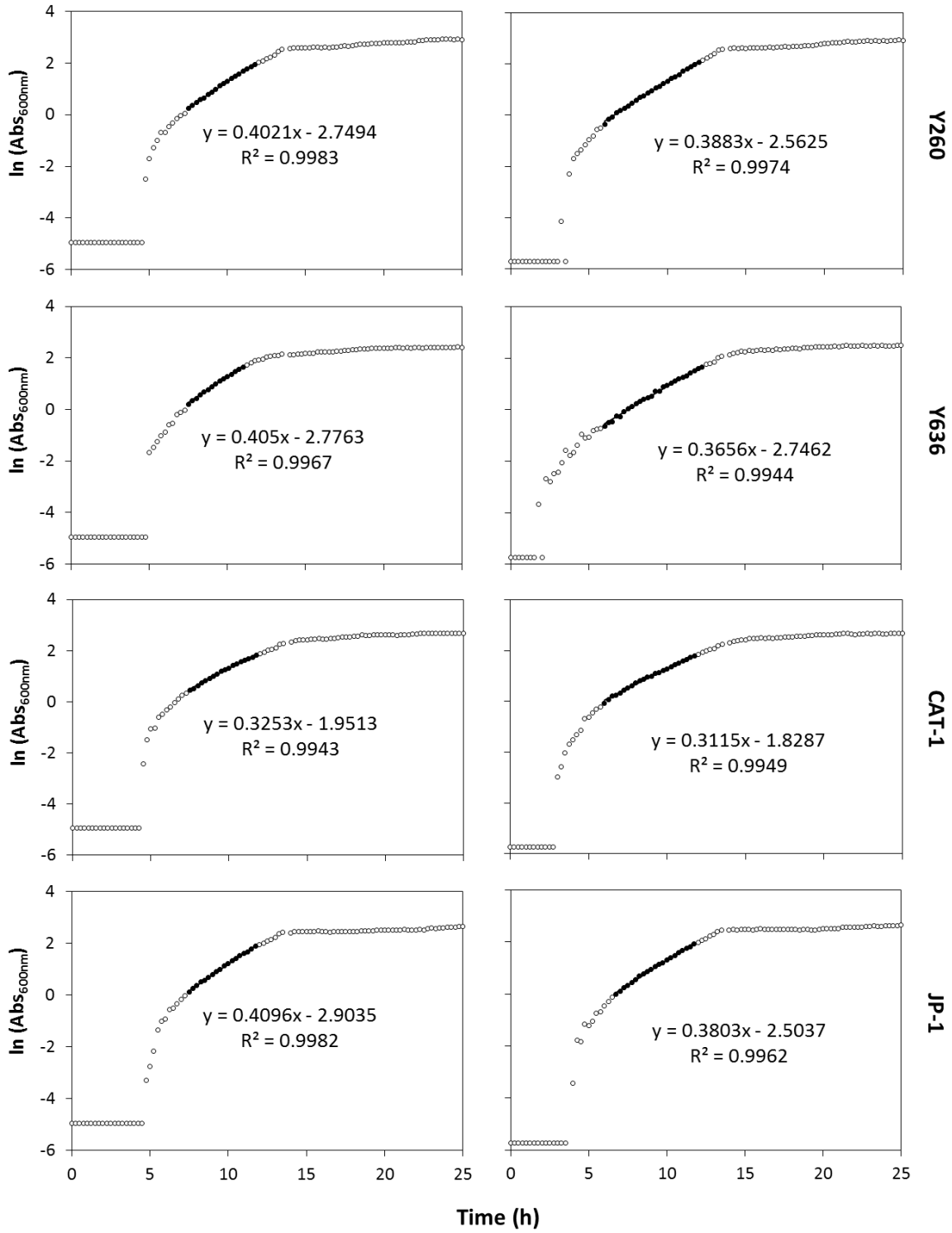

Figure 36 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○). (○).

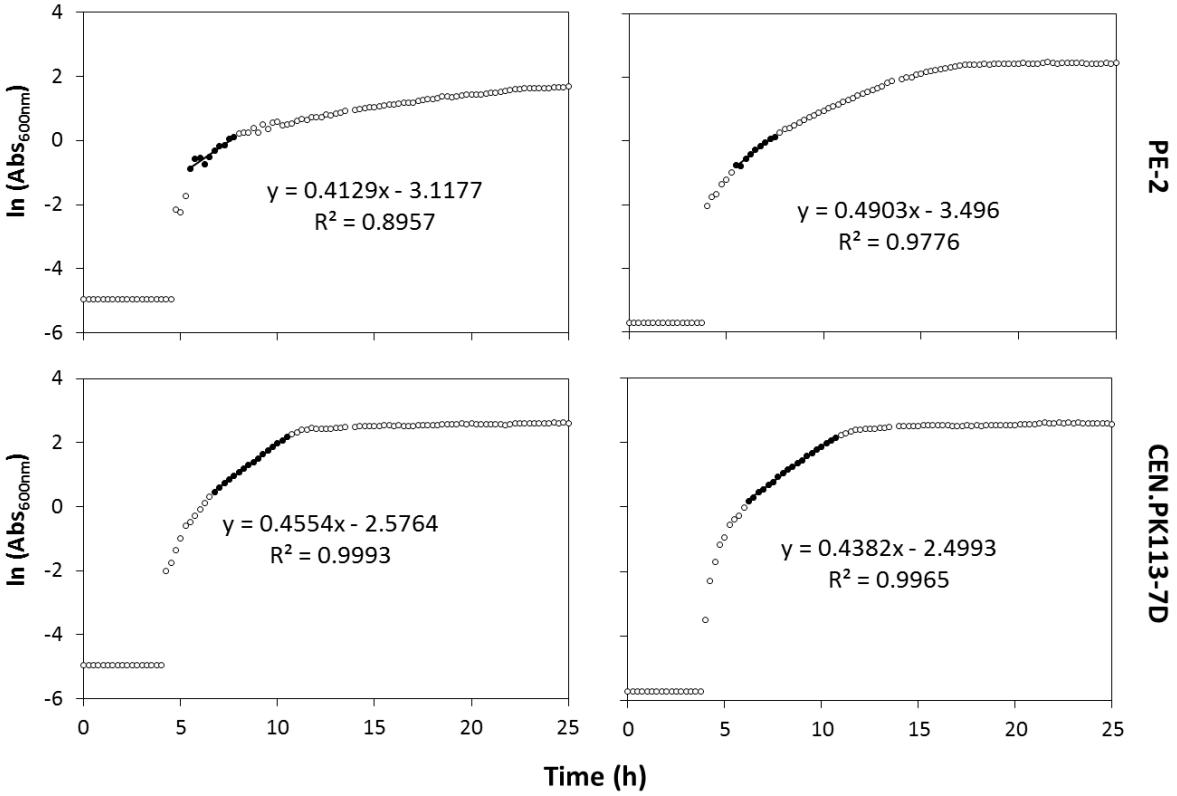

Figure 37 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under low pH stress condition (pH 3.0, 30 °C and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

