**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Avaliação dos efeitos causados por contaminações bacterianas em
fermentações etanólicas de primeira geração utilizando
metodologias multi-ômicas**

Lucas Souza Lopes

Tese apresentada para obtenção do título de Doutor
em Ciências. Área de concentração: Bioenergia

**Piracicaba
2019**

Lucas Souza Lopes
Bacharel em Ciências Biológicas

**Avaliação dos efeitos causados por contaminações bacterianas em
fermentações etanólicas de primeira geração utilizando metodologias
multi-ômicas**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:

Prof. Dr. **CARLOS ALBERTO LABATE**

Co-orientador:

Prof. Dr. **DIEGO MAURÍCIO RIAÑO-PACHÓN**

Tese apresentada para obtenção do título de Doutor
em Ciências. Área de concentração: Bioenergia

Piracicaba
2019

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Lopes, Lucas Souza

Avaliação dos efeitos causados por contaminações bacterianas em fermentações etanólicas de primeira geração utilizando metodologias multi-ômicas / Lucas Souza Lopes. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2019.

93 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz". Universidade Estadual de Campinas. Universidade Estadual Paulista "Julio de Mesquita Filho"

1. Genômica 2. *Lactobacillus* 3. *Acetobacter* 4. Metabolômica I. Título

DEDICATÓRIA

Aos meus pais, Maria José e Miguel:
Pelos conselhos nos momentos de raiva,
Paciência nos momentos de aflição,
Pelo apoio nos momentos de dificuldade,
Pelo amor nos momentos de sofrimento,
Sem vocês este sonho não seria possível.

Aos meus tios, Glória e Geraldo:
Pela receptividade e acolhimento em São Paulo,
Pelas sugestões nos momentos de incerteza,
Pelo carinho demonstrado no convívio diário,
Vocês tiveram um papel fundamental nesta realização.

AGRADECIMENTOS

Um agradecimento especial ao Prof. Carlos Alberto Labate. Sua generosidade, capacidade de liderança, inteligência e apoio incondicional foram verdadeiros motivadores na minha vida. Terei um enorme orgulho de dizer que um dia fiz parte da grande família Max Feffer.

Ao Prof. Diego Maurício Riaño-Pachón, meu co-orientador e amigo. Muito obrigado pelos ensinamentos e conselhos. Todos esses anos de colaboração me ensinaram o verdadeiro sentido de trabalho em equipe.

Á Profa. Jaciane Lutz Ienczak, muito obrigado pela supervisão nas tão desafiadoras fermentações. Sua qualidade profissional, humor e doçura me ensinaram a ser um profissional competente e ético sem perder a minha humanidade.

Ás pesquisadoras e amigas Thaís Regiani Cataldi e Gabriela Félix Persinoti. A prontidão e apoio irrestritos nos inúmeros momentos de dificuldades foram fundamentais na conquista desta vitória. O apoio de vocês ficará eternamente na minha memória. Muito obrigado!

Á minha amiga e irmã (de coração) Emília Estanislau Muniz. Uma amizade como a sua é algo que acontece poucas vezes em nossas vidas. Muito obrigado por tudo, principalmente pela alegria proporcionada ao me convidar para ser padrinho da tão linda e amada Luísa.

Ao apoio financeiro da CAPES, CNPq (141460/2015-6) e FAPESP (2015/13431-5). A conclusão com êxito deste projeto só foi possível com as suas parcerias.

Ao Cleilton, amigo e colega de trabalho. Obrigado pelo apoio nas noites frias de Campinas durante as fermentações. Com certeza, sem seu apoio este trabalho não estaria sendo defendido. Muito obrigado de coração!

Ao Fábio Jarret, terapeuta e *coaching* profissional. Muito obrigado pelo apoio nos momentos de dificuldade. Suas orientações foram fundamentais no meu processo de amadurecimento e decisivas na entrega deste trabalho.

Á todos os funcionários, técnicos e pesquisadores que tive o prazer de conviver e trabalhar no período que estive no CTBE/CNPEM. Muito obrigado pela ajuda, orientações, conselhos e organização do ambiente de trabalho.

Á todos os meus colegas do laboratório Max Feffer de Genética de Plantas. A acolhida que tive de vocês foi maravilhosa. Muito obrigado pelos momentos de descontração e risadas, protagonizados principalmente na hora do almoço.

Aos meus amigos e colegas do grupo “MasterChefBR”. O almoço do CNPEM sempre ficava mais saboroso e requintado quando em companhia de vocês.

Aos meus tios Geraldo Magela e Glória Lage. Foram mais de quatro anos morando juntos, cheios de momentos vividos em família. Obrigado pelos ensinamentos, carinho e apoio. Vocês se tornaram segundos pais para mim.

Aos meus irmãos Vitor e Miguel Jr. A distância de quase mil quilômetros não foi capaz de separar nossos corações.

Á minha mãe Maria José Lopes e meu pai Miguel Salatiel Lopes. Não existem palavras que descrevem o que sinto por vocês. Obrigado por me darem esta vida tão feliz e repleta de realizações. Sem vocês, verdadeiramente, eu não estaria aqui hoje.

EPÍGRAFE

“A chave da ética é o controle da vaidade.

Temos que ser humildes, principalmente quando ocupamos cargos de poder.

Quanto maior meu cargo, mais eu sirvo.

Grandes titulações significam que eu tenho mais responsabilidades e pessoas para cuidar”.

Leandro Karnal

SUMÁRIO

RESUMO.....	7
ABSTRACT	8
1. INTRODUÇÃO.....	9
2. REVISÃO DE LITERATURA	13
2.1. Leveduras de <i>S. cerevisiae</i>	13
2.2. Bactérias lácticas de <i>Lactobacillus</i>	14
2.3. Bactérias acéticas de <i>Acetobacter</i>	15
2.4. Os processos industriais de produção de etanol presentes no Brasil	15
2.5. Análises multi-ômicas aplicadas a interações biológicas e biologia de sistemas	17
3. OBJETIVOS	19
4. MATERIAL E MÉTODOS	21
4.1. Material biológico	21
4.2. Cinética de crescimento e condições de cultivo das bactérias	21
4.3. Estudos filogenéticos do gene 16S do rDNA das bactérias	22
4.4. Cariotipagem da levedura <i>S. cerevisiae</i> CAT-1	23
4.5. Obtenção e preparo do melaço de cana-de-açúcar	23
4.6. Propagação celular da levedura <i>S. cerevisiae</i> CAT-1	23
4.7. Ensaio fermentativos em reatores	24
4.8. Coleta de amostras para metabolômica	26
4.9. Sequenciamento, montagem e anotação dos genomas bacterianos	26
4.10. Análises taxonômicas e filogenéticas das linhagens bacterianas	27
4.11. Estudos pangenômicos	28
4.12. Extração de metabólitos intracelulares e derivatização	28
4.13. Análises em GC-MS	29
4.14. Processamento dos dados, identificação dos metabólitos e análises estatísticas	29
5. RESULTADOS E DISCUSSÕES	31
5.1. Cinética de crescimento das bactérias	31
5.2. Filogenia do gene 16S do rDNA das bactérias	32
5.3. Estudos eletroforéticos <i>fingerprint</i> de <i>S. cerevisiae</i> CAT-1	34
5.4. Clarificação do melaço de cana-de-açúcar	35
5.5. Ensaio fermentativos	36
5.6. Sequenciamento, montagem e anotação dos genomas bacterianos	44
5.7. Estudos taxonômicos e filogenéticos multigênicos	48
5.8. Estudos pangenômicos	52
5.9. Identificação dos metabólitos intracelulares totais por GC-MS	56
5 CONCLUSÕES	67
REFERÊNCIAS.....	69
APÊNDICES	75

RESUMO

Avaliação dos efeitos causados por contaminações bacterianas em fermentações etanólicas de primeira geração utilizando metodologias multi-ômicas

O etanol possui grande importância estratégica para o setor energético brasileiro devido a sua natureza renovável e grande capacidade de geração de emprego e renda. Atualmente um dos principais desafios enfrentados pelas indústrias brasileiras de produção de etanol é a presença de bactérias contaminando as fermentações. Essas atuam modificando o sistema, competindo por nutrientes essenciais, produzindo compostos tóxicos para a levedura e induzindo processos indesejáveis para a indústria como a floculação. Para equilibrar as populações e controlar os efeitos negativos é necessário uma melhor compreensão dos mecanismos genéticos e fisiológicos envolvidos nas interações de bactérias e leveduras em fermentações em fermentações etanólicas de primeira geração (1G). Neste sentido o presente trabalho realizou o sequenciamento, montagem e anotação completa dos genomas de três linhagens bacterianas isoladas diretamente de usinas brasileiras de produção de etanol: *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3. Além disso, foram feitas simulações da condição de co-existência entre bactérias lácticas e Acéticas com as leveduras de *S. cerevisiae* em fermentações semi-industriais em batelada alimentada (“*Melle-Boinol*”) com tratamento ácido e reciclo de células. Três condições experimentais foram avaliadas: (i) fermentações contendo apenas leveduras de *S. cerevisiae* CAT-1, (ii) fermentações contendo leveduras de *S. cerevisiae* CAT-1 e bactérias de *Lactobacillus* sp. I-2, e (iii) fermentações contendo leveduras de *S. cerevisiae* CAT-1 e bactérias de *Acetobacter* sp. A-3. Por fim, os metabolomas das bactérias e leveduras da condição de coexistência nos ensaios fermentativos foram obtidos por metodologias de cromatografia gasosa acoplada a espectrometria de massas (GC-MS). Foram identificados 3056, 2112 e 3557 genes nos genomas de *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente. As leveduras de *S. cerevisiae*, independente da condição experimental avaliada, consumiram totalmente os açúcares redutores totais (ART) presentes no melaço de cana-de-açúcar ao longo das fermentações em batelada alimentada. A glicina, aminoácido polar contendo apenas um hidrogênio na cadeia lateral, apresentou maior abundância durante as fermentações contaminadas com bactérias lácticas e acéticas em detrimento da condição controle sem contaminação bacteriana. Especula-se que o acúmulo intracelular deste aminoácido seja resultado do desvio de intermediários das vias Glicólise e Ciclo do Ácido Cítrico para a produção de compostos nitrogenados. Provavelmente, ambos os microrganismos, bactérias e leveduras, acumularam glicina no ambiente intracelular como alternativa a deficiência de nitrogênio no melaço de cana-de-açúcar. O sequenciamento dos genomas das bactérias e os ensaios fermentativos em biorreatores foram realizados no Laboratório Nacional de Biorrenováveis (LNBR) do Centro Nacional de Pesquisas em Energias e Materiais (CNPEM). As análises de metabolômica foram realizadas no Laboratório Max Feffer de Genética de Plantas da Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ). Espera-se que os resultados apresentados neste trabalho agreguem mais informações a cerca das interações entre bactérias e leveduras e contribua, num futuro próximo, no apontamento de novas estratégias de controle bacteriano que sejam mais eficientes e sustentáveis que as usadas atualmente em fermentações industriais brasileiras.

Palavras-chave: Genômica; *Lactobacillus*; *Acetobacter*; Metabolômica

ABSTRACT

Multi-omics analyzes of bacterial contaminations in first-generation ethanolic fermentations

Ethanol has great strategic importance for the Brazilian energy sector due to its renewable nature and great capacity to generate jobs and income. Currently, one of the main challenges faced by the Brazilian ethanol industry is the presence of contaminating bacteria in the fermentation tanks. They act by modifying the system, competing for essential nutrients, producing toxic compounds to the yeast and inducing undesirable processes for the industry like flocculation. To better balance populations and to control the negative effects, it is necessary a better understanding of the genetic and physiological mechanisms involved in the interactions among bacteria and yeasts from *Saccharomyces cerevisiae*. In this sense, the present work carried out the sequencing, assembly and complete genome annotation of three bacterial strains isolated directly from Brazilian ethanol production plants: *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 and *Acetobacter* sp. A-3. In addition, simulations of the co-existence condition between lactic and acetic bacteria with yeasts from *S. cerevisiae* were carried out in semi-industrial fed-batch (Melle-Boinot) fermentations with acid treatment and cell recycling. Three experimental conditions were evaluated: (i) fermentations containing only yeasts from *S. cerevisiae* CAT-1, (ii) fermentations containing yeasts of *S. cerevisiae* CAT-1 and bacteria from *Lactobacillus* sp. I-2, and (iii) fermentations containing yeasts from *S. cerevisiae* CAT-1 and bacteria from *Acetobacter* sp. A-3. Finally, the metabolomes of bacteria and yeasts in the coexistence condition on fermentation assays were obtained by gas chromatography coupled to mass spectrometry (GC-MS) and liquid chromatography coupled to mass spectrometry (LC-MS). 3056, 2112 and 3557 genes were identified in the genomes of *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 and *Acetobacter* sp. A-3, respectively. The yeasts from *S. cerevisiae*, regardless of the experimental condition evaluated, all the total reducing sugars (ART) present in sugarcane molasses were consumed and the ethanol production was consistent with that observed in first generation (1G) ethanolic fermentations. Glycine, an apolar amino acid containing only one hydrogen in the lateral chain, showed greater abundance during fermentations contaminated with lactic and acetic bacteria. It is speculated that the intracellular accumulation of this amino acid results from the shift of intermediates from the Glycolysis/Glycogenogenesis and the Citric Acid Cycle pathways to the production of nitrogen compounds. Probably, both microorganisms, bacteria and yeasts, used the intracellular accumulation of glycine as an alternative to nitrogen deficiency in sugarcane molasses in the fermentations carried out in bioreactors. The Bacterial genomes sequencing and fermentation assays in bioreactors were performed at the Brazilian Bioethanol Science and Technology Laboratory (CTBE) of the Brazilian Center for Research in Energy and Materials (CNPEM). The metabolomics analyzes were performed at the Max Feffer Laboratory of Plant Genetics at the "Luiz de Queiroz" School of Agriculture (ESALQ). It is hoped that the results presented in this work will add more information about the interactions among bacteria and yeasts and will contribute, in the near future, to new bacterial control strategies in Brazilian industrial fermentations that are more efficient and sustainable than those currently used.

Keywords: Genomics; *Lactobacillus*; *Acetobacter*; Metabolomics

1. INTRODUÇÃO

A utilização do etanol como combustível veicular direto tem alcançado bastante sucesso no Brasil nos últimos anos. No primeiro congresso nacional em aplicações industriais do álcool, realizado em 1903, começou a ser discutido o desenvolvimento de tecnologias que permitissem a utilização de etanol diretamente como fonte de energia automotora (MOREIRA & GOLDEMBERG, 1996). Porém, somente na década de 1970, políticas nacionais concretas foram iniciadas com a crise internacional do petróleo. O Programa Nacional de Incentivo ao uso do Álcool (Proálcool) foi proposto com o intuito de diminuir a dependência do Brasil dos combustíveis fósseis (TOURINHO & ANDRADE, 1998). Como resultado principal, houve um aumento de produção de mais de 27 bilhões de litros de etanol em quase 45 anos, passando de 500 milhões em 1975 para cerca de 28 bilhões em 2017/2018 (UNICA, 2018).

A fermentação alcoólica é o principal processo utilizado para produção de etanol no mundo, sendo a mediada por leveduras de *S. cerevisiae* a mais comum em indústrias (BASSO et al., 2008). Neste processo o microrganismo é capaz de obter energia química na forma de ATP, degradando parcialmente compostos carbonados como glicose e frutose. Dentre os produtos finais gerados temos etanol, CO₂ e glicerol (PELCZAR et al., 1997; BASSO et al., 2008).

O processo de produção de etanol em batelada alimentada (“*Melle-Boinol*”) com tratamento ácido e reciclo de células é o mais utilizado em indústrias no Brasil (MAIORELLA, BLANCH & WILKE, 1981). A principal característica desse sistema é a adição continuada do mosto de cana-de-açúcar sobre um inóculo contendo uma alta concentração celular de *S. cerevisiae*. Quando o volume final da dorna é alcançado, por centrifugação, realiza-se uma separação do inóculo celular e vinho delevedurado que seguem para um tanque de tratamento e colunas de destilação, respectivamente. As células de leveduras recuperadas sofrem um tratamento com ácido sulfúrico (H₂SO₄) para remoção de impurezas e diminuição da concentração de microrganismos contaminantes (VIEGAS, 1999; BORGES, 2008). A principal vantagem adquirida no modelo de batelada alimentada é a diminuição do tempo ocioso entre os ciclos fermentativos através da fermentação dos açúcares presentes no meio concomitante ao processo de enchimento da dorna (MAIORELLA, BLANCH & WILKE, 1981).

As indústrias brasileiras de produção de etanol não são ambientes assépticos, sendo que o mosto de cana-de-açúcar é a principal entrada de microrganismos contaminantes nas fermentações (GALLO, 1989). A presença de bactérias no ambiente industrial pode causar a perda da eficiência fermentativa, redução da produção de etanol, competição por nutrientes essenciais, aumento dos teores de compostos tóxicos para a levedura e indução do processo de floculação (LUDWIG, OLIVA-NETO & ANGELIS, 2001). Os mecanismos de controle de bactérias contaminantes em fermentações utilizados atualmente são o uso do tratamento ácido e aplicações periódicas de antibióticos. Contudo, a aplicação indiscriminada destas práticas envolve um grande aporte de recursos e pode gerar problemas graves para a indústria, espécie humana e meio ambiente (LUDWIG, OLIVA-NETO & ANGELIS, 2001). Carvalho-Netto et al. (2015) e Bonatelli et al. (2017) mostram que bactérias lácticas e acéticas estão entre os microrganismos mais abundantes do sistema, sendo uma das principais causadoras das perdas de produtividade e geração de estresses bióticos.

Bactérias do gênero *Lactobacillus* são as mais abundantes em fermentações industriais brasileiras, sendo as da espécie *L. fermentum* uma das principais causadoras do processo de floculação (CARVALHO-NETTO et al., 2015). Bactérias acéticas são comumente utilizadas em indústrias produtoras de vinagre devido a sua capacidade de produzir ácido acético a partir de açúcares como glicose e frutose (NDOYE et al., 2007). Dentre as espécies conhecidas tem-se *Acetobacter senegalensis*, bactéria termotolerante primeiramente isolada de manga (*Mangifera indica* L.) em Senegal na África subsariana (NDOYE et al., 2007).

A genômica é o estudo da composição genética contida nas sequências de DNA de um organismo (ALBERTS et al., 2010). Nos últimos anos, um grande volume de dados genômicos tem sido gerados com a aplicação de tecnologias como o sequenciamento da nova geração (NGS). Através da genômica, comparações dos conteúdos gênicos preditos de diferentes organismos tem revelado novos conceitos de genes e até mesmo de espécies (GOFFEAU et al., 1996; TETTELIN et al., 2015; VARGHESE et al., 2015). A pangenômica é o estudo integrado dos genomas de diferentes organismos. De maneira geral, os genes apontados num pangenoma podem ser classificados em *core*, acessórios e exclusivos para cada um dos organismos comparados (TETTELIN et al., 2015). Os genes classificados como *core* estão presentes em todos os genomas, sendo em sua maioria, associados a funções essenciais como os de *housekeeping* e transporte de proteínas (TETTELIN et al., 2015). Genes acessórios estão presentes em pelo menos dois, mas não em todos os genomas comparados. Genes exclusivos só podem ser encontrados em um único genoma estudado. De maneira geral, esses estão associados a funções biológicas não essenciais com importância adaptativa para os portadores ao ambiente em que vivem (TETTELIN et al., 2015).

A metabolômica é o estudo em larga escala de pequenas moléculas altamente diversas conhecidas como metabólitos (FIEHN, et al., 2002). O metaboloma é caracterizado pelo conjunto total de metabólitos de um ou mais organismos vivos num determinado tempo e condição ambiental (FIEHN, et al., 2002). De maneira geral, o metaboloma apresenta alta proximidade com o fenótipo, fazendo dos metabólitos compostos altamente interessantes em estudos de alterações fisiológicas em decorrência de modificações ambientais (FIEHN, et al., 2002). Uma das principais formas de se acessar o metaboloma de um organismo é por meio da utilização de métodos cromatográficos de alta eficiência acoplados à espectrometria de massas (WOLFENDER, RODRIGUEZ & HOSTETTMANN, 1998). Por meio dessa fusão de técnicas é possível realizar a separação de compostos presentes em amostras biológicas complexas, obter a razão massa/carga (m/z) dos íons gerados e identificar cada um dos compostos utilizando informações presentes em bancos de dados (WOLFENDER, RODRIGUEZ & HOSTETTMANN, 1998). Estudos de cromatografia gasosa aplicada a espectrometria de massas (GC-MS) são bastante interessantes quando se objetiva a detecção de componentes presentes no metabolismo primário dos organismos. Desta maneira, comparações relativas nos teores de metabólitos de baixa massa molecular, entre diferentes condições experimentais podem ser realizadas (WOLFENDER, RODRIGUEZ & HOSTETTMANN, 1998).

O presente trabalho possui três grandes objetivos principais: (i) caracterizar geneticamente três linhagens de bactérias isoladas diretamente de indústrias brasileiras de produção de etanol: *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3; (ii) simular os efeitos de coexistência entre bactérias lácticas e acéticas em fermentações etanólicas de primeira geração em batelada alimentada com tratamento ácido e reciclo de células; e (iii) avaliar potenciais alterações metabólicas intracelulares em leveduras de *S. cerevisiae* causadas pelas contaminações bacterianas nas fermentações. Os genomas das bactérias foram sequenciados através de metodologias de sequenciamento da nova geração (NGS). Espera-se a obtenção do conjunto total de genes destes organismos, relações filogenéticas com outras espécies de bactérias lácticas e acéticas e identificação de potenciais mecanismos genéticos que possam ser associados a sobrevivência destes organismos à fermentações etanólicas 1G. As simulações dos efeitos de coexistência foram realizadas em reatores de médio porte com controle de aeração, agitação e temperatura. Espera-se que a presença das bactérias contaminantes altere os padrões de consumo de açúcares redutores totais (ARTs), produção de álcoois (Etanol e Glicerol), produção de ácidos orgânicos (ácido láctico, ácido acético e ácido succínico) e alteração nos índices de viabilidade celular médio e taxa de brotamento celular médio das leveduras de *S. cerevisiae*. A avaliação das alterações metabólicas intracelulares das leveduras foi realizada a partir de cromatografia gasosa aplicada a espectrometria de massas (GC-MS). Acredita-se que a presença de ambos os microrganismos nas fermentações, bactérias e leveduras, irá aumentar a competição por nutrientes

presentes no fermentado, produzir compostos tóxicos para as leveduras e potencialmente induzir processos como floculação. Comparações entre condições experimentais controle (sem contaminação bacteriana) e testes (contaminadas com bactérias lácticas e acéticas) permitirão o apontamento de metabólitos diferencialmente abundantes no metabolismo das leveduras que refletem à presença das bactérias. Espera-se que este trabalho contribua para um melhor entendimento da composição genética e evolutiva de alguns dos principais grupos de bactérias que habitam sistemas fermentativos industriais brasileiros. Além disso, espera-se que as simulações das fermentações industriais em biorreatores e os estudos de metabolômica intracelular por GC-MS elucidem os principais efeitos da presença de bactérias de *Lactobacillus* e *Acetobacter* nas fermentações e no metabolismo primário de leveduras de *S. cerevisiae*.

2. REVISÃO DE LITERATURA

2.1. Leveduras de *S. cerevisiae*

As leveduras da espécie *S. cerevisiae* são microrganismos unicelulares eucarióticos com ampla aplicabilidade industrial (PELCZAR et al., 1997; ANDRIETTA, STECKELBERG & ANDRIETTA, 2006). Uma das formas mais comuns de reprodução das leveduras é a assexuada por brotamento. Após eventos de divisão nuclear e migração de organelas citoplasmáticas para a região de brotamento na célula mãe, essa se desprende originando dois organismos independentes (TORTORA et al., 2008). O ciclo sexuado é caracterizado pela formação de quatro células haploides a partir de uma indução natural de meiose em uma célula inicial diploide (TORTORA et al., 2008). Estas células filhas são liberadas no ambiente através do rompimento de estruturas de armazenamento denominadas ascas (TORTORA et al., 2008). A formação de células haploides por sucessivas mitoses de células iniciais n podem ocorrer através da extensão da haplófase em linhagens heterotáticas de *S. cerevisiae* (TORTORA et al., 2008). A união de esporos de diferentes reações sexuais (“*mating types*”), podem formar híbridos diploides, originando a diplófase. Assim, a partir destas células, tanto organismos diploides como haploides podem ser novamente produzidos e armazenadas nos ascas (TORTORA et al., 2008). Roman (1981) mostra que condições de estresse relacionadas à falta de nutrientes no ambiente podem induzir o processo de reprodução sexuada em *S. cerevisiae*.

As leveduras do gênero *Saccharomyces* conseguem gerar energia na forma de ATP tanto na presença como na ausência de oxigênio (PELCZAR et al., 1997). Quando em condições de anaerobiose, açúcares como glicose e frutose são parcialmente degradados em etanol, CO₂ e glicerol através de um processo denominado fermentação (ANDRIETTA, STECKELBERG & ANDRIETTA, 2006). Devido a isso, estes microrganismos têm sido bastante utilizados em processos industriais que visem à produção de bebidas e combustíveis automotores (ANDRIETTA, STECKELBERG & ANDRIETTA, 2006).

O sequenciamento do genoma da espécie *S. cerevisiae* foi um importante marco científico no ano de 1996. O trabalho pioneiro de Goffeau et al. (1996), numa parceria com mais de 600 pesquisadores ao redor do mundo, permitiu a obtenção do primeiro genoma completo de um eucarioto (GOFFEAU et al., 1996; ØSTERGAARD, OLSSON & NIELSEN, 2000). Este trabalho abriu portas para que os estudos genômicos futuros por meio da identificação e comparação de genes nos anos subsequentes (GOFFEAU et al., 1996; ØSTERGAARD, OLSSON & NIELSEN, 2000).

No Brasil, em meados da década de 1980 iniciou-se a busca por diferentes linhagens selvagens de *S. cerevisiae* que fossem capazes de tolerar as condições de estresses presentes nas usinas brasileiras de produção de etanol sem perder viabilidade e produtividade. A grande velocidade de isolamento e purificação de microrganismos resultou na obtenção de mais de 700 linhagens no final da década de 1990 (STEWART & RUSSEL, 1998). As linhagens *S. cerevisiae* CAT-1 e *S. cerevisiae* PE-2 estão entre as mais utilizadas em usinas brasileiras de produção de etanol. Essas apresentam alta tolerância a estresses causados por aumentos de temperatura, teores de açúcares e etanol. Especificamente, a linhagem *S. cerevisiae* CAT-1 foi isolada na usina de Catanduva no interior do estado de São Paulo no final da década de 1990 (BASSO et al., 2008). Características como baixa capacidade floculante, produção de pouca espuma e manutenção de alta viabilidade celular ao longo de sucessivos ciclos fermentativos são algumas vantagens atribuídas a esta linhagem em fermentações industriais brasileiras (BASSO et al., 2008). Amorim Neto et al. (2009) demonstraram que *S. cerevisiae* CAT-1 consegue fermentar açúcares como glicose, frutose e sacarose em temperaturas superiores a 40 °C, algo almejado em países tropicais com o Brasil. Acredita-se que a termotolerância

pode ocorrer, dentre outras coisas, devido aos altos níveis intracelulares de trealose e glicerol, compostos associados à proteção celular contra estresses osmóticos, oxidativos e térmicos em leveduras de *S. cerevisiae* (PANEK, 1963; MATTENBERG et al., 2016). Por fim, o sequenciamento do genoma da linhagem *S. cerevisiae* CAT-1 permitiu a identificação de aproximadamente 6652 potenciais genes codificadores de proteínas. Este trabalho de Babrzadeh et al. (2002) abriu portas para estudos genômicos comparativos com outras linhagens de *S. cerevisiae*, como a S288c.

2.2. Bactérias lácticas de *Lactobacillus*

A presença de bactérias contaminantes é algo comum em fermentações etanólicas industriais brasileiras (GALLO, 1989). Devido a não ocorrência de uma assepsia completa do mosto de cana-de-açúcar, observa-se o adentramento contínuo de bactérias contaminantes nas fermentações (GALLO, 1989). Dentre as várias consequências negativas apontadas pela presença destes microrganismos destacam-se as reduções na eficiência fermentativa e produção de etanol, competição por nutrientes essenciais, aumentos dos teores de compostos tóxicos para a levedura e indução de processos como floculação (LUDWIG, OLIVA-NETO & ANGELIS, 2001).

