UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
Programa de pós-graduação em Bioquímica

João Renato Souza Negrão e Silva

**Construção de vetores para superexpressão
da proteína L1 do HPV16 em *Pichia pastoris***

São Paulo
2010

João Renato Souza Negrão e Silva

**Construção de vetores para superexpressão da
proteína L1 do HPV16 em *Pichia pastoris***

*Dissertação apresentada ao Instituto
de Química da Universidade de São
Paulo para a obtenção do título de
Mestre em Ciências.*

Área de Concentração: Bioquímica.

Orientador: Prof. Dr. Paulo Lee Ho

São Paulo
2010

João Renato Souza Negrão e Silva

**Construção de vetores para superexpressão da proteína L1
do HPV16 em *Pichia pastoris***

*Dissertação apresentada ao Instituto de
Química da Universidade de São Paulo
para a obtenção do título de Mestre em
Ciências.*

Área de Concentração: Bioquímica

Data da Aprovação: _____

Banca Examinadora

Prof. Dr. Paulo Lee Ho

Instituição: Instituto Butantan - Divisão de Biotecnologia

Julgamento: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Julgamento: _____

Assinatura: _____

AGRADECIMENTOS

Ao Dr. Paulo Lee Ho, que contribuiu significativamente para meu desenvolvimento científico e que sempre foi compreensivo nos momentos difíceis.

Aos integrantes do laboratório de biotecnologia do Instituto Butantan, em especial a Agatha Chaves, pelas amizades e auxílios durante esses anos.

Ao Instituto de Química da Universidade de São Paulo, pelo meu crescimento intelectual e pelos excelentes docentes.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de mestrado.

A minha família, que me apoiou em todos os momentos e que me deu o suporte necessário para realizar o Mestrado.

À Clarissa Willets, minha namorada, que me dá o amor necessário para tudo valer a pena.

RESUMO

Silva, JRSN. **Construção de vetores para superexpressão da proteína L1 do HPV16 em *Pichia pastoris***. [Dissertação de Mestrado. Programa de pós-graduação em Bioquímica]. Instituto de Química, Universidade de São Paulo. São Paulo, 2010.

O papilomavirus humano (HPV) é o agente etiológico da infecção sexualmente transmissível mais comum na população mundial. A prevalência da infecção por HPV é estimada em 660 milhões de pessoas e mais de 50% das mulheres serão acometidas pela infecção ao menos uma vez na vida. O HPV é responsável por virtualmente todos os casos de câncer de colo de útero (99%), além de causar muitos outros tipos de câncer de mucosa. O câncer cervical afeta aproximadamente 1.4 milhões de mulheres e 500 mil novos casos surgem anualmente, resultando em 250 mil mortes por ano. Os maiores índices de incidência são observados na América Latina e na África. Duas vacinas contra o HPV são comercializadas desde 2006 por empresas estrangeiras. Entretanto, o custo mínimo da vacinação é superior a 600 reais. Este preço torna sua incorporação pelo sistema de saúde um processo economicamente inviável. Dessa forma, devem ser buscadas alternativas para a produção de uma vacina eficaz, de qualidade e de baixo custo no Brasil. A proteína L1, principal proteína do capsídeo do HPV, tem a propriedade de se auto agregar em partículas semelhantes às virais (VLPs), que são o principal componente das vacinas atuais e possuem muita semelhança com os vírions nativos, introduzindo inclusive, uma imunização predominantemente tipo-específica. Este projeto se propôs a construir e avaliar vetores de expressão visando à produção em larga escala da proteína L1 do HPV 16, o tipo de HPV de maior risco e incidência na população mundial. Foram construídos cinco plasmídeos para *Pichia pastoris*, uma das principais plataformas industriais de expressão. Eles se diferenciam pelos marcadores de seleção e pelas sequências de manutenção do vetor, mas têm exatamente o mesmo objetivo: gerar sistemas que possuam múltiplas cópias do gene da L1 e que não necessitem da utilização de antibióticos. Essa estratégia foi escolhida porque a produtividade de muitas das proteínas heterólogas tem grande dependência da dosagem gênica e porque o uso de antibióticos encarece muito o custo da produção. Na primeira etapa do projeto foi avaliado um sistema de expressão episomal auxotrófico, mas ele não se manteve estável ao longo de 60 gerações. Na segunda parte, dois vetores de integração ao sítio do DNA ribossomal foram testados. O sistema mais estável e produtivo foi caracterizado quanto ao número de cópias integradas ao genoma da *P. pastoris* e ao nível de expressão. O sistema é capaz de produzir clones com mais de 30 cópias do gene da L1 e dois clones expressaram cerca de 20 a 30 miligramas/litro de L1. Adicionalmente outros dois vetores integrativos que utilizam sequências teloméricas para integração foram construídos. Ainda são necessários estudos conjuntos de fermentação em larga escala e purificação da proteína para verificar a viabilidade do sistema.

Palavras-chave: Proteína L1. HPV. Expressão recombinante. *Pichia pastoris*.

ABSTRACT

Silva, JRSN. **Construction of vectors for overexpression of the HPV16 L1 protein in *Pichia pastoris***. [Dissertation. Graduate program in biochemistry]. Chemistry Institute, University of São Paulo. São Paulo, 2010.

The human papillomavirus (HPV) is the etiologic agent of the sexually transmitted infection most common worldwide. The prevalence of HPV infection is estimated at 660 million people and over 50% of women will be affected by the infection at least once in their lifetime. HPV is responsible for virtually all cervical cancers cases (99%), and causes many other types of mucosal cancer. Cervical cancer affects approximately 1.4 million women, and 500.000 new cases occur annually, resulting in 250.000 deaths per year. The highest incidence rates are seen in Latin America and Africa. Two HPV vaccines are marketed since 2006 by foreign companies. However, the minimum cost of vaccination is higher than 360 dollar. This price makes its incorporation by the Brazilian Health System economically unfeasible. Thus, alternatives must be found to produce an effective vaccine, with low cost and high-quality in Brazil. The L1 protein, the major capsid protein of HPV, has the ability to self aggregate into virus-like particles (VLPs) which are the main component of current vaccines. The VLPs are very similar to the native virions and can introduce an immunization predominantly type-specific. This project aimed to construct and evaluate expression vectors to produce, at large scale, the L1 protein of HPV 16, which is the HPV type at greater risk and incidence in the world. Five plasmids were constructed for *Pichia pastoris*, one of the major industrial expression platforms. They differ in the selection markers and the vector maintenance sequences, but they have exactly the same goal: to create systems with multiple copies of the L1 gene and that don't require the use of antibiotics. This strategy was chosen because the productivity of many heterologous proteins have heavy dependence on gene dosage and because the use of antibiotics is extremely expensive for its production. In the first stage of the project, it was tested an episomal auxotrophic expression system, but it was unstable after a period of 60 generations. In the second part, two vectors for integration in the ribosomal DNA locus were tested. The most stable and productive system was featured on the number of copies integrated into the *P. pastoris* genome and on the expression level. It was proved that the system is able to produce clones with more than 30 copies of the L1 gene and two clones expressed approximately 20-30 milligrams/liter of L1. Additionally, two other integrative vectors for integration in telomeric sequences were constructed for future testing. More studies on large-scale fermentation and protein purification will be necessary to determine the feasibility of the system.

Keywords: L1 protein. HPV. Recombinant expression. *Pichia pastoris*.

LISTA DE ILUSTRAÇÕES

Ilustração 1. Dados estatísticos de incidência e mortalidade do câncer cervical no Brasil e na América do Sul.	13
Ilustração 2. Os dez tipos mais prevalentes de HPV entre mulheres com câncer cervical no Brasil, na América do Sul e no Mundo.	15
Ilustração 3. Estrutura do Genoma do HPV16.	16
Ilustração 4. Estrutura das VLPs.	17
Ilustração 5. Mapa dos vetores pPICHOLI-1 e pPIC3,5K.	34
Ilustração 6. Regiões dos DNAs ribossomais de <i>P. pastoris</i> amplificadas por PCR.	48
Ilustração 7. Fluxograma de construção do vetor episomal pEPI-AL1.	55
Ilustração 8. Mapa do vetor pEPI-AL1.	56
Ilustração 9. Primeira análise de expressão da proteína L1 no sistema pEPI-AL1.	57
Ilustração 10. Ensaio de expressão dos clones A e B sob pressões seletivas diferentes.	59
Ilustração 11. Fluxograma de construção do vetor integrativo pINT-AL1.	61
Ilustração 12. Mapa do vetor integrativo pINT-AL1.	62
Ilustração 13. Ensaio de expressão com os clones de 1ª geração do sistema integrativo pINT-AL1.	63
Ilustração 14. Ensaio de expressão com os clones de 60ª geração do sistema integrativo pINT-AL1.	65
Ilustração 15. Comparação entre expressão da L1 e dosagem gênica.	68

Ilustração 16. Resultados do experimento de amplificação das regiões do rDNA de <i>P. pastoris</i>	69
Ilustração 17. Quantificação da expressão da proteína L1 por Western Blot.	70
Ilustração 18. Fluxograma de construção do vetor integrativo pINT-BL1.	71
Ilustração 19. Mapa do vetor integrativo pINT-BL1.	72
Ilustração 20. Ensaio de expressão com os clones de 1ª geração do sistema integrativo pINT-BL1.	73
Ilustração 21. Ensaio de expressão com os clones de 60ª geração do sistema integrativo pINT-BL1.	75
Ilustração 22. Fluxograma de construção dos vetores pINT-CL1 e pINT-DL1. Primeira etapa.	77
Ilustração 23 Fluxograma de construção dos vetores pINT-CL1 e pINT-DL1. Segunda etapa.	78
Ilustração 24. Mapa dos vetores pINT-CL1 e pINT-DL1.	79
Ilustração 25. Função cooperativa das proteínas E6 e E7 no desenvolvimento do câncer cervical.	94
Ilustração 26. Ciclo de vida do vírus HPV.	96
Ilustração 27. Estrutura da unidade do DNA ribossomal de <i>H. polymorpha</i>	100
Ilustração 28. Mecanismos de integração por recombinação homóloga.	102

LISTA DE TABELAS

Tabela 1. - Sequências dos olinucleotídeos de DNA usados no projeto.	32
Tabela 2. - Protocolo das reações de PCR para a construção dos vetores pINT-C e pINT-D.	40
Tabela 3. - Protocolo das reações de qPCR para a quantificação do número de cópias do gene. L1.	46
Tabela 4. - Protocolo das reações de PCR para a amplificação de regiões do rDNA de P.pastoris.	49
Tabela 5. - Histórico experimental e condições dos ensaios de expressão dos clones da 60ª geração transformados com o vetor pEPI-AL1.	58
Tabela 6. - Resultado do experimento de qPCR para quantificação do número de cópias do gene L1 em clones da 60ª geração transformados com o vetor pINT-AL1.	67