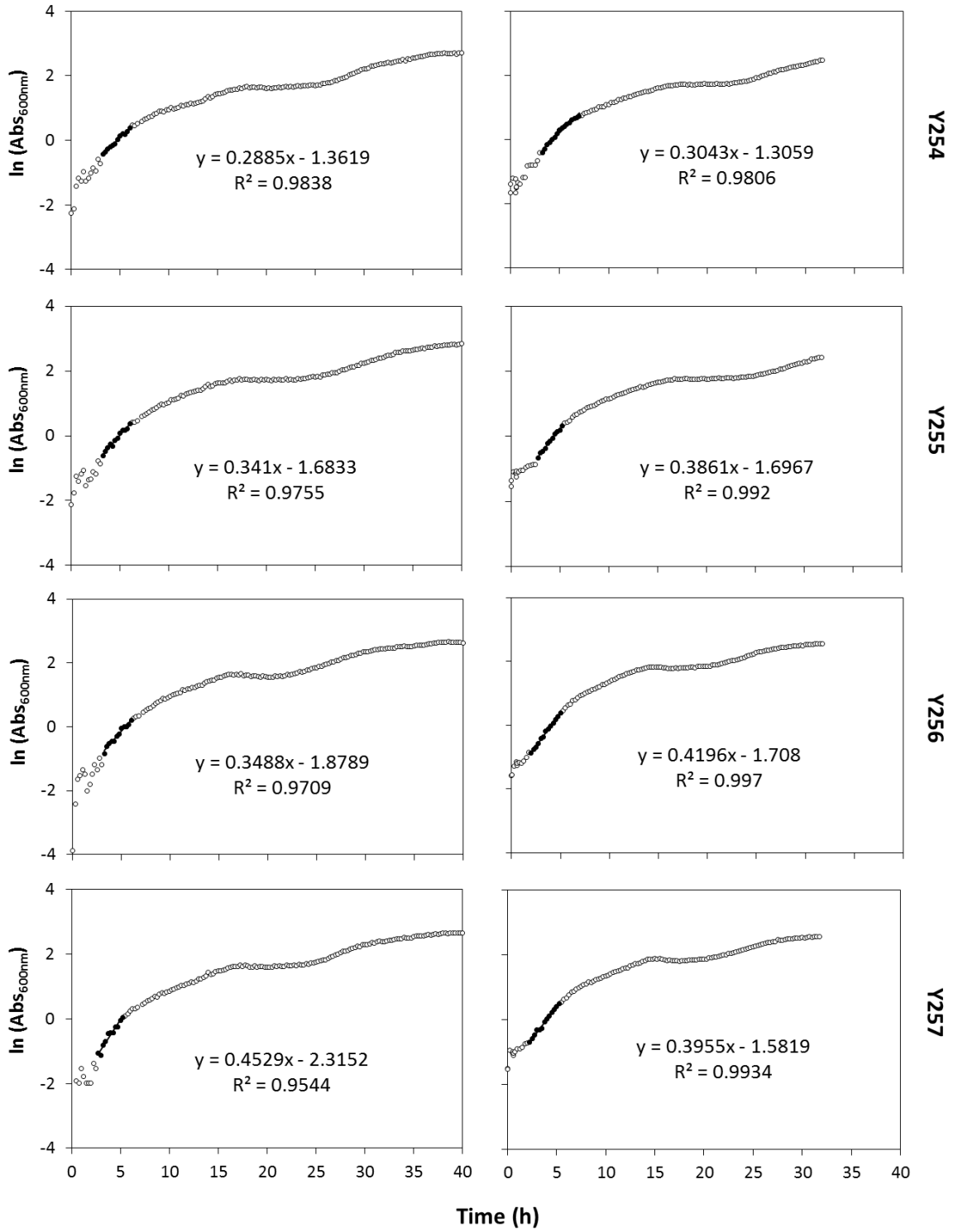

Figure 38 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

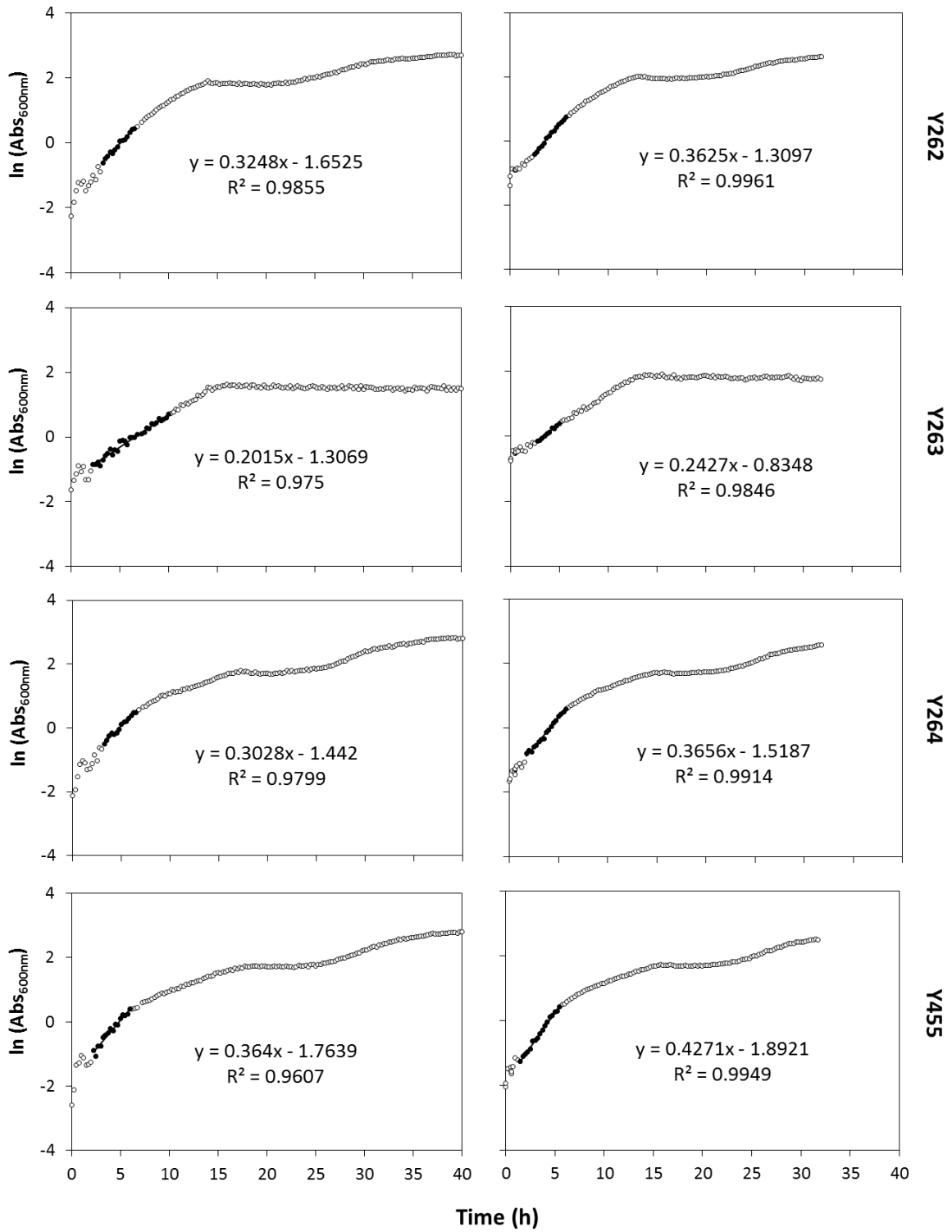

Figure 39 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

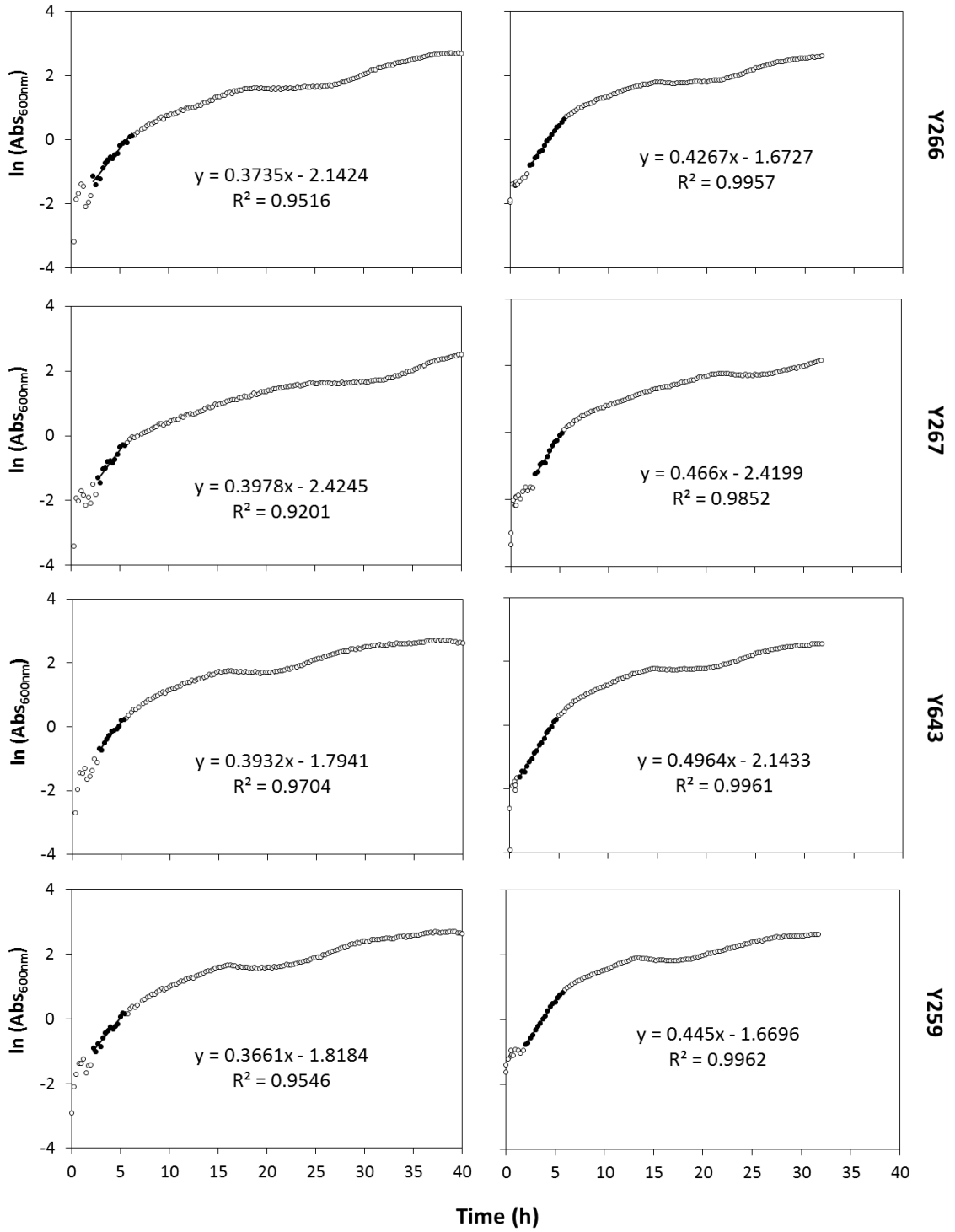

Figure 40 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

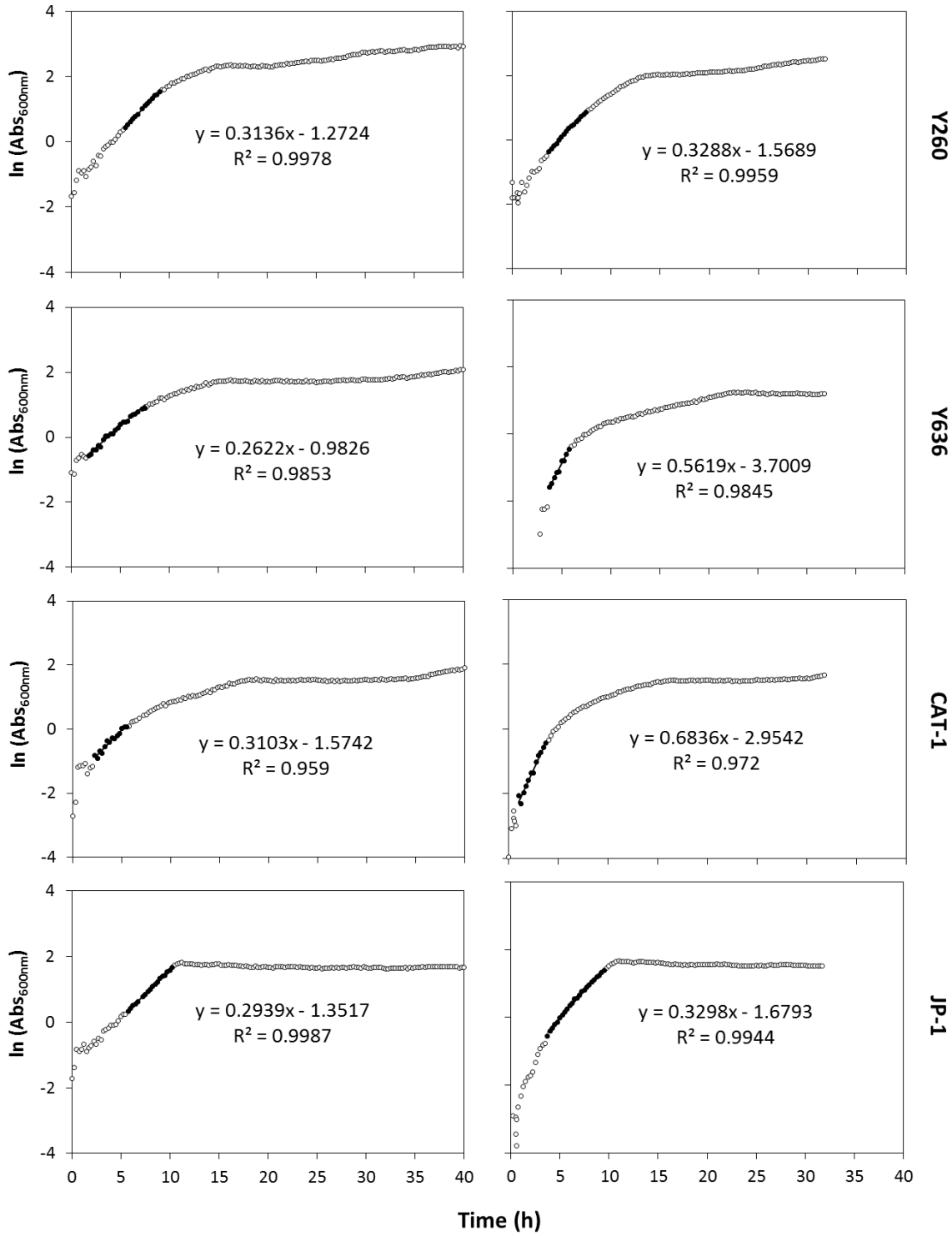

Figure 41 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

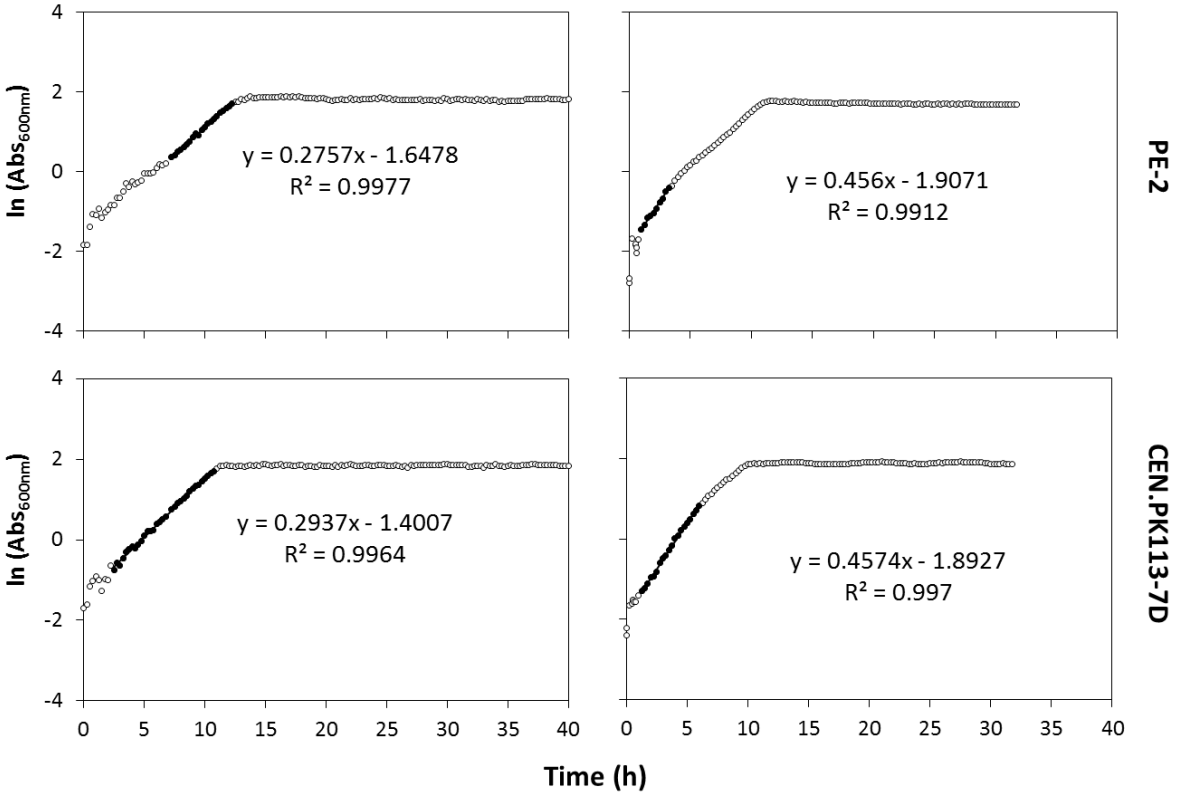