Bactérias ácido lácticas (“*Lactic Acid Bacteria*” – LAB) são classificadas como organismos gram-positivos, anaeróbicos aerotolerantes, com formato de bastonete e sem capacidade de esporulação (HOLZAPFEL et al., 2004). LAB podem ser classificadas como bactérias homo- ou heterofermentativas, baseada nas vias metabólicas utilizadas para degradar hexoses (KANDLER, 1983; COSTA et al., 2008). LAB heterofermentativas podem produzir ácido láctico, ácido acético, etanol e dióxido de carbono como compostos resultantes do metabolismo de moléculas carbonadas. Bactérias homofermentativas são capazes de produzir apenas ácido láctico a partir do consumo de açúcares, principalmente glicose e frutose (COSTA et al., 2008). Bactérias das espécies *L. fermentum*, *L. reuteri*, *L. brevis* e *L. buchneri* são exemplos de LAB heterofermentativas. Bactérias das espécies *L. acidophilus*, *L. gasseri*, *L. crispatus*, *L. johnsonii* são exemplos de LAB homofermentativas.

A grande maioria das bactérias contaminantes de fermentações etanólicas de primeira geração no Brasil são LAB, sendo que entre as espécies mais abundantes destacam-se *L. fermentum* e *L. plantarum* (CARVALHO-NETTO et al., 2015; BONATELLI, 2017). Acredita-se que a capacidade de tolerar altas concentrações de etanol e baixos índices de pH sejam algumas das explicações da presença destes microrganismos em fermentações industriais brasileiras (AMORIM et al., 2011). Bactérias de *L. fermentum* estão entre as principais agentes causadoras de coagregados celulares entre bactérias e leveduras em fermentações, processo conhecido como floculação (CARVALHO-NETTO et al., 2015). Especula-se que proteínas Adesinas Manose-específicas (“*Mannose specific adhesin*” – Mas), presentes no envoltório celular destes microrganismos, podem estar envolvidas no processo de justaposição celular e formação dos agredados celulares de bactérias e leveduras. Carvalho-Netto et al. (2015) mostram que a presença de populações de 10^5 UFC.ml⁻¹ de *L. fermentum* foram capazes de induzir floculação em fermentações contendo leveduras de *S. cerevisiae* PE-2. O trabalho conclui que bactérias desta espécie são as principais causadoras deste processo em fermentações industriais brasileiras (CARVALHO-NETTO et al., 2015).

A coexistência entre bactérias do gênero *Lactobacillus* e leveduras de *Sacharomyces* é algo comum em ambientes naturais (GARCIA et al., 2016). Garcia et al. (2016) discutem os efeitos da coevulação destes microrganismos, manifestadas em interações benéficas tanto para as bactérias lácticas como para as leveduras. De acordo com eles, a produção do ácido láctico pelas bactérias tem papel fundamental na biossíntese intracelular da proteína prion [GAR+]. Neste sentido, através do estímulo causado pelo ácido láctico, leveduras poderiam modular

seu metabolismo de moléculas carbonadas diminuindo a preferência por glicose e metabolizando outras fontes de carbono concomitantemente (GARCIA et al., 2016).

2.3. Bactérias acéticas de *Acetobacter*

Indústrias brasileiras de produção de etanol precisam equilibrar as populações de bactérias e leveduras no processo de fermentação industrial para minimizar os efeitos de perda de produtividade e problemas como o excesso de ácidos orgânicos no meio (CECCATO-ANTONINI & SILVA, 2000). Como medidas essenciais de controle bacteriano temos a utilização de linhagens de leveduras mais competitivas, maior rigor asséptico nos processos de colheita, transporte e moagem da cana-de-açúcar e realização de manutenções periódicas na infraestrutura das usinas (GALLO, 1989). Além disso, um maior controle de fatores abióticos que propiciam o crescimento de bactérias (temperatura e pH do mosto) e adição de antibióticos no meio quando necessário tem propiciado resultados satisfatórios no controle destes microrganismos no ambiente industrial (CECCATO-ANTONINI & SILVA, 2000; CARVALHO-NETTO et al., 2015). No entanto, o uso indiscriminado de antibióticos podem gerar problemas para indústria, espécie humana e meio ambiente, descaracterizando a natureza sustentável deste combustível (LUDWIG, OLIVA-NETO & ANGELIS, 2001).

Como discutido anteriormente, estudos de diversidade genética e caracterização das populações de bactérias em fermentações apontam microrganismos do gênero *Lactobacillus* como os mais abundantes em sistemas industriais brasileiros de produção de etanol (CARVALHO-NETTO et al., 2015; BONATELLI, 2017). Dentre os grupos menos abundantes temos bactérias dos gêneros *Acetobacter*, microrganismos aeróbicos restritos, gram-negativos e com capacidade de produção de ácido acético a partir de etanol, quando em condições de aerobiose (MATSUSHITA et al., 2016; BONATELLI, 2017). Bactérias ácido acéticas (“*Acetic Acid Bacteria*” - AAB) são classificadas como membros do filo Proteobactéria e família *Acetobacteraceae*. Espécies como *A. pasteurianus*, *A. tropicalis* e *A. senegalensis* estão entre as mais estudadas do grupo (MATSUSHITA et al., 2016). Devido a sua alta capacidade de produção de ácido acético, AAB têm sido bastante utilizadas para a produção de vinagre (TRCEK & TEUBER, 2002). Ndoye et al. (2007) descrevem a descoberta de uma nova espécie de AAB, *A. senegalensis*, uma bactéria termotolerante isolada da Manga (*Mangifera indica* L.) em Senegal na África. Esses microrganismos são oxidases negativas, catalase positiva e apresentam tolerância a altas temperaturas (NDOYE et al., 2007). Estes diferem de *A. tropicalis*, espécie filogeneticamente mais semelhante, por serem capazes de crescerem em meios contendo altos teores de etanol e Glicose no meio de cultivo (NDOYE et al., 2007). Além disso, Ndoye et al. (2007) descrevem que, diferente de *A. tropicalis*, *A. senegalensis* não são capazes de crescer em meios contendo maltose como única fonte de carbonos.

2.4. Os processos industriais de produção de etanol presentes no Brasil

As primeiras discussões sobre a produção industrial e utilização do etanol como combustível no Brasil datam o início do século XX no primeiro Congresso Nacional de Aplicações Industriais de Álcoois. Neste foi discutido, dentre outras coisas, tecnologias que permitissem a viabilidade da produção industrial de etanol e sua utilização com combustível veicular direto (MOREIRA & GOLDEMBERG, 1996). Contudo, apenas em 1970, após a criação do Programa Nacional de Incentivo ao uso do Alcool (Proálcool), como tentativa de minimizar os

problemas oriundos da crise mundial do petróleo, observaram-se avanços significativos na produção deste biocombustível no país (ZANIN et al., 2000). Apesar da ampla diversidade de tecnologias de produção de etanol no Brasil, três processos destacam-se por estarem presentes na ampla maioria das indústrias brasileiras: batelada, batelada alimentada (“*Melle-Boinol*”) e contínuo (MAIORELLA, BLANCH & WILKE, 1981; ANDRIETTA & STUPIELLO, 1990).

O processo de batelada foi bastante utilizado até a década de 1960, mas perdeu espaço nos anos subsequentes. Esse é caracterizado pela propagação laboratorial de inóculos de leveduras de *S. cerevisiae*, seguido da transferência para dornas contendo mosto de cana-de-açúcar. O fim da fermentação é determinado quando não se detecta mais aumentos nos teores de etanol. O combustível é recuperado via destilação e a dorna é sanitizada para a etapa subsequente. As principais críticas apontadas para este processo são lentidão da fermentação e a utilização de um novo inóculo a cada ciclo fermentativo (MAIORELLA, BLANCH & WILKE, 1981).

Em meados da década de 1960 começou a ser empregado o sistema de batelada alimentada ou fermentação “*Melle-Boinol*”. Dentre as muitas variações propostas em relação ao modelo em batelada, destaca-se a alimentação crescente e contínua da dorna de fermentação com mosto de cana-de-açúcar. Neste sentido, o inóculo de leveduras deve ser adicionado previamente caracterizando o pé-de-cuba. Esta modificação permitiu a redução da lentidão da fermentação e a inclusão de mais de um ciclo industrial por dia, uma vez que o processo fermentativo é iniciado ainda durante o processo de alimentação do sistema (MAIORELLA, BLANCH & WILKE, 1981). Além disso, a adoção da técnica de lavagem do inóculo de leveduras com ácido sulfúrico diluído (H_2SO_4) permitiu a remoção de impurezas e redução das populações de bactérias contaminantes. A combinação dos processos de batelada alimentada, tratamento com ácido diluído e reciclo de células criou, pela primeira vez no Brasil, um processo de fermentação industrial de açúcares em larga escala viável economicamente (MAIORELLA, BLANCH & WILKE, 1981; BASSO et al. 2011) (Figura 1).

O processo contínuo surgiu na década de 1980 com difusão na década seguinte. Este é baseado no acoplamento subsequente de dornas de diferentes volumes com entrada e saída de mosto de cana-de-açúcar e vinho, respectivamente. Apontam-se como vantagens deste processo a redução dos custos de produção, de mão de obra e dos tempos ociosos como carga e descarga da dorna. No entanto, a potencialização do problema de contaminação bacteriana tem sido uma importante desvantagem deste modelo (ANDRIETTA & STUPIELLO, 1990).

Figura 1. Desenho simplificado do processo de batelada alimentada com tratamento ácido e reciclo de células. Sistema comumente utilizado em usinas brasileiras de produção de etanol (LOPES et al., 2016).

2.5. Análises multi-ômicas aplicadas a interações biológicas e biologia de sistemas

Historicamente, a utilização de abordagens multiômicas em estudos de biologia de sistemas sempre foi acompanhada por limitações tecnológicas e financeiras (CULIBRK, CROFT & TEBBUTT, 2016). As dificuldades de acesso limitaram por muitos anos os estudos quantitativos de respostas celulares a estresses ambientais e interações ecológicas entre organismos. Ferramentas como “*western blot*” e “*reverse transcription quantitative PCR*” (RT-qPCR) apresentam alta sensibilidade e especificidade, contudo são dependentes de procedimentos laboriosos e com alta incidência de erros além da capacidade limitada de análise de poucas moléculas simultaneamente (CULIBRK, CROFT & TEBBUTT, 2016).

O desenvolvimento de tecnologias com capacidade de geração de um grande volume de dados permitiu, nos últimos anos, que o mundo científico entrasse definitivamente na era do “*Big data*” (CULIBRK, CROFT & TEBBUTT, 2016). A aplicação de ferramentas desenvolvidas por nanotecnologia tem revolucionado diversas áreas da biologia, notavelmente aquelas associadas à biologia molecular, bioquímica, fisiologia, biologia de sistemas e bioinformática (CULIBRK, CROFT & TEBBUTT, 2016). A utilização de tecnologias como sequenciamento de nova geração (NGS) e espectrometria de massas tem tornado estudos de genômica, transcriptômica, proteômica e metabolômica cada vez mais acessíveis a pesquisadores no mundo todo. Um exemplo desta universalização é a capacidade atual de se realizar o sequenciamento completo de um genoma humano com um custo inferior a US\$1000 (VERITAS GENETICS, 2019).

Tettelin et al. (2005) realizou diversos estudos genômicos em diferentes linhagens de *Streptococcus agalactiae*. A observação de uma grande diversidade de genes entre os organismos de uma mesma espécie o levou a discutir, pela primeira vez, a realização de análises pangenômicas para melhor caracterizar um grupo de linhagens bacterianas. De acordo com ele, análises de um pangenoma permite o levantamento dos genes *core* (compartilhados por todas as linhagens comparadas), acessórios (genes compartilhados por no mínimo duas mas não todas as linhagens estudadas), únicos (genes exclusivos de uma única linhagem – também conhecido com *singletons*) e exclusivos ausentes (genes ausentes em apenas uma das linhagens estudadas) (TETTELIN et al., 2005). Neste sentido, o *core* genome seria responsável pela codificação de aspectos básicos da biologia da espécie, como os genes codificantes de vias bioquímicas essenciais. Os genes identificados como acessórios e únicos podem estar relacionados a vias bioquímicas suplementares que conferem vantagens seletivas como adaptação ao ambiente, mecanismos de virulência, resistência a antibióticos e colonização de um novo hospedeiro (TETTELIN et al., 2005). As análises de genes exclusivos ausentes de uma determinada linhagem podem indicar a não essencialidade deste para a espécie estudada e, dependendo do volume de genes ausentes atribuídos a uma linhagem específica, a identificação de um possível encolhimento do genoma estudado em relação às espécies filogeneticamente mais próximas (CHAUDHARI, GUPTA & DUTTA, 2016). Desta maneira geral, as análises de pangenomas tem se mostrado uma abordagem interessante para definição de espécies de procariotos, organismos marcados por uma grande diversidade genética entre suas diferentes linhagens (CHAUDHARI, GUPTA & DUTTA, 2016).

A metabolômica pode ser entendida como o estudo de todos os compostos de um organismo num determinado instante sob a influência de alguma condição ambiental (COLQUHOUN, 2007). Existe uma grande diversidade de funções biológicas associadas aos metabólitos, como por exemplo, o armazenamento de energia química na forma de ATP. Os metabólitos são considerados os componentes finais do metabolismo celular, uma vez que a presença e concentração desses é reflexo direto da interação destes com o meio ambiente (COLQUHOUN, 2007). Assim, estes são altamente influenciados por alterações em mecanismos celulares regulatórios, como por exemplo, os eventos de transcrição de genes em moléculas de mRNA e tradução desses em proteínas (CULIBRK, CROFT & TEBBUTT, 2016).

De maneira geral, ferramentas aplicadas a estudos de metabolômica permitem a identificação dos grupos de metabólitos primários e secundários. Os metabólitos primários geralmente estão associados às funções fundamentais de uma célula, como por exemplo, geração de energia a partir da oxidação de açúcares, ácidos graxos, proteínas e ácidos nucleicos (VERPOORTE et al., 2007). Diferentemente dos metabólitos primários, metabólitos secundários não são essenciais para sobrevivência de um determinado organismo (DU FALL, SOLOMON, 2011). Estes compostos são estruturalmente diversos e podem estar associados a respostas diferenciadas a estímulos ambientais, vantagens adaptativas e interações ecológicas com demais organismos presentes no ambiente (DU FALL, SOLOMON, 2011).

3. OBJETIVOS

Com o intuito de melhor compreender particularidades genéticas e os principais mecanismos envolvidos na coexistência entre bactérias lácticas, acéticas e leveduras de *S. cerevisiae* em fermentações industriais brasileiras, o presente trabalho possui três grandes objetivos principais: (i) realizar uma caracterização genética, a partir de estudos genômicos NGS, de três linhagens bacterianas isoladas de usinas produtoras de etanol do interior do estado de São Paulo no Brasil. Estas linhagens foram denominadas *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3. Através dos estudos genômicos será possível mapear o conjunto total de genes destes microrganismos, bem como suas relações com espécies filogeneticamente semelhantes. Ademais, a comparação dos dados genéticos obtidos com outras linhagens bacterianas de *Lactobacillus* e *Acetobacter* permitirá o levantamento dos potenciais pangenomas associados a estes grupos de microrganismos. Neste sentido, através da obtenção de genes acessórios e exclusivos associados às linhagens de estudo, espera-se conseguir realizar o apontamento de mecanismos genéticos adaptativos associados a sobrevivência das bactérias em fermentações industriais brasileiras. (ii) realizar uma simulação de fermentações etanólicas industriais em reatores fermentativos de médio porte com reprodução dos processos de batelada alimentada, tratamento ácido e reciclo de células. Estes ensaios terão o controle de temperatura, pressão e aeração. Com a execução de diferentes condições experimentais, espera-se verificar alterações nas fermentações quando se comparar as condições teste (contaminadas com bactérias lácticas de *Lactobacillus* sp. I-2 e bactérias acéticas de *Acetobacter* sp. A-3) com a controle (fermentações sem contaminação bacteriana, contendo apenas leveduras de *S. cerevisiae*). Espera-se que a presença das bactérias alterem os teores de ácidos orgânicos (essencialmente ácido láctico, ácido acético e ácido succínico) presentes no fermentado. Além disso, acredita-se que a presença de concentrações significativas de bactérias nas fermentações irá agravar os estresses bióticos e abióticos no meio, levando a alterações em índices fermentativos clássicos como viabilidade celular e taxa de brotamento celular das leveduras de *S. cerevisiae*. (iii) caracterizar potenciais alterações metabólicas primárias nos metabolomas intracelulares dos microrganismos presentes nas fermentações realizadas em fermentadores biológicos. Para acessar os metabolomas primários serão utilizadas metodologias de cromatografia gasosa aplicada a espectrometria de massas (GC-MS). Provavelmente, a coexistência de bactérias e leveduras nas fermentações irá agravar a competição por nutrientes essenciais, aumentar os teores de compostos tóxicos no fermentado e potencialmente induzir processos como floculação. Estas variações poderão interferir nos metabolismos intracelulares dos microrganismos, alterando os teores de compostos como polióis, aminoácidos e açúcares. Espera-se os estudos apresentados neste trabalho agreguem informações acerca dos principais mecanismos genéticos associados a adaptação de bactérias as fermentações brasileiras e ajudem na elucidação das alterações fisiológicas envolvidas em fermentações que apresentem leveduras de *S. cerevisiae* e bactérias de *Lactobacillus* e *Acetobacter* como microrganismos fermentativos.

4. MATERIAL E MÉTODOS

4.1. Material biológico

Neste trabalho foram utilizadas leveduras e bactérias das seguintes linhagens: *S. cerevisiae* CAT-1, *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3. A levedura foi adquirida junto à LNF Latino Americana, sendo mantida em estoques apropriados nos laboratórios do CTBE/CNPEM. A bactéria láctica foi adquirida junto ao laboratório do professor Luiz Carlos Basso na ESALQ/USP. A bactéria acética foi isolada pela doutora Dra. Maria Letícia Bonatelli em seu trabalho de doutorado no Laboratório Max Feffer de Genética de Plantas da ESALQ/USP.

4.2. Cinética de crescimento e condições de cultivo das bactérias

As bactérias *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 foram transferidas de estoque mantidas em ultra-freezer -80 °C para placas de petri contendo meio MRS (DE MAN, ROGOSA & SHARPE, 1960) a uma concentração de 55 g.l⁻¹ acrescido de ágar 1% (m.v⁻¹). Após o crescimento em estufas a 30 °C por 48 horas, colônias isoladas foram inoculadas em meio MRS₀ e encubadas em shaker a 30 °C por 10 horas.

Bactérias lácticas são organismos anaeróbicos aerotolerantes, ou seja, o oxigênio presente no meio não é utilizado no metabolismo principal de compostos carbonados. Nesse sentido, para o cultivo de *Lactobacillus* sp. I-2, foram adotadas estratégias para minimizar o contato das bactérias com o oxigênio presente no meio e na área de *headspace* dos frascos de cultivo. Assim, padronizou-se a agitação de 50 rpm em *shaker* e utilização de frascos *erlenmeyers* de 25 ml de capacidade volumétrica contendo um volume final de 20 ml de MRS₀.

Bactérias acéticas são organismos aeróbicos obrigatórios, ou seja, utilizam o oxigênio presente no meio nos processos de oxidação de fontes carbonadas e obtenção de energia química na forma de ATP. Nesse sentido, para o cultivo de *Acetobacter* sp. A-3, foram adotadas estratégias para maximizar o contato das bactérias com o oxigênio presente no meio e na área de *headspace* dos frascos de cultivo. Assim, padronizou-se a agitação de 200 rpm em *shaker* e utilização de frascos *erlenmeyers* de 250 ml de capacidade volumétrica contendo um volume final de 20 ml de MRS₀.

Depois de decorrido o tempo de crescimento do pré-inóculo iniciou-se o cultivo das bactérias para a obtenção das cinéticas de crescimento. Para tanto, cerca de 1% do volume final do cultivo experimental era oriundo da cultura do pré-inóculo: 19,8 ml de meio MRS estéril acrescido de 200 µl do cultivo bacteriano do pré-inóculo. Após iniciado o experimento, amostras foram coletadas em intervalos de tempo regulares para mensuração da densidade óptica em comprimento de onda de 600 nanômetros (DO₆₀₀) e quantificação de unidades formadoras de colônia por mililitro (UFC.ml⁻¹). Foram utilizadas metodologias de plaqueamento *pour plate* e *spread plate* para *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente. Todas as análises foram realizadas em triplicata biológica com contagens de UFC.ml⁻¹ em triplicata técnica. Apenas cultivos que continham entre 30 e 300 colônias isoladas em placa foram considerados nas análises finais.

4.3. Estudos filogenéticos do gene 16S do rDNA das bactérias

As linhagens *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 foram cultivadas em MRS por 12 horas para a obtenção de biomassa celular. Na extração do DNA total foi utilizado o “FastDNA® SPIN for Soil kit” seguindo as recomendações do fabricante. Para avaliar a pureza das amostras de DNA extraídas e quantificá-las utilizaram-se análises eletroforéticas em gel de agarose 1% (m.v⁻¹) e obtenção do espectro de absorbância em espectrofotômetro Nanodrop®.

As amplificações do gene 16S do rDNA foram realizadas utilizando o DNA bacteriano extraído como *template*. A “Polymerase Chain Reaction” (PCR) envolveu a adição dos seguintes reagentes: 1 µl do DNA genômico extraído (50 ng.µl⁻¹), 1 µl da solução de dNTPs (10mM.µl⁻¹), 1 µl de primer 16S_F (10pmol.µl⁻¹), 1 µl de primer 16S_R (10pmol.µl⁻¹), 10 µl da solução tampão contendo MgCl₂, 0,5 µl de Taq DNA Polimerase, 1,5 µl Dimetilsulfóxido (DMSO) e água ultrapura estéril para completar o volume final de 50 µl. A reação foi realizada em termociclador, no qual se programou uma etapa inicial a 98 °C por 3 minutos, seguido de 25 ciclos: 98 °C por 30 segundos para desnaturação da dupla fita do DNA, 55°C por 1 minuto para anelamento dos primers e 72 °C por 1 minuto para polimerização nucleotídica. Foi adicionada uma etapa final de extensão a 72 °C por 1 minuto. A reação final foi mantida a 4°C até a retirada do termociclador. O produto da reação foi analisado em gel de agarose 1% (m.v⁻¹) juntamente com marcador de peso molecular 1Kb (Fermentas®). Os fragmentos amplificados foram purificados com o “Wizard SV Gel and PCR Clean-Up System kit” e então sequenciados nos laboratórios do CTBE/CNPEM. As seqüências nucleotídicas dos *primers* utilizados tanto na reação de amplificação como nas reações de sequenciamento são apresentadas na tabela 1.

Tabela 1. *Primers* utilizados no sequenciamento do gene 16S do rDNA (LANE, 1991; TURNER et al., 1997).

Nome	Sequência nucleotídica (5' → 3')
16S_F	5' - AGA GTT TGA TYM TGG CTC AG - 3'
16S_R	5' - CGG TTA CCT TGT TAC GAC TT - 3'
907_R	5' - CCG TCA ATT CMT TTR AGT TT - 3'

A leitura dos cromatogramas resultantes do sequenciamento e obtenção das seqüências nucleotídicas com um valor de qualidade associado a cada pico cromatográfico foi feita com o auxílio do programa Phred (EWING & GREEN, 1998). A trimagem das seqüências de baixa qualidade foi realizada com o auxílio do programa trimseq (RICE & BLESBY, 2000). A montagem do DNA foi feita pelo programa CAP3 (HUANG & MADAN, 1999). Utilizando os *contigs* gerados e o *webserver* EZBioCloud, foram obtidas seqüências do gene 16S de diferentes linhagens de *Lactobacillus* e *Acetobacter* (YON et al., 2017). O programa MXSCARNA foi utilizado para geração dos alinhamentos nucleotídicos, com predições das estruturas primárias e secundárias do rDNA consenso (TABELI et al., 2008). A filogenia foi realizada através do programa RaxML (STAMATAKIS et al., 2006) com adoção dos modelos evolutivos “Generalised time-reversible” (GTR) para a regiões sem a presença de estruturas secundárias e S16A para regiões com estruturas secundárias e abordagens de máxima verossimilhança (STAMATAKIS et al., 2006). A visualização gráfica final foi feita com o auxílio do *webserver* Itol destacando os valores de *bootstraps* associados a cada ramo. Apenas valores acima de 50 foram mostrados na árvore filogenética final (LETUNIC & BORK, 2016).

4.4. Cariotipagem da levedura *S. cerevisiae* CAT-1

As leveduras de *S. cerevisiae* utilizadas nas fermentações tiveram o seu perfil cromossômico comparado com diferentes linhagens comerciais utilizadas em usinas brasileiras de produção de etanol: SA-1, CAT-1, PE-2, BG e FT-858. Para tanto, amostras de cultivo da linhagem de *S. cerevisiae* presente nos estoques dos laboratórios do CTBE/CNPEM foram enviadas ao Centro Pluridisciplinar de Pesquisas Químicas, Biológicas e Agrícolas da Universidade Estadual de Campinas (CPQBA/UNICAMP). O relatório final apresentou número de processo: CPQBA 1270/16 DBP.

4.5. Obtenção e preparo do melaço de cana-de-açúcar

Cerca de 120 Kg de melaço de cana-de-açúcar *in natura* a 85 graus Brix (°Bx) foi doado pela empresa Raízen® - Piracicaba, São Paulo. A clarificação do melaço foi realizada na Planta Piloto para Desenvolvimento de Processos (PPDP) do CTBE/CNPEM. Para tanto, o melaço foi diluído a 55 °Bx obtendo um volume de trabalho de aproximadamente 160 litros. Inicialmente, em tanque com capacidade de aquecimento e agitação, foram adicionados volumes controlados de ácido fosfórico (H_3PO_4) até alcançar o pH 5,40. Em seguida, adicionou-se solução de óxido de cálcio (CaO) até o pH 6,4 ser estabelecido. Foi realizado o aquecimento do tanque a 60°C por cerca de 5 minutos, mantendo agitação suficiente para garantir a completa homogeneização do sistema. Em seguida, adicionou-se o polímero Biocane Floc DC02 (BIOCANE®) (1 g.l^{-1}) obedecendo à proporção de 4 mililitros de solução polimérica por litro de melaço clarificado. Este polímero é composto por poliácridamida aniônica e sua interação com o H_3PO_4 gera um agrupamento dos flocos de impurezas presentes no melaço. O processo de decantação deste material é potencializado pela adição de CaO. Após 5 minutos de agitação, iniciou-se o processo de separação das impurezas particuladas com o auxílio de equipamento filtro prensa Bomax, modelo Prensamax. O melaço clarificado foi esterilizado em reator BioFlo® 610 a 121 °C por 15 minutos e estocado em *carboys* de 50 litros. Por fim, amostras foram coletadas e tiveram as concentrações de ácidos orgânicos (ácido acético, ácido láctico e ácido succínico), açúcares (sacarose, glicose e frutose) e etanol mensurados via cromatografia líquida de alta eficiência (HPLC) por índice de refração, usando uma coluna Aminex HPX-87H (300 mm x1,00UM, 7,8 mm x 9 UM, Bio-rad®) com temperatura de detector de coluna igual a 35°C e H_2SO_4 a 5 mM como fase móvel.

4.6. Propagação celular da levedura *S. cerevisiae* CAT-1

O cultivo do inóculo foi realizado em 180 ml de meio YPD contendo 20 g.l^{-1} de glicose, 10 g.l^{-1} de extrato de levedura e 20 g.l^{-1} de peptona. As leveduras de *S. cerevisiae* CAT-1, mantidas em placas de petri contendo YPD acrescido de agar 1% (m.v^{-1}), foram transferidas para o meio líquido de forma asséptica e encubadas em *shaker* por 12h, 200 rpm a 30 °C. Metodologia adaptada de Silva et al. (2012).