SUMÁRIO

INTRODUÇÃO	12
O Papilomavírus Humano	12
Epidemiologia.....	12
Classificação e estrutura viral	14
A proteína L1	16
Vacinas	17
Sistemas de Expressão	18
Leveduras	18
Proteínas em larga escala	20
Vetores de expressão.....	21
Vetor Epissomal	22
Vetores Integrativos.....	23
Estabilidade Clonal	24
OBJETIVOS.....	26
MATERIAIS.....	27
Meios de Cultura	27
Soluções-estoque	27
Meios de cultura	27
Materiais para DNA	28
Extração e visualização de DNA plasmidial	28
Extração de DNA genômico de <i>P. pastoris</i>	29
Enzimas para DNA.....	29
Materiais para Análise de Proteínas	30
Materiais para SDS-PAGE	30
Materiais para transferência e Western Blot.....	30
Materiais para preparar leveduras competentes	31
Materiais para lise celular de <i>P. pastoris</i>.....	31
Oligonucleotídeos.....	32
Vetores.....	32
Linhagens de microorganismos.....	35

Antibióticos	31
MÉTODOS	36
Manipulação de <i>E. coli</i> DH5 α ^{TM}	36
Desenvolvimento dos novos vetores	36
Extração e purificação de DNA plasmidial	36
Construção e modificação de plasmídeos.....	37
Construção do vetor pEPI-AL1	37
Construção do vetor pINT-AL1.....	38
Construção do vetor pINT-BL1.....	39
Construção dos vetores pINT-CL1 e pINT-DL1	39
Manipulação de <i>P. pastoris</i> KM71.....	41
Protocolo de Eletro-competência	42
Protocolo de Transformação.....	42
Protocolo de expressão da proteína L1	43
Protocolo de Lise Celular.....	43
Protocolo de análise de estabilidade	44
Protocolo de Extração de DNA genômico	44
Análise da expressão da L1	45
Quantificação do número de cópias integradas	46
Localização da integração	48
Ensaio com o sistema episomal pEPI-AL1.....	49
Ensaio com o sistema integrativo pINT-AL1.....	50
Ensaio com o vetor integrativo pINT-BL1.....	52
RESULTADOS.....	54
Sistema Episomal pEPI-AL1.....	54
Construção do vetor pEPI-AL1	54
Influência da dose de Zeocina na expressão da L1	56
Expressão da L1 sob pressões seletivas diferentes	57
Sistema integrativo pINT-AL1.....	60
Construção do vetor pINT-AL1	60
Expressão da L1 a partir de vetores integrativos	62
Expressão da L1 na ausência de pressão seletiva	64
Quantificação do número de cópias	66
Localização do sítio de integração	68
Quantificação da Expressão da L1.....	69
Sistema Integrativo pINT-BL1.....	70

Construção do vetor pINT-BL1	70
Expressão da L1 nos clones de 1ª geração	72
Expressão da L1 nos clones de 60ª geração	73
Vetores pINT-CL1 e pINT-DL1	75
Construção dos vetores pINT-CL1 e pINT-DL1	76
DISCUSSÃO	80
Sistema de expressão episomal.....	80
Vetores integrativos com o rDNA de <i>H. polymorpha</i>	81
Sistema pINT-AL1	82
Sistema pINT-BL1	83
CONCLUSÕES	85
REFERÊNCIAS	86
APÊNDICE A – BIOLOGIA MOLECULAR DA INFECÇÃO DO CÓLO UTERINO POR HPV.	93
APÊNDICE B – RESPOSTA IMUNE À INFECÇÃO POR HPV DO COLO UTERINO.....	98
APÊNDICE C – SEQUÊNCIA DE DNA DO GENE DA L1 COM CÓDONS OTIMIZADOS PARA <i>P. PASTORIS</i>.....	100
APÊNDICE D – SEQUÊNCIAS DE DNA RIBOSSOMAL DE LEVEDURAS	101
APÊNDICE D – MECANISMOS DE INTEGRAÇÃO HOMÓLOGA.....	103

INTRODUÇÃO

O Papilomavírus Humano

Epidemiologia

As informações a seguir foram retiradas de dois documentos da Organização Mundial da Saúde (WHO, 2005, IARC, 2007).

O *Human papillomavirus* (HPV) é o agente etiológico da infecção sexualmente transmissível mais comum da população. A prevalência mundial da infecção por HPV é estimada em 660 milhões de pessoas e mais de 50% das mulheres serão acometidas pela infecção ao menos uma vez na vida. O pico de incidência da infecção ocorre na adolescência e em mulheres jovens de até 25 anos. Apesar da maioria das infecções serem assintomáticas e se resolverem espontaneamente, as que persistirem, quando não tratadas, podem progredir para um estado de câncer dentro de um período de duas a três décadas.

Virtualmente todos os casos de câncer cervical (99%) estão ligados a infecções genitais com o Papilomavirus. Além desse câncer, o HPV está relacionado com inúmeros outros tipos de câncer (Stanley, 2006).

O câncer cervical é a maior causa de morte feminina por câncer nos países em desenvolvimento e a segunda maior causa nos países desenvolvidos. Mundialmente, o câncer cervical afeta aproximadamente 1.4 milhões de mulheres e 500 mil novos casos surgem anualmente, dos quais 80% ocorrem em países em desenvolvimento que não dispõem nem de uma triagem de rotina (Papanicolau) e nem de um tratamento eficaz. A maior incidência do câncer cervical é observada na América Latina e na África. O câncer cervical é responsável por 250 mil mortes ao ano.

A seguir, a **Ilustração 1** mostra dados estatísticos específicos da população brasileira. As informações foram extraídas do Centro de Informações de HPV e Câncer Cervical (WHO/ICO, 2010).

Ilustração 1. Dados estatísticos de incidência e mortalidade do câncer cervical no Brasil e América do Sul.

Um programa de triagem que possibilite uma detecção prévia e um tratamento médico eficaz das lesões pré-cancerosas é fundamental para evitar o surgimento de câncer, mas esses programas são caros e de difícil implementação em lugares pouco desenvolvidos.

Classificação e estrutura viral

O Papilomavirus Humano é um vírus pertencente à família Papillomaviridae, formada por mais de 120 espécies, das quais cerca de cem infectam o ser humano e as demais infectam um amplo espectro de mamíferos e aves. Essa desproporção decorre do fato do ser humano ser o único hospedeiro investigado intensivamente. A classificação das espécies de Papilomavirus é feita através da sequência do gene da proteína L1 (Villers et al., 2004).

Além da classificação filogenética, os HPVs podem ser amplamente agrupados em tipos cutâneos ou mucosotrópicos, dependendo do tropismo tecidual. Os vírus cutâneos são os causadores das verrugas na pele. Os tipos mais comuns são o HPV1 e HPV2, que causam as verrugas benignas. Já o HPV5 e o HPV8 são considerados carcinogênicos e podem causar a doença *epidermodisplasia verruciforme* (Villers et al., 2004, Scheurer et al., 2005).

Os HPVs mucosotrópicos podem ainda ser divididos em HPVs de alto risco e HPVs de baixo risco, de acordo com as chances de causar um câncer cervical. Os tipos de baixo risco mais frequentes são o HPV 6 e o HPV 11, que juntos correspondem a mais de 90% das verrugas genitais benignas (*condyloma acuminata*) (Muñoz et al., 2003, Brown et al., 1999). Em todas as regiões, os HPVs tipo 16 e 18 são os responsáveis majoritários pelos cânceres de mucosa. O HPV 16 sozinho é responsável por mais de 50% dos casos de câncer cervical e junto com o HPV 18 totalizam mais de 70% dos casos mundiais de câncer cervical (Smith et al., 2007, Clifford et al., 2003, Muñoz et al., 2003).

A **Ilustração 2** (WHO/ICO, 2010) mostra a prevalência das principais espécies de HPV em mulher com câncer cervical.

Ilustração 2. Os dez tipos mais prevalentes de HPV entre mulheres com câncer cervical no Brasil, na América do Sul e no Mundo.

Os principais tipos de HPV pertencem ao gênero *Alphapapillomavirus*. Os vírions desse gênero possuem um núcleo e um capsídeo não envelopado, esférico, de simetria icosaédrica e de diâmetro de 40nm-55nm. O capsídeo é formado por 72 unidades funcionais denominadas capsômeros. Os capsômeros são formados por pentâmeros da proteína L1, e seis capsômeros se interligam através de uma proteína L2 central. A superfície possui pequenas projeções que conferem um aspecto rugoso ao capsídeo. As populações de vírions apresentam-se na forma de partículas de tamanho uniforme, mas algumas formas anormais, como tubular ou filamentosa, são resultado de um processo aberrante de maturação (ICTVdB Management, 2006).

Todos os HPVs contêm um genoma de DNA dupla fita circular de aproximadamente 8 kb, que pode ser dividido em três regiões: inicial (EARLY), tardia (LATE) e não codificante de controle (LCR ou NCR). As três regiões são separadas por sítios de poliadenilação. A região inicial ocupa 50% do genoma e codifica seis fases de leitura (E1, E2, E4, E5, E6, E7) que representam seis proteínas expressas pelo vírus. As outras duas ORFs, E3 e E8, aparentemente não produzem proteínas. As proteínas E1 e E2 estão envolvidas na replicação do vírus e na regulação da transcrição da região inicial. A proteína E4 esta relacionada ao colapso dos filamentos de colágeno e auxilia na liberação dos vírions. As proteínas E5, E6 e E7 são oncogenes e estão relacionadas com a imortalização e transformação das células. A região tardia ocupa 40% do genoma e possui as ORFs L1 e L2, que transcrevem, respectivamente, a proteína majoritária do capsídeo L1 e a proteína

secundária do capsídeo L2. E, por último, a região não transcrita ocupa 10% do genoma e possui a origem de replicação do DNA e múltiplos sítios de ligação de fatores de transcrição que modulam as atividades dos promotores do vírus (Zheng, Baker, 2006). O modelo do genoma do HPV esta representado na **Ilustração 3** (Doorbar, 2006).