Figure 42 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under acetic acid stress condition (10 g L^{-1} acetic acid, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

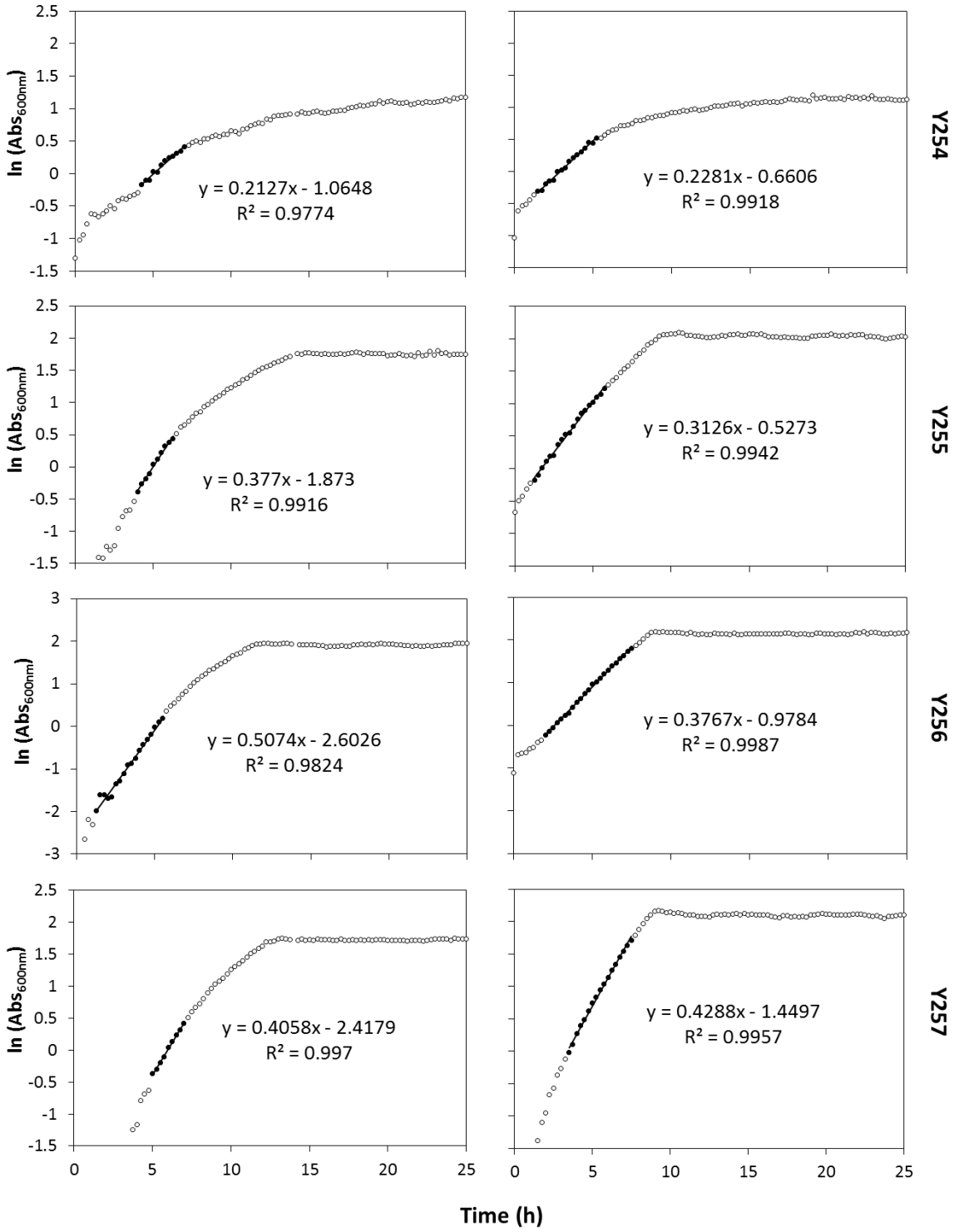

Figure 43 - Growth of *S. cerevisiae* strains Y254, Y255, Y256 and Y257 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

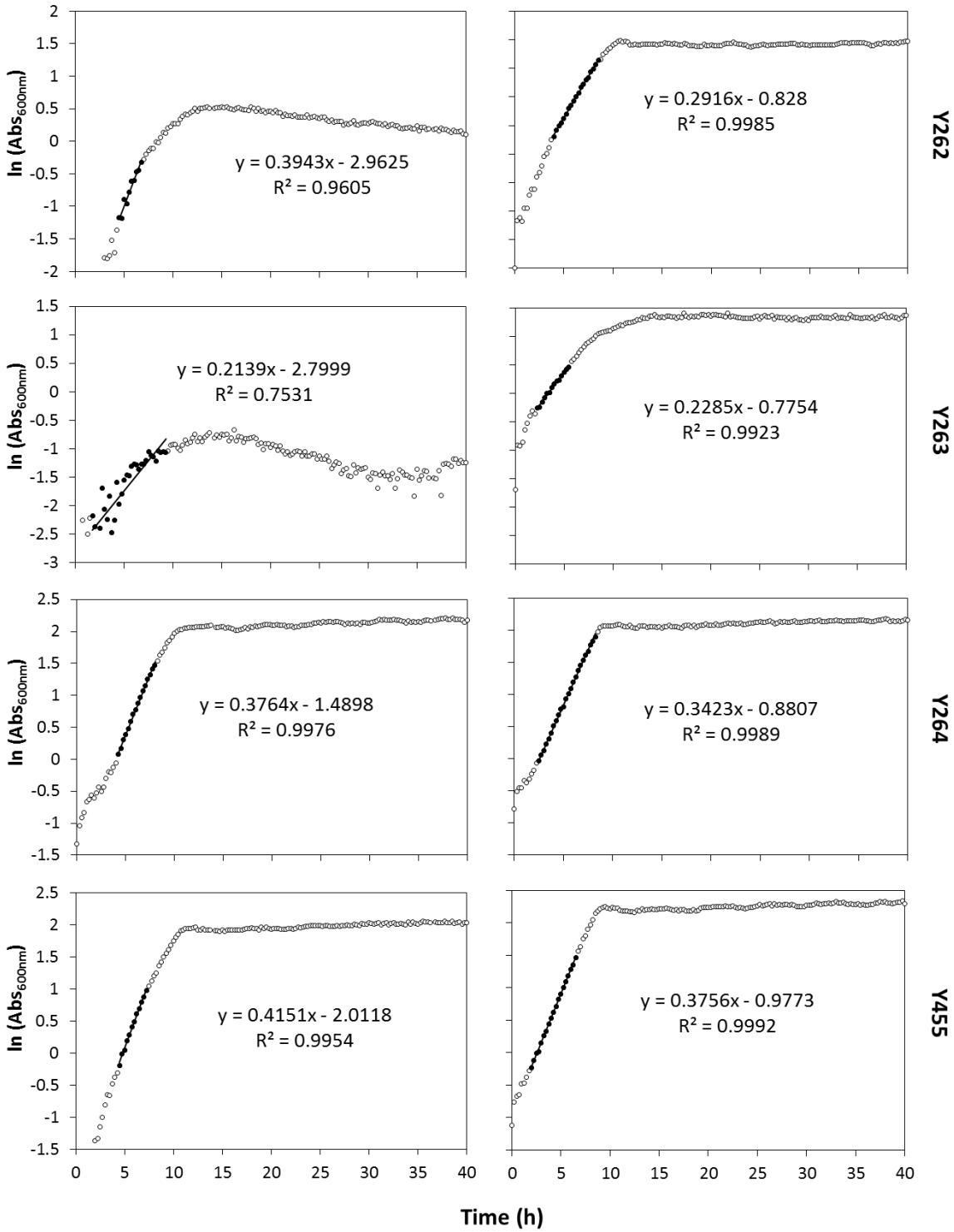

Figure 44 - Growth of *S. cerevisiae* strains Y262, Y263, Y264 and Y455 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