O processo de propagação celular ocorreu em reator New Brunswick BioFlo®/CelliGen® 115 de 3 litros de capacidade volumétrica esterilizado a 121 °C por 45 minutos. Após o tempo de cultivo do inóculo, todo o volume foi transferido para o reator juntamente com melaço de cana-de-açúcar, fosfato monopotássico (KH_2PO_4) e ureia (CH_4N_2O). Estes dois últimos compostos são fontes suplementares de fósforo e nitrogênio, uma vez que o melaço de cana-de-açúcar apresenta baixos teores destes nutrientes (SANTOS et al., 2015). A propagação celular da levedura

foi dividida em duas etapas distintas: uma batelada de 12 horas seguida de uma batelada alimentada de 24 horas (Figura 1). A composição, volumes e concentrações do meio de cultivo da etapa de batelada foram apresentados na tabela 2 (SANTOS et al., 2015).

Depois de decorrido o tempo da batelada, foi adicionado no reator 60 ml de fosfato monopotássico (1 g.l^{-1}) e 54 ml de ureia (3 g.l^{-1}). Em seguida, iniciou-se a fase de batelada alimentada com alimentação de 400 ml de melaço (30°Bx) em 24 horas. O volume final de propagação celular da levedura foi de aproximadamente 1,8 litros (SANTOS et al., 2015).

O reator foi programado para manter o percentual de oxigênio dissolvido (OD) no meio de cultivo em torno de 50%. Esse controle objetivou assegurar o crescimento dos microrganismos em condições de aerobiose. Neste sentido, sempre que necessário, houve alterações nas intensidades de injeção de ar e velocidades de agitação do sistema. A propagação ocorreu à temperatura de 30°C , agitação e aeração inicial de 200 rpm e entrada de 0,5 litros de ar por minuto (LPM), respectivamente. A agitação poderia alcançar o limite máximo de 700 rpm de acordo com a demanda por homogeneização dos gases no cultivo microbiano. A aeração foi uma combinação de ar atmosférico e oxigênio molecular puro e teve o percentual de mistura definido automaticamente de acordo com o OD do sistema. Neste sentido, os percentuais mínimos programadas foram de 95% e 5% para ar atmosférico e oxigênio molecular puro, respectivamente. No entanto, quando necessário, alterações na injeção da mistura poderiam ocorrer até o limite de composição alcançar 50% de oxigênio molecular puro. As metodologias utilizadas na propagação celular das leveduras de *S. cerevisiae* CAT-1 foram adaptadas de Santos et al. (2015).

Tabela 2. Meio de cultivo utilizado na fase batelada da propagação celular de *S. cerevisiae* CAT-1.

Reagente	Volume (ml)	Concentração final (g.l^{-1})
Melaço (30°Bx)	171,5	4
KH_2PO_4	128,6	3
$\text{CH}_4\text{N}_2\text{O}$	64,3	5
H_2O	741,6	-
Total	1286	-

4.7. Ensaios fermentativos em reatores

Três condições fermentativas foram executadas neste trabalho: (i) fermentações contendo leveduras de *S. cerevisiae* CAT-1; (ii) fermentações contendo leveduras de *S. cerevisiae* CAT-1 e bactérias de *Lactobacillus* sp. I-2; e (iii) fermentações contendo leveduras de *S. cerevisiae* CAT-1 e bactérias de *Acetobacter* sp. A-3. Os experimentos foram executados em triplicata biológica com adoção dos processos de batelada alimentada em alta densidade celular com tratamento ácido com ácido diluído e reciclo de células (Figura 2).

Depois de decorridas 36 horas da propagação celular, o material biológico contido no reator foi submetido à centrifugação em equipamento Beckman Coulter J26-P a 7000 RPM por 30 minutos. Cerca de 150 gramas do precipitado úmido (aproximadamente 90 g.l^{-1} de massa seca) foi ressuscitado em água estéril até ser alcançado o volume final de 400 ml. O reator utilizado nesta etapa é o mesmo da propagação celular da levedura alterando-se apenas o volume final de trabalho para 1,2 litros. As fermentações foram executadas imediatamente após a conclusão do processo de propagação celular, sem armazenamento intermediário em câmara fria. A adoção desta estratégia teve como objetivo evitar variações experimentais entre tratamentos em decorrência da idade celular

do inóculo e tempo de exposição a diferentes estresses, principalmente os causados por baixas temperaturas (Figura 2).

As fermentações caracterizaram-se pelo bombeamento de 400 ml da suspensão de leveduras (pé-de-cuba) para o reator, um tratamento ácido com a adição de H_2SO_4 (até se alcançar o pH 2,0) e alimentação do sistema com aproximadamente 800 ml de melaço de cana-de-açúcar a 20 °Bx, medidos em refratômetro digital (SANTOS et al., 2015). A alimentação foi realizada com bomba peristáltica. A etapa de tratamento ácido foi iniciada imediatamente após a conclusão da transferência do pé-de-cuba para o reator e durou cerca de 30 minutos a 30 °C, agitação de 200 rpm e injeção de ar de 0,5 LPM (SANTOS et al., 2015). Depois de decorrido o tempo do tratamento ácido, iniciou-se a etapa de alimentação com melaço por 6 horas. Cada ciclo fermentativo apresentou duração total de 8h com agitação de 200 rpm e temperatura de 30 °C. Não houve injeção de ar durante as fermentações (Figura 2) (SANTOS et al., 2015).

Uma vez concluída a fermentação, o vinho foi centrifugado a 7000 rpm por 30 minutos para recuperação do precipitado celular. Esse foi ressuspensionado em água estéril até a obtenção de 400 ml de suspensão microbiana e, em seguida, iniciou-se uma nova batelada alimentada. Foram realizados três ciclos fermentativos consecutivos (Figura 2) (SANTOS et al., 2015).

A infecção com as bactérias ocorreu no tempo de 2h da primeira batelada alimentada com a adição de bactérias suficientes para se obter 10^8 UFC.ml⁻¹ em 1,2 litros de cultivo microbiano. Para tanto, a obtenção de biomassa bacteriana foi realizada simultaneamente a propagação celular da levedura. A determinação da quantidade celular de bactérias a ser inoculada no reator foi calculada utilizando dados de DO₆₀₀ medidos em espectrofotômetro. Para tanto, utilizaram-se as equações lineares obtidas nas cinéticas de crescimento bacterianas descritas no item 3.2.

No final do processo de propagação celular da levedura *S. cerevisiae* CAT-1 foi mensurada a massa seca para avaliações da concentração celular deste microrganismo. Nas fermentações foram realizadas amostragens a cada 4 horas com determinação das concentrações de açúcares (sacarose, frutose e glicose), álcoois (etanol e glicerol) e ácidos orgânicos (ácido láctico, ácido acético e ácido succínico). Além disso, foi mensurada a concentração de biomassa celular por massa seca nos pontos de início (0h), meio (4h) e final (8h) de cada ciclo fermentativo. A contagem relativa em câmara de *Neubauer* de células vivas, mortas, brotos vivos e mortos foi realizada no início (0h) e final (8h) de cada batelada alimentada. As colorações das células mortas foram feitas com azul de metileno. A média e desvio padrão foram calculados para cada uma das amostras avaliadas. Valores que ultrapassaram a média cerca de duas vezes o desvio padrão foram descartados das análises de índice de viabilidade celular médio e taxa de brotamento celular médio. Na terceira batelada alimentada, em todos os pontos amostrados, além dos dados anteriores, foram realizadas contagens de UFC.ml⁻¹ de leveduras em meio YPD acrescido de ampicilina (200 µg.ml⁻¹) e bactérias em meio MRS acrescido de cicloheximida (200 g.ml⁻¹). Análises de variância (ANOVA) e o teste de Tukey foram utilizados para verificar possíveis diferenças estatísticas entre os tratamentos estudados.

A determinação da concentração celular feita por massa seca caracterizou-se por análise gravimétrica, em que um mililitro do precipitado obtido nas amostragens foi centrifugado, lavado e submetido à secagem em estufa a 80°C por 48 horas. A massa final foi obtida fazendo-se a diferença entre as massas dos tubos iniciais vazios e as massas dos tubos acrescidos do precipitado seco após 48 horas de estufa. As determinações das concentrações dos açúcares, álcoois e ácidos orgânicos foram feitas por meio de HPLC por índice de refração, usando uma coluna Aminex HPX-87H (300 mm x1,00UM, 7,8 mm x 9 UM, Bio-rad®) com temperatura de detector de coluna igual a 35°C e H_2SO_4 a 5 mM como fase móvel.

4.8. Coleta de amostras para metabolômica

Por fim, amostras foram coletadas para análises de metabolômica em todos os pontos da terceira batelada alimentada. Nesta etapa, realizou-se uma centrifugação inicial a 3220 g por 10 minutos a 4°C para separação do precipitado celular e sobrenadante. O sobrenadante foi aliqotado em microtubos previamente tratados com metanol e congelados em nitrogênio líquido. O precipitado foi ressuspensionado em solução salina 0,9% NaCl (0°C), aliqotado em microtubos previamente tratados com metanol e centrifugado novamente a 16000 g por 5 minutos a 4°C. Após o descarte do sobrenadante, o precipitado foi congelado em nitrogênio líquido e armazenado em ultrafreezer -80 °C (Figura 2).

Figura 2. Modelo esquemático do desenho experimental executado nos estudos fermentativos. Três condições foram avaliadas: (i) fermentações contendo leveduras de *S. cerevisiae* CAT-1, (ii) fermentações contendo leveduras de *S. cerevisiae* CAT-1 e *Lactobacillus* sp. I-2 e (iii) fermentações contendo leveduras de *S. cerevisiae* CAT-1 e *Acetobacter* sp. A-3. Cada condição apresentou um tempo total de 60 horas de cultivo microbiano, divididos em 36 horas de propagação celular e 24 horas de fermentações. A propagação celular foi dividida em duas etapas: 12 horas em fase de batelada e 24 horas em fase de batelada alimentada. Cada ciclo fermentativo teve duração de 8 horas com alimentação do reator com melaço de cana-de-açúcar nas 6 primeiras horas. A contaminação bacteriana, quando necessária, aconteceu na primeira batelada alimentada em 2 horas de fermentação. Foram realizadas amostragens no início (0h), meio (4h) e final (8h) de cada ciclo fermentativo.

4.9. Sequenciamento, montagem e anotação dos genomas bacterianos

As linhagens bacterianas *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3 foram transferidas de estoques mantidos em ultra freezer -80 °C para placas de petri contendo meio MRS (DE MAN, ROGOSA & SHARPE, 1960) a uma concentração de 55 g.l⁻¹ acrescido de ágar 1% (m.v⁻¹). Após o crescimento em estufas a 30 °C

por 48 horas, colônias isoladas foram inoculadas em MRS₀ e encubadas em shaker a 30 °C por 10 horas para obtenção da biomassa celular. O processo de extração do DNA total foi realizado com o “FastDNA[®] SPIN for Soil kit” seguindo as recomendações do fabricante. Para avaliar a pureza das amostras de DNA extraídas e quantificá-las utilizaram-se análises eletroforéticas em gel de agarose 1% (m.v⁻¹) e obtenção do espectro de absorbância em espectrofotômetro Nanodrop[®].

Os genomas das bactérias foram sequenciados por meio de tecnologia de sequenciamento de nova geração (NGS). Bibliotecas genômicas foram sintetizadas utilizando o “Nextera DNA Library Preparation Kit” (Illumina[®]), 58 °C de temperatura de tagmentação e 30 nanogramas (ng) do DNA total extraído. A reação de sequenciamento foi realizada em equipamento HiSeq 2500 produzindo 2X100 *reads* nos laboratórios do CTBE/CNPEN.

As qualidades das reações de sequenciamento foram avaliadas com o auxílio do programa FastQC (ANDREWS, 2010). A trimagem dos dados foi feita com o programa Trimmomatic (BOLGER, LOHSE & USADEL., 2014). Com o programa Kmergenie foi possível estimar o melhor kmer para montagem dos genomas (CHIKHI & MEDVEDEV, 2013). As montagens dos genomas foram realizadas com a combinação dos programas “*St. Petersburg genome assembler*” (SPADES) e Velvet (ZERBINO & BIRNEY, 2008; BANKEVICH et al., 2012). O programa SPADES estimou o tamanho do inserto das bibliotecas genômicas (BANKEVICH et al., 2012). As estatísticas das diferentes montagens foram obtidas com o programa QUAST (GUREVICH et al., 2013). Para identificar o quão completo os genomas estavam em termos de composição gênica, a presença de diferentes genes cópia única de bactérias foi verificada com o programa BUSCO. (SIMÃO et al., 2015). A presença de possíveis *reads* de organismos contaminantes foi verificada com “*Taxon-annotated GC-coverage plot*” (TAGC) (KUMAR et al., 2013). O preenchimento de *gaps* e correções de erros de alinhamento de *contigs* foram feitos com o programa Pilon (WALKER et al., 2014). O programa PEP_scaffolder foi utilizado na obtenção de *scaffolds* a partir dos *reads* originais (ZHU et al., 2016). Fragmentos menores que 1 kb foram removidos e então os *scaffolds* finais foram anotados automaticamente pelo programa Prokka (SEEMANN, 2014).

4.10. Análises taxonômicas e filogenéticas das linhagens bacterianas

Konstantinidis & Tiedje (2005) discutem que uma das estratégias mais robustas para identificação de similaridades genéticas entre dois genomas de procariotos é a mensuração da identidade média nucleotídica (“*genome-wide Average Nucleotide Identity*” - gANI). Além disso, a identificação da fração de alinhamento de genes ortólogos (“*Alignment Fraction*” - AF) pode ser utilizada como medida de mensuração de parentesco genético entre conteúdos gênicos de um par de genomas (VARGHESE et al., 2015). Nesse sentido, foi realizada a metodologia proposta por Varghese et al. (2015) nos estudos taxonômicos de diferentes linhagens de *Lactobacillus* e *Acetobacter*, incluindo *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3.

Potenciais proteínas preditas (PPPs) de 52 linhagens de *Lactobacillus* e 28 de *Acetobacter* foram utilizadas nos estudos filogenéticos, incluindo *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3. Essas estão disponíveis no banco “*Reference Sequence*” (RefSeq) do “*National Center for Biotechnology Information*” (NCBI). O alinhamento das sequências foi feito com o programa Diamond, via algoritmo BLASTp (BUCHFINK, XIE & HUSON, 2015). A identificação dos grupos de genes ortólogos foi realizada com o programa FastOrtho (<http://enews.patricbrc.org/fastortho/>). Somente PPPs presentes em cópia única em todas as linhagens comparadas foram utilizadas como marcadores nos estudos filogenéticos. Com o programa Mafft foi possível obter os múltiplos

alinhamentos de sequências, bem como os modelos evolutivos associados a cada marcador filogenético (KATO & STANDLEY, 2013). A filogenia foi realizada por meio do programa RaxML utilizando abordagens de máxima verossimilhança (STAMATAKIS et al., 2006). A visualização gráfica foi feita com o auxílio do *webserver* ItoI destacando os valores de *bootstraps* associados a cada ramo. Apenas valores acima de 50 foram mostrados na árvore filogenética final (LETUNIC & BORK, 2016).

4.11. Estudos pangenômicos

Potenciais proteínas preditas (PPPs) de diferentes linhagens de *Lactobacillus* e *Acetobacter* obtidas no banco de dados RefSeq do NCBI foram utilizadas nos estudos pangenômicos. Com o programa BPGA foram comparados 26 linhagens de *L. fermentum*, incluindo *Lactobacillus* sp. I-2; 13 linhagens de *L. plantarum* e *L. paraplantarum*, incluindo *Lactobacillus* sp. L-10; e 8 linhagens de *A. tropicalis* e *A. senegalensis*, incluindo *Acetobacter* sp. A-3. O pré-processamento com a formação de *clusters* de sequências proteicas foi realizado com o uso do algoritmo USEARCH, adotando o percentual de 50% como limite de identidade mínima (CHAUDHARI, GUPTA & DUTTA, 2016). PPPs *core*, acessórias, únicas e exclusivas ausentes foram determinadas para cada uma das linhagens estudadas. As proteínas classificadas como *core* possivelmente estão presentes em todas as linhagens comparadas. São classificadas como acessórias PPPs encontradas, no mínimo, em duas linhagens de bactérias, mas não em todas as analisadas (CHAUDHARI, GUPTA & DUTTA, 2016). As proteínas classificadas como únicas potencialmente são codificadas por genes presentes em uma única linhagem e as PPPs exclusivas ausentes são aquelas que não foram encontradas nos potenciais proteomas de apenas uma linhagem do pangenoma. Por fim, “*Clusters of Orthologous Groups*” (COGs) foram apontados para PPPs exclusivas e acessórias de *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3 (CHAUDHARI, GUPTA & DUTTA, 2016).

4.12. Extração de metabólitos intracelulares e derivatização

Para os estudos de metabolômica foram utilizadas as amostras de precipitado celular coletadas durante o terceiro ciclo fermentativo dos ensaios fermentativos realizados em fermentadores biológicos. Foi realizada a extração de metabólitos intracelulares das amostras de início (0h), meio (4h) e final (8h) das três repetições biológicas de cada uma das três condições experimentais testadas. Um total de 27 amostras foram analisadas por cromatografia gasosa aplicada à espectrometria de massas (GC-MS).

A extração dos metabólitos intracelulares foi baseada na metodologia proposta por Gullberg et al., 2004. As amostras para precipitado celular estocadas em ultra freezer -80 °C foram liofilizadas a -50 °C (Thermo Eletron Corporation – ModulyoD Freeze Dryer) por 48h. Cerca de 100 mg deste material foi triturado em almofariz com nitrogênio líquido e transferido para microtubos de 2 ml previamente tratados com metanol. Foram adicionados 750 µl de metanol contendo ácido succínico como padrão interno (50 pmol). Em seguida, adicionaram-se 750 µl de água ultrapura gelada, sempre realizando homogeneizações periódicas em *vortex*. As amostras foram sonicadas em banho de gelo por 30 minutos. Após a conclusão da etapa de sonicação, as amostras foram mantidas em ultrafreezer -80°C por 30 minutos, seguida de homogeneização em *vortex*. As etapas de estoque em ultrafreezer e homogeneização foram repetidas três vezes. O material foi centrifugado a 15000 g por 15 minutos e teve o seu sobrenadante coletado

para um novo tubo. A solução coletada foi filtrada com auxílio de filtro Millex (MilliQ) de “polyvinylidene difluoride” (PVDF) com 0,22 µm de porosidade e armazenadas em tubos de vidro (*vial* 1.5 ml) em ultra freezer -80 °C.

Cerca de 250 µl da amostra de metabólitos extraída foi transferida para outro *vial* e seca em equipamento liofilizador (Thermo Eletron Corporation – ModulyoD Freeze Dryer) a -50 °C por 6 horas. Em seguida, iniciou-se o processo de derivatização com a adição de 30 µl de solução de metoxiamina em pirimidina (15 mg.ml⁻¹). As amostras foram homogeneizadas em *vortex* e mantidas em repouso, em condições de escuro, por 16 horas a temperatura ambiente. Em seguida foi realizada a etapa de silição com a adição de 30 µl de N-metil-trimetilsilil-trifluoroacetamida (MSTFA) com 1% de trimetilclorosilano (TMCS). Após nova homogeneização em *vortex*, essas foram repousadas a temperatura ambiente por 1 hora. Por fim, foram adicionados 30 µl de Heptano e as amostras foram injetadas em cromatógrafo gasoso acoplada à espectrometria de massas (GC-TOFMS) de acordo com a metodologia descrita em Jonsson et al. (2005). Para o cálculo dos índices de retenção foram utilizadas amostras controles (brancos) e um *mix* de alcanos (C12 – C40) (SCHAUER et al., 2005).

4.13. Análises em GC-MS

Foi injetado automaticamente 1 µl das amostras derivatizadas em cromatógrafo gasoso (7890A Agilent Technologie), modo *splitless*. Esse utiliza colunas de sílica fundida com 20 m de comprimento x 0,18 mm de diâmetro interno x 0,10 µm de filme (AGILENT J & W SCIENTIFIC). A temperatura de injeção foi de 280 °C e após 250 segundos do início da aquisição dos dados iniciou o fluxo de hélio a 1 ml.min⁻¹. A temperatura inicial da coluna foi de 80 °C, mantida por 2 minutos e aumentada 15 °C por minuto até atingir 305 °C, a qual foi mantida por mais 10 minutos. O efluente da coluna foi introduzido na fonte de íons (impacto de elétrons) do equipamento GC x GC-TOFMS (Pegasus 4D, Leco). Utilizou-se 250 °C e 70-eV como temperatura da fonte de íons e feixe de elétrons, respectivamente. A corrente de ionização foi de 2 mA. Foram registrados espectros na faixa de massa-carga (*m/z*) 45-800 e a voltagem do detector utilizada foi a de 1500V.

4.14. Processamento dos dados, identificação dos metabólitos e análises estatísticas

Após a conclusão das análises em GC-MS, os cromatogramas gerados foram processados no programa ChromaTOF, versão 4.32 (LECO®). Neste foi realizada a correção da linha de base (*threshold*) para remoção dos ruídos gerados na corrida. Em seguida, utilizando o pacote TargetSearch e o programa R, foram feitas as etapas de deconvolução, obtenção do índice de retenção (IR), correção do tempo de retenção (RT), identificação e alinhamento dos picos cromatográficos e identificação dos metabólitos (CUADROS-INOSTROZA, et al., 2009; R DEVELOPMENT CORE TEAM, 2008). A identificação final dos metabólitos foi feita com o auxílio do banco de dados “The Golm Metabolome Database” (GMD), disponível em (<http://gmd.mpimp-golm.mpg.de/>). Apenas foram considerados metabólitos válidos aqueles presentes em duas ou mais repetições técnicas. A intensidade de cada metabólito foi normalizada pela contagem iônica total (TIC) de cada amostra. Por fim, informações sobre os números de referências no banco de dados “Kyoto Encyclopedia of Genes and Genomes” (KEGG) foram coletadas para os metabólitos identificados previamente nas comparações com o banco GMD.

As análises estatísticas comparativas entre os tratamentos foram realizadas no *webserver* MetaboAnalyst (XIA et al., 2015). Foram feitas comparações par a par das amostras controle e teste coletadas nos pontos de início

(0h), meio (4h) e final (8h) das fermentações. As condições comparadas foram: (i) “*S. cerevisiae* CAT-1 vs *S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2”; e (ii) “*S. cerevisiae* CAT-1 vs *S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3”. Para tanto, procedeu-se a filtragem dos dados por “intervalo interquartil” (“*Interquantile Range*” – IQR), redimensionamento dos dados para $\log_2$ e escalonamento por pareto (“*Pareto scaling*”). Os dados pré-processados foram analisados por teste t usando um p -valor menor ou igual a 0,05. Uma tabela volcano foi gerada utilizando “*Fold Change threshold*” mínimo de 4 e p -valor (teste t) menor ou igual a 0.05. Foi utilizada a análise multivariada de componentes principais (“*Principal Component Analysis*” – PCA) para cada uma das comparações par a par realizadas. Essa permitiu a identificação de potenciais repetições *outliers*, homogeneidade e tendência de separação dos grupos teste e controle comparados (SENGUPTA et al., 2011).

5. RESULTADOS E DISCUSSÕES

5.1. Cinética de crescimento das bactérias

Os estudos das cinéticas de crescimento das linhagens *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, em meio MRS, exigiram uma série de adaptações experimentais para sua execução laboratorial. Devido ao fato de serem anaeróbicas facultativas, as bactérias lácticas de *Lactobacillus* sp. I-2 apresentam melhor crescimento em cultivos com pouca aeração, manifestadas em adaptações como reduções na velocidade de agitação em *shaker* e área de *headspace* nos frascos de cultivo. Em contraste, as bactérias acéticas apresentam as melhores taxas de crescimento quando submetidas às condições de alta agitação em *shaker* e grande área de *headspace* nos frascos de cultivo. Além disso, somente foi possível mensurar o tamanho das populações destas bactérias quando adotei as metodologias de plaqueamento, distintas para cada um dos organismos: “*pour plate*” e “*spread plate*” para *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente.

O meio de cultura microbiano MRS é um dos mais utilizados para o cultivo de bactérias lácticas com poucos relatos de cultivo de bactérias acéticas (DE MAN, ROGOSA & SHARPE, 1960). Os estudos de cinética microbiana mostraram que as bactérias *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 conseguem crescer bem neste meio. De maneira geral, observou-se uma fase lag curta (2h a 4h) seguida de uma fase log com duração entre 8h e 12h (Figura 3A). Além disso, os dados de plaqueamento mostraram que na fase estacionária as bactérias alcançaram tamanhos populacionais em torno de 10^9 (UFC.ml⁻¹) (Figura 3B). Quando comparadas, às contagens de células viáveis (UFC.ml⁻¹) e dados de medida de absorbância (OD₆₀₀), mostraram correlação linear durante a fase log da cinética microbiana com coeficiente de correlação (r^2) em torno de 95% (Figura 3C). Desta maneira, foi utilizada esta fase da cinética microbiana para estimar equações lineares capazes de determinar a concentração celular bacteriana, em UFC.ml⁻¹, presentes numa determinada medida de OD₆₀₀ (Figura 3D).

Figura 3. Cinética de crescimento das linhagens *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 em meio MRS. A) Valores de absorvância (OD₆₀₀) mensurados ao longo do tempo. B) Número de células viáveis (UFC.ml⁻¹) contadas ao longo do tempo. C) Correlação linear entre os dados de absorvância e contagem de células viáveis na fase logarítmica. D) Equações lineares utilizadas para estimar quantitativamente, por dados de absorvância, o número de células viáveis presentes no cultivo. As contagens de células viáveis foram feitas utilizando metodologias de plaqueamento por profundidade (“pour plate”) e superfície (“spread plate”) para *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente.

5.2. Filogenia do gene 16S do rDNA das bactérias

A realização do sequenciamento pela plataforma Sanger do gene 16S do rDNA das linhagens *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3 gerou *contigs* de cerca de 1500 pares de base (pb). A comparação das seqüências obtidas no sequenciamento com as de diferentes linhagens de *Lactobacillus* e *Acetobacter* depositadas no banco de dados do *webserver* EzBioCloud permitiu o levantamento das espécies filogeneticamente mais semelhantes as linhagens bacterianas. De maneira geral, os estudos filogenéticos com o gene 16S do rDNA mostraram pouca resolução na identificação das espécies bacterianas mais semelhantes às linhagens *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3. Em constraste, para a bactéria láctica *Lactobacillus* sp. I-2, houve o agrupamento de apenas uma espécie no mesmo ramo filogenético. Assumindo o gene 16S do rDNA e os modelos evolutivos GTR e S16A foi possível apontar as espécies bacterianas mais próximas de *Lactobacillus* sp. L-10: *L. paraplantarum*, *L. plantarum*, *L. pentosus*, *L. plajomi*, *L. modestisalitolans* e *L. herbarum*. Diferentemente de *Lactobacillus* sp. L-10, a linhagem *Lactobacillus* sp. I-2

apresentou a espécie *L. fermentum* como a mais próxima filogeneticamente. As espécies *A. siceræ*, *A. tropicalis* e *A. senegalensis* foram as que se mostraram mais semelhantes à linhagem *Acetobacter* sp. A-3. As linhagens *Komagataeibacter xylinus* NBRC 15237 e *Pediococcus pentosaceus* DSM 20336 foram utilizadas como grupos externos nas filogenias de *Acetobacter* e *Lactobacillus*, respectivamente (Figura 4).

Figura 4. Filogenias do gene 16S do rDNA para as linhagens bacterianas A) *Lactobacillus* sp. L-10 B) *Acetobacter* sp. A-3 e C) *Lactobacillus* sp. I-2. Fragmentos de aproximadamente 1500 pb foram sequenciados usando a plataforma Sanger. *L. paraplantarum*, *L. plantarum*, *L. pentosus*, *L. plajomi*, *L. modestisalitolterans* e *L. herbarum* são filogeneticamente próximas de *Lactobacillus* sp. L-10. *L. fermentum* é a única espécie presente no ramo filogenético de *Lactobacillus* sp. I-2. As espécies *A. siceræ*, *A. tropicalis* e *A. senegalensis* são as mais semelhantes à *Acetobacter* sp. A-3. As linhagens *Komagataeibacter xylinus* NBRC 15237 e *Pediococcus pentosaceus* DSM 20336 foram utilizadas como grupos externos nas filogenias de *Acetobacter* e *Lactobacillus*, respectivamente.