Ilustração 3. Genoma do HPV16 (7904 bp). As setas representam o promotor de início (p97) e o promotor tardio (p670). As seis ORFs iniciais [E1, E2, E4 e E5 (em verde) e E6 e E7 (em vermelho)] são expressas por ambos os promotores em estágios diferentes da diferenciação epitelial. As ORFs tardias [L1 e L2 (em amarelo)] são expressas pelo promotor p670. Todas as ORFs são codificadas pela mesma fita. A região regulatória está ampliada e permite a visualização dos diferentes sítios de ligação de fatores de transcrição e o elemento TATA do promotor p97.

A proteína L1

A L1 é uma proteína de 55kDa que compõe o capsídeo viral. Este é formado pela união de 72 capsômeros, que por sua vez são formados por pentâmeros da proteína L1. Os capsômeros se ligam à proteína menor L2 na proporção de seis pentâmeros para cada molécula de L2 (Xu et al., 2006). A proteína L1 representa aproximadamente 80% da proteína viral total.

A proteína L1 sozinha ou junto com a proteína L2 possui a capacidade de se auto-organizar em uma estrutura similar ao capsídeo viral, denominada “*virus-like particles*”

(VLPs), quando expressa em sistemas heterólogos (Zhang et al., 1998). A estrutura morfológica (Chen et al., 2000) e as características imunológicas das VLPs são muito similares às do HPV natural (Stanley et al., 2006). A **Ilustração 4** mostra o processo de formação de uma VLP.

Ilustração 4. Estrutura das VLPs. Cinco proteínas L1 se associam em pentâmeros denominados capsômeros e 72 deles se associam para formar a VLP.

Vacinas

As vacinas contra HPV têm como base o uso de VLPs formadas por proteína L1 (Mach et al., 2006, Stanley, 2006; Lowy, Schiller, 2006). Essas partículas não possuem o material genético do vírus e são incapazes de se multiplicar, o que assegura que não sejam infecciosas. Embora a proteína menor do capsídeo (L2) apresente alguns epítomos capazes de induzir a produção de anticorpos neutralizantes, a maioria dos anticorpos neutralizantes induzidos por VLPs formadas por L1 e L2 reconhecem predominantemente epítomos de L1 (Lowy et al., 1994), demonstrando o potencial imunogênico e protetor de VLPs formadas apenas pela proteína principal do capsídeo do HPV. Anticorpos neutralizantes contra o papilomavírus são altamente tipo-específicos (Roden et al., 1996).

Duas vacinas foram aprovadas pela FDA e pela Comissão Européia (EC) e demonstraram uma eficácia excelente (Stanley et al., 2006; Lowy et al., 2006; Harper et al.,

2004; Schmiedeskamp et al., 2006). No entanto, não substituem as triagens, são medidas complementares.

Ambas possuem a proteína L1 do HPV16 e HPV18, que são responsáveis por mais de 70% dos casos de câncer cervical (WHO, 2006, 2008, IARC, 2008). Além disso, uma das vacinas possui também a proteína L1 do HPV6 e do HPV11, que são responsáveis por 90% das verrugas genitais (Stanley et al., 2006; WHO, 2006, 2008).

A eficácia dessas vacinas foi testada em mulheres jovens (16-26 anos de idade) que nunca foram expostas a nenhum dos quatro tipos de HPV da vacina (WHO, 2006, 2008). Essas pesquisas clínicas demonstraram quase 100% de eficácia em prevenir lesões causadas pelos vírus alvos da vacina e 95% de eficácia em prevenir as verrugas genitais causadas pelos vírus alvos da vacina (Lowy et al 2006; Harper et al 2004; Schmiedeskamp et al., 2006).

A vacina comercializada pela Merck no Brasil possui um custo mínimo de aproximadamente R\$ 250 reais a dose. Este preço elevado torna sua incorporação pelo Sistema Único de Saúde um processo economicamente inviável. Dessa forma, devem ser buscadas alternativas para a produção de uma vacina eficaz e de baixo custo no Brasil.

Sistemas de Expressão

Leveduras

O primeiro sistema desenvolvido para a expressão de genes heterólogos em fungos foi baseado na levedura do pão, *Saccharomyces cerevisiae*. Essa plataforma possui o status de GRAS e foi utilizada com sucesso na produção de vários fármacos aprovados pela FDA incluindo a insulina (Melmer, 2005) e a primeira vacina recombinante, o antígeno viral da hepatite B (Harford et al., 1987).

A quantidade massiva de informações sobre esse microorganismo nas áreas da genética, biologia molecular, fermentação, bioquímica, microbiologia e fisiologia a tornou a espécie mais bem caracterizada do sistema eucariótico. Esse conhecimento possibilitou a utilização da engenharia genética para modificar o microorganismo como um todo e alterar

o fluxo metabólico visando a produção de proteínas específicas. Por esses motivos a *S. cerevisiae* se tornou um dos principais sistemas de expressão (Gellissen et al., 1997).

Contudo, o uso desse sistema tem certas limitações e inconvenientes: A *S. cerevisiae* hiperglicosila proteínas recombinantes; as cadeias de carboidratos N-ligadas terminam com uma manose anexada à cadeia por uma ligação $\alpha 1,3$, que é considerada alergênica; na fermentação ocorre perda de biomassa pela formação de etanol; a capacidade de secreção não é alta e se limita a proteínas de até 30 kDa; algumas linhagens recombinantes apresentam instabilidade mitótica; entre outras limitações conseqüentes da variedade limitada de fontes de carbono que a espécie pode utilizar (Smidt, 2004; Gellissen et al., 1997, 2005).

Por essas razões novos sistemas de expressão estão sendo desenvolvidos baseados em leveduras não convencionais. Um número limitado de espécies pertencentes ao gênero *Hansenula*, *Cândida*, *Torulopsis* e *Pichia* é capaz de utilizar metanol como fonte única de carbono. Entre essas leveduras metilotróficas facultativas, as espécies *H. polymorpha* e *P. pastoris* fazem parte dos sistemas industriais em ascensão (Hollenberg et al., 1997).

As leveduras metilotróficas partilham da mesma cadeia metabólica para a utilização do metanol. As reações acontecem inicialmente em organelas especializadas, os peroxissomos, e depois no citoplasma. O metanol entra nos peroxissomos e é oxidado por oxidases específicas em formaldeído e peróxido de hidrogênio, que é decomposto em água e oxigênio molecular por uma catalase. O formaldeído gerado pelas oxidases entra tanto na via catabólica para obtenção de energia, quanto na via anabólica para obtenção de biomassa. Os genes que codificam essas oxidases foram identificados e clonados. A *P. pastoris* possui dois genes, AOX1 (alcool oxidase1) e AOX2 (alcool oxidase2). A enzima AOX2 possui a mesma atividade específica que a AOX1, porém seu nível de expressão é muito menor, devido ao fato do seu promotor ser mais fraco. No caso da *H. polymorpha* foi identificado somente um gene, denominado MOX (metanol oxidase). As enzimas AOX1 e MOX possuem baixa afinidade por oxigênio e esta característica é compensada pela superexpressão dessas enzimas, cuja abundância pode chegar até 30% do conteúdo de proteína celular total. Sendo assim, os fortes promotores da AOX1 e MOX podem ser utilizados para direcionar a expressão de proteínas recombinantes a altos níveis. Os promotores são estritamente regulados pelas fontes de carbono disponíveis no meio de

cultura e tal fenômeno possibilita o controle preciso da expressão de proteínas recombinantes sob o controle destes promotores.

A *Pichia pastoris* foi a levedura escolhida porque possui diversas linhagens e ferramentas moleculares disponíveis comercialmente (Invitrogen, San Diego) e seu uso é muito difundido na comunidade científica. Além disso, a *P. pastoris* é metilotrófica, capaz de utilizar metanol como fonte única de carbono e indutor, e o Instituto Butantan está familiarizado com a produção de vacinas recombinantes em leveduras deste tipo.

Até o começo dos anos 90 poucas proteínas heterólogas haviam sido expressas em *P. pastoris* (Russel et al., 1991), hoje já há mais de 400 proteínas expressas registradas (Cereghino et al., 2000, 2002).

Proteínas em larga escala

A produção industrial demanda tanto alta produtividade das proteínas como dispositivos de segurança e autenticidade dos compostos (Gellisen et al., 2005, Gellisen, 2005). A agência FDA (“Food and Drug Administration”), por exemplo, recomenda o uso de linhagens reconhecidas como seguras para a produção. Esses organismos recebem o status de GRAS (“Generally Recognized As Safe”).

Devido a todos esses fatores, os fungos vêm se tornando os sistemas favoritos para a produção das proteínas em larga escala. Os sistemas de expressão em leveduras combinam a fácil manipulação e crescimento de organismos unicelulares com o processamento de modificações pós-traducionais dos eucariotos. As leveduras não são pirogênicas; têm menor risco de contaminação por patógenos humanos e vírus; possuem genoma de pequeno tamanho - cerca de 200 vezes menor que o de mamíferos, o que simplifica análises moleculares e genéticas; e crescem em meio de cultura não-complexo. Além disso, os custos de triagem e produção são muito menores do que em outros sistemas eucariotos.

Vetores de expressão

Diversos aspectos em relação a cada elemento de um vetor devem ser analisados para a criação de um vetor de expressão que atenda adequadamente as necessidades do projeto. Basicamente, um vetor de expressão em leveduras é formado pelos seguintes elementos: um **cassete de expressão**, (que é constituído por um promotor constitutivo ou indutível, um gene acoplado ou não a seqüências sinalizadoras, com codons selvagens ou otimizados e um fragmento de terminação da transcrição); uma **seqüência de replicação** autônoma ou por integração (homóloga ou não-homóloga); um **marcador de seleção** (auxotrofismo, antibióticos e outros); e, para facilitar o armazenamento e amplificação do vetor, uma **origem de replicação bacteriana** e um **marcador de seleção bacteriana**.

Neste projeto foram construídos cinco novos vetores, um vetor episomal e quatro integrativos.

Todos os vetores possuem um cassete de expressão constituído por três elementos: o promotor indutível do gene AOX1; o gene sintético da proteína L1 do HPV16 com códons otimizados para *P. pastoris* e o fragmento de terminação da transcrição do gene AOX1. O AOX1 é o gene de uma das enzimas que participa do metabolismo do metanol. Esse gene é induzido por metanol e sua abundância pode chegar até 30% do conteúdo de proteína celular total (Higgins, Cregg, 1998). Sendo assim, o seu promotor pode ser utilizado para direcionar a expressão de proteínas recombinantes a altos níveis.

Foi descartado o uso de promotores constitutivos porque, de maneira geral, a expressão constitutiva de proteínas heterólogas causa estresse celular, dificultando o crescimento do microorganismo e também o próprio controle dos níveis de expressão, principalmente em sistemas de expressão intracelular.