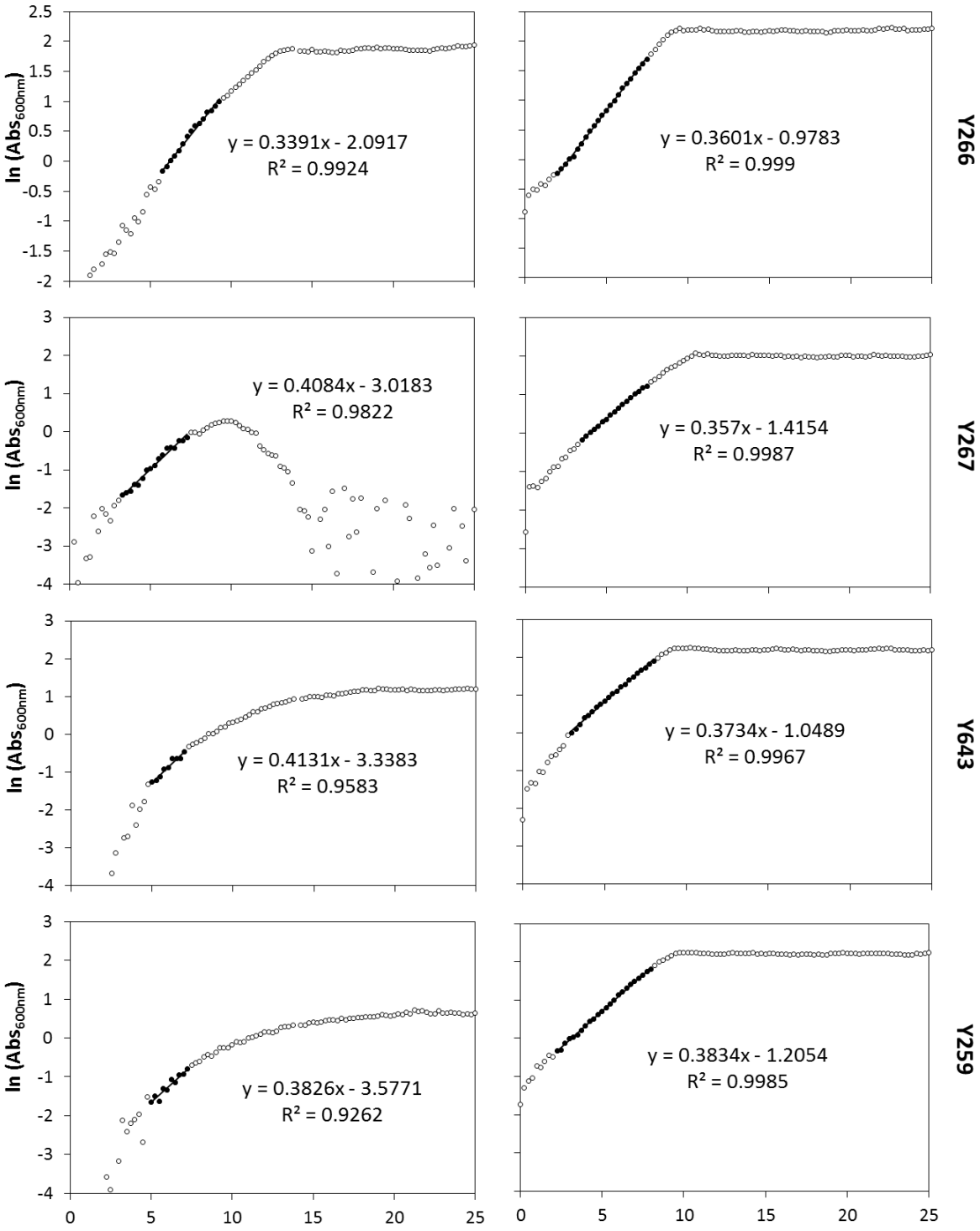

Figure 45 - Growth of *S. cerevisiae* strains Y266, Y267, Y643 and Y259 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

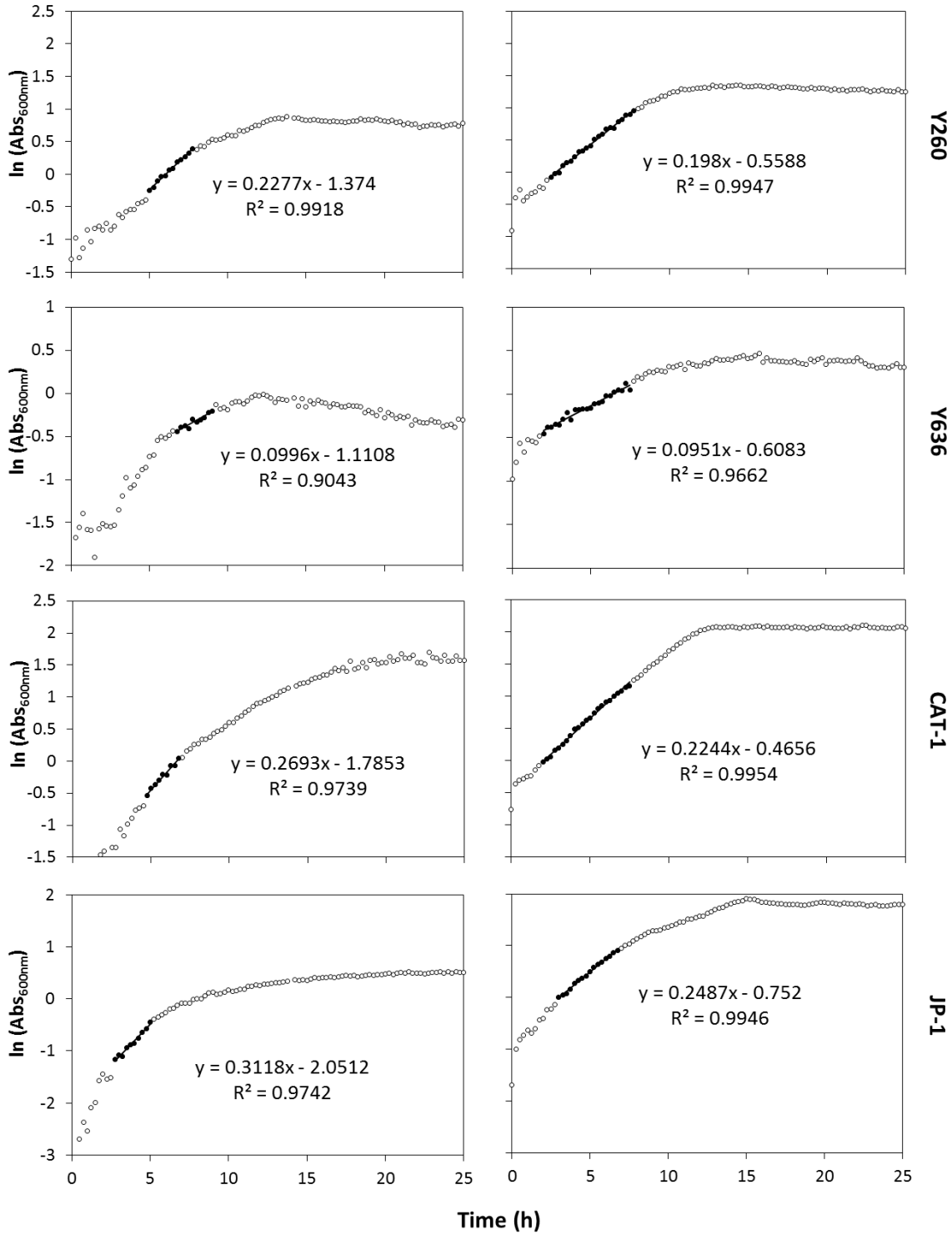

Figure 46 - Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

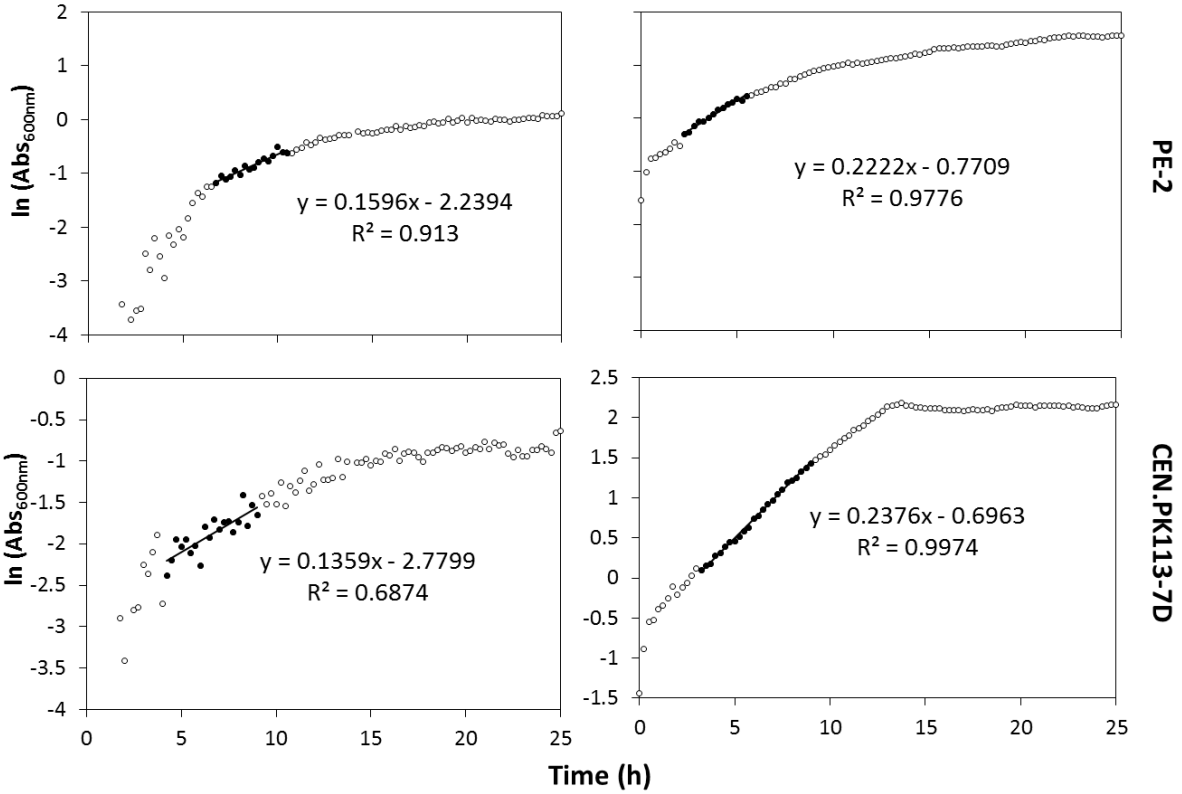

Figure 47 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under high temperature stress condition (40 °C, pH 5.5 and 0 g L^{-1} acetic acid or ethanol) using 20 g L^{-1} glucose as sole carbon/energy source. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

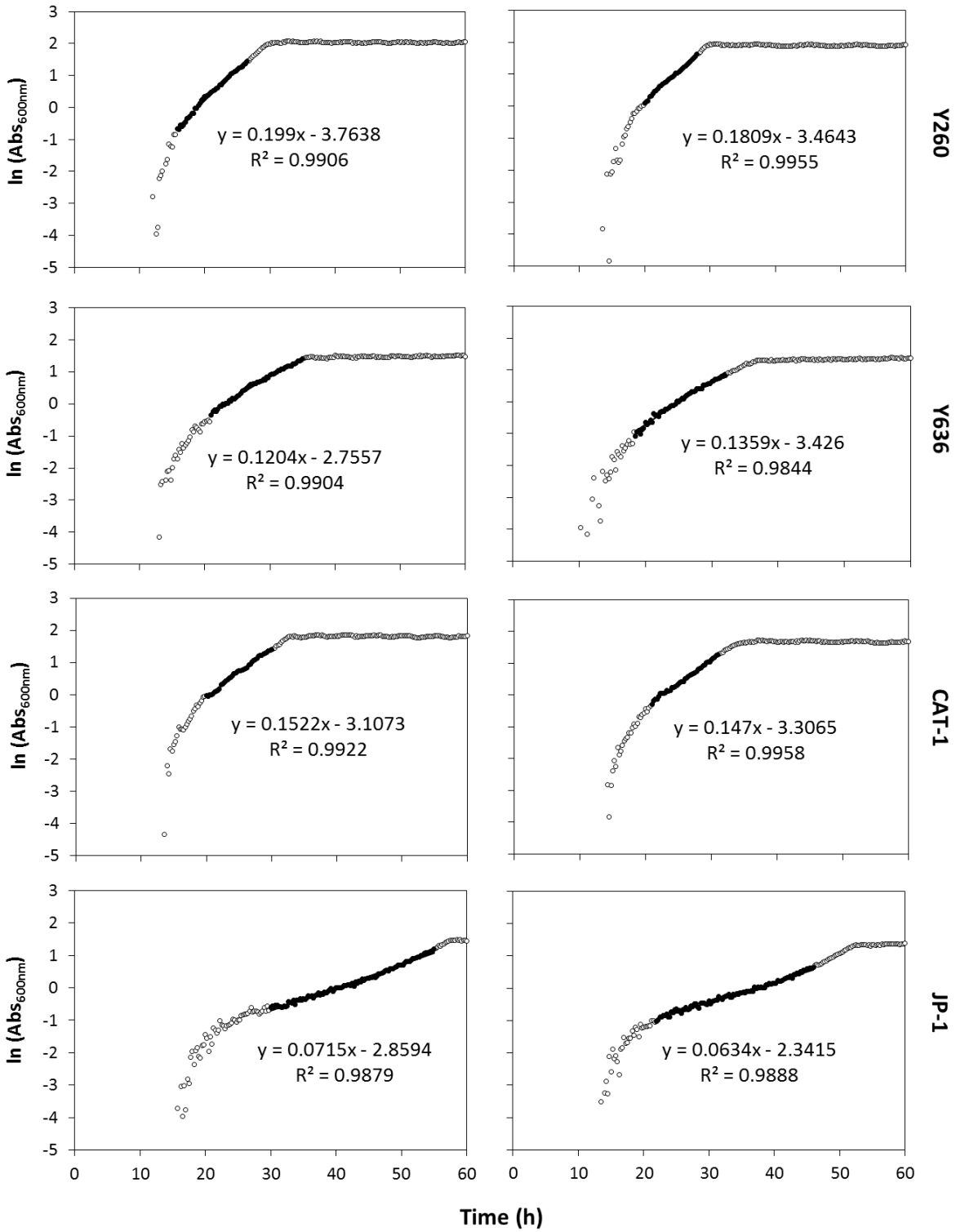

Figure 48 Growth of *S. cerevisiae* strains Y260, Y636, CAT-1 and JP-1 under ethanol stress condition (80 g L^{-1} ethanol, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