Figura 4. Filogenias do gene 16S do rDNA para as linhagens bacterianas A) *Lactobacillus* sp. L-10 B) *Acetobacter* sp. A-3 e C) *Lactobacillus* sp. I-2. Fragmentos de aproximadamente 1500 pb foram sequenciados usando a plataforma Sanger. *L. paraplantarum*, *L. plantarum*, *L. pentosus*, *L. plajomi*, *L. modestisalitolterans* e *L. herbarum* são filogeneticamente próximas de *Lactobacillus* sp. L-10. *L. fermentum* é a única espécie presente no ramo filogenético de *Lactobacillus* sp. I-2. As espécies *A. siceae*, *A. tropicalis* e *A. senegalensis* são as mais semelhantes à *Acetobacter* sp. A-3. As linhagens *Komagateibacter xylinus* NBRC 15237 e *Pediococcus pentosaceus* DSM 20336 foram utilizadas como grupos externos nas filogenias de *Acetobacter* e *Lactobacillus*, respectivamente.(Continuação).

5.3. Estudos eletroforéticos *fingerprint* de *S. cerevisiae* CAT-1

As análises de eletroforese em campo pulsado para as leveduras utilizadas nos ensaios fermentativos mostraram perfis cromossômicos semelhantes à linhagem comercial *S. cerevisiae* CAT-1. A figura 5 apresenta o gel obtido durante a realização do processo de cariotipagem pelo CPQBA/UNICAMP. O laudo final apresentado discriminou a linhagem utilizada nos laboratórios do CTBE/CNPEM como detentora do perfil característico de *S. cerevisiae* CAT-1.

Figura 5. Análises eletroforéticas de campo pulsado de seis linhagens de *S. cerevisiae*. Da esquerda para direita: linhagem teste (CTBE/CNPEM), *S. cerevisiae* SA, *S. cerevisiae* PE-2, *S. cerevisiae* CAT-1, *S. cerevisiae* BG e *S. cerevisiae* FT 858. O laudo final emitido pelo CPQBA/UNICAMP afirmou que a linhagem de *S. cerevisiae* utilizada neste trabalho apresenta perfil cromossômico característico de *S. cerevisiae* CAT-1.

5.4. Clarificação do melaço de cana-de-açúcar

O processo de clarificação objetivou a remoção de impurezas e materiais particulados presentes no melaço de cana-de-açúcar doado pela Raízen®. Após a conclusão dos procedimentos foram obtidos aproximadamente 90 litros de material estéril, dividido em dois lotes e estocados em câmara fria (6°C) nos laboratórios do CTBE/CNPEM. As quantificações dos teores de ácidos orgânicos (ácido acético, ácido láctico e ácido succínico), açúcares (sacarose, glicose e frutose) e etanol, via HPLC, podem ser observadas na tabela 3. Não foram encontrados valores detectáveis de etanol no material final. As baixas concentrações de ácidos orgânicos são equivalentes às observadas no melaço clarificado de usinas brasileiras de produção de etanol. A concentração final de açúcares redutores totais (ART) foi de 390 gramas de açúcares por quilograma de melaço clarificado (g.Kg⁻¹).

Tabela 3. Quantificação final dos lotes clarificados de melaço de cana-de-açúcar em g.Kg⁻¹.

Lote	Ac. Acético	Ac. Láctico	Ac. Succínico	Etanol	Sacarose	Glicose	Frutose
01	2,60	1,34	0,91	0,00	292,64	50,27	48,07
02	2,59	1,28	0,81	0,00	293,18	49,83	47,18

5.5. Ensaios fermentativos

Os experimentos objetivaram a avaliação de contaminações bacterianas em fermentações etanólicas semi-industriais com reprodução dos processos tecnológicos de batelada alimentada em alta densidade celular com tratamento ácido e reciclo de células em sucessivos ciclos fermentativos. As propagações celulares realizadas previamente às fermentações resultaram na obtenção de cerca de 40 g.l⁻¹ de biomassa celular (base seca), em 1,8 litros, no final de 36 horas de cultivo em reator. As primeiras 12 horas desta etapa foram realizadas em batelada e tiveram como principal objetivo a adaptação do inóculo recém-propagado em meio YPD às condições de cultivo em melaço de cana-de-açúcar suplementado como fosfato monopotássico e ureia. Para a realização do processo de propagação celular é essencial uma alta disponibilidade de oxigênio para garantir a produção de biomassa em detrimento de etanol. Os controles de agitação e injeção de ar no sistema foram dependentes de medições dos teores de oxigênio dissolvidos no meio de cultivo dentro do biorreator por sondas eletrolíticas. Neste sentido, nas primeiras horas de propagação observou-se uma baixa necessidade de injeção de ar resultando em agitações mínimas de 200 rpm e injeções de uma mistura de ar contendo 95% de ar atmosférico e 5% de oxigênio molecular puro. Contudo, ao final de 12 horas de propagação em batelada os índices já estavam em seus valores máximos: agitação de 700 rpm e injeção de mistura de ar contendo 50% de ar atmosférico e 50% de oxigênio molecular puro. A segunda etapa da fase de propagação foi realizada em batelada alimentada com alimentação de 400 ml de melaço de cana-de-açúcar (30°Bx) em 24 horas, totalizando um volume de 1,8 litros de propagação celular em biorreator e 36 horas de cultivo microbiano (Figura 2). Nesta fase especificamente, limitou-se a vazão de entrada da fonte de carbono para potencializar o direcionamento do metabolismo celular das leveduras no sentido da produção de biomassa em detrimento de etanol. Imediatamente após a conclusão do processo de propagação iniciaram-se as fermentações em batelada alimentada. A realização sequencial das fases de propagação (36 horas) e três ciclos fermentativos de 8 horas totalizaram 60 horas de cultivo microbiano em reator para cada uma das nove fermentações executadas (Figura 2).

As análises de ART e álcoois presentes no vinho delevedurado podem ser observadas na Tabela 4. Praticamente todo o açúcar alimentado foi consumido ao longo de 8 horas de fermentação, havendo baixos teores de ART residuais, essencialmente frutose, durante a fase de enchimento do reator, nas fermentações contaminadas com *Lactobacillus* sp. I-2. A produção final de etanol variou entre 40 g.kg⁻¹ e 60 g.kg⁻¹, gerando um rendimento fermentativo final superior a 90% em todas as condições experimentais avaliadas. Os resultados finais das fermentações mostraram não haver diferenças estatísticas entre os perfis de consumo de açúcares e produção de etanol em decorrência das contaminações com *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3. Acerca do Glicerol, metabólito tipicamente produzido por leveduras para equilibrar o balanço redox celular, foram detectados aumentos significativos em fermentações contaminadas com *Lactobacillus* sp. I-2 (Tabela 4). As quantificações dos teores de ácido láctico mostraram maior abundância nas fermentações contaminadas com *Lactobacillus* sp. I-2 quando comparadas com a condição controle sem contaminação (Tabela 5). Este ácido orgânico é um subproduto do processo de oxidação parcial de compostos carbonados para obtenção de ATP em bactérias lácticas. Especula-se que a presença da bactéria *Lactobacillus* sp. I-2 gerou um maior acúmulo de ácido láctico no fermentado, podendo ter agravado os estresses nas leveduras de *S. cerevisiae*.

Surpreendentemente, não foram detectadas variações significativas nos teores de ácidos orgânicos nas fermentações contaminadas com *Acetobacter* sp. A-3 (Tabela 5). Assim como nos ensaios realizados com a bactéria láctica, as inoculações iniciais foram de 10⁸ UFC.ml⁻¹ de bactérias acéticas nas fermentações. Contudo, após a realização de três ciclos fermentativos consecutivos, observou-se uma alta mortalidade de bactérias resultando numa

concentração celular final de 10^5 UFC.ml⁻¹ no final da terceira batelada alimentada (Tabela 6). De fato, algo semelhante ocorre em usinas brasileiras de produção de etanol nas quais se observam populações de bactérias acéticas em concentrações bem inferiores às bactérias lácticas (BONATELLI et al., 2017). Acredita-se que a ocorrência de uma redução nas concentrações da população da bactéria *Acetobacter* sp. A-3 possa ter influenciado na interação dessa com as leveduras de *S. cerevisiae* CAT-1. Contudo, a observação de uma redução e não da completa eliminação da população destes microrganismos sugere a presença de mecanismos de tolerância à baixa disponibilidade de oxigênio, altas concentrações de açúcares, etanol e estresses presentes no meio.

A mensuração do tamanho da população de leveduras no terceiro ciclo fermentativo em batelada alimentada revelaram 10^9 UFC.ml⁻¹. Interessantemente, apesar do processo de batelada alimentada ter apresentado um pé-de-cuba com 1/3 do volume final do fermentado, ou seja, uma diluição final teórica de três vezes quando o reator estivesse completamente cheio, observou-se que a diluição no número de células viáveis não seguiu o mesmo fator. Os resultados mensurados para as populações de leveduras no final do terceiro ciclo fermentativo sugerem a ocorrência de um baixo crescimento celular destes microrganismos, mesmo em condições adversas como baixa disponibilidade de oxigênio no meio (Tabela 6).

A mensuração do tamanho da população da bactéria láctica no terceiro ciclo fermentativo em batelada alimentada revelaram 10^8 UFC.ml⁻¹. Estes resultados mostram uma viabilidade celular estável ao longo das fermentações, uma vez que essa foi à concentração inoculada no início do primeiro ciclo fermentativo em batelada alimentada. Provavelmente, as condições de microaerofilia foram determinantes neste resultado. Carvalho-Netto et al (2015) mencionam a alta atividade floculante de bactérias da espécie *Lactobacillus fermentum* quando presentes em fermentações na concentração de 10^5 UFC.ml⁻¹. Esses mostraram que a interação entre bactérias e leveduras de *S. cerevisiae* PE-2 resultaram na adesão celular entre as células dos diferentes microrganismos, caracterizando a floculação (CARVALHO-NETTO et al., 2015). Interessantemente, mesmo em altas concentrações celulares, 10^8 UFC.ml⁻¹, não se observou nenhuma floculação nos ensaios fermentativos em reatores. Diferente do trabalho descrito por Carvalho-Netto et al (2015), os ensaios fermentativos apresentados neste trabalho utilizaram leveduras da linhagem *S. cerevisiae* CAT-1. Especula-se que diferenças genéticas e fisiológicas entre as linhagens estudada por Carvalho-Netto et al (2015) e as apresentadas neste trabalho podem ter influenciado nos resultados finais. Contudo, apenas estudos genéticos e fisiológicos comparativos aprofundados entre as diferentes linhagens de bactérias lácticas e leveduras auxiliará numa melhor compreensão destas variações.

A mensuração do tamanho da população da bactéria acética no terceiro ciclo fermentativo em batelada alimentada revelaram 10^5 UFC.ml⁻¹. Provavelmente, as competições com as leveduras e as condições de baixa disponibilidade de oxigênio no fermentado contribuíram para a alta mortalidade celular ao longo das fermentações. Contudo, a redução e não a completa eliminação da bactéria acética evidencia a capacidade destes microrganismos tolerarem as condições adversas e sobreviverem em fermentações etanólicas de primeira geração.

Os índices de viabilidade celular médio e taxa de brotamento celular médio de leveduras mostraram valores percentuais diferentes aos comumente encontrados em fermentações etanólicas de primeira geração no Brasil (Tabela 7). Provavelmente, a utilização de uma biomassa celular jovem e recém-propagada refletiu na observação de índices de viabilidade superiores a 95% e taxas de brotamento inferiores a 2% em todas as condições experimentais avaliadas. Estes resultados evidenciam que a idade celular do inóculo, e por consequência, o tempo de exposição às condições de estresses (frio e “starvation”) refletem no desempenho da levedura em fermentações de maneira geral (Tabela 7).

Cada ciclo fermentativo teve um tempo total de 8 horas com alimentação do sistema com melaço de cana-de-açúcar nas primeiras 6 horas de cultivo. Neste sentido, os pontos amostrais de 0h e 4h representaram o pé-de-cuba e a fase de enchimento do reator, respectivamente. O ponto amostral de 8h foi coletado em condições de volume completo de fermentado imediatamente antes do início do processo de reciclo celular. Por fim, foram obtidas concentrações celulares finais em torno de 30 g.l⁻¹ (base seca) em todas as condições experimentais avaliadas, reproduzindo a característica de alta densidade celular presentes em fermentações industriais brasileiras (Tabela 8).

Por fim, observou-se uma intensa atividade respiratória manifestada na formação de bolhas de gases na condição contaminada com *Acetobacter* sp. A-3. Quantificações preliminares acerca do tamanho da população bacteriana mostraram uma menor concentração celular desta bactéria no fundo em detrimento da superfície do fermentado presente no reator. Especula-se que o consumo do oxigênio dissolvido no meio de cultivo pela bactéria acética possa ter ocasionado à migração destes microrganismos na coluna do fermentado em busca do ar presente no *headspace* do reator. Possivelmente, as bolhas de gases podem ser consequência da intensa liberação de CO₂ pelo metabolismo aeróbico da bactéria. A condição de microaerofilia presentes em fermentações 1G pode estar influenciando a ocorrência de variações nas concentrações das populações de bactérias ao longo da coluna do fermentado.

Tabela 4. Quantificação dos teores de açúcares redutores totais (ART) e álcoois presentes no sobrenadante do fermentado gerado nas condições experimentais avaliadas. Resultados mostram valores médios $\pm$ desvio padrão de três repetições biológicas em triplicata técnica. Valores apresentados em g.Kg⁻¹.

Condições experimentais avaliadas	Batelada alimentada 1								
	ART			Etanol			Glicerol		
	0h	4h	8h	0h	4h	8h	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	119,841 $\pm$ 1,851	0,244 $\pm$ 0,141	0,000 $\pm$ 0,000	3,965 $\pm$ 0,857	38,345 $\pm$ 7,303	47,432 $\pm$ 5,428	0,181 $\pm$ 0,034	3,266 $\pm$ 0,906	3,876 $\pm$ 0,794
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	140,954 $\pm$ 5,066	1,819 $\pm$ 0,345	0,379 $\pm$ 0,219	4,130 $\pm$ 3,609	38,378 $\pm$ 4,315	45,358 $\pm$ 4,023	0,233 $\pm$ 0,211	4,104 $\pm$ 0,447	4,815 $\pm$ 0,401
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	145,693 $\pm$ 2,097	0,000 $\pm$ 0,000	0,000 $\pm$ 0,000	10,116 $\pm$ 5,021	48,811 $\pm$ 2,884	52,104 $\pm$ 7,841	0,664 $\pm$ 0,413	3,577 $\pm$ 0,775	3,813 $\pm$ 0,575
Condições experimentais avaliadas	Batelada alimentada 2								
	ART			Etanol			Glicerol		
	0h	4h	8h	0h	4h	8h	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	119,841 $\pm$ 1,851	0,311 $\pm$ 0,179	0,000 $\pm$ 0,000	13,268 $\pm$ 1,380	42,792 $\pm$ 4,600	50,784 $\pm$ 5,860	1,120 $\pm$ 0,178	2,429 $\pm$ 0,287	3,009 $\pm$ 0,478
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	140,954 $\pm$ 5,066	1,417 $\pm$ 0,375	0,163 $\pm$ 0,094	13,310 $\pm$ 1,131	40,719 $\pm$ 3,920	47,807 $\pm$ 3,770	1,427 $\pm$ 0,150	3,054 $\pm$ 0,626	3,843 $\pm$ 0,814
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	145,693 $\pm$ 2,097	0,221 $\pm$ 0,128	0,000 $\pm$ 0,000	16,125 $\pm$ 3,228	52,255 $\pm$ 2,905	52,600 $\pm$ 8,678	1,379 $\pm$ 0,304	2,884 $\pm$ 0,139	3,153 $\pm$ 0,076
Condições experimentais avaliadas	Batelada alimentada 3								
	ART			Etanol			Glicerol		
	0h	4h	8h	0h	4h	8h	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	119,841 $\pm$ 1,851	0,272 $\pm$ 0,157	0,000 $\pm$ 0,000	13,694 $\pm$ 1,971	45,156 $\pm$ 7,614	50,663 $\pm$ 5,464	1,043 $\pm$ 0,230	2,518 $\pm$ 0,480	2,935 $\pm$ 0,404
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	140,954 $\pm$ 5,066	1,873 $\pm$ 0,349	0,058 $\pm$ 0,033	11,863 $\pm$ 0,340	43,872 $\pm$ 2,878	51,145 $\pm$ 4,182	1,157 $\pm$ 0,332	3,954 $\pm$ 0,340	4,747 $\pm$ 0,295
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	145,693 $\pm$ 2,097	0,276 $\pm$ 0,159	0,000 $\pm$ 0,000	13,500 $\pm$ 7,827	50,703 $\pm$ 6,132	54,289 $\pm$ 4,641	0,943 $\pm$ 0,594	3,173 $\pm$ 0,799	3,544 $\pm$ 0,679

Tabela 5. Quantificação dos teores de ácidos orgânicos presentes no sobrenadante do fermentado gerado nas três condições experimentais avaliadas. Os resultados mostram valores médios $\pm$ desvio padrão de três condições biológicas em triplicata técnica. Valores apresentados em g.Kg⁻¹.

Condições experimentais avaliadas	Batelada alimentada 1								
	Ácido Láctico			Ácido Acético			Ácido Succínico		
	0h	4h	8h	0h	4h	8h	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	0,000 $\pm$ 0,000	0,959 $\pm$ 0,655	1,348 $\pm$ 1,131	0,250 $\pm$ 0,236	0,746 $\pm$ 0,193	0,987 $\pm$ 0,157	0,407 $\pm$ 0,085	0,638 $\pm$ 0,012	0,621 $\pm$ 0,044
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	0,332 $\pm$ 0,576	1,580 $\pm$ 0,087	2,040 $\pm$ 0,363	0,000 $\pm$ 0,000	1,011 $\pm$ 0,145	1,368 $\pm$ 0,048	0,419 $\pm$ 0,408	0,702 $\pm$ 0,124	0,750 $\pm$ 0,197
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	0,146 $\pm$ 0,253	0,632 $\pm$ 0,034	0,416 $\pm$ 0,359	0,600 $\pm$ 0,601	1,097 $\pm$ 0,606	1,269 $\pm$ 0,546	0,598 $\pm$ 0,617	0,950 $\pm$ 0,409	0,892 $\pm$ 0,373
Condições experimentais avaliadas	Batelada alimentada 2								
	Ácido Láctico			Ácido Acético			Ácido Succínico		
	0h	4h	8h	0h	4h	8h	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	0,130 $\pm$ 0,225	1,282 $\pm$ 1,206	1,596 $\pm$ 1,496	0,480 $\pm$ 0,089	0,757 $\pm$ 0,098	0,993 $\pm$ 0,141	0,055 $\pm$ 0,048	0,401 $\pm$ 0,074	0,450 $\pm$ 0,089
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	0,582 $\pm$ 0,397	2,162 $\pm$ 0,484	2,552 $\pm$ 0,797	0,724 $\pm$ 0,140	1,447 $\pm$ 0,212	1,898 $\pm$ 0,235	0,077 $\pm$ 0,074	0,487 $\pm$ 0,094	0,534 $\pm$ 0,096
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	0,273 $\pm$ 0,291	0,667 $\pm$ 0,037	0,488 $\pm$ 0,386	0,953 $\pm$ 0,589	1,300 $\pm$ 0,482	1,375 $\pm$ 0,498	0,295 $\pm$ 0,403	0,667 $\pm$ 0,333	0,682 $\pm$ 0,360
Condições experimentais avaliadas	Batelada alimentada 3								
	Ácido Láctico			Ácido Acético			Ácido Succínico		
	0h	4h	8h	0h	4h	8h	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	0,000 $\pm$ 0,000	1,419 $\pm$ 1,353	1,683 $\pm$ 1,634	0,237 $\pm$ 0,084	0,505 $\pm$ 0,098	0,724 $\pm$ 0,117	0,037 $\pm$ 0,033	0,508 $\pm$ 0,142	0,521 $\pm$ 0,125
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	0,611 $\pm$ 0,386	2,397 $\pm$ 0,698	2,793 $\pm$ 0,971	0,739 $\pm$ 0,222	1,524 $\pm$ 0,166	1,852 $\pm$ 0,207	0,021 $\pm$ 0,037	0,528 $\pm$ 0,059	0,570 $\pm$ 0,080
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	0,186 $\pm$ 0,323	0,637 $\pm$ 0,047	0,722 $\pm$ 0,084	0,721 $\pm$ 0,630	1,064 $\pm$ 0,443	1,264 $\pm$ 0,398	0,490 $\pm$ 0,424	0,894 $\pm$ 0,351	0,856 $\pm$ 0,303

Tabela 6. Contagem das unidades formadoras de colônia (UFC.ml⁻¹) de bactérias e leveduras presentes na terceira batelada alimentada de cada condição experimental. Para a mensuração das células de levedura foi utilizado a metodologia de plaqueamento por superfície em meio YPD acrescido de ampicilina. A mensuração das células viáveis de bactérias foi feita mediante cultivo em meio MRS acrescido de cicloheximida. As contagens de UFC de *Lactobacillus* e *Acetobacter* foram feitas através das metodologias de plaqueamento por profundidade e superfície, respectivamente. Os resultados mostram valores médios $\pm$ desvio padrão de três condições biológicas.

Condições experimentais avaliadas	Organismo	UFC.ml ⁻¹		
		0h	4h	8h
<i>S. cerevisiae</i> CAT-1	<i>S. cerevisiae</i> CAT-1	2,52E+009 $\pm$ 1,70E+008	1,55E+009 $\pm$ 1,70E+008	1,27E+009 $\pm$ 7,26E+007
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	<i>S. cerevisiae</i> CAT-1	4,07E+009 $\pm$ 1,10E+009	1,41E+009 $\pm$ 1,70E+008	1,27E+009 $\pm$ 2,08E+008
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	<i>S. cerevisiae</i> CAT-1	4,05E+009 $\pm$ 1,78E+008	1,66E+009 $\pm$ 6,15E+008	1,24E+009 $\pm$ 3,74E+008
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	<i>Lactobacillus</i> sp. I-2	9,83E+008 $\pm$ 3,31E+008	2,25E+008 $\pm$ 1,17E+008	3,81E+008 $\pm$ 1,68E+008
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	<i>Acetobacter</i> sp. A-3	2,72E+006 $\pm$ 1,36E+005	8,93E+005 $\pm$ 4,04E+004	5,83E+005 $\pm$ 6,11E+004

Tabela 7. Análise de viabilidade celular e taxa de brotamento celular médio de leveduras baseado em contagem de células coradas com azul de metileno em câmara de *Neubauer*. A mensuração foi realizada em todas as condições experimentais testadas nos pontos iniciais (0h) e finais (8h) de cada ciclo fermentativo.

Batelada alimentada 1				
Condições experimentais avaliadas	Índice de viabilidade celular médio (%)		Taxa de brotamento celular médio (%)	
	0h	8h	0h	8h
<i>S. cerevisiae</i> CAT-1	97,72	97,52	0,46	0,90
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	99,39	99,24	0,61	0,00
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	98,85	99,24	0,86	0,76
Batelada alimentada 2				
Condições experimentais avaliadas	Índice de viabilidade celular médio (%)		Taxa de brotamento celular médio (%)	
	0h	8h	0h	8h
<i>S. cerevisiae</i> CAT-1	96,72	93,73	1,28	0,60
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	97,99	98,31	0,67	0,00
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	98,73	98,12	0,41	0,00
Batelada alimentada 3				
Condições experimentais avaliadas	Índice de viabilidade celular médio (%)		Taxa de brotamento celular médio (%)	
	0h	8h	0h	8h
<i>S. cerevisiae</i> CAT-1	96,02	75,53	1,49	21,91
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	97,16	89,89	0,95	7,92
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	97,33	82,78	0,28	16,51

Tabela 8. Valores de biomassa celular (base seca) mensurada ao longo das três bateladas alimentadas consecutivas realizadas para cada uma das condições experimentais testadas. Os resultados mostram valores médios $\pm$ desvio padrão de três condições biológicas. Valores apresentados em g.l⁻¹.

Condições experimentais avaliadas	Batelada alimentada 1		
	Massa Seca		
	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	81,657 $\pm$ 9,114	40,000 $\pm$ 3,281	32,733 $\pm$ 2,346
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	90,943 $\pm$ 2,800	40,933 $\pm$ 2,226	34,913 $\pm$ 3,891
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	90,203 $\pm$ 6,483	38,677 $\pm$ 2,408	32,470 $\pm$ 1,039
Condições experimentais avaliadas	Batelada alimentada 2		
	Massa Seca		
	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	82,990 $\pm$ 7,397	42,377 $\pm$ 1,744	33,957 $\pm$ 1,194
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	85,533 $\pm$ 1,582	40,700 $\pm$ 1,300	32,243 $\pm$ 0,792
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	84,500 $\pm$ 3,822	40,620 $\pm$ 1,846	33,277 $\pm$ 2,166
Condições experimentais avaliadas	Batelada alimentada 3		
	Massa Seca		
	0h	4h	8h
<i>S. cerevisiae</i> CAT-1	81,490 $\pm$ 8,136	40,190 $\pm$ 3,800	32,500 $\pm$ 2,488
<i>S. cerevisiae</i> CAT-1 + <i>Lactobacillus</i> sp. I-2	82,033 $\pm$ 2,033	37,957 $\pm$ 1,231	29,880 $\pm$ 1,146
<i>S. cerevisiae</i> CAT-1 + <i>Acetobacter</i> sp. A-3	85,643 $\pm$ 5,023	43,090 $\pm$ 1,725	34,680 $\pm$ 2,335

5.6. Sequenciamento, montagem e anotação dos genomas bacterianos

O sequenciamento dos genomas gerou 3.415.112, 5.972.252 e 4.818.977 fragmentos de DNA com tamanhos variando entre 250pb a 500pb para *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente. Após a realização da trimagem das sequências de baixa qualidade, mais de 90% dos *reads* originais foram mantidos e utilizados na etapa de montagem dos genomas (Figura 6).

Figura 6. Pontuação atribuída a cada base dos *reads* gerados durante o sequenciamento dos genomas das bactérias. Dados de qualidade baseados na escala Phred para A) *Lactobacillus* sp. L-10, B) *Lactobacillus* sp. I-2 e C) *Acetobacter* sp. A-3. As linhas vermelhas, caixas amarelas e linhas azuis representam os valores de mediana, amplitude interquartil e valores de média, respectivamente. Nos eixos das ordenadas e abcissas observam-se a escala Phred e a posição de cada base sequenciada nos *reads* gerados, respectivamente. Figura gerada pelo programa FastQC.

Figura 6. Pontuação atribuída a cada base dos *reads* gerados durante o sequenciamento dos genomas das bactérias. Dados de qualidade baseados na escala Phred para A) *Lactobacillus* sp. L-10, B) *Lactobacillus* sp. I-2 e C) *Acetobacter* sp. A-3. As linhas vermelhas, caixas amarelas e linhas azuis representam os valores de mediana, amplitude interquartil e valores de média, respectivamente. Nos eixos das ordenadas e abcissas observam-se a escala Phred e a posição de cada base sequenciada nos *reads* gerados, respectivamente. Figura gerada pelo programa FastQC. (Continuação).

A reconstrução do cromossomo circular original das bactérias a partir dos *reads* obtidos no sequenciamento NGS é conhecida como montagem do genoma. Utilizando dois programas distintos, SPADes e Velvet, foi possível obter genomas de alta qualidade e pouco fragmentados com 60, 180 e 147 *scaffolds* para *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente (Tabela 9). Estudos da composição gênica das bactérias mostraram que 97% dos genes ortólogos presentes no banco de dados do programa BUSCO estavam em cópia única e não fragmentados nos genomas gerados de *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 (Tabela 8). Com a utilização de algoritmos *ab initio* para predição computacional de genes foi possível identificar 3.056, 2.112 e 3.557 genes para *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3, respectivamente (Tabela 9).

Tabela 9. Dados gerais obtidos nas montagens e anotações dos genomas bacterianos.

Informações	<i>Lactobacillus</i> sp. L-10	<i>Lactobacillus</i> sp. I-2	<i>Acetobacter</i> sp. A-3
Número de <i>Scaffolds</i>	60	180	147
Comprimento (pb)	3220008	2067020	3871619
Maior <i>scaffold</i> (pb)	388517	149893	235927
Valor de N50 (pb)	145723	25784	69340
Cobertura média	370x	495x	347x
Número de genes codificantes	3056	2112	3557
Cobertura gênica (%) *	97	97	97

* Dados oriundos das análises do programa BUSCO com 40 genes ortólogos cópia única de bactérias.