Todos os vetores também possuem uma mesma origem de replicação bacteriana e um marcador de seleção para facilitar a armazenagem e propagação dos plasmídeo em *E. coli*.

Os cinco vetores se diferenciam pelos marcadores de seleção e pelas seqüências de replicação em levedura, que serão discutidos com mais detalhes posteriormente.

Após a construção e validação dos vetores foi imprescindível desenvolver uma metodologia de obtenção de clones com alta produtividade. Uma vez que a produtividade das proteínas heterólogas intracelulares tem grande dependência da dosagem gênica (Clare et al., 1991), diversas técnicas vêm sendo empregadas na tentativa de obter recombinantes com múltiplas cópias do gene heterólogo, dentre elas: padronização do processo de transformação (Gietz et al., 2001); seleção com altas doses de antibióticos (Scorer et al., 1994); exposição à doses crescentes de antibiótico (Marx et al., 2009); enfraquecimento dos marcadores de seleção (Sohn et al., 1999), multimerização in vitro (Mansur et al., 2005); e utilização de sequências de múltiplas integrações ao genoma (Klabunde et al., 2002, 2003; Sohn et al 1996). A partir desses conhecimentos foram desenhados ensaios experimentais para a obtenção de um clone com alta produção da proteína L1 do HPV16.

Vetor Epissomal

Os processos de transformação, seleção e expressão apresentam menor complexidade quando são utilizados vetores epissomais em vez de vetores integrativos. Em contrapartida, como os vetores epissomais não estão integrados no genoma, eles são facilmente perdidos na ausência de uma pressão seletiva.

Na maioria dos casos os vetores epissomais possuem um gene de resistência a antibiótico, que funciona como marcador de seleção em um meio contendo o determinado antibiótico. Entretanto, na produção em escala industrial, o uso de antibióticos gera um grande impacto ambiental, tem um custo excessivamente elevado e pode estar em discordância com as recomendações das agências regulatórias.

Por esses motivos, neste projeto foi construído um plasmídeo contendo um marcador auxotrófico com o objetivo de eliminar o uso de antibióticos durante expressão em larga escala.

O gene histidinol dehydrogenase de *Pichia pastoris* (HIS4), essencial para a biossíntese do aminoácido histidina, funciona como marcador de seleção em linhagens de *Pichia pastoris* His⁻. Essas linhagens, defectivas em HIS4, crescem normalmente em meio complexo, mas são incapazes de proliferar em meio de cultura sem histidina.

Nesse modelo, a pressão seletiva exercida pelo meio mínimo é baixa e uma única cópia do gene HIS4 é suficiente para restabelecer as funções normais da célula, sendo que somente 1 a 10% dos transformantes possuem mais que uma cópia do gene (Higgins, Cregg, 1998). Isso significa que é necessário fazer uma triagem de centenas de clones para achar alguns com múltiplas cópias.

Para contornar esse problema, além do marcador de seleção auxotrófico, foram adicionados também dois marcadores de resistência aos antibióticos Zeocina e Geneticina. A idéia desse sistema foi utilizar marcadores de resistência à antibiótico para seleção inicial dos clones com múltiplas cópias, e posteriormente, utilizar a marcação auxotrófica para fazer a manutenção do número de plasmídeos.

A característica epissomal do vetor foi conferida pela seqüência de replicação autônoma PARS1 de *P. pastoris*. Esse elemento faz com que os plasmídeos mantenham-se em aproximadamente 10 cópias por célula (Cregg, 1985).

Vetores Integrativos

Por se integrarem ao genoma do hospedeiro, os vetores integrativos demonstram, em teoria, grande estabilidade na ausência de uma pressão seletiva, até mesmo em transformantes com múltiplas cópias (Cregg, 1985). Essa característica os torna ideais para a produção de clones industriais, porque dispensam o uso de marcadores de seleção no cultivo em larga escala.

Os vetores comerciais atuais se integram por recombinação homóloga entre os segmentos distais do vetor linearizado e uma região específica do genoma do organismo. Na maioria dos casos, essa recombinação ocorre em sítios de cópia única no genoma. Esse método gera transformantes com uma ou apenas poucas cópias do cassete de expressão.

Já os quatro vetores construídos neste projeto utilizam elementos que promovem a recombinação com sequências genômicas de múltiplas cópias, possibilitando a integração de muitas cópias do vetor de expressão.

Um dos elementos é um fragmento da unidade do DNA ribossomal de *Hansenula polymorpha* (Klabunde et al., 2002, 2003). Diversos trabalhos descrevem a integração de

múltiplas cópias de vetores portadores de fragmentos do DNA ribossomal de *H. polymorpha*, inclusive em diferentes leveduras (Liu et al., 2005; Steinborn et al 2005, 2006).

O outro elemento é um fragmento da sequência de replicação autônoma de *H. polymorpha* denominada HARS36 (Sohn et al., 1996, 1999). Esse fragmento, além de possuir uma região de replicação autônoma, também possui várias cópias repetidas em tandem de sequências teloméricas. Essas cópias, auxiliadas pela exposição prolongada do vetor na forma episomal ao genoma, promovem a integração de dezenas de cópias do vetor nas regiões teloméricas da *H. polymorpha*.

O marcador de seleção utilizado nos vetores integrativos pode ser o gene *Sh ble*, que confere resistência à Zeocina, ou o gene da Aminoglicosídeo-3-fosfotransferase (APH), que confere resistência à Geneticina e Kanamicina. Ambos os genes são ativos em leveduras e em *E.coli*.

Adicionalmente o gene APH foi modificado para promover a seleção de clones com um maior número cópias dos plasmídeos. O promotor do gene foi trocado por um fragmento do promotor do gene gliceraldeído-3-fosfato desidrogenase (GADPH61) de *H. polymorpha* (Sohn et al., 1999). Esse fragmento de promotor funciona como um promotor fraco, e faz com que seja necessário um maior número de cópias do gene para conferir o mesmo nível de resistência ao antibiótico.

Até o momento, nenhum trabalho descreveu a utilização de algum desses elementos de *H. polymorpha* em *P. pastoris*.

Estabilidade Clonal

Em processos industriais, a estabilidade clonal é tão importante quanto a obtenção de um clone super produtor. A exatidão e a reprodutibilidade imposta por esse tipo de processo exigem um comportamento constante da expressão, sobretudo em escalas muito grandes de produção, onde as células são cultivadas por longos períodos.

Em condições experimentais controladas, uma das principais causas de diminuição da expressão da proteína recombinante é a perda da estabilidade clonal. Na ausência de

pressão seletiva, as células têm a tendência de perder o DNA exógeno ao longo das gerações, a menos que esse DNA confira uma vantagem adaptativa. Instabilidades diretamente relacionadas ao plasmídeo incluem instabilidade estrutural e instabilidade segregacional. A instabilidade estrutural ocorre por causa de recombinações, inserções ou deleções, enquanto a segregacional ocorre pela partição desigual dos vetores episomais durante a divisão celular. Instabilidades também podem ser geradas por fatores indiretos que diminuem o *fitness* da célula. Essa instabilidade pode ser causada por um fenômeno denominado fardo metabólico, que resulta em uma baixa taxa de crescimento celular e desregulação de algumas vias metabólicas durante a expressão de proteínas recombinantes em grandes quantidades (Zhu et al., 2009).

OBJETIVOS

Desenvolver um sistema de expressão em levedura para a produção em larga escala da proteína L1 do papilomavirus humano tipo 16. O sistema tem como base a levedura metilotrófica amplamente utilizada como plataforma industrial para produção de proteínas heterólogas, *Pichia pastoris*. O sistema necessita de alta produtividade, baixo custo, reprodutibilidade em larga escala além de autenticidade da proteína L1.

O desenvolvimento de um sistema de expressão em larga escala da proteína L1 pretende viabilizar a implementação de uma vacina eficiente de baixo custo contra o HPV no o sistema de saúde.

MATERIAIS

Meios de Cultura

Soluções-estoque

- **Solução-estoque 10X YNB (13,4%)**: Dissolver 34g de base nitrogenada de levedura (YNB) e 100g de sulfato de amônio em 1L de água Milli-Q e filtrar em poro de 0.22µm. Reservar a 4°C.
- **Solução-estoque 500X Biotina (0,02%)**: Dissolver 20mg de biotina em um volume final de 100ml de água Milli-Q e filtrar em poro de 0.22µm. Reservar a 4°C.
- **Solução-estoque 100X Histidina (0,4%)**: Dissolver 400mg de L-histidina em um volume final de 100ml de água Milli-Q e filtrar em poro de 0.22µm. Reservar a 4°C.
- **Solução-estoque 10X Dextrose (20%)**: Dissolver 200g de D-glicose em um volume final de 1L de água Milli-Q e autoclavar. Reservar a 4°C.
- **Solução-estoque 10X Metanol (5%)**: Diluir 5 ml de metanol em um volume final de 100ml de água Milli-Q e filtrar em poro de 0.22µm. Reservar a 4°C.
- **Solução-estoque tampão fosfato de potássio 1M pH6.0**: Misturar 132ml de 1M K₂HPO₄ com 868ml de 1M KH₂PO₄. Regular o pH para 6.0 com ácido fosfórico ou hidróxido de potássio. Autoclavar.
- **Solução estoque de glicerol**: Solução de glicerol 50%(v/v) em água destilada. Filtrar em poro de 0.22µm. Reservar a 4°C.

Meios de cultura

- **Meio LB (Luria-Bertani)**: Preparar solução de 1% de triptona, 1% de NaCl, 0,5% de extrato de levedura, (1,5% de ágar para meio sólido), ajustar o pH para 7.5 com NaOH e autoclavar. Reservar a 4°C.
- **Meio LSLB (Low Salt Luria-Bertani)**: Preparar solução de 1% de triptona, 0,5% de NaCl, 0,5% de extrato de levedura, (1,5% de ágar para meio sólido), ajustar o pH para 7.5 com NaOH e autoclavar. Reservar a 4°C.
- **Meio YPD (Yeast Extract Peptone Dextrose)**: Preparar solução de 1% de extrato de levedura, 2% de peptona, 2% de dextrose e (1,5% de ágar para meio sólido). Autoclavar. Reservar a 4°C.
- **Meio MD (Minimal Dextrose)**: Autoclavar água destilada (adicionar 1,5% de ágar para meio sólido) e adicionar as soluções-estoque para uma concentração final de 1,34% de YNB, 4×10^{-5} % de biotina e 2% de dextrose. Reservar a 4°C.