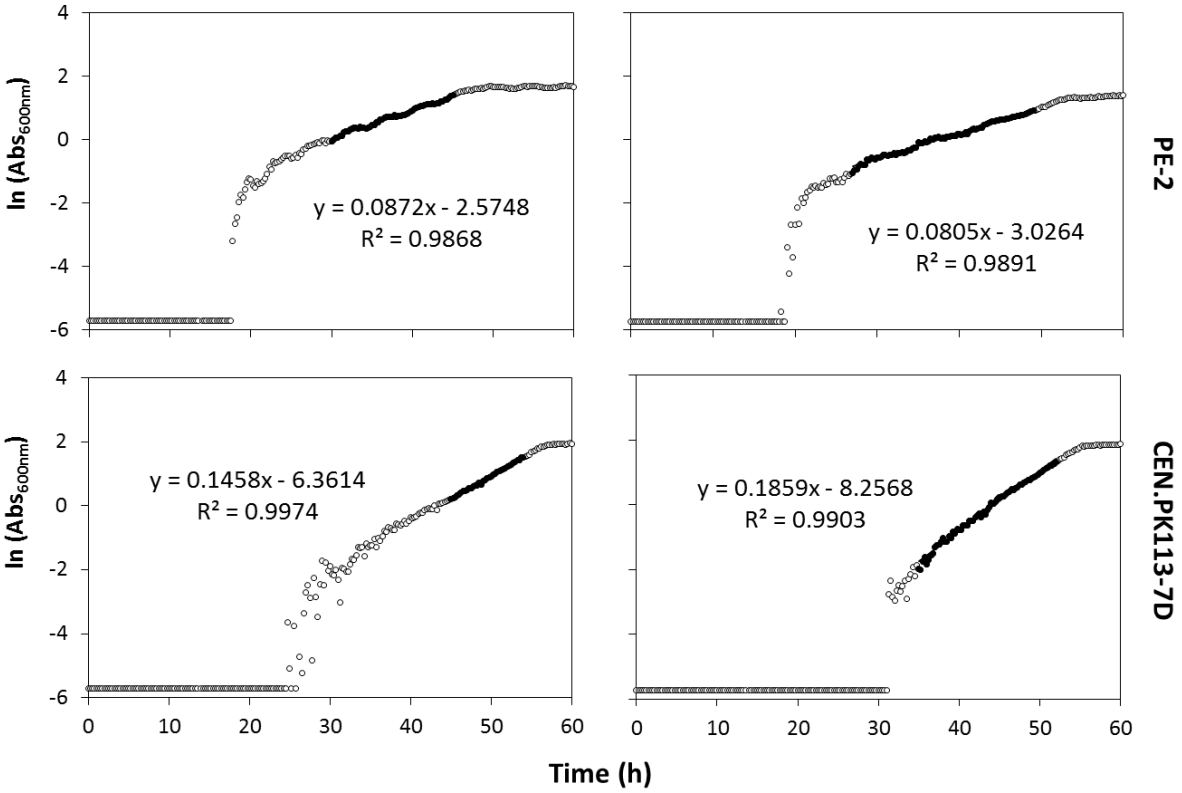

Figure 49 - Growth of *S. cerevisiae* strains PE-2 and CEN.PK113-7D under ethanol stress condition (80 g L^{-1} ethanol, pH 5.5, and 30 °C) using 20 g L^{-1} glucose. Strains were cultivated in duplicates using the automated plate reader Growth Profiler 1152 (Enzyscreen). The maximum specific growth rates (μ_{max}; in h^{-1}) were obtained as the slope of the straight line corresponding to the exponential growth phase (●) from the plot of the natural logarithm of absorbance values against time (○).

A.3 Ethanol yield calculation

Ethanol yield (in terms of moles of ethanol produced divided by moles of hexose equivalents consumed) was obtained as the slope of the straight line from the plot of ethanol concentration against hexose-equivalent concentration **(Figure 50, Figure 51, Figure 52 and Figure 53).**

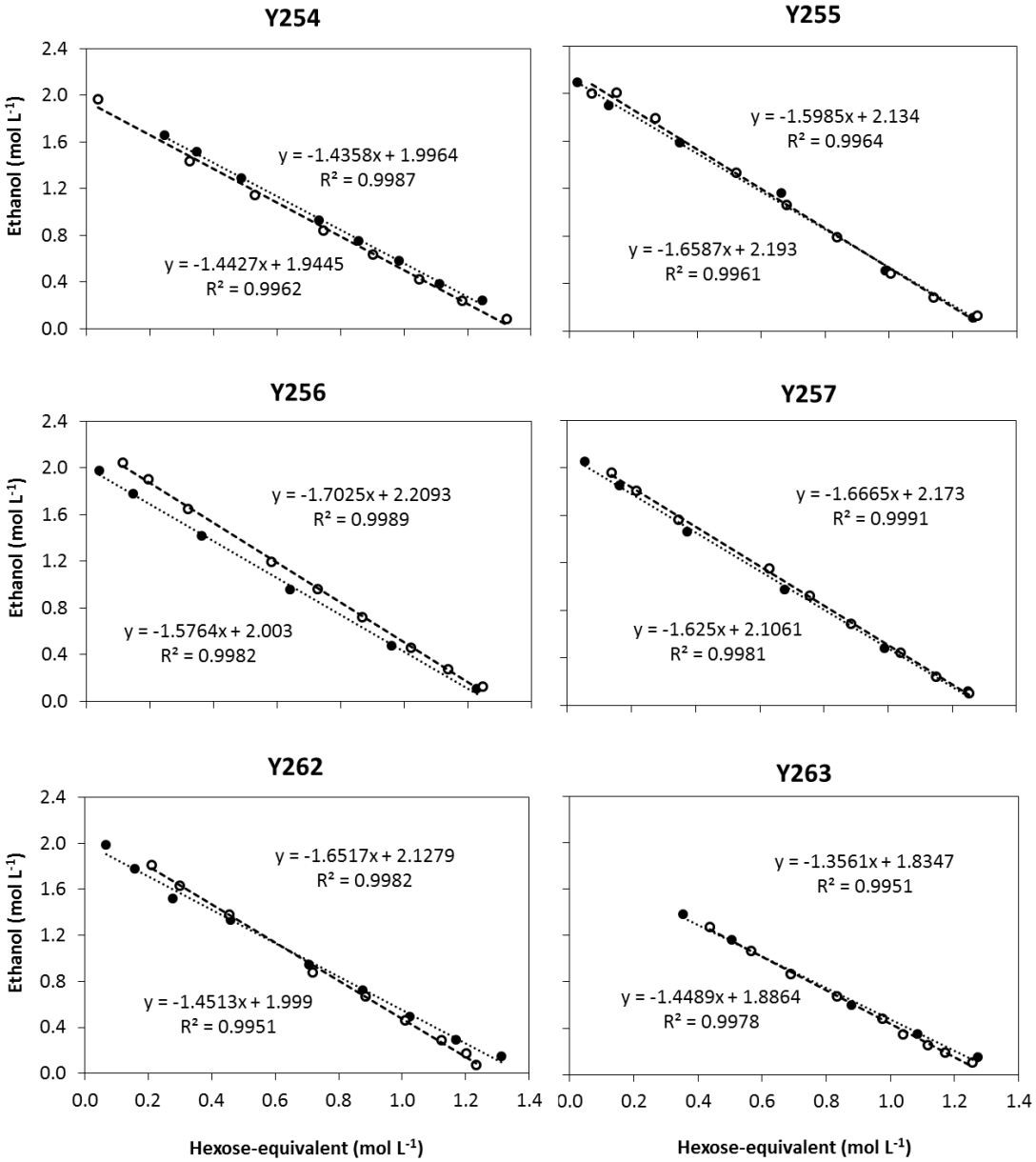

Figure 50 - Calculation of ethanol yield for strains UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y262 and UFMG-CM-Y263, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●).

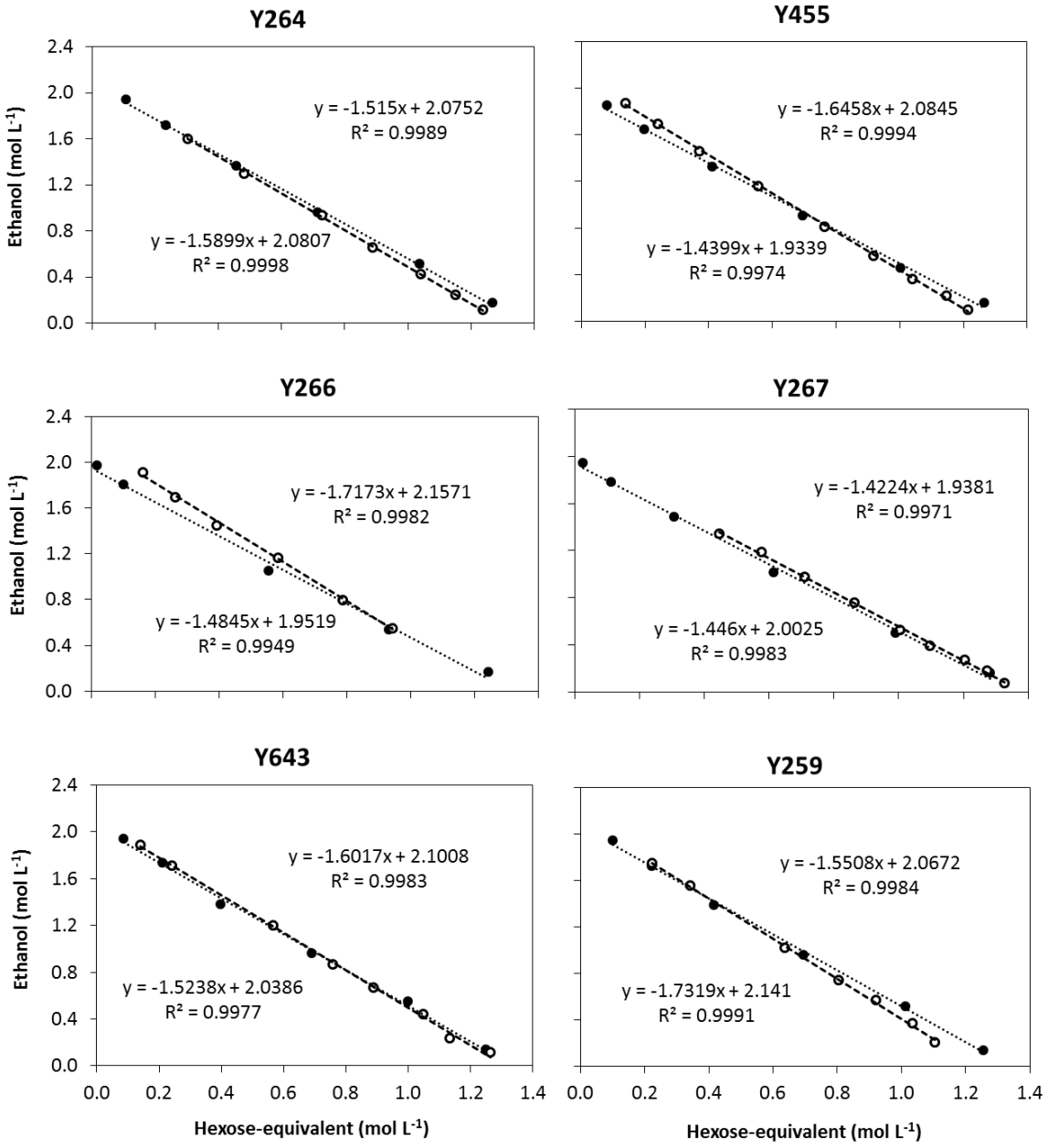

Figure 51 - Calculation of ethanol yield for strains UFMG-CM-Y264, UFMG-CM-Y455, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y643 and UFMG-CM-Y259, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●).