A checagem da presença de DNA contaminante nos *reads* sequenciados visou evitar a ocorrência de erros na mensuração do tamanho do genoma ou no número e composição dos genes associados às linhagens de bactérias.

A detecção de *scaffolds* com coberturas médias de sequenciamentos similares, faixas estreitas entre os teores médios de GC e anotação taxonômica apontado para *Lactobacillus* e *Acetobacter* sugerem ausência de contaminação das amostras com DNA de outros microrganismos que não *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 (Figura 7).

Figura 7. Verificação da presença de sequências de organismos contaminantes nos *reads* gerados pelo sequenciamento NGS das linhagens A) *Lactobacillus* sp. L-10, B) *Lactobacillus* sp. I-2 e C) *Acetobacter* sp. A-3. Gráficos gerados pelo programa TAGC plot destacando informações como cobertura média de sequenciamento, proporção de GC e anotação taxonômica para cada *scaffold*.

Figura 7. Verificação da presença de seqüências de organismos contaminantes nos *reads* gerados pelo sequenciamento NGS das linhagens A) *Lactobacillus* sp. L-10, B) *Lactobacillus* sp. I-2 e C) *Acetobacter* sp. A-3. Gráficos gerados pelo programa TAGC plot destacando informações como cobertura média de sequenciamento, proporção de GC e anotação taxonômica para cada *scaffold*. (Continuação).

Figura 7. Verificação da presença de sequências de organismos contaminantes nos *reads* gerados pelo sequenciamento NGS das linhagens A) *Lactobacillus* sp. L-10, B) *Lactobacillus* sp. I-2 e C) *Acetobacter* sp. A-3. Gráficos gerados pelo programa TAGC plot destacando informações como cobertura média de sequenciamento, proporção de GC e anotação taxonômica para cada *scaffold*. (Continuação).

5.7. Estudos taxonômicos e filogenéticos multigênicos

Os estudos taxonômicos e filogenéticos multigênicos comparativos de diferentes linhagens de *Lactobacillus* e *Acetobacter* permitiram identificar as linhagens de bactérias mais semelhantes à *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3. A comparação dos genomas pelas métricas gANI e AF de 64 linhagens de *Lactobacillus* distribuídas em 8 espécies diferentes através revelaram que as espécies *L. fermentum* e *L. plantarum* são as mais semelhantes à *Lactobacillus* sp. I-2 e *Lactobacillus* sp. L-10, respectivamente (Figura 8A). A comparação dos genomas de 62 linhagens de *Acetobacter* distribuídas em 17 espécies diferentes apontaram as espécies *A. tropicalis* e *A. senegalensis* como as mais semelhantes à *Acetobacter* sp. A-3 (Figura 8B).

A

Figura 8. Heatmaps contendo os valores probabilísticos resultantes da combinação das métricas gANI e AF de diferentes genomas para linhagens de A) *Lactobacillus* e B) *Acetobacter*. As análises comparativas par a par foram feitas independentemente em ambas as direções: do genoma A em relação ao genoma B e do genoma B em relação ao genoma A. As linhagens *Lactobacillus* sp. I2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3 foram destacadas em vermelho.

B

Figura 8. Heatmaps contendo os valores probabilísticos resultantes da combinação das métricas gANI e AF de diferentes genomas para linhagens de A) *Lactobacillus* e B) *Acetobacter*. As análises comparativas par a par foram feitas independentemente em ambas as direções: do genoma A em relação ao genoma B e do genoma B em relação ao genoma A. As linhagens *Lactobacillus* sp. 12, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3 foram destacadas em vermelho. (Continuação).

As buscas por potenciais proteínas preditas (PPPs) codificadas a partir de genes em cópia única em diferentes genomas de *Lactobacillus* resultaram na identificação de 236 potenciais marcadores filogenéticos em 52 linhagens de bactérias lácticas, incluindo *Lactobacillus* sp. L-10 e *Lactobacillus* sp. I-2. Utilizando diversos modelos evolutivos e estimativas por máxima verossimilhança foi possível identificar as espécies *L. plantarum*, *L. paraplantarum* e *L. pentosus* como as mais semelhantes à *Lactobacillus* sp. L-10. As espécies *L. fermentum*, *L. equigenersi* e *L. gastricus* foram as mais semelhantes à *Lactobacillus* sp. I-2 (Figura 9A). Ademais, estes estudos apontaram as linhagens *L. plantarum* subesp. *argenteratensis* DSM 16365 e *L. fermentum* SNUV 175 como as bactérias filogeneticamente mais semelhantes às linhagens *Lactobacillus* sp. L-10 e *Lactobacillus* sp. I-2, respectivamente (Figura 9A).

As buscas por PPPs codificadas a partir de genes em cópia única em diferentes genomas de *Acetobacter* resultaram na identificação de 1020 potenciais marcadores filogenéticos em 28 linhagens de bactérias acéticas, incluindo *Acetobacter* sp. A-3. Utilizando diversos modelos evolutivos e estimativas por máxima verossimilhança foi possível identificar as espécies *A. tropicalis* e *A. senegalensis* como as mais semelhantes à *Acetobacter* sp. A-3 (Figura 9B). Ademais, estes estudos apontaram as linhagens *A. senegalensis* DmL 050, *A. senegalensis* LMG 23690 e *A. senegalensis* 108B como as bactérias filogeneticamente mais semelhantes à *Acetobacter* sp. A-3 (Figura 9B). A presença da bactéria *A. tropicalis* NBRC 101654 no mesmo ramo sinaliza um possível erro na sua classificação taxonômica (Figura 9B). Os estudos realizados com as bactérias acéticas mostram que diferentes espécies podem ter identidades nucleotídicas e fração de alinhamento de genes ortólogos superiores a 80%, caso especificamente observado entre *Acetobacter* sp. A-3, *A. tropicalis* e *A. senegalensis*. Por fim, os estudos das métricas gANI e AF combinados com as análises filogenéticas multigênicas identificaram as espécies mais semelhantes à *Lactobacillus* sp. L-10, *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3: *L. plantarum*, *L. fermentum* e *A. senegalensis*, respectivamente.

Figura 9. Filogenias multigênicas para A) *Lactobacillus* sp. L-10 e *Lactobacillus* sp. I-2 e B) *Acetobacter* sp. A-3. Foram utilizadas 236 PPPs como marcadores filogenéticos para *Lactobacillus* e 1020 para *Acetobacter*. *L. plantarum* subsp. *argentorantensis* DSM 16365^T é a mais semelhante à *Lactobacillus* sp. L-10. *L. fermentum* SNUV 175 se assemelha mais a *Lactobacillus* sp. I-2 e *A. senegalensis* 108B está entre as mais próximas evolutivamente de *Acetobacter* sp. A-3.

Figura 9. Filogenias multigênicas para A) *Lactobacillus* sp. L-10 e *Lactobacillus* sp. I-2 e B) *Acetobacter* sp. A-3. Foram utilizadas 236 PPPs como marcadores filogenéticos para *Lactobacillus* e 1020 para *Acetobacter*. *L. plantarum* subsp. *argenterantensis* DSM 16365^T é a mais semelhante à *Lactobacillus* sp. L-10. *L. fermentum* SNUV 175 se assemelha mais a *Lactobacillus* sp. I-2 e *A. senegalensis* 108B está entre as mais próximas de *Acetobacter* sp. A-3. (Continuação).

5.8. Estudos pangenômicos

Os estudos pangenômicos permitiram comparar potenciais proteínas preditas (PPPs) de diferentes linhagens de *Lactobacillus* e *Acetobacter*. Os estudos comparativos de 26 linhagens de *L. fermentum* revelaram 772 e 62 PPPs classificadas como acessórias e únicas para *Lactobacillus* sp. I-2, respectivamente (Tabela 10). Os processos “*Transcription*”, “*Defense mechanisms*”, “*Mobilome: prophages and transposons*” e “*Carbohydrate transport and metabolism*” estão entre os comumente apontados nas análises de COGs para PPPs desta linhagem bacteriana (Figura 10A, Figura 11A).

O COG1309 (“*DNA-binding transcriptional regulator, AcrR family*”) foi encontrado tanto no conjunto de PPPs acessórias como exclusivas de *Lactobacillus* sp. I-2 (Apêndice A). Em *Escherichia coli*, o gene *acrR* está localizado *upstream* ao gene A do Operon AB, um importante agente regulatório de mecanismos associados a tolerância a antibióticos e bombeamento de drogas para fora da célula bacteriana (JELLEN-RITTER & KERN, 2001; WANG et al., 2001; HIRAKAWA et al., 2008). Acredita-se que a proteína AcrR atue reprimindo os genes presentes no Operon AB, uma vez que a deleção ou inativação do gene *acrR* resulta no aumento da expressão dos seus constituintes conferindo uma maior tolerância da bactéria a antibióticos, principalmente Fluoroquinolononas (JELLEN-RITTER & KERN, 2001; WANG et al., 2001). A utilização de antibióticos de amplo espectro para controle das populações de bactérias contaminantes é algo comum em fermentações industriais. Neste sentido, a identificação da PPPs acessórias e exclusivas com funções preditas associadas ao COG1309 sinaliza a presença de mecanismos de tolerância a antibióticos em *Lactobacillus* sp. I-2. Estes mecanismos de tolerância a compostos citotóxicos podem estar

contribuindo para a sobrevivência destes microrganismos nas dornas de fermentação em indústrias brasileiras de produção de etanol.

Diferentes COGs associados à transposases, integrases e proteínas virais foram encontradas tanto no conjunto de PPPs acessórias como exclusivas de *Lactobacillus* sp. I-2 (Apêndice A). Elementos transponíveis são segmentos de DNA móveis compostos por sequências repetitivas (palindrômicas) e genes codificantes de transposases, enzimas capazes de clivar a dupla fita de DNA em regiões específicas (ALBERTS et al., 2010). O deslocamento destes elementos ao longo de um genoma pode acarretar alterações na expressão gênica do organismo, uma vez que a sua inserção ou remoção em sequências gênicas pode alterar os seus quadros abertos de leitura (“*Open reading frames*”- ORF) (ALBERTS et al., 2010). Ademais, alguns transposons podem transportar genes, como os associados à tolerância a antibióticos em bactérias (ALBERTS et al., 2010). Integrases são transposases que permitem, dentre outras coisas, a inserção de um fragmento de DNA viral num genoma bacteriano (ALBERTS et al., 2010). A presença de potenciais genes codificadores de elementos genéticos móveis, integrases e proteínas virais sinalizam as ocorrências de infecções virais pretéritas e mecanismos adaptativos de tolerância a drogas e toxinas em *Lactobacillus* sp. I-2.

Os estudos pangenômicos realizados com 13 linhagens de *L. plantarum* e *L. paraplantarum* indicaram 579 e 142 PPPs acessórias e únicas para *Lactobacillus* sp. L-10 (Tabela 10). Os processos “*Carbohydrate transport and metabolism*”, “*Transcription*”, “*Defense mechanisms*”, “*Mobilome: prophages and transposons*” e “*Replication, recombination and repair*” estão entre os comumente apontados nas análises de COGs para PPPs desta linhagem bacteriana (Figura 10B, Figura 11B).

Diferentes COGs associados à família de transportadores ABC (“*ATP-binding cassette*”), glicosidases e fatores de transcrição envolvidos na regulação e metabolismo de açúcares foram encontradas tanto no conjunto de PPPs acessórias como exclusivas de *Lactobacillus* sp. L-10 (Apêndice B). A oxidação de carboidratos em organismos vivos é uma das principais formas de se obter energia química para execução de processos biológicos como biossíntese de macromoléculas, motilidade e transporte de solutos através da membrana (KONINGS, POOLMAN & VAN VEEN, 1994). Essencialmente, essa energia pode ser armazenada em ligações químicas em moléculas de ATP ou num gradiente eletroquímico de prótons ou íons de sódio (KONINGS, POOLMAN & VAN VEEN, 1994). Ademais, transportadores ABC são proteínas capazes de transportar diferentes substratos (sais, íons, açúcares, antibióticos, dentre outros) pela membrana celular, com gasto de energia (KONINGS, POOLMAN & VAN VEEN, 1994). Em bactérias lácticas estes estão associados a mecanismos de tolerância a drogas, através do bombeamento de compostos citotóxicos para o ambiente extracelular (KONINGS, POOLMAN & VAN VEEN, 1994). Assim como em *Lactobacillus* sp. I-2, especula-se que a identificação de PPPs com funções preditas associadas a COGs acessórios e exclusivos do mecanismo de transporte ABC sinaliza a presença de mecanismos de tolerância a estresses causados pela presença de antibióticos, sais, metais pesados e outros compostos para as células de *Lactobacillus* sp. L-10.

Os estudos pangenômicos realizados com 8 linhagens de *A. tropicalis* e *A. senegalensis* indicaram 705 e 156 PPPs acessórias e únicas para *Acetobacter* sp. A-3 (Tabela 10). Os processos “*General function only*”, “*Transcription*”, “*Defense mechanisms*”, “*Inorganic ion transport and metabolism*” e “*Function unknown*” estão entre os comumente apontados nas análises de COGs para PPPs desta linhagem bacteriana (Figura 10C, Figura 11C). Especificamente para a linhagem *Acetobacter* sp. A-3 foram encontradas diferentes PPPs com funções pobremente descritas ou desconhecidas (Apêndice C). Estes resultados sinalizam um menor conhecimento a cerca da composição genética e genômica de bactérias acéticas em detrimento de bactérias lácticas. Diferentes PPPs acessórias e exclusivas de *Acetobacter* sp. A-3 apresentaram potenciais domínios proteicos relacionados à hidrolases, ATPases e permeases, estruturas também

identificadas em *Lactobacillus* sp. I-2 e *Lactobacillus* sp. L-10. Possivelmente, a semelhança entre os ambientes naturais destes microrganismos pode ter resultado em pressões seletivas similares e por consequência na adoção de estratégias adaptativas comuns.

O processo “*Transcription*” é o segundo mais abundante em PPPs acessórias, perdendo apenas para “*General function prediction only*” em *Acetobacter* sp. A-3 (Figura 10C). Assim como *Lactobacillus* sp. I-2, foram observados diferentes COGs associados a reguladores transcricionais com domínios de ligação a molécula de DNA das famílias AcrR, HXIR e LysR (Apêndice C). Os genomas de organismos procarióticos são caracterizados pela organização de genes em operons com poucas sequências intergênicas (ALBERTS, et al. 2010). Ademais, os processos de transcrição e tradução geralmente ocorrem de maneira simultânea com a formação de mRNAs policistrônicos e proteínas (ALBERTS, et al. 2010). A adoção destes mecanismos genéticos implica num maior controle e refinamento da expressão e silenciamento de genes, em procariotos, ainda no nível transcricional (ALBERTS, et al. 2010). A identificação de PPPs acessórias com função predita associada a COGs de reguladores transcricionais das famílias AcrR, HXIR e LysR destaca esta importância em bactérias acéticas de *Acetobacter* sp. A-3. Especificamente para os reguladores da família AcrR, assim como em *Lactobacillus* sp. I-2, a presença de uma proteína AcrR pode estar contribuindo para o refinamento dos mecanismos de tolerância de *Acetobacter* sp. A-3 a antibióticos comumente utilizados nas fermentações industriais.

Tabela 10. Número de potenciais proteínas preditas (PPPs) *core*, acessórias, únicas e exclusivas ausentes obtidas nas comparações de 26 linhagens de *L. fermentum*, incluindo *Lactobacillus* sp. I-2; 13 linhagens de *L. plantarum* e *L. paraplantarum*, incluindo *Lactobacillus* sp. L-10; e 8 linhagens de *A. senegalensis* e *A. tropicalis*, incluindo *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs.

Organismo	Core	Acessórias	Exclusivas	Exclusivas ausentes
<i>Lactobacillus</i> sp. I-2	1075	772	62	17
<i>Lactobacillus</i> sp. L-10	2159	579	142	18
<i>Acetobacter</i> sp. A-3	2268	705	156	9

Figura 10. Apontamento de COGs para as PPPs classificadas como acessórias pelo programa BPGA. A) *Lactobacillus* sp. I-2, B) *Lactobacillus* sp. L-10, e C) *Acetobacter* sp. A-3.

Figura 11. Apontamento de COGs para as PPPs classificadas como exclusivas pelo programa BPGA. A) *Lactobacillus* sp. I-2, B) *Lactobacillus* sp. L-10, e C) *Acetobacter* sp. A-3

5.9. Identificação dos metabólitos intracelulares totais por GC-MS

As análises de cromatografia gasosa acoplada à espectrometria de massas (GC-MS) das amostras biológicas de leveduras e bactérias permitiram a identificação de 101 metabólitos intracelulares presentes em pelo menos uma condição experimental avaliada. Os identificadores do banco de dados GMD, fórmulas moleculares e, quando possível, os números de referência no KEGG (“*KEGG Compound*”) para cada metabólito pode ser observado no Apêndice D. Cerca de 38% dos metabólitos identificados foram classificados como desconhecidos (“*Unknown*”) pelo banco de dados GMD (Apêndice D).

As análises multivariadas por PCA revelaram amostras homogêneas e com boa separação entre os tratamentos (Figura 1). Foram identificadas duas repetições técnicas *outliers* na amostra “*S. cerevisiae* CAT-1 0h” que foram removidas das análises estatísticas subsequentes com os tratamentos referentes às contaminações com bactérias lácticas e acéticas (Figura 1A e 1B). Nas análises comparativas de meio de fermentação (4h) foram removidas uma e duas repetições técnicas nas amostras “*S. cerevisiae* CAT-1 4h” quando comparadas com os tratamentos “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 4h” e “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 4h”, respectivamente (Figura 1C e 1D). Por fim, nas análises de final de fermentação (8h) foi identificada uma repetição técnica *outlier* na amostra “*S. cerevisiae* CAT-1 8h” que foi removida das análises comparativas finais (Figura 1E e 1F).

Os metabólitos diferencialmente abundantes nas comparações par a par entre os tratamentos “*S. cerevisiae* CAT-1” e “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2” nos pontos de início (0h), meio (4h) e final (8h) de fermentação podem ser vistos na tabela 11. Os metabólitos diferencialmente abundantes nas comparações par a par entre os tratamentos “*S. cerevisiae* CAT-1” e “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3” nos pontos de início (0h), meio (4h) e final (8h) de fermentação podem ser vistos na tabela 12.

Figura 12. PCA contendo as análises estatísticas par a par das amostras coletadas durante as fermentações. A) “*S. cerevisiae* CAT-1 0h” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 0h”; B) “*S. cerevisiae* CAT-1 0h” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 0h”; C) “*S. cerevisiae* CAT-1 4h” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 4h”; D) “*S. cerevisiae* CAT-1 4h” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 4h”; E) “*S. cerevisiae* CAT-1 8h” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 8h”; F) “*S. cerevisiae* CAT-1 8h” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 8h”.

Figura 12. PCA contendo as análises estatísticas par a par das amostras coletadas durante as fermentações. A) “*S. cerevisiae* CAT-1 0h” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 0h”; B) “*S. cerevisiae* CAT-1 0h” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 0h”; C) “*S. cerevisiae* CAT-1 4h” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 4h”; D) “*S. cerevisiae* CAT-1 4h” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 4h”; E) “*S. cerevisiae* CAT-1 8h” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2 8h”; F) “*S. cerevisiae* CAT-1 8h” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3 8h”. (Continuação).

O glicerol foi apontado entre os metabólitos com diferença de abundância significativa nas comparações par a par de final de fermentação (8h) entre os tratamentos controle e contaminado com a bactéria acética (Tabela 12). De maneira geral, a produção de glicerol envolve a redução de uma molécula de dihidroxiacetona com recuperação de uma molécula de NAD⁺ (ALBERTYN et al., 1991; NORBECK et al., 1996; ANSELL et al., 1997). Essa estratégia implica na redução da produção final de ATP, mas assume papel fundamental na proteção contra

estresses, principalmente os oriundos de desequilíbrio osmótico e desbalanço do potencial redox (BROWN, 1990). De fato, diversos trabalhos discutem os efeitos da não acumulação intracelular de glicerol em leveduras de *S. cerevisiae* (ALBERTYN et al., 1991; NEVOIG & STAHL, 1997; NISSEN et al., 2000; BIDEAUX et al., 2006). O melaço de cana-de-açúcar apresenta uma alta concentração de sais e açúcares, o que resulta na necessidade dos microrganismos equilibrarem o fluxo de água e balanço redox celular com o meio hiperosmótico. Aparentemente, a diminuição da produção de glicerol em leveduras pode aumentar a produção de etanol em condições de anaerobiose, porém acaba resultando em alta sensibilidade dos microrganismos a presença de sais, açúcares e etanol presentes no meio (ALBERTYN et al., 1991; NEVOIG & STAHL, 1997; NISSEN et al., 2000; BIDEAUX et al., 2006). Bactérias do gênero *Acetobacter* podem consumir glicerol em seus processos de fermentação oxidativa (MATSUSHITA et al., 2016). Especula-se que uma maior abundância de glicerol no tratamento sem contaminação bacteriana seja reflexo da imposição dos stresses osmóticos e oxidativos nas fermentações. Possivelmente, a presença da bactéria *Acetobacter* sp. A-3 resultou num maior consumo de glicerol e por consequência na potencialização das condições de stresses, principalmente os causados pelo desbalanço redox, em leveduras de *S. cerevisiae* CAT-1.

O sorbitol foi apontado entre os metabólitos com diferença de abundância significativa nas comparações par a par de final de fermentação (8h) entre os tratamentos controle e o contaminado com bactérias acéticas (Tabelas 2). Assim como o glicerol, o acúmulo intracelular de sorbitol em leveduras pode proteger a célula de stresses, principalmente os associados à desbalanço redox e desequilíbrio osmótico (SHEN et al., 1999). A redução de uma molécula de glicose ou frutose em sorbitol resulta na recuperação de uma molécula de NAD⁺ que pode ser utilizada novamente nos processos de oxidação de compostos carbonados para biossíntese de ATP (SHEN et al., 1999). Contudo, uma provável distribuição intracelular não homogênea deste poliol resulta num efeito osmoprotetor menos efetivo que o observado pelo acúmulo de glicerol em leveduras (SHEN et al., 1999). A enzima sorbitol desidrogenase (“*sorbitol dehydrogenase*” – SLDH) está presente nas membranas de bactérias acéticas do gênero *Gluconobacter* conferindo alta capacidade de oxidação deste poliol para produção de energia na forma de ATP (MATSUSHITA et al., 2016). Em contraste, bactérias dos gêneros *Acetobacter* e *Komagataeibacter* apresentam baixa capacidade de oxidação deste poliol em detrimento de etanol (MATSUSHITA et al., 2016). Especula-se que o acúmulo de sorbitol em células de levedura no tratamento com contaminação da bactéria acética seja reflexo, principalmente, da necessidade de osmoproteção destes microrganismos. Como discutido anteriormente, um provável consumo de glicerol por *Acetobacter* sp. A-3 pode ter resultado na potencialização do desequilíbrio redox nas leveduras de *S. cerevisiae*. Assim, a diminuição dos teores de glicerol pode ter levado a um maior acúmulo intracelular de sorbitol em leveduras como mecanismo de proteção dos stresses causados pelas fermentações.

O sorbitol foi apontado entre os metabólitos com diferença de abundância significativa nas comparações par a par de final de fermentação (8h) entre os tratamentos controle e o contaminado com bactérias lácticas (Tabelas 1). Bactérias lácticas (“*Lactic acid Bacteria*” – LAB) não são incapazes de produzirem sorbitol, exceto nas condições em que foram metabolicamente engenheiradas como nos casos descritos por Nissen, Pérez-Martínez & Yebra (2005) e Ladero et al. (2007) para *L. casei* e *L. plantarum*, respectivamente. Ambas as espécies foram geneticamente modificadas para produzirem sorbitol através da redução de frutose-6-fosfato pela enzima sorbitol-6-fosfato desidrogenase (“*sorbitol-6-phosphate dehydrogenase*” – S6PDH) (NISSEN, PÉREZ-MARTÍNEZ, & YEBRA, 2005; LADERO et al., 2007). Bactérias de *L. plantarum* engenheiradas apresentaram alta atividade da enzima S6PDH, contudo a modificação genética resultou em deficiência na atividade da enzima Lactato desidrogenase (LADERO et al., 2007). Em *L. casei*, o primeiro passo no consumo de D-sorbitol ocorre quando esse é convertido em sorbitol-6-fosfato por uma fosfotransferase dependente de fosfoenolpiruvato (YEBRA & PÉREZ-MARTÍNEZ, 2002). Em

seguida a atuação da enzima sorbitol-6-fosfato desidrogenase permite a produção de frutose-6-fosfato e entrada desse metabólito na via glicolítica (YEBRA & PÉREZ-MARTÍNEZ, 2002). Especula-se que o acúmulo intracelular de sorbitol em leveduras no tratamento controle pode ser reflexo dos mecanismos de tolerância deste microrganismo contra os estresses osmóticos e oxidativos causados pelas fermentações. Ademais, um provável consumo deste poliol pelas bactérias lácticas pode ter potencializado os efeitos de desequilíbrio osmótico nos microrganismos de *S. cerevisiae* no tratamento contaminado com *Lactobacillus* sp. I-2.

A trealose foi apontada entre os metabólitos com diferença de abundância significativa nas comparações par a par de início de fermentação (0h) entre os tratamentos controle e o contaminado com bactérias acéticas (Tabelas 12). A trealose é um dissacarídeo formado por duas moléculas de glicose amplamente encontrada na natureza em bactérias, fungos e plantas (PETITJEAN et al., 2015). A ampla distribuição da trealose se deve a sua estrutura peculiar: açúcar não redutor, alta hidrofiliçidade, alta estabilidade química, baixo potencial de cristalização e ausência de ligações de hidrogênio em sua estrutura interna (AGÜELLES, 2000). Em organismos procarióticos, a trealose pode atuar como componente estrutural, agente protetor contra estresses e fonte de carbono e energia na biossíntese de compostos celulares e ATP (SPARGO et al., 1991; STROM & KAASEN, 1993). Em leveduras, este dissacarídeo pode atuar como fonte secundária de carbono e energia e fornecer proteção celular contra estresses, principalmente os causados por desidratação (PARROU, TESTE & FRANÇOIS, 1997; AGÜELLES, 2000; PETITJEAN et al., 2015). Acredita-se que a sua atuação como mecanismo protetor das células ocorra devido à capacidade desse metabólito se ligar a grupos polares de proteínas e fosfolipídeos de membrana estabilizando-os e evitando o rompimento celular (SINGER & LINDQUIST, 1998). Bactérias do gênero *Acetobacter* são capazes de realizar fermentação oxidativa da trealose com geração de ácido acético e ATP (KADERE et al., 2008). Além disso, leveduras de *Saccharomyces cerevisiae* modificadas geneticamente para aumentar o acúmulo intracelular de trealose apresentaram maior tolerância a ácido acético (YOSHIYAMA et al., 2015). Especula-se que a identificação de altos teores intracelulares de trealose na condição contaminada com *Acetobacter* sp. A-3 seja reflexo do agravamento de estresses sobre as membranas causados pela presença da bactéria. A identificação da trealose em altas concentrações em início de fermentação (0h) sinaliza uma provável não utilização deste dissacarídeo como fonte adicional de energia no final do ciclo fermentativo anterior.

A trealose foi apontada entre os metabólitos com diferença de abundância significativa nas comparações par a par de final de fermentação (8h) entre os tratamentos controle e o contaminado com bactérias lácticas (Tabela 11). Carvalho et al. (2011) descrevem os efeitos do acúmulo intracelular de trealose em *Lactococcus lactis*. De acordo com eles, as linhagens modificadas apresentaram alta tolerância a estresses, principalmente os causados por tratamento com ácidos e aumentos de temperatura (CARVALHO et al., 2011). Especula-se que uma maior abundância de trealose nas amostras contaminadas com *Lactobacillus* sp. I-2 seja reflexo da potencialização dos estresses causados pela presença da bactéria nas fermentações. O aumento dos teores de ácido láctico no fermentado podem ter potencializado os efeitos de estresse nas membranas de bactérias e leveduras e levado a um maior acúmulo intracelular deste dissacarídeo.