- **Meio MDH (Minimal Dextrose - Histidine):** Autoclavar água destilada (adicionar 1,5% de ágar para meio sólido) e adicionar as soluções-estoque para uma concentração final de 1,34% de YNB, 4×10^{-5} % de biotina, 2% de dextrose e 4×10^{-3} % de histidina. Reservar a 4°C.
- **Meio BMM (Buffered Minimal Metanol):** Autoclavar água destilada e adicionar as soluções-estoque para uma concentração final de 100 mM de tampão fosfato de potássio pH6.0, 1,34% de YNB, 4×10^{-5} % de biotina e 0,5% metanol. Reservar a 4°C.
- **Meio BMM-H (Buffered Minimal Metanol - Histidine):** Autoclavar água destilada e adicionar as soluções-estoque para uma concentração final de 100 mM de tampão fosfato de potássio pH6.0, 1,34% de YNB, 4×10^{-5} % de biotina, 0,5% metanol, e 4×10^{-3} % de histidina. Reservar a 4°C.
- **Meio BMMY (Buffered Complex Methanol):** Solução 1% de extrato de levedura, 2% de peptona, 100mM de tampão fosfato de potássio pH6.0, 1,34% de YNB, 4×10^{-5} % de biotina, 0,5% metanol, e 4×10^{-3} % de histidina]. Para 1L de solução dissolver 10g de extrato de levedura e 20g de peptona em 700 ml de água destilada e autoclavar. Em seguida, adicionar as soluções-estoque de acordo com as concentrações finais da descritas.

Materiais para DNA

Extração e visualização de DNA plasmidial

- **Solução tampão 10X TAE (tris-acetato-EDTA):** Tris base 400 mM, ácido acético 190 mM, EDTA 10mM, pH 7.6.
- **Gel de Ágarose:** 0,5% -3,0% de ágarose em solução 1X de TAE.
- **Solução tampão de amostras de DNA 10X Ficoll Dye:** 0,2% (m/v) de azul de bromofenol, 50%(v/v) de ficollem TAE 1X.
- **Marcador de massa molecular de DNA:** Lambda DNA/HindIII Fermentas (23130, 9416, 6557, 4361, 2322, 2027, 564,125); GeneRuler™ DNA Ladder – Low Range (700, 500, 400, 300, 200, 150, 100, 75, 50, 25); 1kb Plus DNA Ladder (12000 até 1000, de mil em mil; e também 1650, 850, 650, 500, 400, 300, 200, 100).
- **Solução estoque de brometo de etídio:** 10mg/ml em água destilada (aplicar 1ul / 50 ml de gel de ágarose).
- **Solução de fenol equilibrado pH8.0.**
- **Clorofórmio.**
- **Kit de Extração de DNA:** GFX Micro Plasmid Prep Kit (Amersham Biosciences).
- **Kit de Purificação de DNA:** GFX PCR DNA and Band Purification kit (Amersham Biosciences).

Extração de DNA genômico de *P. pastoris*

- **Solução de lise:** Preparar solução tampão de 50 mM Tris-HCl pH 7.2, 50mM EDTA, 3% SDS, 1% β -mercaptoetanol.
- **Solução tampão de acetato de sódio 3M.**
- **Isopropanol.**
- **Solução de RNase:** Preparar solução estoque de 100mg/ml de RNase A.

Enzimas para DNA

- **Enzimas de Restrição:** Endonucleases de restrição do tipo II fornecidas pela New England Biolabs, Invitrogen e GE Healthcare. As enzimas reconhecem seqüências específicas de DNA e clivam pontos constantes dentro da seqüência.
- **CIP:** “Calf Intestinal alkaline phosphatase” (CIP) fornecida pela New England Biolabs. As fosfatases alcalinas catalisam a remoção do grupo fosfato do DNA e do RNA.
- **Klenow:** A “DNA Polymerase I large fragment” (Klenow), fornecida pela New England Biolabs, é fragmento da DNA polimerase I de *E. coli* que mantém a atividade 3' $\rightarrow$ 5' exonuclease, mas não possui a atividade 5' $\rightarrow$ 3' exonuclease.
- **Ligase:** A enzima T4 DNA Ligase de *E. coli*., fornecida pela New England Biolabs, catalisa a formação de uma ligação fosfodiéster entre o grupo fosfato5' e o grupo hidroxil3' do DNA ou RNA.
- **DNA Polymerase:** Platinum® Taq DNA Polymerase High Fidelity, fornecida pela invitrogen, catalisa a polimerização de moléculas de DNA de até 20kb e possui atividade *proofreading*.
- **Kit de Real-Time PCR:** Platinum® SYBR® Green qPCR SuperMix-UDG, fornecido pela invitrogen. O kit possui todos os componentes necessários para amplificação e detecção do DNA através de PCR em tempo real.
- **Solução de MgSO₄:** Solução estoque de 50mM de MgSO₄ para reações de PCR.
- **dNTP MIX:** Solução 0,6mM de Tris-HCl pH 7.5, com 10mM de cada nucleotídeo.

Materiais para Análise de Proteínas

Materiais para SDS-PAGE

- **Solução de Bradford:** Solução de corante Brilliant Blue G usada para quantificação de proteínas (Sigma Aldrich).
- **Solução tampão de amostra para SDS-PAGE 10X:** Tris-HCL 1M pH6.8, 10%(m/v) de SDS, 0,5%(m/v) de azul de bromofenol e/ou xilenocianol, 50%(v/v) de glicerol e β -mercaptoetanol 1M.
- **Solução tampão 5X Tris-glicina:** 1,5% (m/v) de Tris base, 9,4%(m/v) de glicina, 0,5%(m/v) de SDS, pH 8.3.
- **Gel de poliacrilamida para SDS-PAGE:** fase de separação 15% e fase de concentração 5%.
- **Solução de Coomassie:** Solução 40% de etanol, 10% de ácido acético glacial e 0,25% de Coomassie Brilliant Blue R-250.
- **Solução descorante:** Solução de 40% de etanol e 10% de ácido acético glacial.
- **Marcador de peso molecular para Proteína:** LMW (97, 66, 45, 30, 20.1, 14.4) ou MW (66, 45, 34.7, 18.4, 14.3).

Materiais para transferência e Western Blot

- **Solução de transferência:** Tris-glicina 1X, 1%(m/v) de SDS, etanol 20%.
- **Membrana de Nitrocelulose:** Hybond ECL Nitrocellulose Membranes (Amersham Biosciences).
- **Solução de tampão fosfato (PBS):** NaCl 3mM, KCl 127mM, Na_2HPO_4 100mM, KH_2PO_4 14mM, pH7.4.
- **Solução de tampão fosfato Tween (PBST):** solução tampão PBS, 0,05%(v/v) de Tween® 20.
- **Solução de tampão fosfato Tween (PBST-Leite):** solução PBST com leite desnatado 5% (m/v).
- **Solução de tampão fosfato Tween (PBST-BSA):** solução PBST com BSA 3% (m/v).
- **Kit de Western Blot:** ECL Western Blotting Detection Reagents and Analysis System (Amersham Biosciences).
- **Filme fotográfico:** Hyperfilm ECL (Amersham Biosciences).
- **Anticorpo Primário:** anticorpo IgG anti-L1 do HPV16 produzido em camundongo (CamVir® ABcam). Reservar a -20°C .
- **Anticorpo Secundário:** anticorpo anti-IgG de camundongo conjugado a peroxidase (Sigma Aldrich). Reservar a -20°C .

Materiais para preparar leveduras competentes

- **Solução de DTT** : Preparar solução aquosa de DTT 1M. Filtrar em poro de 0.22µm e reservar a -20°C.
- **Solução de HEPES**: Preparar solução de HEPES 1M pH 8.0, filtrar em poro de 0.22µm e reservar a 4°C.
- **Solução de Sorbitol**: Preparar solução aquosa de sorbitol 1M. Autoclavar e reservar a 4°C.

Materiais para lise celular de *P. pastoris*

- **Solução de lise celular para leveduras**: Preparar solução tampão PBS 1X, EDTA 1mM e glicerol 5% (v/v). Adicionar PMSF para uma concentração final de 1 mM na hora do uso.
- **Solução de PMSF**: Solução 100 mM do inibidor de protease PMSF dissolvido em isopropanol.
- **Esferas de vidro de 0.5mm** (Sigma Aldrich).

Antibióticos

Para que se preservem suas propriedades, as drogas são estocadas a -20°C e adicionadas a soluções cuja temperatura não ultrapasse 50°C. O estoque, os meios de cultura e as placas contendo Zeocina são armazenados protegidos da luz, porque este antibiótico é fotossensível.

- **Ampicilina (Invitrogen)**: Estoque de 100mg/ml. Dose para seleção de 100µg/ml para *E. coli*.
- **Kanamicina (Invitrogen)**: Estoque de 100mg/ml. Dose para seleção de 100µg/ml para *E. coli*.
- **Geneticina (Invitrogen)**: Estoque de 100mg/ml. Dose para seleção de 250µg/ml a 4000µg/ml para *P. pastoris*.
- **Zeocina (Invitrogen)**: Estoque de 100mg/ml. Dose para seleção de 25µg/ml para *E. coli* e 100µg/ml a 1000µg/ml para *P. pastoris*.

Oligonucleotídeos

Neste projeto, foram utilizados 14 oligonucleotídeos de DNA em experimentos de sequenciamento e reações de PCR. Todos eles foram sintetizados pela empresa Invitrogen. A sequência de cada um deles está descrita a seguir, na **Tabela 1**.

Nome	Sequência
<i>pINT-A forward</i>	CATATGTCGGATCCCCACACACCATA
<i>pINT-A reverse</i>	CATATGGCCTGACTGCGTTAGCAATT
<i>pINT-B forward</i>	CATATGACCCCATACGCCTTAGAAAA
<i>pINT-B reverse</i>	CATATGATCGATAAGCTTGACAAAC
<i>HIS4 forward</i>	CCTGGTGAACCGCAATTCTGC
<i>HIS4 reverse</i>	TGCAACCAGCAACTTTGGCAGG
<i>L1 forward</i>	TTCCGGTCACCCATTGTTGAACA
<i>L1 reverse</i>	CACCAGCGTTAGCAGCGTAAGCA
<i>18S-25S forward</i>	GGCTGCTGGCACCAGACTTGC
<i>18S-25S reverse</i>	ACTGGCTTGTGGCAGTCAAGCG
<i>5S forward</i>	GCAGAAAGCACCGTTTCCCGTCC
<i>5S reverse</i>	TGCAGCACCTGAGTTTCGCGTATGG
<i>AOX5'</i>	TTGCGACTGGTTCCAATTGACAAG
<i>AOX3'</i>	CATCTCTCAGGCAAATGGCATTCTG

Tabela 1. Sequências dos oligonucleotídeos de DNA usados no projeto.