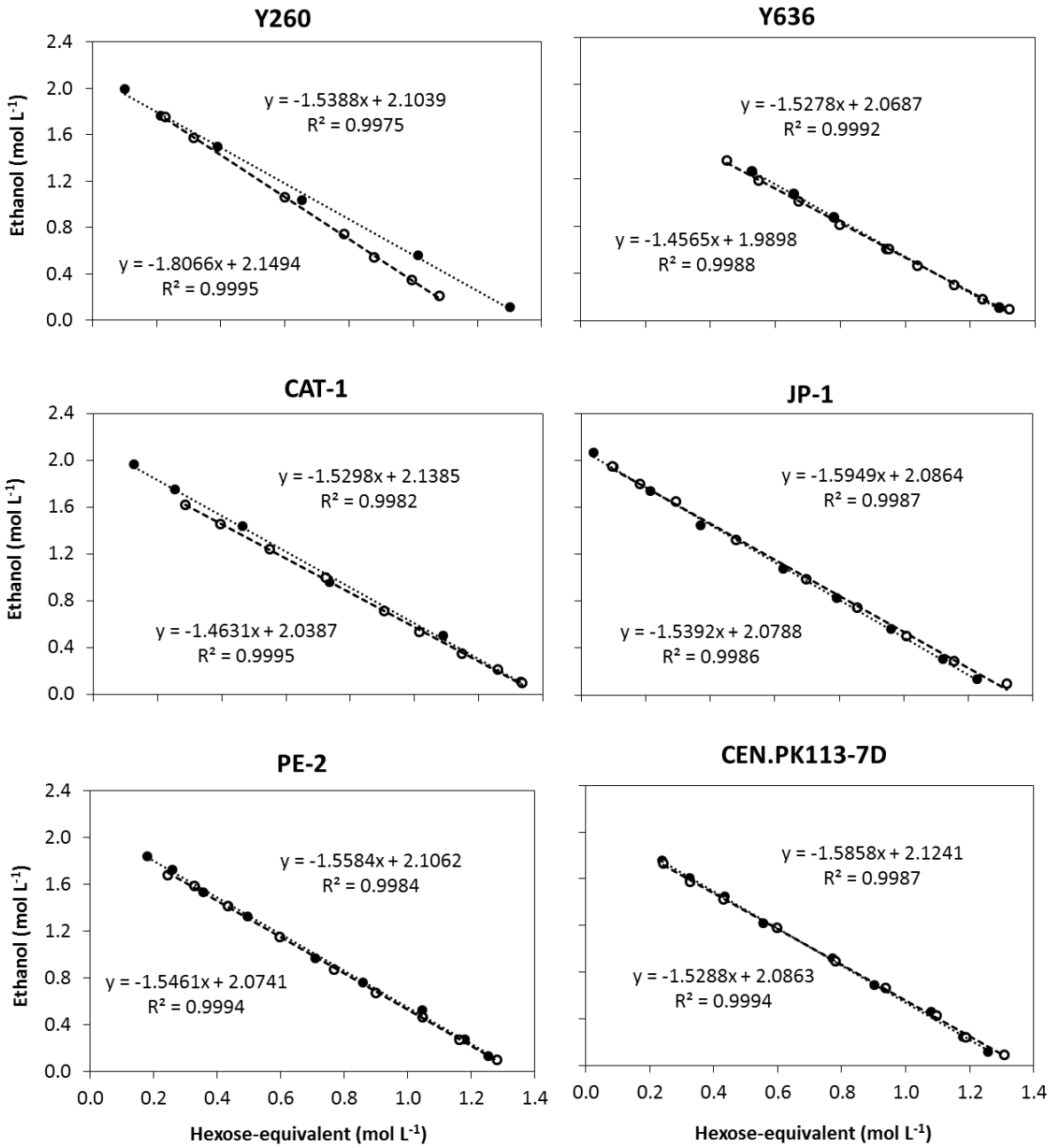

Figure 52 - Calculation of ethanol yield for strains UFMG-CM-Y260, UFMG-CM-Y636, CAT-1, JP-1, PE-2 and CEN.PK113-7D, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●).

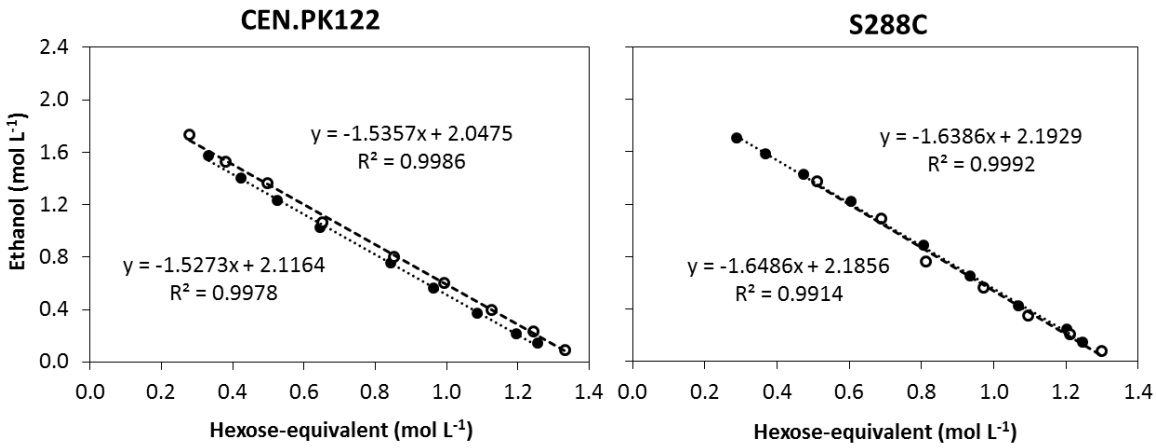

Figure 53 - Calculation of ethanol yield for strains CEN.PK122 and S288C, during cultivations mimicking the 1G fuel ethanol process. Ethanol yield (mol_{ETH} $mol_{Hex.\ Eq}^{-1}$) was obtained as the absolute value of the slope of the straight line from the plot of ethanol concentration (mol L^{-1}) against hexose-equivalent concentration (mol L^{-1}). Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●).

A.4 Relative biomass increase

Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the experimental data at 10 h or 0 h were classified as outliers. **Figure 54, Figure 55, Figure 56 and Figure 57** show the absorbance values obtained for each strain.

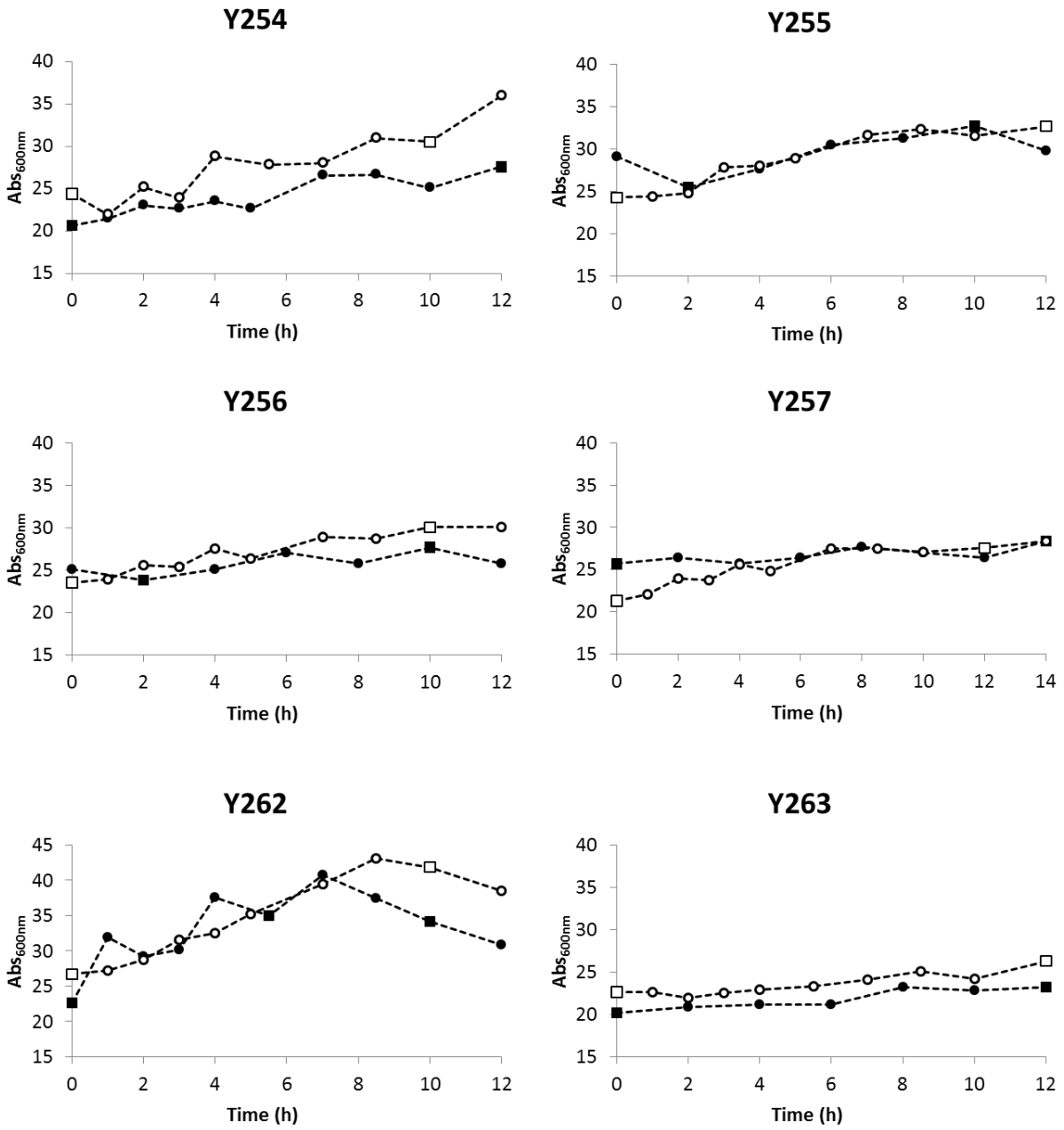

Figure 54 - Growth of strains UFMG-CM-Y254, UFMG-CM-Y255, UFMG-CM-Y256, UFMG-CM-Y257, UFMG-CM-Y262 and UFMG-CM-Y263 under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■).