A Glicina foi apontada entre os metabólitos com diferença de abundância significativa nas comparações par a par de meio de fermentação (4h) entre o tratamento controle e os contaminados com bactérias lácticas e acéticas (Tabelas 11 e 12). A Glicina é um aminoácido apolar contendo um átomo de hidrogênio na sua cadeia lateral (ALBERTS, et al., 2010). Em *S. cerevisiae* existem duas principais vias de biossíntese de glicina: a que utiliza intermediários fosfatos da via glicolítica e a que desvia glioxalato e isocitrato, intermediários do ciclo do ácido cítrico e ciclo do glioxalato, para biossíntese de aminoácidos (DEBOISO & STOPPANI, 1967; ULANE & OGUR, 1972).

(FIGURA X- ref. glicina 2). Em condições de alta disponibilidade de glicose, a via de biossíntese do glioxalato apresenta baixa atividade, sendo que quase a totalidade da glicina encontrada nas células é oriunda da oxidação de 3-fosfoglicerato (ULANE & OGUR, 1972). Em contrapartida, quando a glicose passa a ser um nutriente limitante, a via do glioxalato se torna a principal fonte de biossíntese de glicina em *S. cerevisiae* (ULANE & OGUR, 1972). Apesar de possuir apenas um átomo de nitrogênio na sua estrutura, a glicina é um das principais fontes de nitrogênio para biossíntese de purinas, colinas, timidilato e outros compostos nitrogenados (ULANE & OGUR, 1972; SINCLAIR & DAWES, 1995). Ademais, o mosto de cana-de-açúcar apresenta baixos teores de nitrogênio sendo muitas vezes necessária a realização de suplementação com sais amoniacais para redução do tempo de fermentação total (BELY, SABLAYROLLES & BARRE, 1990). As análises de GC-MS apontaram uma maior abundância de glicina nas amostras contaminadas com *Lactobacillus* sp. I-2 e *Acetobacter* sp. A-3 em detrimento do tratamento controle (sem contaminação) (Tabelas 11 e 12). Fermentações etanólicas em batelada alimentada caracterizam-se pela adição continuada de mosto de cana-de-açúcar até que o volume final da dorna seja alcançado (MAIORELLA, BLANCH & WILKE, 1981). A grande vantagem deste bioprocesso em detrimento de fermentações em batelada está na redução do tempo ocioso através da realização do enchimento da dorna e fermentação dos compostos carbonados de maneira concomitante (MAIORELLA, BLANCH & WILKE, 1981). Provavelmente, a grande densidade celular de leveduras e bactérias presentes nas fermentações etanólicas somada à simulação do bioprocessos de batelada alimentada tornaram limitante a absorção de glicose e frutose pelos microrganismos. Especula-se desta maneira, que a biossíntese de glicina possa ser oriunda, principalmente, da via do glioxalato em detrimento da utilização de compostos intermediários da via glicolítica. Ademais, a maior abundância deste aminoácido nas fermentações contaminadas com as bactérias, provavelmente reflete a necessidade de todos os microrganismos presentes no meio criarem alternativas para aquisição de nutrientes nitrogenados. Especula-se que ambos os microrganismos, bactérias e leveduras, acumularam glicina no ambiente intracelular com alternativa à baixa disponibilidade de nitrogênio no melaço de cana-de-açúcar.

O ácido málico foi apontado entre os metabólitos com diferença de abundância nas comparações par a par de meio de fermentação (4h) entre o tratamento controle e o contaminado com bactérias lácticas (Tabela 11). Esse metabólito é um componente do ciclo do ácido cítrico, importante via metabólica associada à oxidação de compostos carbonados para produzir moléculas como ATP e NADH₂ (ALBERTS et al., 2010). Em leveduras, a respiração aeróbica é ativada em condições de disponibilidade de oxigênio resultando na produção de biomassa celular (BASSO et al., 2008). Ademais, bactérias lácticas não produzem ácido málico, mas podem utilizar a energia resultante da sua degradação para acelerar o seu crescimento (BRANEN & KEENAN, 1970; SALOU et al., 1991; PASSOS et al., 2003). As fermentações realizadas em reator descritas previamente foram realizadas em condição de microaerofilia (oxigênio molecular limitante, porém não completamente ausente) e revelaram aumentos discretos de biomassa ao longo dos ciclos fermentativos de 8 horas. A presença do ácido málico entre os metabólitos mais abundantes em meio de fermentação (4h) pode ser reflexo da ativação do ciclo do ácido cítrico em leveduras durante as fermentações. Provavelmente uma maior abundância deste metabólito na condição controle em detrimento da contaminada com *Lactobacillus* sp. I-2 seja resultado do seu consumo pelas bactérias presentes no sistema.

Tabela 11. Tabela Volcano contendo os metabólitos resultantes das comparações estatísticas par a par de “*S. cerevisiae* CAT-1” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2”. Apenas foram considerados metabólitos que apresentaram “*Fold Change threshold*” mínimo de 4 e *p*-valor (teste *t*) menor ou igual a 0.05.

Metabolite	GMD ID	Statistic comparation	log2FC	-log10(p)	Most abundant group
beta-D-Fructofuranosyl-2,1-D-Fructofuranose	A267008	CAT-1 0h vs CAT-1 + I-2 0h	-8.9143	5.0691	CAT-1 0h
1-methyl-Tryptamine	A216012	CAT-1 0h vs CAT-1 + I-2 0h	-10.836	5.0088	CAT-1 0h
Unknown	A194007	CAT-1 0h vs CAT-1 + I-2 0h	11.334	4.4889	CAT-1 + I-2 0h
Galactonic acid-1,4-lactone	A189003	CAT-1 0h vs CAT-1 + I-2 0h	10.091	3.3689	CAT-1 + I-2 0h
Lysine	A186002	CAT-1 0h vs CAT-1 + I-2 0h	-11.3	3.2092	CAT-1 0h
Unknown	A286005	CAT-1 0h vs CAT-1 + I-2 0h	10.128	2.4466	CAT-1 + I-2 0h
Isomaltose	A287001	CAT-1 0h vs CAT-1 + I-2 0h	10.241	2.4339	CAT-1 + I-2 0h
Unknown	A235009	CAT-1 0h vs CAT-1 + I-2 0h	-2.2164	2.4142	CAT-1 0h
2-methyl-Maleic acid	A137003	CAT-1 0h vs CAT-1 + I-2 0h	10.638	2.3916	CAT-1 + I-2 0h
Hexaconasoic acid	A304003	CAT-1 0h vs CAT-1 + I-2 0h	10.576	1.8161	CAT-1 + I-2 0h
Unknown	A304003	CAT-1 0h vs CAT-1 + I-2 0h	9.2499	1.8142	CAT-1 + I-2 0h
Unknown	A272004	CAT-1 0h vs CAT-1 + I-2 0h	9.296	1.764	CAT-1 + I-2 0h
Unknown	A308005	CAT-1 0h vs CAT-1 + I-2 0h	12.512	1.7401	CAT-1 + I-2 0h
Unknown	A186010	CAT-1 4h vs CAT-1 + I-2 4h	-10.358	26.531	CAT-1 4h
Idose	A188011	CAT-1 4h vs CAT-1 + I-2 4h	-10.135	20.558	CAT-1 4h
Allose	A190004	CAT-1 4h vs CAT-1 + I-2 4h	10.322	15.908	CAT-1 + I-2 4h
Unknown	A300006	CAT-1 4h vs CAT-1 + I-2 4h	-9.7662	15.734	CAT-1 4h
trans-Ferulic acid	A210001	CAT-1 4h vs CAT-1 + I-2 4h	-2.4764	12.039	CAT-1 4h
Talose	A188010	CAT-1 4h vs CAT-1 + I-2 4h	-9.7301	4.6615	CAT-1 4h
Unknown	A283016	CAT-1 4h vs CAT-1 + I-2 4h	-10.011	4.6456	CAT-1 4h
Malic acid	A149001	CAT-1 4h vs CAT-1 + I-2 4h	-9.5428	4.4138	CAT-1 4h
Unknown	A272004	CAT-1 4h vs CAT-1 + I-2 4h	-11.043	4.3853	CAT-1 4h
Glycine	A133001	CAT-1 4h vs CAT-1 + I-2 4h	10.908	4.1568	CAT-1 + I-2 4h
Glycinamide	A143025	CAT-1 4h vs CAT-1 + I-2 4h	9.8428	3.0241	CAT-1 + I-2 4h
N-2-oxocaproyl-Homoserine lactone	A186014	CAT-1 4h vs CAT-1 + I-2 4h	10.67	2.9982	CAT-1 + I-2 4h
Unknown	A194007	CAT-1 4h vs CAT-1 + I-2 4h	-3.9301	2.4827	CAT-1 4h
Gluconic acid	A200001	CAT-1 4h vs CAT-1 + I-2 4h	9.4119	2.139	CAT-1 + I-2 4h

Tabela 11. Tabela Volcano contendo os metabólitos resultantes das comparações estatísticas par a par de “*S. cerevisiae* CAT-1” vs “*S. cerevisiae* CAT-1 + *Lactobacillus* sp. I-2”. Apenas foram considerados metabólitos que apresentaram “*Fold Change threshold*” mínimo de 4 e *p*-valor (teste *t*) menor ou igual a 0.05. (Continuação).

Metabolite	GMD ID	Statistic comparison	log2FC	-log10(p)	Most abundant group
Maltitol	A284001	CAT-1 4h vs CAT-1 + I-2 4h	9.9022	2.1386	CAT-1 + I-2 4h
Alanine	A138002	CAT-1 4h vs CAT-1 + I-2 4h	9.4957	2.0493	CAT-1 + I-2 4h
Unknown	A289002	CAT-1 4h vs CAT-1 + I-2 4h	-9.4877	2.0231	CAT-1 4h
Unknown	A286005	CAT-1 4h vs CAT-1 + I-2 4h	9.0455	1.5444	CAT-1 + I-2 4h
Galactonic acid-1,4-lactone	A189003	CAT-1 4h vs CAT-1 + I-2 4h	10.184	1.5402	CAT-1 + I-2 4h
alpha-Lactose	A269001	CAT-1 4h vs CAT-1 + I-2 4h	10.344	1.5169	CAT-1 + I-2 4h
Unknown	A300006	CAT-1 8h vs CAT-1 + I-2 8h	-12.302	28.937	CAT-1 + I-2 8h
trans-Ferulic acid	A210001	CAT-1 8h vs CAT-1 + I-2 8h	-13.17	24.828	CAT-1 + I-2 8h
beta-D-Fructofuranosyl-2,1-D-Fructofuranose	A267008	CAT-1 8h vs CAT-1 + I-2 8h	-11.69	21.515	CAT-1 + I-2 8h
2-methyl-Maleic acid	A137003	CAT-1 8h vs CAT-1 + I-2 8h	-11.001	13.177	CAT-1 + I-2 8h
Unknown	A282009	CAT-1 8h vs CAT-1 + I-2 8h	11.692	13.033	CAT-1 8h
Ribonic acid	A177001	CAT-1 8h vs CAT-1 + I-2 8h	-11.447	5.6764	CAT-1 + I-2 8h
Beta beta-Trehalose	A275017	CAT-1 8h vs CAT-1 + I-2 8h	-11.524	5.6761	CAT-1 + I-2 8h
Talose	A188010	CAT-1 8h vs CAT-1 + I-2 8h	-12.5	5.6145	CAT-1 + I-2 8h
Gluconic acid-1,4-lactone	A189013	CAT-1 8h vs CAT-1 + I-2 8h	-12.204	3.6973	CAT-1 + I-2 8h
Sorbitol	A193001	CAT-1 8h vs CAT-1 + I-2 8h	12.049	3.607	CAT-1 8h
Galactonic acid-1,4-lactone	A189003	CAT-1 8h vs CAT-1 + I-2 8h	12.266	3.4213	CAT-1 8h
Idose	A188011	CAT-1 8h vs CAT-1 + I-2 8h	-11.464	2.614	CAT-1 + I-2 8h
3-amino Propane-1,2-diol	A139011	CAT-1 8h vs CAT-1 + I-2 8h	-10.128	2.6035	CAT-1 + I-2 8h
Lactulose	A261012	CAT-1 8h vs CAT-1 + I-2 8h	-12.073	2.6025	CAT-1 + I-2 8h
Gulose	A190003	CAT-1 8h vs CAT-1 + I-2 8h	-11.052	2.5939	CAT-1 + I-2 8h
N-dansyl-Aziridine	A250011	CAT-1 8h vs CAT-1 + I-2 8h	-2.1507	2.4699	CAT-1 + I-2 8h
Phosphoric acid	A129001	CAT-1 8h vs CAT-1 + I-2 8h	-2.46	2.0821	CAT-1 + I-2 8h
Fumaric acid	A137001	CAT-1 8h vs CAT-1 + I-2 8h	-2.4568	1.5635	CAT-1 + I-2 8h
Unknown	A144007	CAT-1 8h vs CAT-1 + I-2 8h	-2.7055	1.3818	CAT-1 + I-2 8h

Tabela 12. Tabela Volcano contendo os metabólitos resultantes das comparações estatísticas par a par de “*S. cerevisiae* CAT-1” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3”. Apenas foram considerados metabólitos que apresentaram “*Fold Change threshold*” mínimo de 4 e *p*-valor (teste *t*) menor ou igual a 0.05.

Metabolite	GMD ID	Statistic comparison	log2FC	-log10(p)	Most abundant group
Homogentisic acid	A182016	CAT-1 0h vs CAT-1 + A-3 0h	-13.328	16.798	CAT-1 0h
Isomaltose	A287001	CAT-1 0h vs CAT-1 + A-3 0h	12.527	11.15	CAT-1 + A-3 0h
Galactonic acid-1,4-lactone	A189003	CAT-1 0h vs CAT-1 + A-3 0h	11.779	8.8365	CAT-1 + A-3 0h
Unknown	A262001	CAT-1 0h vs CAT-1 + A-3 0h	12.144	5.632	CAT-1 + A-3 0h
2-methyl-Maleic acid	A137003	CAT-1 0h vs CAT-1 + A-3 0h	12.462	5.0056	CAT-1 + A-3 0h
Idose	A188011	CAT-1 0h vs CAT-1 + A-3 0h	-11.189	3.6548	CAT-1 0h
Unknown	A194007	CAT-1 0h vs CAT-1 + A-3 0h	10.748	3.5415	CAT-1 + A-3 0h
Lysine	A186002	CAT-1 0h vs CAT-1 + A-3 0h	-12.78	3.5014	CAT-1 0h
Unknown	A289002	CAT-1 0h vs CAT-1 + A-3 0h	12.301	2.0205	CAT-1 + A-3 0h
Ribonic acid	A177001	CAT-1 0h vs CAT-1 + A-3 0h	10.459	1.9775	CAT-1 + A-3 0h
beta beta-Trehalose	A275017	CAT-1 0h vs CAT-1 + A-3 0h	11.009	1.9547	CAT-1 + A-3 0h
Unknown	A283016	CAT-1 4h vs CAT-1 + A-3 4h	-8.2182	5.6574	CAT-1 4h
Galactonic acid-1,4-lactone	A189003	CAT-1 4h vs CAT-1 + A-3 4h	8.4929	5.402	CAT-1 + A-3 4h
Unknown	A272004	CAT-1 4h vs CAT-1 + A-3 4h	-9.1196	5.1338	CAT-1 4h
Gluconic acid-1,4-lactone	A189013	CAT-1 4h vs CAT-1 + A-3 4h	8.3837	3.7436	CAT-1 + A-3 4h
D-Glucopyranose	A189005	CAT-1 4h vs CAT-1 + A-3 4h	-8.5325	3.6941	CAT-1 4h
Glycine	A133001	CAT-1 4h vs CAT-1 + A-3 4h	9.0651	3.4389	CAT-1 + A-3 4h
Unknown	A192001	CAT-1 4h vs CAT-1 + A-3 4h	-7.8717	2.612	CAT-1 4h
Allose	A190004	CAT-1 4h vs CAT-1 + A-3 4h	7.2835	2.0209	CAT-1 + A-3 4h
Gluconic acid	A200001	CAT-1 4h vs CAT-1 + A-3 4h	6.7828	2.0204	CAT-1 + A-3 4h
Maltitol	A284001	CAT-1 4h vs CAT-1 + A-3 4h	7.5903	2.0191	CAT-1 + A-3 4h
Alanine	A138002	CAT-1 4h vs CAT-1 + A-3 4h	6.3818	2.0146	CAT-1 + A-3 4h
Glycinamide	A143025	CAT-1 4h vs CAT-1 + A-3 4h	7.9177	2.0094	CAT-1 + A-3 4h

Tabela 12. Tabela Volcano contendo os metabólitos resultantes das comparações estatísticas par a par de “*S. cerevisiae* CAT-1” vs “*S. cerevisiae* CAT-1 + *Acetobacter* sp. A-3”. Apenas foram considerados metabólitos que apresentaram “*Fold Change threshold*” mínimo de 4 e *p*-valor (teste *t*) menor ou igual a 0.05. (Continuação).

Metabolite	GMD ID	Statistic comparison	log2FC	-log10(p)	Most abundant group
Homogentisic acid	A182016	CAT-1 8h vs CAT-1 + A-3 8h	-11.557	18.593	CAT-1 8h
Unknown	A286005	CAT-1 8h vs CAT-1 + A-3 8h	-11.363	17.974	CAT-1 8h
Gluconic acid	A200001	CAT-1 8h vs CAT-1 + A-3 8h	-10.784	5.6843	CAT-1 8h
Ribonic acid	A177001	CAT-1 8h vs CAT-1 + A-3 8h	-11.157	5.6756	CAT-1 8h
Unknown	A308005	CAT-1 8h vs CAT-1 + A-3 8h	-10.87	5.6711	CAT-1 8h
alpha-Lactose	A269001	CAT-1 8h vs CAT-1 + A-3 8h	-10.457	5.6698	CAT-1 8h
Unknown	A287005	CAT-1 8h vs CAT-1 + A-3 8h	-11.019	5.6343	CAT-1 8h
Unknown	A282009	CAT-1 8h vs CAT-1 + A-3 8h	11.518	5.615	CAT-1 + A-3 8h
Galactonic acid-1,4-lactone	A189003	CAT-1 8h vs CAT-1 + A-3 8h	11.652	2.6909	CAT-1 + A-3 8h
Glycinamide	A143025	CAT-1 8h vs CAT-1 + A-3 8h	-11.455	2.6127	CAT-1 8h
Unknown	A277007	CAT-1 8h vs CAT-1 + A-3 8h	-10.769	2.6074	CAT-1 8h
Unknown	A272004	CAT-1 8h vs CAT-1 + A-3 8h	-9.5284	2.6057	CAT-1 8h
3-amino Propane-1,2-diol	A139011	CAT-1 8h vs CAT-1 + A-3 8h	-9.8389	2.6028	CAT-1 8h
N-3-aminopropyl-Morpholine	A165009	CAT-1 8h vs CAT-1 + A-3 8h	10.671	2.0189	CAT-1 + A-3 8h
Unknown	A192001	CAT-1 8h vs CAT-1 + A-3 8h	11.343	2.0182	CAT-1 + A-3 8h
Sorbitol	A193001	CAT-1 8h vs CAT-1 + A-3 8h	11.898	1.9996	CAT-1 + A-3 8h
Glycerol	A129003	CAT-1 8h vs CAT-1 + A-3 8h	-2.4412	1.7594	CAT-1 8h

5 CONCLUSÕES

O presente trabalho objetivou levantar o conteúdo total de genes e mensurar as principais relações filogenéticas de três linhagens bacterianas isoladas de fermentações brasileiras para produção de etanol. Os estudos genômicos e filogenéticos permitiram identificar as espécies *L. fermentum*, *L. plantarum* e *A. senegalensis* como as mais semelhantes à *Lactobacillus* sp. I-2, *Lactobacillus* sp. L-10 e *Acetobacter* sp. A-3, respectivamente. A denominação final destes microrganismos, após a conclusão dos estudos filogenéticos multigênicos, foi *Lactobacillus fermentum* I-2, *Lactobacillus plantarum* L-10 e *Acetobacter senegalensis* A-3.

Os estudos pangenômicos das linhagens bacterianas permitiram o levantamento de potenciais adaptações genéticas às fermentações etanólicas industriais brasileiras. De posse do conteúdo gênico de diferentes linhagens de *Lactobacillus* e *Acetobacter* foi possível identificar Potenciais Proteínas Preditas (PPPs) *core*, acessórias e exclusivas para *L. fermentum* I-2, *L. plantarum* L-10 e *A. senegalensis* A-3. O apontamento de PPPs acessórias e exclusivas com funções associadas a diferentes “*Clusters of Orthologous Genes*” (COGs) permitiu a identificação de genes associados à transposases, proteínas virais, reguladores transcricionais associados à tolerância de bactérias a antibióticos e transportadores do tipo ABC. As presenças destes genes nos conjuntos acessórios e exclusivos das bactérias sugerem a ocorrência de adaptações genéticas nestes microrganismos associadas a uma maior tolerância a antibióticos e outros compostos citotóxicos.

Os ensaios fermentativos executados em reatores biológicos resultaram em alterações significativas nas concentrações de ácido láctico e glicerol no fermentado nas condições experimentais contaminadas com *L. fermentum* I-2. Interessantemente, mesmo com populações superiores a 10^8 UFC.mL⁻¹, não foi detectado a ocorrência do processo de floculação. Carvalho-Netto et al. (2015) reportou floculação pode ocorrer em fermentações contendo 10^5 UFC.mL⁻¹ de *L. fermentum* em fermentações contendo leveduras de *S. cerevisiae* PE-2. Provavelmente, a diferença das linhagens de bactérias e leveduras entre os trabalhos influenciaram nas divergências finais.

A ocorrência de uma alta mortalidade celular da bactéria acética *A. senegalensis* A-3 sugere que a limitação de oxigênio presente no meio é determinante na sobrevivência e proliferação deste microrganismo. As populações finais variaram de 10^8 UFC.mL⁻¹ no primeiro ciclo fermentativo em batelada alimentada para 10^5 UFC.mL⁻¹ no final do terceiro ciclo.

Os estudos dos metabólitos intracelulares de bactérias e leveduras coletadas nos ensaios fermentativos revelaram que a presença das bactérias foi capaz de alterar o metabolismo intracelular das leveduras de *S. cerevisiae*. Um dos principais resultados foi a detecção de uma maior abundância de glicina, um aminoácido polar, no meio das fermentações contaminadas com *L. fermentum* I-2 e *A. senegalensis* A-3. Provavelmente, o acúmulo deste aminoácido reflete uma maior competição das bactérias por nutrientes nitrogenados presentes no melaço de cana-de-açúcar. Ambos os microrganismos, bactérias e leveduras, provavelmente necessitaram ativar mecanismos intracelulares de produção de fontes alternativas de nitrogênio para sobreviver às condições ambientais impostas nas fermentações.

REFERÊNCIAS

- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2010) *Biologia Molecular da célula*. **Artmed** 5.ed. Porto Alegre 1268 p.
- Albertyn N, Hohmann S, Thevelein JM, Prior BA (1994) GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathways. **Mol. Cell Biol** 6:135-144
- Amorim HV, Lopes ML, de Castro Oliveira JV, Buckridge MS, Goldman GH (2011) Scientific challenges of bioethanol production in Brazil. **Appl Microbiol Biotechnol** 91:1267-75
- Amorim-Neto HB, Yohanann BK, Bringhurst TA, Brosnan JM, Pearson SY, Walker W, Walker GM (2009) Evaluation of a Brazilian fuel alcohol yeast strain for Scotch whisky fermentations. **J I Brewing** 115:198-207
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. **Disponível em:** <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>>. Acesso em 23/04/2015
- Andrietta, MGS, Steckelberg C, Andrietta SR (2006) Bioetanol – Brasil, 30 anos na vanguarda. **MultiCiência: Construindo a história dos produtos naturais** 7:1-16
- Andrietta SR, Stupiello JP (1990) Simulação e modelagem para processos de fermentação alcoólica contínua. **STAB Açúcar, Álcool e Subprodutos** 8:36-40
- Ansell R, Granath K, Hohmann S, Thevelein JM, Adler L (1997) The two isoenzymes for yeast NAD-dependent glycerol-3-phosphate dehydrogenase encoded by GPD1 and GPD2 have distinct roles in osmoadaptation and redox regulation. **EMBO J** 9:2179-2187
- Argüelles JC (2000) Physiology roles of trehalose in bacteria and yeasts: a comparative analysis. **Arch Microbiol** 4:217-224
- Babrzadeh F, Jalili R, Wang C, Shokralla S, Pierce S, Robinson-Mosher A, Nyren P, Shafer RW, Basso LC, de Amorim HV, de Oliveira AJ, Davis RW, Ronaghi M, Gharizadeh B, Stambuk BU (2012) Whole-genome sequencing of the eficiente industrial fuel-ethanol fermentative *Saccharomyces cerevisiae* strain CAT-1. **Mol Genet Genomics** 287:485-495
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijbelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevsner PA (2012) SPAdes: A new Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. **J Comput Biol** 19:455-477
- Basso LC, Amorim HV, De Oliveira AJ, Lopes ML (2008) Yeast selection for fuel ethanol production in Brazil. **FEMS Yeast Res** 8:1155-1163
- Basso LC, Basso TO, Rocha SN (2011) Ethanol production in Brazil: the industrial process and its impact on yeast fermentation. **Biofuel Production – Recent Developments and Prospects** 1:85-100
- Belly M, Sablayrolles JM, Barre P (1990) Description of alcoholic fermentation kinetics: its variability and significance. **Am J Enol Vitic** 159:25-32
- Bideaux C, Alfenore S, Cameleyre X, Molina-Jouve C, Uribe Larrea JL, Guillouet SE (2006) Minimization of glycerol production during the high-performance fed-batch ethanolic fermentation in *Saccharomyces cerevisiae*, using a metabolic model as a prediction tool. **Appl Environ Microbiol** 72:2134-2140
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics** 30:2114-2120

- Bonatelli ML, Quecine MC, Silva MS, Labate CA (2017) Characterization of the contaminant bacterial communities in sugarcane first-generation industrial ethanol production. **FEMS Microbiol Lett** 364:1-17
- Borges PCS (2008) Otimização da dinâmica de fermentação alcoólica no processo em batelada alimentada. **Dissertação** (UFU Engenharias) 162 p.
- Branen AL, Keenan TW (1970) Growth stimulation of *Lactobacillus casei* by sodium citrate. **J Dairy Sci** 53:593-597
- Brown AD (1990) Microbial water stress physiology, principles and perspectives. **John Wiley & Sons** 313p.
- Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. **Nature Methods** 12:59-60
- Carvalho AL, Cardoso FS, Bohn A, Neves AR, Santos H (2011) Engineering Trehalose Synthesis in *Lactococcus lactis* for Improved Stress Tolerance. **Appl Environ Microbiol** 12:4189-4199
- Carvalho-Netto OV, Carazzolle MF, Mofatto LS, Teixeira PJPL, Noronha MF, Calderón LAL, Mieczkowski PA, Argueso JL, Pereira GAG (2015) *Saccharomyces cerevisiae* transcriptional reprogramming due to bacterial contamination during industrial scale bioethanol production. **Microbial Cell Factories** 14:1-13
- Ceccato-Antonini SR, Silva DF (2000) Eficiência de meios diferenciais no isolamento de cepas de leveduras de processos industriais de fermentação alcoólica. **STAB Açúcar Álcool e Subprodutos** 18:40-46
- Chikhi R, Medvedev P (2013) Informed and automated k-mer size selection for genome assembly. **Bioinformatics** 30:31-37
- Chaudhari NM, Gupta VK, Dutta C (2016) BPGA – an ultra-fast pan-genome analysis pipeline. **Sci Rep** 6:1-10
- Colquhoun IJ (2007) Use of NMR for metabolic profiling in plant systems. **Journal of Pesticide Science** 32:200-212
- Costa VM, Basso TO, Angeloni LHP, Oetterer M, Basso LC (2008) Production of acetic acid, ethanol and optical isomers of lactic acid by *Lactobacillus* strains isolated from industrial ethanol fermentations. **Ciência e Agrotecnologia** 32(2):503-509.
- Cuadros-Inostroza A, Caldana C, Redestig H, Kusano M, Lisec J, Peña-Cortés H, Willmitzer L, Hannah MA (2009) TargetSearch: a bioconductor package for the eficiente preprocessing of GC-MS metabolite profiling data. **BMC Bioinformatics** 10:1-12
- Culibrk L, Croft CA, Tebbutt SJ (2016) Systems Biology Approaches for Host-Fungal Interactions: na expading multiomics frontier. **OMICS: a Journal of Integrative Biology** 20:127-138
- De Boiso JF, Stoppani AO (1967) Metabolism of serine and glycine in baker's yeast. **Biochim Biophys Acta** 148:48-59
- De Man JC, Rogosa M, Sharpe ME (1960) A medium for the cultivation of *Lactobacilli*. **J Appl Microbiol** 37:511-513
- Du Fall LA, Solomon PS (2011) Role of cereal secondary metabolites involved in mediating the outcome of plant: pathogen interactions. **Metabolite** 1:64-78
- Ewing B, Green P (1998) Base-Calling of Automated Sequencer Traces Using Phred. **Genome Res** 8:186-194
- Fiehn O (2002) Metabolomics -- The link between Genotypes and Phenotypes. **Plant Mol Biol** 48:155-171
- Gallo CR (1989) Determinação da microbiota bacteriana de mosto e de dornas de fermentação alcoólica. **Tese** (Unicamp) 388 p.
- Garcia DM, Dietrich D, Clardy J, Jarosz D (2016) A common bacterial metabolite elicits prion-based by-pass of glucose repression. **Microbiol Infect Disease** 5:1-17