Vetores

O vetor pGEM®-T Easy (Promega) foi utilizado para armazenar três elementos importantes na construção dos vetores de expressão. Os vetores receberam um sufixo de acordo com as sequências neles clonadas.

- O pGEM-Teasy-L1 contém o gene sintético da proteína L1 otimizada para *P. pastoris*. O gene foi doado pelo Dr. Robert L. Garcea da Colorado University EUA.
- O pGEM-Teasy-rDNA contém o fragmento do DNA ribossomal de *H. polymorpha* (Klabunde et al, 2002).

- O pGEM-Teasy-G61APH contém o gene de resistência à Geneticina, controlado pelo promotor fraco GAPDH61 (Sohn et al, 1999).

O vetor pHARS36JR, fornecido pela empresa GENEART, é um vetor de clonagem que possui a sequência sintética do fragmento HARS36 de *H. polymorpha* (Sohn et al, 1996).

Os vetores de expressão pPICHOLI® Shuttle vector (Mobitec) e pPIC3.5K® *Pichia pastoris* vector (invitrogen) foram utilizados como base para a construção dos vetores deste projeto. Os mapas desses dois vetores são descritos na **Ilustração 5**.

Ilustração 5. A. Vetor pPICHOLI-1 (3579 bp). Expressão em *P. pastoris* e em *E. coli*. Contém um cassete de expressão controlado pelo promotor do gene AOX1 de *P. pastoris* e o promotor T7; uma seqüência de replicação autônoma (PARS1) de *P. pastoris*; um marcador de resistência à Zeocina para *P. pastoris* e *E. coli*; e uma origem de replicação bacteriana ColE1. **B.** Vetor pPIC3.5K (9004 bp). Expressão intracelular em *P. pastoris* sob controle de um marcador auxotrófico ou de resistência à Geneticina. Este plasmídeo contém um cassete de expressão controlado pelo promotor do gene AOX1; o gene *his4* de *P. pastoris* para seleção auxotrófica de linhagens defectivas na síntese da histidina; o gene de resistência a Kanamicina em *E. coli*, que também confere resistência à Geneticina em leveduras; um fragmento da extremidade 3' do gene AOX1 para integração por substituição homóloga; uma origem de replicação bacteriana pMB1 e um gene de resistência à Ampicilina para *E. coli*.

Linhagens de microorganismos

- ***Escherichia coli* DH5α (Life Technologies®)**: genótipo [F⁻, Φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA1*, *endA1*, *hsdR17*(r_K⁻, m_K⁺), *phoA*, *supE44*, λ⁻, *thi-1*, *gyrA96*, *relA1*]. Esta linhagem foi utilizada para armazenamento, clonagem e amplificação de DNAs plasmidiais.
- ***Pichia pastoris* (Invitrogen)**: A linhagem KM71 (*his4*, *arg4*, *aox1*Δ*ARG4*), patenteada pela Invitrogen, possui mutações no gene histidinol desidrogenase (*his4*, o que impede a síntese do aminoácido histidina). A cepa KM71 contém o gene AOX1 não funcional e depende da enzima álcool oxidase produzida pelo gene AOX2. As taxas de consumo de metanol nesta linhagem são baixas, portanto seu fenótipo é denominado “methanol utilization slow” (Mut^s). A linhagem parental da KM71 possui o gene argininosuccinato liase (*arg4*) mutado, porém o gene ARG4 de *S. cerevisiae* foi utilizado para causar a ruptura do gene AOX1, criando uma linhagem Mut^s Arg⁺ His⁻.

MÉTODOS

Manipulação de *E. coli* DH5 α ^{TM}

A linhagem DH5 α ^{TM} de *Escherichia coli* foi utilizada como organismo de armazenamento e amplificação de plasmídeos. A transferência dos plasmídeos para a DH5 α ^{TM} foi realizada a partir dos protocolos de preparação de células competentes e de transformação por choque térmico descritos no manual Molecular Cloning (Sambrook et al., 2002). Para o cultivo de DH5 α ^{TM}, células foram inoculadas em meio LSLB com antibiótico ao qual o plasmídeo confere resistência e cultivadas por 16 horas a 37°C e 250rpm. Para a preservação das células, foram misturados 500 μ l de cultura e 500 μ l de glicerol 50% em eppendorf e esse estocado a -80°C. Novas culturas foram feitas a partir dos estoques, conforme a necessidade.

Desenvolvimento dos novos vetores

Extração e purificação de DNA plasmidial

A obtenção de plasmídeos clonados em DH5 α ^{TM} foi realizada com o kit de extração de DNA e a purificação de fragmentos de DNA de bandas de géis de agarose, feita com o kit de purificação de DNA, ambos fornecidos pela Amersham Biosciences. Ocasionalmente, quando era necessário submeter os plasmídeos somente a experimentos de eletroforese, foi utilizado o método fenol/clorofórmio para extrair os plasmídeos (Sambrook et al., 2002).

Construção e modificação de plasmídeos

Os vetores foram construídos a partir de técnicas de reestruturação de seqüências de DNA por endonucleases de restrição, fosfatase alcalina, DNA ligase e pela enzima Klenow, com atividade exonuclease 3'→5' (Sambrook et al.; 2002).

Após cada etapa de construção dos vetores, o material foi clonado em células DH5α. Após cada etapa de modificação, a integridade da molécula de DNA foi verificada através de ensaios de restrição e eletroforese em gel de agarose. Para verificar o sentido de inserção do gene da proteína L1 nos cassetes de expressão foram feitos seqüenciamentos automáticos com os iniciadores AOX 5' e AOX3', que flanqueiam o sítio de clonagem, e também ensaios de restrição.

Construção do vetor pEPI-AL1

O vetor pPIC3,5K foi digerido pelas enzimas SphI e Klenow, seguida pela clivagem com SacI. O vetor pPICHOLI-1 foi digerido pelas enzimas PvuII e SacI. Em seguida, os materiais foram submetidos a uma separação por eletroforese. O fragmento de 5552bp do vetor pPIC3,5K e o de 2400bp do vetor pPICHOLI-1 foram purificados, ligados, e o produto foi transformado em DH5α™.

As células foram plaqueadas em meio LSLB-Zeocina/Kanamicina. Parte das colônias obtidas possuía o plasmídeo desejado de 7952bp. Foram realizados ensaios de restrição para verificar a integridade do plasmídeo e os padrões de fragmentos gerados estavam de acordo com o mapa de restrição do vetor planejado. O novo vetor foi denominado **pEPI-A**.

Para fazer a inserção da proteína L1 no vetor pEPI-A foram realizadas novas digestões dos plasmídeos. O pGEM-TL1 e o pEPI-A foram digeridos com a enzima EcoRI e submetidos a uma separação por eletroforese. O fragmento de aproximadamente 1550bp contendo o gene da proteína L1 foi inserido no vetor pEPI-A. O vetor resultante de 9511bp foi analisado por sequenciamento automático e o vetor pEPI-A com o inserto do gene da proteína L1 no sentido correto foi denominado **pEPI-AL1**.

Células competentes de *P. pastoris* KM71 foram transformadas com o vetor pEPI-AL1 para confirmar a eficácia dos marcadores de seleção. Os transformantes foram plaqueados

em todos os meios seletivos: Geneticina (250µg/ml), Zeocina(100µg/ml) e meio MD. Como controle negativo foram plaqueadas células de *P. pastoris* KM71 sem o vetor nos mesmos meios seletivos.

Construção do vetor pINT-AL1

O vetor pPICHOLI-1 foi digerido com a enzima Sall, tratado com a enzima Klenow, purificado, digerido pela enzima Sspl e por último desfosforilado pela CIAP. O vetor pGEM-Teasy-rDNA foi digerido com a enzima EcoRI e tratado pela enzima Klenow. O fragmento de 3119 bp do vetor pPICHOLI-1 e o de 2.455 bp do vetor pGEM-Teasy-rDNA foram purificados, ligados e transformados em DH5α™.

Os clones contendo vetores de aproximadamente 5400bp foram selecionados e seus plasmídios foram extraídos, seqüenciados e submetidos a um ensaio de restrição para verificar a integridade do construto. O plasmídeo gerado recebeu o nome de pPICHOLI-rDNA.

Os vetores pPICHOLI-rDNA e pPIC3.5K foram utilizados na segunda etapa de construção. O vetor pPICHOLI-rDNA foi digerido com Accl e BglII e o vetor pPIC3.5K digerido com BglII e EcoRV. O fragmento de 1478bp do vetor pPIC3,5K e o fragmento de 4480 do vetor pPICHOLI-rDNA foram purificados, ligados e transformados em células competentes de DH5α™.

Os colônias geradas tiveram seus plasmídeos extraídos, e os vetores com o tamanho de 5899 bp foram submetidos a ensaios de restrição. O vetor que apresentou o padrão de fragmentos correto e que possuía os sites de clonagem (EcoRI) e de linearização (ScaI ou XhoI) íntegros foi selecionado. Este novo vetor recebeu o nome de **pINT-A**.

Para fazer a inserção do gene L1 no vetor pINT-A foram realizados os mesmos procedimentos usados para a inserção no vetor pEPI-A. O pGEM-Teasy-L1 e o pINT-A foram digeridos com a enzima EcoRI. O fragmento de aproximadamente 1550 bp contendo o gene da proteína L1 foi inserido no vetor pINT-A. O vetor resultante de 7400 bp teve sua sequência verificada por sequenciamento automático e foi denominado **pINT-AL1**.

Construção do vetor pINT-BL1

Os vetores pINT-A e pGEM-Teasy-G61GAPH foram utilizados como base para a construção do vetor integrativo para *P. pastoris* pINT-B.

O vetor pINT-A foi digerido com as enzimas MluI e NdeI e tratado com a enzima Klenow. O vetor pGEM-Teasy-G61APH foi digerido pelas enzimas NheI e NarI e depois tratado com a enzima Klenow. O fragmento de 897bp do vetor pGemT-G61APH e o fragmento de 4723bp do vetor pINT-A foram purificados, ligados e transformados em *E.coli* DH5 α ^{TM}.

As colônias formadas tiveram seus plasmídeos extraídos e, aqueles que possuíam o tamanho de 5624bp, foram selecionados e tiveram seus vetores submetidos a um ensaio de restrição. O vetor que apresentou o padrão de fragmentos correto e que possuía os sítios de clonagem (NotI) e de linearização (NaeI) íntegros foi selecionado. Este novo vetor recebeu o nome de **pINT-B**.