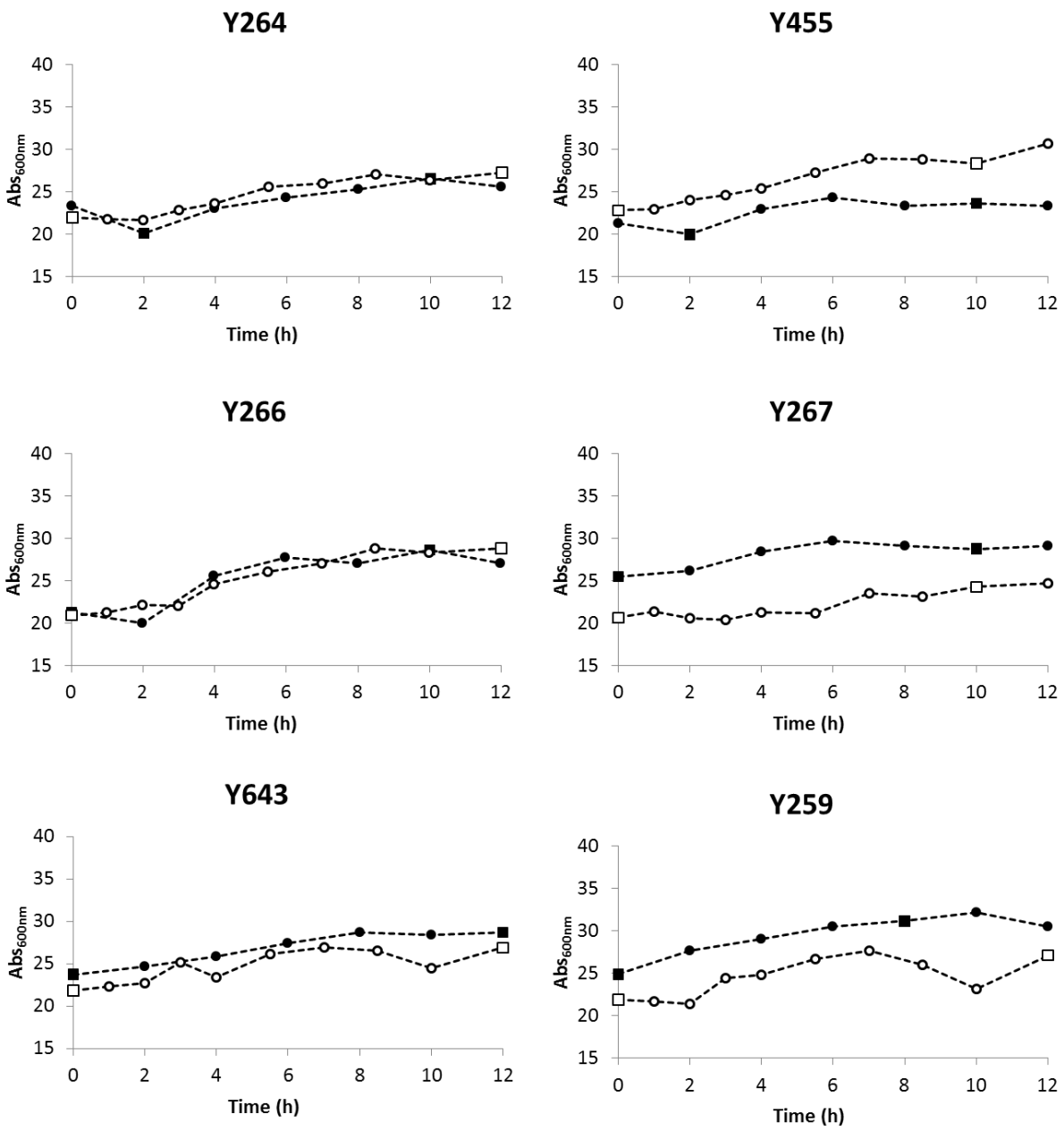

Figure 55 - Growth of strains UFMG-CM-Y264, UFMG-CM-Y455, UFMG-CM-Y266, UFMG-CM-Y267, UFMG-CM-Y643 and UFMG-CM-Y259 under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■).

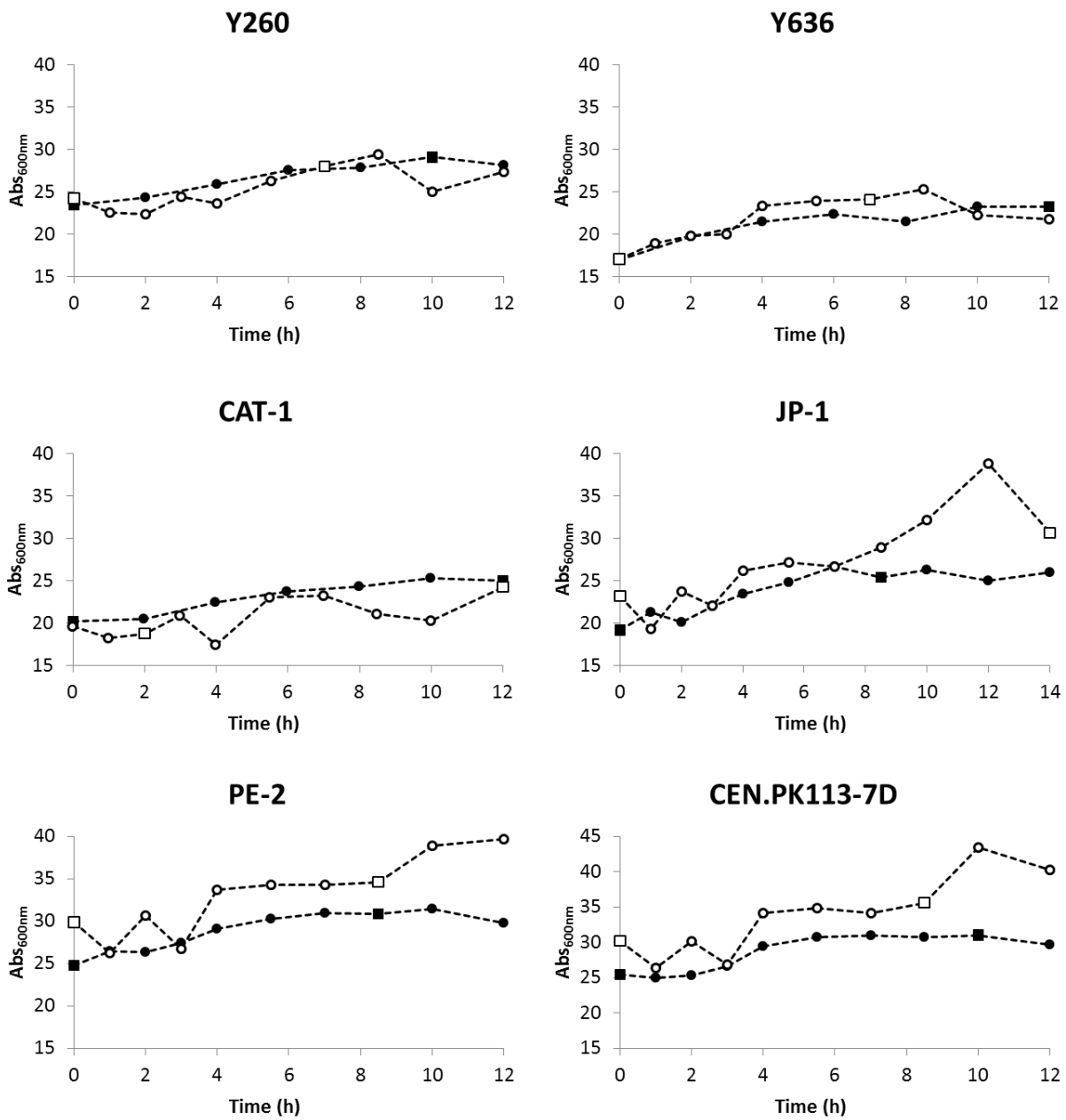

Figure 56 - Growth of strains UFMG-CM-Y260, UFMG-CM-Y636, CAT-1, JP-1, PE-2 and CEN.PK113-7D under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■).

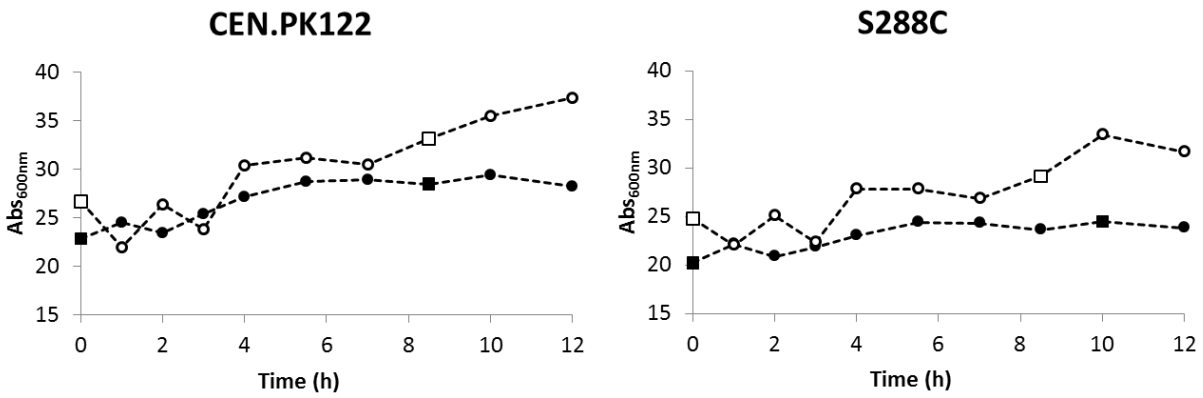

Figure 57 - Growth of strains CEN.PK122 and S288C under conditions mimicking the 1G fuel ethanol process. Abs_{600nm} (measured using a 96-microwell plate) values are shown against cultivation time. Relative biomass increase (ΔX%) was taken as the difference in the absorbance values obtained at 10 h and 0 h, divided by the absorbance value at time 0 h. Different data points were used when the data at 10 h or 0 h were classified as outliers. Data were obtained from two independent culture replicates, represented by the open (○) and solid circles (●). The squares represent the data points chosen for calculation of ΔX (□ and ■).