- Goffeau A, Barrel BG, Bussey H, Davis RW, Dujon B, Fieldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. **Science** 546:563-567
- Gullberg J, Jonsson P, Nordström A, Sjöström M, Moritz T (2004) Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of *Arabidopsis thaliana* samples in metabolomics studies with gas chromatography/mass spectrometry. **Analytical Biochemistry** 331(2):283-295.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUASt: quality assessment tool for genome assemblies. **Bioinformatics** 15:1072-1075
- Holzapfel WH, Haberer P, Geisen R, Björkroth J, Schilinger U (2004) Taxonomy and important features of probiotic microorganisms in food nutrition. **Am J Clin Nutr** 73:1-66
- Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. **Genomic Research** 9:868-877
- Jellen-Ritter AS, Kern WV (2001) Enhanced Expression of the multidrug efflux pumps AcrAB e AcrEF associated with insertion element transposition in *Escherichia coli* mutants selected with Fluoroquinolone. **Antimicrob Agents Chemother** 45:1467-1472
- Jonsson P, Johansson AI, Gullberg J, Trygg JAJ, Grung B, Marklund S, Sjöström M, Antti H, Moritz T (2005) High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analysis. **Anal Chem** 77:5635-5642
- Kadere TT, Miyamoto T, Oniang RK, Kutima PM, Njoroge SM (2008) Isolation and identification of the genera *Acetobacter* and *Gluconobacter* in coconut toddy (mnazi). **African J Biotechnol** 7: 2963-2971
- Katoh K, Standley DM (2013) MAFFT Multiple Sequence Alignment Software version 7: Improvements in Performance and Usability. **Mol Biol Evol** 4:772-780
- Kandler O (1983) Carbohydrate metabolism in lactic acid bacteria. **Antonie Van Leeuwenhoek Journal of Microbiology** 49(3): 209-224
- Konings WN, Poolman B, Van Veen HM (1994) Solute transport and energy transduction in bacteria. **J Microbiol** 4:369-380
- Konstantinidis KT, Tiedje JM (2005) Genomic insights that advances the species definition por prokaryotes. **Proc Natl Acad Sci USA** 7:2567-2572
- Kumar S, Jones M, Koutsovoulos G, Clarke M, Blaxter M (2013) Blobology: exploring raw genome data for contaminants, symbionts, and parasites using taxon-annotated GC-coverage plots. **Front Genet** 4:1-12
- Ladero V, Ramos A, Wiersma A, Goffin P, Schanck A, Kleerebezem M, Hugenholtz J, Smid EJ, Hols P (2007) High-level production of the low-calorie sugar sorbitol by *Lactobacillus plantarum* through metabolic engineering. **Appl Environ Microbiol** 73:1864-1872
- Lane DJ (1991) 16S/23S rRNA Sequencing. **Nucleic Acid techniques in bacterial systematics** 1:115-175
- Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. **Nucleic Acids Research** 44:242-245
- Lopes ML, Paulillo SCL, Godoy A, Cherubin RA, Lorenzi MS, Giometti FHC, Bernadino CD, de Amorim-Neto HB, de Amorim HV (2016) Ethanol production in Brazil: a bridge between science and industry. **Biotech Indus Microbiol** 47:64-76
- Lucena BTL, Santos BM, Moreira JLS, Moreira APB, Nunes AC, Azevedo V, Miyoshi A, Thompson FL, Junior MAM (2010) Diversity of lactic acid bacteria of the bioethanol process. **BMC Microbiol** 23:1-8

- Ludwig KM, Oliva-Neto P, Angelis DF (2001) Quantificação da floculação de *S. cerevisiae* por bactérias contaminantes da fermentação alcoólica. **Ciência e Tecnologia de Alimentos** 21:63-66
- Maiorella BL, Blanch HW, Wilke CR (1981) Advances in biochemical engineering/ Biotechnology. **Spring Verlag** 1:43-92
- Matsushita K, Toyama H, Tonouchi N, Okamoto-Kainuma A (2016) Acetic Acid Bacteria – Ecology and Physiology. **Springer** 350 p
- Mattenberger F, Sabater-Muñoz A B, Hallsworth JE, Fares MA (2016) Glycerol stress in *Saccharomyces cerevisiae*: cellular responses and evolved adaptations. **Environ Microbiol** 19:990-1007
- Moreira JR, Goldemberg, J (1996) The Alcohol Program. **Energy Policy** 27:229-245
- Ndoye B, Cleenwerck I, Engelbeen K, Dubois-Dauphin R, Guirio AT, Trappen SV, Willems A, Thonart P (2007) *Acetobacter senegalensis* sp. nov., a thermotolerant acetic acid bacterium isolated in Senegal (sub-Saharan Africa) from mango fruit (*Mangifera indica* L.). **Int J Syst Evol Microbiol** 57:1576-1581
- Nevoigt E, Stahl U (1997) Osmoregulation and glycerol metabolism in the yeast *Saccharomyces cerevisiae*. **FEMS Microbiol Rev** 21:231-241
- Nissen TL, Hamann CW, Kielland-Brandt MC, Nielsen J, Villadsen J (2000) Anaerobic and aerobic batch cultivations of *Saccharomyces cerevisiae* mutants impaired in glycerol synthesis. **Yeast** 16:463-474
- Nissen L, Pérez-Martínez G, Yebra MJ (2005) Sorbitol synthesis by a engineered *Lactobacillus casei* strain expressing a sorbitol-6-phosphate dehydrogenase gene within the lactose operon. **FEMS Microbiol Lett** 249:177-183
- Norbeck J, Pahlman A, Akhta N, Blomberg A, Adler L (1996) Purification and characterization of two isoenzymes of DL-glycerol-3-phosphate from *Saccharomyces cerevisiae*. Identification of the corresponding GPP1 and GPP2 genes and evidence of osmotic regulation of Gpp2p expression by the osmosensing mitogen-activated protein kinase signal transduction pathway. **J Biol Chem** 271:13875-13881
- Ostergaard S, Olsson L, Nilsen J (2000) Metabolic engineering of *Saccharomyces cerevisiae*. **Microbiol Mol Biol Rev** 64:34-50
- Panek A (1963) Function of trehalose in Baker's yeast (*Saccharomyces cerevisiae*). **Arch Biochem Biophys** 100:422-425
- Parrou JL, Teste MA, François J (1997) Effects of various types of stress on the metabolism of reserve carbohydrates in *Saccharomyces cerevisiae*: genetics evidence of a stress-induced recycling of glycogen and trehalose. **Microbiology** 143:1891-1900
- Passos FV, Fleming HP, Hassan HM, McFeeters RF (2003) Effect of malic acid on the growth kinetics of *Lactobacillus plantarum*. **Appl Microbiol Biotechnol** 63:207-211
- Pelczar JR, Chan ECS, Krieg NR (1997) Microbiologia: conceitos e aplicações. **Makron Books** 2.ed. São Paulo 1324p.
- Petitjean M, Teste MA, François JM, Parrou JL (2015) Yeast tolerance to various stresses relies on the Trehalose-6P Synthase (Tps1) Protein, not on Trehalose. **J Biol Chem** 290:16177-16190
- Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. **Trends in Genet** 16:276-277
- Roman H (1981) Development of yeast as an experimental organisms. **Cold Spring Harbor Monogr Ser** 1:1-10
- Salou P, Leroy MJ, Goma G, Pareilleux A (1991) Influence of pH and malate-glucose ratio on the growth of *Leuconostoc oenos*. **Appl Microbiol Biotechnol** 36:87-91

- Santos SC, Dionisio SR, De Andrade ALD, Roque LR, Da Costa AC, Ienczak JL (2015) Fermentation of xylose and glucose mixture in intensified reactors by *Scheffersomyces stipites* to produce ethanol. **International Journal of Biotechnology and Bioengineering**: 9(5)1-9
- Schauer N, Steinhäuser D, Strelkov S, Schomburg D, Allison G, Moritz T, Lundgren K, Roessner-Tunali U, Forbes MG, Willmitzer L, Fernie AR, Kopka J (2005) GC-MS libraries for the rapid identification of metabolites in complex biological samples **FEBS Lett** 6:1332-1337
- Seemann T (2014) Prokka: rapid prokaryotic genome annotation. **Bioinformatics** 14:2068-2069
- Sengupta A, Ghosh S, Basant A, Malusare S, Johri P, Pathak S, Sharma S, Sonawat HM (2011) Global host metabolic response to *Plasmodium vivax* infection: a HNMR urinary metabolomics study. **Malaria J** 10:1-13
- Shen B, Hohmann S, Jensen RG, Bohnert HJ (1999) Roles of sugar alcohols in osmotic stress adaptation. Replacement of Glycerol by mannitol and sorbitol in yeast. **Plant Physiol** 121:45-52
- Silva JPA, Mussato SI, Roberto IC, Teixeira JA (2012). Ethanol production from xylose by *Pichia stipitis* NRRL Y-7124 in a Stirred tank bioreactor. **Brazilian Journal of Chemical Engineering** 28(1):151-156
- Simão FA, Waterhouse RM, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics** 31:3210-3212
- Sinclair DA, Dawes IW (1995) Genetics of the synthesis of serine from glycine and the utilization of glycine as sole nitrogen source by *Saccharomyces cerevisiae*. **Genetics** 140:13-22
- Singer MA, Lindquist S (1998) Multiple effects of trehalose on protein folding in vivo. **Mol. Cell** 1:639-648
- Spargo BJ, Crowe LM, Itoneda T, Beaman BL, Crowe JH (1991) Cord factor (alpha, alpha-trehalose 6,6'-dimycolate) fusion between phospholipid vesicles. **Proc Natl Acad Sci USA** 3:737-740
- Stamatakis A (2006) RaxML-VI-HP: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics** 22:2688-2690
- Stewart GG, Russel I (1998) An Introduction to Brewing Science & Technology – Brewer's Yeast. **The Institute of Brewing** 108p.
- Strom AR, Kaasen I (1993) Trehalose metabolism in *Escherichia coli*: stress protection and stress regulation of gene expression. **Mol Microbiol** 2:205-210
- Tabai Y, Kiryu H, Kin T, Asai K (2008) A fast structural multiple alignment method for long RNA sequences. **BMC Bioinformatics** 9:1-17
- Tettelin H, Maignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM. (2005) Genomic analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications of microbial "pan-genome". **Proc Natl Acad Sci USA** 102:13950-13955
- Tiukova I, Eberhard T, Passoth V (2014) Interaction of *Lactobacillus vini* with the ethanol-producing yeasts *Dekkera bruxellensis* and *Saccharomyces cerevisiae*. **Biotechnol Appl Biochem** 64:40-44
- Tortora GJ, Funke BR, Case CL (2008) Microbiology. **Artmed** 8.ed. Berkeley 920 p.
- Tourinho OAF, ANDRADE, SC (1998) Cenários para o início do milênio no Brasil. **Ipea** 1:1-12
- Trcek J, Teuber M (2002) Genetic and restriction analysis of the 16S-23S rDNA internal transcribed spacer regions of the acetic acid bacteria. **FEMS Microbiol Lett** 208: 69-75

- Turner S, Pryer KM, Miao VP, Palmer JD (1999) Investigation deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. **J Eukaryot Microbiol** 46(4):327-38
- Ulane R, Ogur M (1972) Genetic and physiological control of serine and glycine biosynthesis in *Saccharomyces*. **J Bacteriol** 109:34-43
- União da indústria de cana-de-açúcar (UNICA) – Moagem de cana-de-açúcar e produção de etanol na safra de 2017/2018. **Disponível em:** <<http://www.unidata.com.br/>>. Acesso em 29 de janeiro de 2019.
- Varghese NJ, Mukherjee S, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pardi A (2015) Microbial species delineation using whole genome sequences. **Nucleic Acids Res** 43:6761-6771
- Veritas Genetics: whole genome sequencing company. **Disponível em** <<http://www.veritasgenetics.com/>>. Acesso em 15 de janeiro de 2019.
- Verpoorte R (2007) NMR-based metabolomics at work in phytochemistry. **Phytochem Rev** 6:3-14
- Viegas MC (1999) Desenvolvimento de um reator de bancada de leito fluidizado para produção de etanol utilizando linhagens de leveduras floculantes. **Tese** (Unicamp) 86p.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. **Plos One** 9:1-14
- Wang H, Dzink-Fox J, Chen M, Levy SB (2001) Genetic characterization of Highly Fluoroquinolone-Resistant clinical *Escherichia coli* strains from China: role of *acrR* mutations. **Antimicrob Agents Chemother** 45:1515-1521
- Wolfender JL, Rodriguez S, Hostettmann K (1998) Liquid chromatography coupled to mass spectrometry and nuclear magnetic resonance spectrometry of the screening of plant constituents. **J Chromatogr A** 794:299-316
- Xia J, Sinelkinov I, Han B, Wishart DS (2015) MetaboAnalyst 3.0: making metabolomics more meaningful. **Nucleic Acids Res** 43:251-257
- Yebra MJ, Pérez-Martínez G (2002) Cross-talk between the L-sorbose and D-sorbitol (D-glucitol) metabolic pathways in *Lactobacillus casei*. **Microbiology** 148:2351-2359
- Yokoya F, Oliva-Neto P (1991) Características de floculação de leveduras por *Lactobacillus fermentum*. **Revista de Microbiologia** 22:12-16
- Yon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J (2017) Introducing EzBioCloud: A taxonomically united database of 16S rRNA and whole genome assemblies. **Int J of Syst Evol Microbiol** 67:1613-1617
- Yoshiyama Y, Tanaka K, Yoshiyama K, Hibi M, Ogawa J, Shima J (2015) Trehalose accumulation enhances tolerance of *Saccharomyces cerevisiae* to acetic acid. **J Biosci Bioeng** 2:172-175
- Zanin GM, Santana CS, Bon EPS, Giordano RCL, Moraes FF, Andrietta SR, Carvalho-Neto CC, Macedo IC, Lahr-Fo D, Ramos LP, Fontana SR (2000) Brazilian Bioethanol Program. **Appl Biochem Biotechnol** 84:1147-1161
- Zerbino DR, Birney E (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. **Genome Research**, 18:821-829
- Zhu BH, Song YN, Xue W, Xu GC, Xiao J, Sun MY, Sun, XW, Li JT (2016) PEP_scaffolder: using (homologous) proteins to scaffold genomes. **Bioinformatics** 20:3193-3195

APÊNDICES

APÊNDICE A. Tabela com a descrição dos *Clusters of Orthologous Genes* (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. I-2. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs.

COG ID	Category	Group	Description
COG2002	Exclusive	Transcription	Bifunctional DNA-binding transcriptional regulator of stationary/ sporulation/ toxin gene expression and antitoxin component of the YhaV-PrfF toxin-antitoxin module
COG3279	Exclusive	Transcription	DNA-binding response regulator, LytR/AlgR family
COG3711	Accessory	Transcription	Transcriptional antiterminator
COG0640	Accessory	Transcription	DNA-binding transcriptional regulator, ArsR family
COG1309	Accessory	Transcription	DNA-binding transcriptional regulator, AcrR family
COG1309	Accessory	Transcription	DNA-binding transcriptional regulator, AcrR family
COG0454	Accessory	Transcription	N-acetyltransferase, GNAT superfamily (includes histone acetyltransferase HPA2)
COG0553	Accessory	Transcription	Superfamily II DNA or RNA helicase, SNF2 family
COG0477	Accessory	Amino acid transport and metabolism	MFS family permease
COG0119	Accessory	Amino acid transport and metabolism	Isopropylmalate/homocitrate/citramalate synthases
COG0128	Accessory	Amino acid transport and metabolism	5-enolpyruvylshikimate-3-phosphate synthase
COG3534	Exclusive	Carbohydrate transport and metabolism	Alpha-L-arabinofuranosidase
COG2273	Exclusive	Carbohydrate transport and metabolism	Beta-glucanase, GH16 family
COG0477	Accessory	Carbohydrate transport and metabolism	MFS family permease
COG5263	Accessory	Carbohydrate transport and metabolism	Glucan-binding domain (YG repeat)
COG0036	Accessory	Carbohydrate transport and metabolism	Pentose-5-phosphate-3-epimerase
COG3414	Accessory	Carbohydrate transport and metabolism	Phosphotransferase system, galactitol-specific IIB component
COG1472	Accessory	Carbohydrate transport and metabolism	Periplasmic beta-glucosidase and related glycosidases
COG0489	Accessory	Cell cycle control, cell division, chromosome partitioning	Chromosome partitioning ATPase, Mrp family, contains Fe-S cluster
COG0463	Accessory	Cell wall/membrane/envelope biogenesis	Glycosyltransferase involved in cell wall biosynthesis
COG0438	Accessory	Cell wall/membrane/envelope biogenesis	Glycosyltransferase involved in cell wall biosynthesis
COG1316	Accessory	Cell wall/membrane/envelope biogenesis	Anionic cell wall polymer biosynthesis enzyme, LytR-Cps2A-Psr (LCP) family
COG0452	Accessory	Coenzyme transport and metabolism	Phosphopantothienoylcysteine synthetase/decarboxylase
COG0610	Exclusive	Defense mechanisms	Type I site-specific restriction-modification system, R (restriction) subunit and related helicases
COG0732	Exclusive	Defense mechanisms	Restriction endonuclease S subunit

APÊNDICE A. Tabela com a descrição dos *Clusters of Orthologous Genes* (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. I-2. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0286	Exclusive	Defense mechanisms	Type I restriction-modification system, DNA methylase subunit
COG1131	Exclusive	Defense mechanisms	ABC-type multidrug transport system, ATPase component
COG2002	Exclusive	Defense mechanisms	Bifunctional DNA-binding transcriptional regulator of stationary/sporulation/toxin gene expression and antitoxin component of the YhaV-PrlF toxin-antitoxin module
COG1668	Accessory	Energy production and conversion	ABC-type Na ⁺ efflux pump, permease component
COG1479	Exclusive	Function unknown	Uncharacterized conserved protein, contains ParB-like and HNH nuclease domains
COG0586	Exclusive	Function unknown	Uncharacterized membrane protein DedA, SNARE-associated domain
COG2323	Exclusive	Function unknown	Uncharacterized membrane protein YcaP, DUF421 family
COG3237	Accessory	Function unknown	Uncharacterized conserved protein YjbJ, UPF0337 family
COG3976	Exclusive	General function prediction only	Uncharacterized protein, contains FMN-binding domain
COG0477	Accessory	General function prediction only	MFS family permease
COG1028	Accessory	General function prediction only	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG1062	Accessory	General function prediction only	Zn-dependent alcohol dehydrogenase
COG0454	Accessory	General function prediction only	N-acetyltransferase, GNAT superfamily (includes histone acetyltransferase HPA2)
COG1028	Accessory	General function prediction only	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG0110	Accessory	General function prediction only	Acetyltransferase (isoleucine patch superfamily)
COG2801	Exclusive	Mobilome: prophages, transposons	Transposase InsO and inactivated derivatives
COG3293	Exclusive	Mobilome: prophages, transposons	Transposase
COG1943	Accessory	Mobilome: prophages, transposons	REP element-mobilizing transposase RayT
COG0582	Accessory	Mobilome: prophages, transposons	Integrase
COG2826	Accessory	Mobilome: prophages, transposons	Transposase and inactivated derivatives, IS30 family
COG2963	Accessory	Mobilome: prophages, transposons	Transposase and inactivated derivatives
COG4626	Accessory	Mobilome: prophages, transposons	Phage terminase-like protein, large subunit, contains N-terminal HTH domain
COG4653	Accessory	Mobilome: prophages, transposons	Predicted phage phi-C31 gp36 major capsid-like protein
COG5614	Accessory	Mobilome: prophages, transposons	Bacteriophage head-tail adaptor

APÊNDICE A. Tabela com a descrição dos *Clusters of Orthologous Genes* (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. I-2. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG4824	Accessory	Mobilome: prophages, transposons	Phage-related holin (Lysis protein)
COG0270	Exclusive	Replication, recombination and repair	Site-specific DNA-cytosine methylase
COG0270	Exclusive	Replication, recombination and repair	Site-specific DNA-cytosine methylase
COG0270	Exclusive	Replication, recombination and repair	Site-specific DNA-cytosine methylase
COG1961	Accessory	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG1961	Accessory	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG0305	Accessory	Replication, recombination and repair	Replicative DNA helicase
COG0582	Accessory	Replication, recombination and repair	Integrase
COG0513	Accessory	Replication, recombination and repair	Superfamily II DNA and RNA helicase
COG3857	Accessory	Replication, recombination and repair	ATP-dependent helicase/DNAse subunit B
COG0553	Accessory	Replication, recombination and repair	Superfamily II DNA or RNA helicase, SNF2 family
COG1570	Accessory	Replication, recombination and repair	Exonuclease VII, large subunit
COG1028	Accessory	Secondary metabolites biosynthesis, transport and catabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG0656	Accessory	Secondary metabolites biosynthesis, transport and catabolism	Aldo/keto reductase, related to diketogulonate reductase
COG1028	Accessory	Secondary metabolites biosynthesis, transport and catabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG0642	Exclusive	Signal transduction mechanisms	Signal transduction histidine kinase
COG3279	Exclusive	Signal transduction mechanisms	DNA-binding response regulator, LytR/AlgR family
COG2501	Accessory	Translation, ribosomal structure and biogenesis	Ribosome-associated protein YbcJ, S4-like RNA binding protein
COG1214	Accessory	Translation, ribosomal structure and biogenesis	tRNA A37 threonylcarbamoyladenosine modification protein TsaB

APÊNDICE B. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. L-10. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs.

COG ID	Category	Group	Description
COG3774	Exclusive	Cell wall/membrane/envelope biogenesis	Mannosyltransferase OCH1 or related enzyme
COG0438	Exclusive	Cell wall/membrane/envelope biogenesis	Glycosyltransferase involved in cell wall biosynthesis
COG2327	Exclusive	Cell wall/membrane/envelope biogenesis	Polysaccharide pyruvyl transferase family protein WcaK
COG2244	Exclusive	Cell wall/membrane/envelope biogenesis	Membrane protein involved in the export of O-antigen and teichoic acid
COG0028	Accessory	Coenzyme transport and metabolism	Acetolactate synthase large subunit or other thiamine pyrophosphate-requiring enzyme
COG0732	Exclusive	Defense mechanisms	Restriction endonuclease S subunit
COG0732	Exclusive	Defense mechanisms	Restriction endonuclease S subunit
COG1715	Exclusive	Defense mechanisms	Restriction endonuclease Mrr
COG0286	Exclusive	Defense mechanisms	Type I restriction-modification system, DNA methylase subunit
COG3587	Exclusive	Defense mechanisms	Restriction endonuclease
COG0732	Exclusive	Defense mechanisms	Restriction endonuclease S subunit
COG1518	Accessory	Defense mechanisms	CRISPR/Cas system-associated endonuclease Cas1
COG2184	Exclusive	Signal transduction mechanisms	Fido, protein-threonine AMPylation domain
COG2071	Exclusive	Amino acid transport and metabolism	Gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD (putrescine degradation), contains GATase1-like domain
COG1063	Exclusive	Amino acid transport and metabolism	Threonine dehydrogenase or related Zn-dependent dehydrogenase
COG0367	Exclusive	Amino acid transport and metabolism	Asparagine synthetase B (glutamine-hydrolyzing)
COG0028	Accessory	Amino acid transport and metabolism	Acetolactate synthase large subunit or other thiamine pyrophosphate-requiring enzyme
COG0395	Exclusive	Carbohydrate transport and metabolism	ABC-type glycerol-3-phosphate transport system, permease component
COG1175	Exclusive	Carbohydrate transport and metabolism	ABC-type sugar transport system, permease component
COG1653	Exclusive	Carbohydrate transport and metabolism	ABC-type glycerol-3-phosphate transport system, periplasmic component
COG0366	Exclusive	Carbohydrate transport and metabolism	Glycosidase
COG1501	Exclusive	Carbohydrate transport and metabolism	Alpha-glucosidase, glycosyl hydrolase family GH31
COG5017	Exclusive	Carbohydrate transport and metabolism	UDP-N-acetylglucosamine transferase subunit ALG13
COG1455	Exclusive	Carbohydrate transport and metabolism	Phosphotransferase system cellobiose-specific component IIC

APÊNDICE B. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. L-10. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG2723	Exclusive	Carbohydrate transport and metabolism	Beta-glucosidase/6-phospho-beta-glucosidase/beta-galactosidase
COG1349	Exclusive	Carbohydrate transport and metabolism	DNA-binding transcriptional regulator of sugar metabolism, DeoR/GlpR family
COG3775	Exclusive	Carbohydrate transport and metabolism	Phosphotransferase system, galactitol-specific IIC component
COG0524	Exclusive	Carbohydrate transport and metabolism	Sugar or nucleoside kinase, ribokinase family
COG0036	Exclusive	Carbohydrate transport and metabolism	Pentose-5-phosphate-3-epimerase
COG3345	Accessory	Carbohydrate transport and metabolism	Alpha-galactosidase
COG1175	Accessory	Carbohydrate transport and metabolism	ABC-type sugar transport system, permease component
COG1762	Accessory	Carbohydrate transport and metabolism	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)
COG3414	Accessory	Carbohydrate transport and metabolism	Phosphotransferase system, galactitol-specific IIB component
COG0383	Accessory	Carbohydrate transport and metabolism	Alpha-mannosidase
COG1929	Accessory	Carbohydrate transport and metabolism	Glycerate kinase
COG5039	Accessory	Carbohydrate transport and metabolism	Exopolysaccharide biosynthesis protein EpsI, predicted pyruvyl transferase
COG0489	Exclusive	Cell cycle control, cell division, chromosome partitioning	Chromosome partitioning ATPase, Mrp family, contains Fe-S cluster
COG1192	Exclusive	Cell motility	Cellulose biosynthesis protein BcsQ
COG1192	Accessory	Cell motility	Cellulose biosynthesis protein BcsQ
COG1732	Accessory	Cell wall/membrane/envelope biogenesis	Periplasmic glycine betaine/choline-binding (lipo)protein of an ABC-type transport system (osmoprotectant binding protein)
COG0562	Accessory	Cell wall/membrane/envelope biogenesis	UDP-galactopyranose mutase
COG5039	Accessory	Cell wall/membrane/envelope biogenesis	Exopolysaccharide biosynthesis protein EpsI, predicted pyruvyl transferase
COG0561	Accessory	Coenzyme transport and metabolism	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases
COG1131	Accessory	Defense mechanisms	ABC-type multidrug transport system, ATPase component
COG0286	Accessory	Defense mechanisms	Type I restriction-modification system, DNA methylase subunit