O vetor pGEM-Teasy-L1 foi usado para obter o gene da proteína L1 do HPV16. O pGEM-Teasy-L1 e o pINT-A foram digeridos com a enzima NotI e o último foi desfosforilado. Em seguida os fragmentos foram ligados e transformados em DH5 α ^{TM}.

Os clones contendo plasmídeos com aproximadamente 7199bp foram selecionados e seus vetores foram submetidos a um ensaio de restrição e analisados por sequenciamento. O vetor com a inserção correta do gene da L1 foi denominado **pINT-BL1**.

Construção dos vetores pINT-CL1 e pINT-DL1

Os vetores pINT-A e pINT-B foram utilizados como base para a construção dos vetores pINT-C e pINT-D, respectivamente.

Os iniciadores *pINT-A forward* e *pINT-A reverse* foram usados para amplificar o vetor pINT-A por PCR formando um produto que possui toda a sequência do plasmídeo, menos a sequência do *rDNA* de *H. polymorpha*. Os iniciadores possuem a sequência CATATG em suas extremidades 5' e, por causa disso, o sítio de restrição para a enzima NdeI foi adicionado nas extremidades do produto de PCR. O mesmo procedimento foi realizado no vetor pINT-B utilizando os iniciadores *pINT-B forward* e *pINT-B reverse*.

Para facilitar a reação de amplificação, os vetores foram previamente linearizados com a enzima Scal, que cliva um sítio interno da sequência do rDNA de *H. polymorpha*. As duas reações de PCR utilizaram o mesmo protocolo e ele está descrito na **Tabela 2**.

Componente	Concentração Final	Protocolo de Amplificação
10X High Fidelity PCR Buffer	1X	94°C por 30 segundos 35 ciclos: 94°C por 30 segundos 58°C por 30 segundos 68°C por 4 minutos 4°C ∞
Nucleotídeos	0.2mM cada	
MgSO ₄	2 mM	
Iniciadores	0,4 mM cada	
Amostra de plasmídeo	0,1 µl de miniprep	
Platinum® Taq High Fidelity	1 unidade	
Água destilada autoclavada	Para completar 50 µl	
Volume Final	50 µl	

Tabela 2. Protocolo da reação de PCR para a construção dos vetores pINT-C e pINT-D

Cada um dos produtos foi purificado, ligado a um vetor pGEM-Teasy aberto e esse, transformado em DH5α™. Os vetores extraídos dos transformantes contendo os tamanhos de 6318 bp e 6074 bp foram denominados, respectivamente, pGEM-Teasy-INTA e pGEM-Teasy-INTB.

Os vetores pHARS36JR, pGEM-Teasy-INTA e pGEM-Teasy-INTB foram utilizados na segunda etapa de construção. O vetor pHARS36JR forneceu a sequência HARS36 e os demais componentes foram fornecidos pelos vetores base: pGEM-Teasy-INTA e pGEM-Teasy-INTB. Todos os vetores foram digeridos com a enzima de restrição NdeI. O fragmento de 540 bp do vetor pHARS36 foi ligado aos fragmentos de 3301bp e 3057bp dos vetores pGEM-Teasy-INTA e pGEM-Teasy-INTB, separadamente. Os dois produtos foram transformados em DH5α™.

Clones transformados com o primeiro produto contendo plasmídeos de 3831bp e clones transformados com o segundo produto contendo plasmídeos de 3587bp foram selecionados e seus plasmídeos, extraídos. Em seguida, os vetores passaram por ensaios de restrição. Os vetores que apresentaram o padrão correto de fragmentação e que possuíam os sítos de clonagem (NotI) e de linearização (Sall) íntegros, foram selecionados. O vetor de 3831bp recebeu o nome de **pINT-C** e o de 3587bp recebeu o nome de **pINT-D**.

O vetor pGEM-Teasy-L1 foi usado para obter o gene da proteína L1 do HPV16. O vetor foi digerido com a enzima NotI e a mesma enzima foi utilizada para linearizar os vetores pINT-C e pINT-D. Os quais foram desfosforilados e depois ligados ao fragmento de

aproximadamente 1550 bp contendo o gene da proteína L1. Os dois produtos foram transformados em DH5 α ^{TM}.

Finalmente, os clones contendo os plasmídeos com aproximadamente 5406bp e 5162bp foram selecionados e analisados por seqüenciamento automático. Clones transformados com o primeiro produto contendo plasmídeos de 5406bp e os clones transformados com o segundo produto contendo plasmídeos de 5162bp foram selecionados, e seus plasmídeos extraídos. Os vetores foram submetidos a um ensaio de restrição e analisados por sequenciamento. Os vetores com a inserção correta do gene da L1 foram denominado, respectivamente, **pINT-CL1** e **pINT-DL1**.

Manipulação de *P. pastoris* KM71

A linhagem KM71 de *Pichia pastoris*, fornecida pela empresa Invitrogen, foi utilizada como organismo base para o desenvolvimento dos sistemas de expressão da proteína L1. As metodologias para a manipulação dela foram extraídas do manual do vetor pPICHOLI-1 (Mobitec, 2003).

Para transferir os vetores para células KM71 por eletroporação, a levedura precisa passar por um procedimento para tornar-se eletro-competente. Para realizar a transformação, os vetores de expressão integrativos precisam ser linearizados por enzimas de restrição que clivam no interior das sequências de integração. Os vetores episomais não necessitam de linearização.

Após a transformação, as células são plaqueadas em meio seletivo sólido. As colônias demoram de 48 a 72 horas para surgir nas placas e, em seguida, são isoladas e inoculadas em meio seletivo líquido. A partir dessas culturas são feitos estoques dos clones para posterior utilização.

Para fazer um estoque, cerca de 2 ml de cultura de DO₆₀₀= 2 são centrifugados por 1 minuto a 2000G e o precipitado é ressuspendido em 1ml de meio YPD 25% de glicerol e estocado a -80°C.

Adicionalmente, todas as culturas de *P. pastoris* são suplementadas com o antibiótico Ampicilina para evitar contaminação por bactérias.

Protocolo de Eletro-competência

O protocolo de eletro-competência foi retirado do manual do vetor pPICHOLI-1 (Mobitec, 2003). Uma colônia de *P. pastoris* é inoculada em 10 ml de meio YPD e cultivada por 16 horas a 30°C e a 250rpm. Essa cultura é inoculada em 500 ml de meio YPD e crescida a 30°C e a 250rpm até uma $DO_{600} = 1.3 - 1.5$. Logo após, a cultura é centrifugada por 10 minutos a 4°C e a 2000G e as células são ressuspensas em meio 100 ml de meio YPD suplementado com 20 ml de HEPES 1M e 2,5ml de DTT 1M. As células são incubadas por 15 minutos a 30°C sem agitação. Em seguida é adicionada água fria para um volume final de 500 ml e a solução é centrifugada por 10 minutos a 4°C e a 2000G. Em seguida as células são lavadas com 250 ml de água fria e novamente são centrifugadas por 10 minutos a 4°C e a 2000G. Depois as células são ressuspensas em 20 ml de Sorbitol 1M e centrifugadas pela última vez nas condições anteriores. Por último as células são ressuspensas em 500 µl de Sorbitol 1M, aliquotadas e estocadas a -80°C.

Protocolo de Transformação

O protocolo foi retirado do manual do vetor pPICHOLI-1 (Mobitec, 2003). Uma amostra de 100 a 1000ng de DNA diluído em 5 µl água destilada estéril é misturada com uma alíquota de 40µl de células competentes. A solução é transferida para uma cubeta de eletroporação de 2 mm de diâmetro previamente resfriada em gelo. As células são eletroporadas no equipamento GENE-PULSER da Bio-Rad (1,500 V, 200 W, 25 µF). Logo em seguida, 1 ml de Sorbitol 1M é adicionado na cubeta e misturado cuidadosamente com as células. A solução é transferida para um eppendorf de 1,5ml e incubada por 1 ou 2 horas a 30°C. Por último, alíquotas de 50µl, 100µl e 200µl da transformação são plaqueadas e incubadas por 2 a 3 dias a 30°C em meio sólido contendo o respectivo marcador de seleção do vetor.

Protocolo de expressão da proteína L1

O protocolo descrito foi retirado do manual do vetor pPICHOLI-1 (Mobitec, 2003) e descreve as linhas gerais do processo de expressão. A composição do meio de cultura varia em função do experimento e será especificada em cada caso. Na etapa de crescimento podem ser usados os meios YPD, MD ou MDH, e na etapa de expressão podem ser utilizados os meios BMMY, BMM ou BMMH.

Uma colônia ou raspa do estoque é inoculada em 5 ml de meio líquido e cultivada a 30°C e 250rpm. Após um período de 24 a 48 horas, a cultura atinge uma OD₆₀₀ de 2-4. Parte da cultura é misturada com um meio de cultura tamponado sem metanol na proporção de 1:10. A cultura é novamente cultivada a 30°C e 250rpm até atingir uma OD₆₀₀ = 1. Nesse momento, é adicionado metanol para uma concentração final de 0,5% para induzir a expressão da proteína heteróloga. A cultura é mantida a 30°C e 250rpm. O metanol é repostado a cada 24 horas. A expressão é mantida por 48 a 72 horas. Após esse período, a cultura é centrifugada por 10 minutos a 4°C e a 2000G. Por último, o sobrenadante é descartado e o precipitado celular é estocado a -80°C. Não foi necessário guardar os sobrenadantes porque só foram utilizados sistemas de expressão intracelular.

Protocolo de Lise Celular

As células de levedura precisaram ser lisadas para se obter a proteína L1. Primeiramente, a massa celular é ressuspensa em tampão de lise (que possui inibidor de protease) a 4°C, em um volume 10 vezes menor do que o utilizado na etapa de expressão. Depois são adicionados 0,5 a 1 volume de esferas de vidro e executados seis ciclos de 1 minuto de agitação em vortex e 1 minuto de incubação no gelo. Por fim, o lisado celular é centrifugado a 10000G e 4°C por 10 minutos, o sobrenadante é separado e congelado a -20°C para posterior análise. Para se obter um extrato protéico menos degradado, deve-se adicionar inibidor de protease também no final do procedimento.

Protocolo de análise de estabilidade

Para estudar a estabilidade mitótica dos clones de *Pichia pastoris*, foram realizadas análises comparativas da expressão da proteína L1 por *Western Blot* entre leveduras da 1ª e da 60ª geração de um mesmo clone.