A.5 Sequences of the D1/D2 domain of the 26S ribosomal DNA

Consensus sequences of the D1/D2 domain of large subunit (26S) ribosomal DNA of *S. cerevisiae* indigenous strains and *Y. lipolytica* strain used in the alignment with *S. cerevisiae* reference strain NRRL Y-12632 (KURTZMAN, ROBNETT, 1998):

>*Y. lipolytica* Po1g

GTCGGTGGCAGTGTGGAGGGTTAGGGGAGAACGCCCGAAAGGCGCTCCCATTTGTAACCCT
CGTCTCGCTATCGATGACTCGGCGTCGGCAGTACACCGCCCACGAGGGGCGGCTGAAACCTCGGCC
ACTCTCCACTCATTTCCTTCCCTATCAACAATTTCACATACTATTTCACTCTCTTTTCAAAGTTCTTTTCA
CCTTTCCTTCACAGTACTTGTTCGCTATCGGTCTCTCACCAGTATTTAGCTTTAGATGGAGTTTACCACC
CACTTTGAGCTGCATTCCCAAACAACTCGACTCTTTGATAAGGCAATACATGGAGAACGGTTAGCCA
GACGGGGTTGTCACCCTCTATGACGTACTATTCCAAGCAACTTGGGTTAGCTTTCTCCAATGCCAAAT
CTTCAAATTACAATCCCGAGGGTTTCAAATTTGAGCTTTTGCCGCTTCACTCGCCGTTACTGAGGCAAT
CCCTGTTGGTTTCTTTTCCTCCGCT

>UFMG-CM-Y254

GTGAAGCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGA
GAGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATC
CCGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAG
CTCTAAGTGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACA
GTGATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGA
AGGGCATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCA
CTGGGCCAGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTA
TAGCCTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAAT
GGTTATATGCCGCC

>UFMG-CM-Y255

GCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAA
CTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGG
CGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTG
GGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGA
AAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTT
GATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCA
GCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGT
GGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATAT
GC

>UFMG-CM-Y256

GCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGG
GGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGA
GTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGT
AAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATG
AAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAG

ACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAG
TTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATA
CTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCCGCC

>UFMG-CM-Y257

AGTGAAGCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGG
AGAGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAAT
CCCGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCA
GCTCTAAGTGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTAC
AGTGATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGG
AAGGGCATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTC
ACTGGGCCAGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATT
ATAGCCTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAA
TGGTTATATGCCGC

>UFMG-CM-Y262

GGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGC
AACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGT
GGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAA
GTGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGAT
GGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGC
ATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGG
CCAGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCC
TGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTA
TATGC

>UFMG-CM-Y263

GCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGG
GGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGA
GTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGT
AAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATG
AAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAG
ACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAG
TTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATA
CTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCCG

>UFMG-CM-Y264

GCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGG
GGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGA
GTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGT
AAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATG
AAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAG
ACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAG
TTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATA
CTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCC

>UFMG-CM-Y455

GCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGG
GGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGA
GTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGT
AAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATG
AAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAG
ACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAG
TTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATA
CTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCCG

>UFMG-CM-Y266

TCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGGGGCCGTTCCTTGTC
TATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTTTGT
AAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAA
GCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAA
AAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGT
GCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAGTTTTGGTGGCAGG
ATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGA
CTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATG

>UFMG-CM-Y267

CTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGGGGCCGTTCCTTGTCT
ATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTTTGT
AAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAA
GCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAA
AAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGT
GCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAGTTTTGGTGGCAGG
ATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGA
CTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCCGCCCGTC

>UFMG-CM-Y643

TGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGGGGCCGTTCCTTGTCTA
TGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTTTGTA
AAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAAG
CTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAAA
AGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGTG
CCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAGTTTTGGTGGCAGGA
TAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGAC
TGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATA

>UFMG-CM-Y259

TGAAGCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAG
AGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCC
CGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCT
CTAAGTGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGT
GATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAG
GGCATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACT

GGGCCAGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATA
GCCTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGG
TTATATG

>UFMG-CM-Y260

GCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGG
GCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCCCGT
GTGGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTA
AGTGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGTGA
TGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGG
CATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGG
GCCAGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGC
CTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTT
ATAT

>UFMG-CM-Y636

AACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTG
CCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAAGCTAAA
TATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAAAAGAG
AGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGTGCCCTC
TGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAGTTTTGGTGGCAGGATAAAT
CCATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGACTGAGG
ACTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCCGC

>CEN.PK113-7D

TGAAGCGGCAAAAGCTCAAATTTGAAATCTGGTACCTTCGGTGCCCGAGTTGTAATTTGGAG
AGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCCTTGGAACAGGACGTCATAGAGGGTGAGAATCC
CGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTGCCTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCT
CTAAGTGGGTGGTAAATTCCATCTAAAGCTAAATATTGGCGAGAGACCGATAGCGAACAAGTACAGT
GATGGAAAGATGAAAAGAACTTTGAAAAGAGAGTGAAAAAGTACGTGAAATTGTTGAAAGGGAAG
GGCATTTGATCAGACATGGTGTTTTGTGCCCTCTGCTCCTTGTGGGTAGGGGAATCTCGCATTTCACT
GGGCCAGCATCAGTTTTGGTGGCAGGATAAATCCATAGGAATGTAGCTTGCCTCGGTAAGTATTATA
GCCTGTGGGAATACTGCCAGCTGGGACTGAGGACTGCGACGTAAGTCAAGGATGCTGGCATAATGG
TTATATG

>NRRL Y-12632 (*S. cerevisiae* reference strain)

AAACCAACCGGGATTGCCTTAGTAACGGCGAGTGAAGCGGCAAAAGCTCAAATTTGAAATC
TGGTACCTTCGGTGCCCGAGTTGTAATTTGGAGAGGGCAACTTTGGGGCCGTTCCTTGTCTATGTTCC
TTGGAACAGGACGTCATAGAGGGTGAGAATCCCGTGTGGCGAGGAGTGCGGTTCTTTGTAAAGTGC
CTTCGAAGAGTCGAGTTGTTTGGGAATGCAGCTCTAAGTGGGTGGTAAATTCCATCTAAAGCTAAAT
ATTGGCGAGAGACCGATAGCGAACAAGTACAGTGATGGAAAGATGAAAAGAACTTTGAAAAGAGA
GTGAAAAAGTACGTGAAATTGTTGAAAGGGAAGGGCATTTGATCAGACATGGTGTTTTGTGCCCTCT
GCTCCTTGTGGGTAGGGGAATCTCGCATTTCACTGGGCCAGCATCAGTTTTGGTGGCAGGATAAATC
CATAGGAATGTAGCTTGCCTCGGTAAGTATTATAGCCTGTGGGAATACTGCCAGCTGGGACTGAGGA
CTGCGACGTAAGTCAAGGATGCTGGCATAATGGTTATATGCCGCCCGTCTTGAAACACGGACCAAGG
AGTCTAACGTCTATGCGAGTGTTTGGGTGTAAAACCCATACGTGTAATGAAAGTGAACGTAGGTTGG
GGCCTCGCAAGAGGTGCACAATCGACCGATCCTGATGTCTTCGGATGGATTTGAGTAAGAGCATAGC
TGTTGGGACCCGAAAGATGGTGAACTATGCCTGAATAGGGTGAAGCCAGAGGAAACTCTGGTGGA

GGCTCGTAGCGGTTCTGACGTGCAAATCGATCGTCGAATTTGGGTATAGGGGCGAAAGACTAATCG
AACCATCTAGTAGCTGGTTCCTGCCGAAGTTTCCCTCAGGATAGCAGAAGCTCGTATCAGTTTTATGA
GGTAAAGCGAATGATTAGAGGTTCCGGGGTCGAAATGACCTTGACCTATTCTCAAACTTTAAATATG
TAAGAAGTCCTTGTTACTTAATTGAACGTGGACATTTGAATGAAGAGCTTTTAGTGGGCCATTTTTGG
TAAGCAGAACTGGCGATGCGGGATGAACCGAACGTAGAGTTAAGGTGCCGGAATACACGCTCATCA
GACACCACAAAAGGTGTTAGTTCATCTAGACAGCCGGACGGTGGCCATGGAAGTCGGAATCCGCTA
AGGAGTGTGTAACAACTCACCGGCCGAATGAACTAGCCCTGAAAATGGATGGCGCTCAAGCGTGTT
ACCTATACTCTACCGTCAGGGTTGATATGATGCCCTGACGAGTAGGCAGGCGTGGAGGTCAGTGAC
GAAGCCTAGACCGTAAGGTCGGGTCGAACGGCCTCTAGTGCAGATCTTGGTGGTAGTAGCAAATAT
TCAAATGAGAACTTTGAAGACTGAAGTGGGGAAAGGTTCCACGTCAACAGCAGTTGGACGTGGGTT
AGTCGATCCTAAGAGATGGGGAAGCTCCGTTTCAAAGGCCTGATTTTATGCAGGCCACCATCGAAAG
GGAATCCGGTTAAGATTCCGGAACCTGGATATGGATTCTTCACGGTAACGTAACTGAATGTGGAGAC
GTCGGCGCGAGCCCTGGGAGGAGTTATCTTTTCTTCTTAACAGCTTATCACCCCGGAATTGGTTTATC
CGGAGATGGGGTCTTATGGCTGGAAGAGGCCAGCACCTTTGCTGGCTCCGGTGCGCTTGTGACGGC
CCGTGAAAATCCACAGGAAGGAATAGTTTTCATGCCAGGTCGTACTGATAACCGCAGCAGGTCTCCA
AGGTGAACAGCCTCTAGTTGATAGAATAATGTAGATAAGGGAAGTCGGCAAAATAGATCCGTAACTT
CGGGATAAGGATTGGCTCTAAGGGTCGGGTAGTGAGGGCCTTGGTCAGACGCAGCGGGCGTGCTT
GTGGACTGCTTGGTGGGGCTTGCTCTGCTAGGCGGACTACTTGCGTGCCTTGTTGTAGACGGTCTTG
GTAGGTCTCTTGTAGACCGTCGCTTGCTACAATTAACGATCAACTTAGAACTGGTACGGACAAGGGG
AATCTGACTGTCTAATTAAAACATAGCATTGCGATGGTCAGAAAGTGATGTTGACGCAATGTGATTTC
TGCCCAGTGCTCTGAATGTCAAAGTGAAGAAATTCAACCAAGCGCGGGTAAACGGCGGGAGTAACT
ATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGATTAACGAG
ATTCCCACTGTCCCTATCTACTATCTAGCGAAACCACAGCCAAGGGAACGGGCTTGGCAGAATCAGC
GGGGAAAGAAGACCCTGTTGAGCTTGACTCTAGTTTGACATTGTGAAGAGACATAGAGGGTGTAGA
ATAAGTGGGAGCTTCGGCGCCAGTGAAATACCACTACCTTTATAGTTTCTTTACTTATTCAATGAAGC
GGAGCTGGAATTCATTTTCCACGTTCTAGCATTCAAGGTCCCATTCGGGGCTGATCCGGGTTGAAGA
CATTGTCAGGTGGGGAGTTTGGCTGGGGCGGCACATCTGTTAAACGATAACGCAGATGTCCTAAGG
GGGGCTCATGGAGAACAGAAATCTCCAGTAGAACAAAAGGGTAAAAGCCCCCTTGATTTTGATTTTC
AGTGTGAATACAAACCATGAAAGTGTGGCCTATCGATCCTTTAGTCCCTCGGAATTTGAGGCTAGAG
GTGCCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCAGTCAAGCGTTCATAGCGACATTGCTTT
TTGATTCTTCGATGTCGGCTCTTCCTATCATACCGAAGCAGAATTCGGTAAGCGTTGGATTGTTCACC
CACTAATAGGGAACGTGAGCTGGGTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAA
TGTTACCGCAATAGTAATTGAACTTAGTACGAGAGGAACAGTTCATTCGGATAATTGGTTTTTGCGGC
TGTCTGATCAGGCATTGCCGCGAAGCTACCATCCGCTGGATTATGGCTGAACGCCTCTAAGTCAGAA
TCCATGCTAGAACGCGGTGATTTCTTTGCTCCACACAATATAGATGGATACGAATAAGGCGTCCTTGT
GGCGTCGCTGAACCATAGCAGGCTAGCAACGGTGCACTTGGCGGAAAGGCCTTGGGTGCTTGCTGG
CGAATTGCAATGTCATTTTGCGTGGGGATAAATCATTTGTATACGACTTAGATGTACAACGGGGTATT
GTAAGCAGTAGAGTAGCCTTGTTGTTACGATCA