APÊNDICE B. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. L-10. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG2337	Accessory	Defense mechanisms	mRNA-degrading endonuclease, toxin component of the MazEF toxin-antitoxin module
COG3077	Accessory	Defense mechanisms	Antitoxin component of the RelBE or YafQ-DinJ toxin-antitoxin module
COG2189	Exclusive	DNA replication, recombination, and repair	Adenine specific DNA methylase Mod
COG3421	Exclusive	Function unknown	Uncharacterized protein
COG3595	Accessory	Function unknown	Uncharacterized conserved protein YvlB, contains DUF4097 and DUF4098 domains
COG4876	Accessory	Function unknown	Uncharacterized protein YdaT
COG4716	Accessory	Function unknown	Myosin-crossreactive antigen (function unknown)
COG1811	Accessory	Function unknown	Uncharacterized membrane protein YqgA, affects biofilm formation
COG1063	Exclusive	General function prediction only	Threonine dehydrogenase or related Zn-dependent dehydrogenase
COG4185	Accessory	General function prediction only	Predicted ABC-type ATPase
COG0693	Accessory	General function prediction only	Putative intracellular protease/amidase
COG0561	Accessory	General function prediction only	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases
COG0609	Accessory	Inorganic ion transport and metabolism	ABC-type Fe ³⁺ -siderophore transport system, permease component
COG1116	Accessory	Inorganic ion transport and metabolism	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component
COG0715	Accessory	Inorganic ion transport and metabolism	ABC-type nitrate/sulfonate/bicarbonate transport system, periplasmic component
COG3559	Exclusive	Intracellular trafficking, secretion, and vesicular transport	Putative exporter of polyketide antibiotics
COG3600	Exclusive	Mobilome: prophages, transposons	Uncharacterized phage-associated protein
COG3316	Exclusive	Mobilome: prophages, transposons	Transposase (or an inactivated derivative)
COG5527	Exclusive	Mobilome: prophages, transposons	Protein involved in initiation of plasmid replication
COG3547	Exclusive	Mobilome: prophages, transposons	Transposase
COG5283	Exclusive	Mobilome: prophages, transposons	Phage-related tail protein

APÊNDICE B. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. L-10. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG2801	Accessory	Mobilome: prophages, transposons	Transposase InsO and inactivated derivatives
COG2826	Accessory	Mobilome: prophages, transposons	Transposase and inactivated derivatives, IS30 family
COG3747	Accessory	Mobilome: prophages, transposons	Phage terminase, small subunit
COG0582	Accessory	Mobilome: prophages, transposons	Integrase
COG1943	Accessory	Mobilome: prophages, transposons	REP element-mobilizing transposase RayT
COG3464	Accessory	Mobilome: prophages, transposons	Transposase
COG1961	Exclusive	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG0419	Exclusive	Replication, recombination and repair	DNA repair exonuclease SbcCD ATPase subunit
COG0433	Exclusive	Replication, recombination and repair	Archaeal DNA helicase HerA or a related bacterial ATPase, contains HAS-barrel and ATPase domains
COG2189	Exclusive	Replication, recombination and repair	Adenine specific DNA methylase Mod
COG0513	Exclusive	Replication, recombination and repair	Superfamily II DNA and RNA helicase
COG0513	Exclusive	Replication, recombination and repair	Superfamily II DNA and RNA helicase
COG0513	Accessory	Replication, recombination and repair	Superfamily II DNA and RNA helicase
COG0389	Accessory	Replication, recombination and repair	Nucleotidyltransferase/DNA polymerase involved in DNA repair
COG1961	Accessory	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG0507	Accessory	Replication, recombination and repair	ATP-dependent exoDNAse (exonuclease V), alpha subunit, helicase superfamily I
COG0582	Accessory	Replication, recombination and repair	Integrase
COG1762	Accessory	Signal transduction mechanisms	Phosphotransferase system mannitol/fructose-specific IIA domain (Ntr-type)
COG0642	Accessory	Signal transduction mechanisms	Signal transduction histidine kinase
COG1309	Exclusive	Transcription	DNA-binding transcriptional regulator, AcrR family
COG1396	Exclusive	Transcription	Transcriptional regulator, contains XRE-family HTH domain
COG0640	Exclusive	Transcription	DNA-binding transcriptional regulator, ArsR family
COG1349	Exclusive	Transcription	DNA-binding transcriptional regulator of sugar metabolism, DeoR/GlpR family
COG1396	Exclusive	Transcription	Transcriptional regulator, contains XRE-family HTH domain
COG1396	Exclusive	Transcription	Transcriptional regulator, contains XRE-family HTH domain

APÊNDICE B. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Lactobacillus* sp. L-10. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG1609	Exclusive	Transcription	DNA-binding transcriptional regulator, LacI/PurR family
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG1737	Accessory	Transcription	DNA-binding transcriptional regulator, MurR/RpiR family, contains HTH and SIS domains
COG1476	Accessory	Transcription	DNA-binding transcriptional regulator, XRE-family HTH domain
COG1309	Accessory	Transcription	DNA-binding transcriptional regulator, AcrR family
COG1813	Exclusive	Translation, ribosomal structure and biogenesis	Archaeal ribosome-binding protein aMBF1, putative translation factor, contains Zn-ribbon and HTH domains

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs.

COG ID	Category	Group	Description
COG0477	Exclusive	Amino acid transport and metabolism	MFS family permease
COG0477	Accessory	Amino acid transport and metabolism	MFS family permease
COG1113	Accessory	Amino acid transport and metabolism	L-asparagine transporter and related permeases
COG0477	Accessory	Amino acid transport and metabolism	MFS family permease
COG0620	Accessory	Amino acid transport and metabolism	Methionine synthase II (cobalamin-independent)
COG0493	Accessory	Amino acid transport and metabolism	NADPH-dependent glutamate synthase beta chain or related oxidoreductase
COG3616	Accessory	Amino acid transport and metabolism	D-serine deaminase, pyridoxal phosphate-dependent
COG0477	Accessory	Amino acid transport and metabolism	MFS family permease
COG0477	Accessory	Amino acid transport and metabolism	MFS family permease
COG0477	Accessory	Amino acid transport and metabolism	MFS family permease
COG1984	Accessory	Amino acid transport and metabolism	Allophanate hydrolase subunit 2
COG0477	Exclusive	Carbohydrate transport and metabolism	MFS family permease
COG2814	Exclusive	Carbohydrate transport and metabolism	Predicted arabinose efflux permease, MFS family
COG0477	Accessory	Carbohydrate transport and metabolism	MFS family permease
COG3386	Accessory	Carbohydrate transport and metabolism	Sugar lactone lactonase YvrE
COG1501	Accessory	Carbohydrate transport and metabolism	Alpha-glucosidase, glycosyl hydrolase family GH31
COG0477	Accessory	Carbohydrate transport and metabolism	MFS family permease
COG0477	Accessory	Carbohydrate transport and metabolism	MFS family permease

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0477	Accessory	Carbohydrate transport and metabolism	MFS family permease
COG0477	Accessory	Carbohydrate transport and metabolism	MFS family permease
COG0463	Exclusive	Cell wall/membrane/envelope biogenesis	Glycosyltransferase involved in cell wall bisynthesis
COG4421	Exclusive	Cell wall/membrane/envelope biogenesis	Capsular polysaccharide biosynthesis protein
COG1538	Accessory	Cell wall/membrane/envelope biogenesis	Outer membrane protein TolC
COG3895	Accessory	Cell wall/membrane/envelope biogenesis	Membrane-bound inhibitor of C-type lysozyme
COG0451	Accessory	Cell wall/membrane/envelope biogenesis	Nucleoside-diphosphate-sugar epimerase
COG2843	Accessory	Cell wall/membrane/envelope biogenesis	Poly-gamma-glutamate biosynthesis protein CapA/YwtB (capsule formation), metallophosphatase superfamily
COG1538	Accessory	Cell wall/membrane/envelope biogenesis	Outer membrane protein TolC
COG3659	Accessory	Cell wall/membrane/envelope biogenesis	Carbohydrate-selective porin OprB
COG0596	Exclusive	Coenzyme transport and metabolism	Pimeloyl-ACP methyl ester carboxylesterase
COG2141	Exclusive	Coenzyme transport and metabolism	Flavin-dependent oxidoreductase, luciferase family (includes alkanesulfonate monooxygenase SsuD and methylene tetrahydromethanopterin reductase)
COG2141	Accessory	Coenzyme transport and metabolism	Flavin-dependent oxidoreductase, luciferase family (includes alkanesulfonate monooxygenase SsuD and methylene tetrahydromethanopterin reductase)
COG0610	Exclusive	Defense mechanisms	Type I site-specific restriction-modification system, R (restriction) subunit and related helicases
COG0732	Exclusive	Defense mechanisms	Restriction endonuclease S subunit
COG0286	Exclusive	Defense mechanisms	Type I restriction-modification system, DNA methylase subunit
COG0732	Exclusive	Defense mechanisms	Restriction endonuclease S subunit
COG0251	Exclusive	Defense mechanisms	Enamine deaminase RidA, house cleaning of reactive enamine intermediates, YigF/YER057c/UK114 family
COG3549	Exclusive	Defense mechanisms	Plasmid maintenance system killer protein
COG3093	Exclusive	Defense mechanisms	Plasmid maintenance system antidote protein VapI, contains XRE-type HTH domain
COG1518	Exclusive	Defense mechanisms	CRISPR/Cas system-associated endonuclease Cas1
COG1203	Exclusive	Defense mechanisms	CRISPR/Cas system-associated endonuclease/helicase Cas3
COG1715	Accessory	Defense mechanisms	Restriction endonuclease Mrr

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0286	Accessory	Defense mechanisms	Type I restriction-modification system, DNA methylase subunit
COG2337	Accessory	Defense mechanisms	mRNA-degrading endonuclease, toxin component of the MazEF toxin-antitoxin module
COG0251	Accessory	Defense mechanisms	Enamine deaminase RidA, house cleaning of reactive enamine intermediates, YigF/YER057c/UK114 family
COG0841	Accessory	Defense mechanisms	Multidrug efflux pump subunit AcrB
COG2076	Accessory	Defense mechanisms	Multidrug transporter EmrE and related cation transporters
COG1294	Accessory	Energy production and conversion	Cytochrome bd-type quinol oxidase, subunit 2
COG1902	Accessory	Energy production and conversion	2,4-dienoyl-CoA reductase or related NADH-dependent reductase, Old Yellow Enzyme (OYE) family
COG1012	Accessory	Energy production and conversion	Acyl-CoA reductase or other NAD-dependent aldehyde dehydrogenase
COG0674	Accessory	Energy production and conversion	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidoreductase, alpha subunit
COG4221	Accessory	Energy production and conversion	NADP-dependent 3-hydroxy acid dehydrogenase YdfG
COG4371	Accessory	Function unknown	Uncharacterized membrane protein
COG5519	Accessory	Function unknown	Uncharacterized protein, DUF927 family
COG2340	Accessory	Function unknown	Uncharacterized conserved protein YkwD, contains CAP (CSP/antigen 5/PR1) domain
COG1479	Accessory	Function unknown	Uncharacterized conserved protein, contains ParB-like and HNH nuclease domains
COG1479	Accessory	Function unknown	Uncharacterized conserved protein, contains ParB-like and HNH nuclease domains
COG3619	Accessory	Function unknown	Uncharacterized membrane protein YoaK, UPF0700 family
COG1512	Accessory	Function unknown	Uncharacterized membrane protein YgcG, contains a TPM-fold domain
COG1289	Accessory	Function unknown	Uncharacterized membrane protein YccC
COG3182	Accessory	Function unknown	Uncharacterized iron-regulated membrane protein
COG4292	Accessory	Function unknown	Low temperature requirement protein LtrA (function unknown)
COG2855	Accessory	Function unknown	Uncharacterized membrane protein YadS
COG0477	Exclusive	General function prediction only	MFS family permease

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0488	Accessory	General function prediction only	ATPase components of ABC transporters with duplicated ATPase domains
COG1848	Accessory	General function prediction only	Predicted nucleic acid-binding protein, contains PIN domain
COG4188	Accessory	General function prediction only	Predicted dienelactone hydrolase
COG0667	Accessory	General function prediction only	Predicted oxidoreductase (related to aryl-alcohol dehydrogenase)
COG1028	Accessory	General function prediction only	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG1028	Accessory	General function prediction only	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG4113	Accessory	General function prediction only	Predicted nucleic acid-binding protein, contains PIN domain
COG3545	Accessory	General function prediction only	Predicted esterase of the alpha/beta hydrolase fold
COG0477	Accessory	General function prediction only	MFS family permease
COG2085	Accessory	General function prediction only	Predicted dinucleotide-binding enzyme
COG4122	Accessory	General function prediction only	Predicted O-methyltransferase YrrM
COG0493	Accessory	General function prediction only	NADPH-dependent glutamate synthase beta chain or related oxidoreductase
COG4310	Accessory	General function prediction only	Uncharacterized protein, contains an aminopeptidase-like domain
COG0500	Accessory	General function prediction only	SAM-dependent methyltransferase
COG5476	Accessory	General function prediction only	Microcystin degradation protein MlrC, contains DUF1485 domain
COG1473	Accessory	General function prediction only	Metal-dependent amidase/aminoacylase/carboxypeptidase
COG0477	Accessory	General function prediction only	MFS family permease
COG0523	Accessory	General function prediction only	GTPase, G3E family
COG0477	Accessory	General function prediction only	MFS family permease
COG1473	Accessory	General function prediction only	Metal-dependent amidase/aminoacylase/carboxypeptidase
COG2374	Accessory	General function prediction only	Predicted extracellular nuclease
COG0477	Accessory	General function prediction only	MFS family permease
COG2141	Accessory	General function prediction only	Flavin-dependent oxidoreductase, luciferase family (includes alkanesulfonate monooxygenase SsuD and methylene tetrahydromethanopterin reductase)
COG0667	Accessory	General function prediction only	Predicted oxidoreductase (related to aryl-alcohol dehydrogenase)
COG0477	Exclusive	Inorganic ion transport and metabolism	MFS family permease

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0715	Exclusive	Inorganic ion transport and metabolism	ABC-type nitrate/sulfonate/bicarbonate transport system, periplasmic component
COG1629	Exclusive	Inorganic ion transport and metabolism	Outer membrane receptor proteins, mostly Fe transport
COG0605	Exclusive	Inorganic ion transport and metabolism	Superoxide dismutase
COG0477	Accessory	Inorganic ion transport and metabolism	MFS family permease
COG0025	Accessory	Inorganic ion transport and metabolism	NhaP-type Na ⁺ /H ⁺ or K ⁺ /H ⁺ antiporter
COG0477	Accessory	Inorganic ion transport and metabolism	MFS family permease
COG0477	Accessory	Inorganic ion transport and metabolism	MFS family permease
COG1629	Accessory	Inorganic ion transport and metabolism	Outer membrane receptor proteins, mostly Fe transport
COG0477	Accessory	Inorganic ion transport and metabolism	MFS family permease
COG2375	Accessory	Inorganic ion transport and metabolism	NADPH-dependent ferric siderophore reductase, contains FAD-binding and SIP domains
COG1629	Accessory	Inorganic ion transport and metabolism	Outer membrane receptor proteins, mostly Fe transport
COG0477	Accessory	Inorganic ion transport and metabolism	MFS family permease
COG3712	Accessory	Inorganic ion transport and metabolism	periplasmic ferric-dicitrate binding protein FecR, regulates iron transport through sigma-19
COG1116	Accessory	Inorganic ion transport and metabolism	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component
COG1629	Accessory	Inorganic ion transport and metabolism	Outer membrane receptor proteins, mostly Fe transport
COG0609	Accessory	Inorganic ion transport and metabolism	ABC-type Fe ³⁺ -siderophore transport system, permease component
COG3843	Accessory	Intracellular trafficking, secretion, and vesicular transport	Type IV secretory pathway, VirD2 components (relaxase)
COG2303	Exclusive	Lipid transport and metabolism	Choline dehydrogenase or related flavoprotein
COG1028	Accessory	Lipid transport and metabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG1960	Accessory	Lipid transport and metabolism	Acyl-CoA dehydrogenase related to the alkylation response protein AidB
COG0657	Accessory	Lipid transport and metabolism	Acetyl esterase/lipase
COG1028	Accessory	Lipid transport and metabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG1028	Accessory	Lipid transport and metabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG2084	Accessory	Lipid transport and metabolism	3-hydroxyisobutyrate dehydrogenase or related beta-hydroxyacid dehydrogenase
COG1960	Accessory	Lipid transport and metabolism	Acyl-CoA dehydrogenase related to the alkylation response protein AidB
COG0446	Accessory	Lipid transport and metabolism	NADPH-dependent 2,4-dienoyl-CoA reductase, sulfur reductase, or a related oxidoreductase
COG3415	Exclusive	Mobilome: prophages, transposons	Transposase
COG2801	Exclusive	Mobilome: prophages, transposons	Transposase InsO and inactivated derivatives
COG3668	Accessory	Mobilome: prophages, transposons	Plasmid stabilization system protein ParE
COG3547	Accessory	Mobilome: prophages, transposons	Transposase [Mobilome: prophages, transposons]
COG3335	Accessory	Mobilome: prophages, transposons	Transposase
COG0582	Accessory	Mobilome: prophages, transposons	Integrase
COG3039	Accessory	Mobilome: prophages, transposons	Transposase and inactivated derivatives, IS5 family
COG3293	Accessory	Mobilome: prophages, transposons	Transposase
COG0582	Accessory	Mobilome: prophages, transposons	Integrase
COG3293	Accessory	Mobilome: prophages, transposons	Transposase
COG3415	Accessory	Mobilome: prophages, transposons	Transposase
COG0044	Accessory	Nucleotide transport and metabolism	Dihydroorotase or related cyclic amidohydrolase
COG0625	Accessory	Posttranslational modification, protein turnover, chaperones	Glutathione S-transferase
COG0419	Accessory	Posttranslational modification, protein turnover, chaperones	Periplasmic serine protease, ClpP class
COG0616	Accessory	Posttranslational modification, protein turnover, chaperones	Periplasmic serine protease, ClpP class
COG1765	Accessory	Posttranslational modification, protein turnover, chaperones	Predicted redox protein, regulator of disulfide bond formation
COG1961	Exclusive	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG1961	Accessory	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG4973	Accessory	Replication, recombination and repair	Site-specific recombinase XerC
COG0582	Accessory	Replication, recombination and repair	Integrase
COG1961	Accessory	Replication, recombination and repair	Site-specific DNA recombinase related to the DNA invertase Pin
COG0553	Accessory	Replication, recombination and repair	Superfamily II DNA or RNA helicase, SNF2 family
COG1743	Accessory	Replication, recombination and repair	Adenine-specific DNA methylase, contains a Zn-ribbon domain

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0507	Accessory	Replication, recombination and repair	ATP-dependent exoDNAse (exonuclease V), alpha subunit, helicase superfamily I
COG0582	Accessory	Replication, recombination and repair	Integrase
COG1028	Accessory	Secondary metabolites biosynthesis, transport and catabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG1028	Accessory	Secondary metabolites biosynthesis, transport and catabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG1028	Accessory	Secondary metabolites biosynthesis, transport and catabolism	NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family
COG0500	Accessory	Secondary metabolites biosynthesis, transport and catabolism	SAM-dependent methyltransferase
COG5635	Exclusive	Signal transduction mechanisms	Predicted NTPase, NACHT family domain
COG3707	Accessory	Signal transduction mechanisms	Two-component response regulator, AmiR/NasT family, consists of REC and RNA-binding antiterminator (ANTAR) domains
COG3712	Accessory	Signal transduction mechanisms	periplasmic ferric-dicitrate binding protein FecR, regulates iron transport through sigma-19
COG3905	Exclusive	Transcription	Predicted transcriptional regulator
COG3177	Exclusive	Transcription	Fic family protein
COG0789	Exclusive	Transcription	DNA-binding transcriptional regulator, MerR family
COG1595	Exclusive	Transcription	DNA-directed RNA polymerase specialized sigma subunit, sigma24 family
COG2378	Exclusive	Transcription	Predicted DNA-binding transcriptional regulator YafY, contains an HTH and WYL domains
COG0583	Exclusive	Transcription	DNA-binding transcriptional regulator, LysR family
COG0583	Exclusive	Transcription	DNA-binding transcriptional regulator, LysR family
COG0583	Exclusive	Transcription	DNA-binding transcriptional regulator, LysR family
COG0640	Exclusive	Transcription	DNA-binding transcriptional regulator, ArsR family
COG1309	Exclusive	Transcription	DNA-binding transcriptional regulator, AcrR family
COG1522	Accessory	Transcription	DNA-binding transcriptional regulator, Lrp family
COG1309	Accessory	Transcription	DNA-binding transcriptional regulator, AcrR family
COG1733	Accessory	Transcription	DNA-binding transcriptional regulator, HxlR family
COG2207	Accessory	Transcription	AraC-type DNA-binding domain and AraC-containing proteins

APÊNDICE C. Tabela com a descrição dos “*Clusters of Orthologous Genes*” (COGs) apontados para PPPs acessórias e exclusivas observadas na bactéria *Acetobacter* sp. A-3. Dados gerados pelo programa BPGA utilizando 50% de limite de identidade mínima nos alinhamentos das PPPs. (Continuação).

COG ID	Category	Group	Description
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG1396	Accessory	Transcription	Transcriptional regulator, contains XRE-family HTH domain
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG3177	Accessory	Transcription	Fic family protein
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG1733	Accessory	Transcription	DNA-binding transcriptional regulator, HxlR family
COG0553	Accessory	Transcription	Superfamily II DNA or RNA helicase, SNF2 family
COG3707	Accessory	Transcription	Two-component response regulator, AmiR/NasT family, consists of REC and RNA-binding antiterminator (ANTAR) domains
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG2207	Accessory	Transcription	AraC-type DNA-binding domain and AraC-containing proteins
COG1737	Accessory	Transcription	DNA-binding transcriptional regulator, MurR/RpiR family, contains HTH and SIS domains
COG1595	Accessory	Transcription	DNA-directed RNA polymerase specialized sigma subunit, sigma24 family
COG3712	Accessory	Transcription	periplasmic ferric-dicitrate binding protein FecR, regulates iron transport through sigma-19
COG1595	Accessory	Transcription	DNA-directed RNA polymerase specialized sigma subunit, sigma24 family
COG0583	Accessory	Transcription	DNA-binding transcriptional regulator, LysR family
COG0162	Exclusive	Translation, ribosomal structure and biogenesis	Tyrosyl-tRNA synthetase

APÊNDICE D. Tabela com a descrição dos metabólitos identificados por cromatografia gasosa acoplada a espectrometria de massas (GC-MS). Metabólitos não identificados pelo banco de dados GMD foram classificados como “*Unknown*”. Foram descritos, para cada metabólito, sempre que possível, o código de identificação no banco de dados GMD, fórmula molecular e “*KEGG Compound*”.

Metabolite	GMD ID	Molecular formula	KEGG Compound
Unknown	A114009	-	-
Unknown	A114008	-	-
Unknown	A124007	-	-
Glycerol	A129003	C3H8O3	C00116
Phosphoric acid	A129001	H3PO4	C00009
Unknown	A128013	-	-
Glycine	A133001	C2H5NO2	C00037
Uracil	A136001	C4H4N2O2	C00106
Maleic acid, 2-methyl	A137003	C5H6O4	C02226
Fumaric acid	A137001	C4H4O4	C00122
Alanine	A138002	C3H7NO2	C00041
Unknown (similar to Aspartic acid)	A144007	-	-
Glycinamide	A143025	C2H6N2O	-
Malic acid	A149001	C4H6O5	C00149
Propane-1,2-diol, 3-amino	A139011	C3H9NO2	C06057
Aspartic acid, 3-hydroxy-, threo-	A164016	C4H7NO5	C03961
Morpholine, N-(3-aminopropyl)-	A165009	C7H16N2O	-
Unknown	A168015	-	-
Unknown	A169012	-	-
3-Hydroxytridecan-4-one	A169014	C13H26O2	-
Ribonic acid	A177001	C5H10O6	C01685
Arabinonic acid	A179001	-	-
Homogentisic acid	A182016	C8H8O4	C00544
Lysine	A186002	C6H14N2O2	C00047
Homoserine lactone, N-2-oxocaproyl-	A186014	C4H7NO2	-
Allose	A187008	C6H12O6	C01487
Unknown	A186010	-	-
3-Hydroxytetradecanoic acid	A187020	C14H28O3	C13790
Idose	A188011	C6H12O6	C06466
Talose	A188010	C6H12O6	C06467
Galactonic acid-1,4-lactone	A189003	C6H10O6	C03383
Glucopyranose, D-	A189005	C6H12O6	C00031
Gluconic acid-1,4-lactone	A189013	C6H10O6	C00198
Allose	A190004	C6H12O6	C01487
Altrose	A190005	C6H12O6	C06464
Gulose	A190003	C6H12O6	C15923
Unknown	A192018	-	-
Mannitol	A193002	C6H14O6	C00392
Unknown	A192001	-	-
Sorbitol	A193001	C6H14O6	C00794
Sedoheptulose, 2,7-anhydro-, beta-	A212013	C7H12O6	-
Galactitol	A194001	C6H14O6	C01697
Unknown	A194007	-	-
Unknown	A197007	-	-
Gluconic acid	A200001	C6H12O7	C00257
Hexadecenoic acid, n-	A202001	C16H32O2	C00249

APÊNDICE D. Tabela com a descrição dos metabólitos identificados por cromatografia gasosa acoplada a espectrometria de massas (GC-MS). Metabólitos não identificados pelo banco de dados GMD foram classificados como “*Unknown*”. Foram descritos, para cada metabólito, sempre que possível, o código de identificação no banco de dados GMD, fórmula molecular e “*KEGG Compound*”. (Continuação).

Metabolite	GMD ID	Molecular formula	KEGG Compound
Hexadecanoic acid	A205001	C16H32O2	C00249
Ribose-5-phosphate	A211004	C5H11O8P	C00117
Ferulic acid, trans-	A210001	C10H10O4	C01494
Glucoheptonic acid-1,4-lactone	A225012	C7H12O7	-
Unknown	A216005	-	-
Tryptamine, 1-methyl-	A216012	C11H14N2	C06213
Unknown	A218006	-	-
Unknown	A220003	-	-
Octadecanoic acid	A225002	C18H36O2	C01530
Unknown	A225016	-	-
Muramic acid, N-acetyl-	A226012	C11H19NO8	C02713
Unknown	A228016	-	-
Tricosane, n-	A230004	C23H48	C17433
Galactose-6-phosphate	A232001	C6H13O9P	C01113
Unknown	A234005	-	-
Unknown	A235009	-	-
Unknown	A237001	-	-
Glucose-6-phosphate	A239001	C6H13O9P	C00668
Tetracosane, n-	A240003	C24H50	-
Unknown	A242009	-	-
Inositol-2-phosphate, myo-	A249003	C11H18O16P2R2	C04549
Pentacosane, n-	A250004	C25H52	-
Aziridine, N-dansyl-	A250011	C14H16N2O2S	-
Unknown	A257001	-	-
Unknown	A256008	-	-
Unknown	A257017	-	-
Hexacosane, n-	A260001	C26H54	-
Lactulose	A261012	C12H22O11	C07064
Unknown	A262001	-	-
Lactose, alpha-	A269001	C12H22O11	C00243
beta-D-Fructofuranosyl-(2,1)-beta-D-Fructofuranose	A267008	-	-
Unknown	A272004	-	-
Trehalose, alpha,beta-	A274008	C12H22O11	C01083
Trehalose, beta,beta'-	A275017	C12H22O11	C01083
Unknown	A277007	-	-
Maltitol	A284001	C12H24O11	-
Unknown	A282009	-	-
Unknown	A284004	-	-
Unknown	A283016	-	-
Adenosine, alpha-	A284013	C10H13N5O4	C00212
Unknown	A286002	-	-
Isomaltose	A287001	C12H22O11	C00252
Unknown	A286005	-	-
Unknown	A287005	-	-
Unknown	A289002	-	-
Unknown	A300006	-	-

APÊNDICE D. Tabela com a descrição dos metabólitos identificados por cromatografia gasosa acoplada a espectrometria de massas (GC-MS). Metabólitos não identificados pelo banco de dados GMD foram classificados como “*Unknown*”. Foram descritos, para cada metabólito, sempre que possível, o código de identificação no banco de dados GMD, fórmula molecular e “*KEGG Compound*”. (Continuação).

Metabolite	GMD ID	Molecular formula	KEGG Compound
Hexacosanoic acid	A304003	C ₂₆ H ₅₂ O ₂	C21931
Unknown	A308005	-	-
Hentriacontane, n-	A310004	C ₃₁ H ₆₄	C08376
Dotriacontane, n-	A320001	C ₃₂ H ₆₆	-
Unknown	A323004	-	-
Tetratriacontane	A340005	C ₃₄ H ₇₀	-
Unknown	A348003	-	-
Maltotriose	A355003	C ₁₈ H ₃₂ O ₁₆	C01835
Hexatriacontane, n-	A360001	C ₃₆ H ₇₄	-