Para a obtenção de um clone da 60ª geração contendo o vetor episomal com marcador auxotrófico, foi feito um monitoramento da proliferação do clone de 1ª geração através de diluições seriadas e medições da densidade óptica das culturas. Cada clone foi cultivado até a 60ª geração em dois tipos de meio mínimo: MD (sem histidina) e MDH (com histidina). Após esse período, as culturas foram plaqueadas em seus respectivos meios sólidos para obtenção de colônias isoladas. Uma colônia foi inoculada novamente em meio líquido, cultivada e estocada.

Esse processo de monitoramento é trabalhoso e tem um risco considerável de contaminação, por isso foi utilizado um método diferente para cultivar os vários clones com vetores integrativos.

Para a obtenção da 60ª geração de um clone contendo um vetor integrativo, foram feitos plaqueamentos sucessivos em meio sólido não seletivo. As colônias obtidas pela transformação foram repicadas a cada 24 horas por 7 dias (foi estipulado que cada passagem de placa representa aproximadamente 10 gerações, porque o tempo médio de duplicação da *P. pastoris* em meio líquido é de 2 horas). Depois, as colônias foram inoculadas em meio líquido seletivo e as culturas que cresceram foram estocadas.

Protocolo de Extração de DNA genômico

Neste projeto, foram feitas extrações do DNA genômico de *Pichia pastoris* para realizar experimentos de PCR. Os protocolos convencionais de extração usam hidrolase, como Zymolyase e Lyticase, para degradar a parede celular, mas essas enzimas são muito caras. Por esse motivo, as extrações foram feitas a partir de um protocolo de lise química (Tomita et al., 2002).

Análise da expressão da L1

Para detectar a expressão da proteína L1 foi realizada uma análise do extrato celular por SDS-PAGE e Western blot (Sambrook et al., 2002). A eletroforese foi realizada em um gel de poliacrilamida 15%. A quantidade de proteína total das amostras foi normalizada pelo método de Bradford para possibilitar uma análise comparativa da expressão da L1 entre os clones. Foram aplicadas amostras de 20 µg de proteína total no sistema de SDS-PAGE.

Todos os experimentos de SDS-PAGE foram feitos em duplicata. Um dos géis foi usado para corar as proteínas com o reagente Commassie Brilliant Blue e o outro para transferir as proteínas para uma membrana de nitrocelulose através do sistema Bio-Rad Trans-Blot.

Em seguida, as membranas foram bloqueadas com PBST-Leite, por 1 hora, sob agitação. Após a remoção da solução de bloqueio, as membranas foram incubadas por 1 hora, sob agitação, com o anticorpo primário anti-L1 de HPV16 CamVir® em uma diluição de 1:3000 em PBST-BSA. Após três lavagens de 10 min cada com PBST, as membranas foram incubadas por 1 hora e sob agitação, com o anticorpo secundário anti-IgG de camundongo conjugado com peroxidase (Sigma Aldrich) em uma diluição de 1:10000 em PBST-Leite. Após as lavagens, as membranas foram reveladas em filme de raio-X *Hyperfilm* (GE Healthcare), utilizando o kit *ECL* (GE Healthcare), conforme as instruções do fabricante. Os filmes foram expostos por 5 minutos.

Na etapa final do projeto, também foi realizado um experimento de quantificação da expressão da proteína L1. Esse experimento foi possível graças à gentil doação de uma amostra da L1 purificada feita pela Dra. Luisa Lina Villa do Instituto Ludwig.

A quantificação foi realizada a partir uma análise comparativa das intensidades das bandas de um experimento de Western Blot. A leitura das intensidades das bandas formadas foi realizada pelo instrumento Eagle Eye II da Stratagene. Como a amostra possuía apenas alguns nanogramas, não foi possível fazer uma curva padrão da L1 completa e, por isso, os cálculos foram feitos a partir de apenas dois pontos.

Quantificação do número de cópias integradas

O número de cópias do gene da L1 do HPV16 integradas no genoma dos clones foi identificado através de experimentos de PCR quantitativo (qPCR).

Foram analisados seis clones da 60ª geração transformados com o vetor pINT-AL1. O DNA genômico de cada clone foi extraído, quantificado utilizando-se o ND-1000 *Spectrophotometer* (NanoDrop Technologies) e diluído em água para uma concentração final de 10ng/μl.

Foram desenhados oligonúcleotídeos específicos para o gene L1 do HPV16 e o gene HIS4 de *P. pastoris* utilizando o programa Vector NTI™. Os iniciadores *L1 forward* e *L1 reverse* amplificam um segmento de 71bp do gene L1 e os iniciadores *HIS4 forward* e *HIS4 reverse* amplificam um segmento de 70bp do gene HIS4 de *P. pastoris*. O gene HIS4 possui somente uma cópia no genoma da levedura, e por isso foi utilizado como controle interno das amostras.

O vetor pEPI-A foi utilizado como calibrador para corrigir a diferença entre os sinais de fluorescência da amplificação do gene L1 e HIS4. Como o vetor possui uma cópia dos dois genes, ele funciona como uma amostra equimolar dos genes L1 e HIS4. Uma solução de 1ng/ul do vetor pEPI-A foi diluída em série, e as diluições foram submetidas ao experimento de qPCR junto com as amostras de DNA genômico. Foi utilizada como calibrador a diluição cuja concentração do gene HIS4 era mais semelhante às das amostras genômicas.

Nesse experimento foi usado o kit Platinum® SYBR® Green qPCR SuperMix-UDG. O protocolo da reação está descrito na **Tabela 3**.

Componente do PCR	Concentração Final	Reação de Amplificação
DNA genômico	2 ng/μl	50°C por 2 minutos 95°C por 10 minutos 40 Ciclos de: 95°C por 15 segundos 60°C por 60 segundos 4°C ∞
Oligonucleotídeo	200 nM cada	
SYBR Mix	6,25 μl	
ROX	500 nM	
Volume Total	12,5 μl	

Tabela 3. Protocolo da reação de qPCR para a quantificação do número de cópias do gene L1.

O experimento foi conduzido no termociclador Applied Biosystems 7300 Real-Time PCR System e seguiu o protocolo padrão estabelecido pela invitrogen. Todas as reações foram feitas em triplicata.

A amplificação por PCR foi medida em tempo real analisando a emissão de fluorescência do SYBR Green, que se intercala apenas entre a fita dupla de DNA, conforme esta é produzida. As emissões de fluorescência a cada ciclo foram coletadas e analisadas pelo software *7300 System SDS RQ Study* e os resultados foram apresentados em gráficos de fluorescência (ΔRn) *versus* ciclos.

A interpretação dos dados brutos, o cálculo da eficiência da amplificação e a determinação do limiar de leitura foram feitos através do programa LingRegPCR (Ramakers et al., 2003). O modelo matemático utilizado para o calculo do número de cópias do gene da L1 se baseia em modelos de quantificação relativa (Pfaffl 2001; Rebrikov e Trofimov 2005):

$$\text{ratio} = \frac{(E_{\text{ref}})^{CP_{\text{Sample}}}}{(E_{\text{target}})^{CP_{\text{Sample}}}} \div \frac{(E_{\text{ref}})^{CP_{\text{Calibrator}}}}{(E_{\text{target}})^{CP_{\text{Calibrator}}}}$$

Onde *ratio* é a razão entre os objetos de estudo; *E* é a eficiência da amplificação; *CP* é o número de ciclos para atingir o limiar pré-estabelecido de fluorescência; *ref* é o gene usado como referência; *target* é o gene alvo; *Sample* é a amostra estudada e *Calibrator* é a amostra de calibração.

Nesse experimento o gene alvo é o gene L1. O gene de referência é o gene HIS4 que possui apenas uma cópia no genoma. A amostra estudada é o DNA genômico de um clone transformado com o vetor pINT-AL1 e a amostra calibradora é o vetor pEPI-A, que possui quantidades equimolares do gene L1 e HIS4. Considerando que o número de cópias do gene referência é um, a razão entre o gene alvo e o gene de referência indica a quantidade absoluta de cópias do gene L1 no genoma.

Localização da integração

Para verificar se a integração dos vetores pINT-A e pINT-B ocorreu nos sítios do DNA ribossomal de *P. pastoris*, foi realizado um experimento de PCR. Foram desenhados dois pares de iniciadores para amplificar duas regiões específicas do DNA ribossomal de *P. pastoris*, onde a integração possa ter ocorrido.

Cada par de oligonucleotídeos flangeia uma região genômica de *P. pastoris* homóloga à região do rDNA de *H. polymorpha* utilizada nos vetores. Os pares de oligonucleotídeos não são capazes de amplificar as sequências do DNA ribossomal do vetor.

Os oligonucleotídeos *18S-25S forward* e *18S-25S reverse* flangeiam a região intergênica dos genes 18S e 25S do DNA ribossomal de *P. pastoris*. Já os oligonucleotídeos *5S forward* e *5S reverse* flangeiam o DNA ribossomal 5S. A **Ilustração 6** mostra as regiões amplificadas.

Ilustração 6. Regiões dos DNAs ribossomais de *P. pastoris* amplificadas por PCR. A integração dos vetores contendo a sequência do rDNA de *H. polymorpha* no DNA da *P. pastoris* deve ocorrer no interior das regiões amplificadas. Em *P. pastoris*, o DNA 5S está separado do resto da unidade do DNA ribossomal.

As reações de PCR foram feitas a partir dos DNAs genômicos de clones da 60ª geração transformados com o vetor pINT-AL1 e seguiram o protocolo descrito na **Tabela 4**. Os produtos de PCR foram analisados por eletroforese em gel de agarose. Caso tivesse ocorrido a integração dos vetores no interior das regiões amplificadas, os produtos de PCR

teriam um tamanho muito maior do que o tamanho da região amplificada. Foi utilizado como controle negativo o DNA genômico de *P. pastoris* KM71.

Componente	Concentração Final	Protocolo de Amplificação
10X High Fidelity PCR Buffer	1X	94°C por 30 segundos 35 ciclos: 94°C por 30 segundos 58°C por 30 segundos 68°C por 20 minutos 4°C ∞
Nucleotídeos	0.2mM cada	
MgSO ₄	2 mM	
Iniciadores	0,2 mM cada	
Amostra de DNA genômico	100 ng (2mM)	
Platinum® Taq High Fidelity	1 unidade	
Água destilada autoclavada	Para completar 50 µl	
Volume Final	50 µl	

Tabela 4. Protocolo da reação de PCR para a amplificação de regiões do rDNA de *P.pastoris*.

Ensaio com o sistema epissomal pEPI-AL1

Células competentes de *P. pastoris* KM71 foram transformadas com o vetor pEPI-AL1 e plaqueadas em meio MDH com antibiótico: uma placa com 100µg/ml de Zeocina e a outra com 600µg/ml de Zeocina. Após três dias, foram selecionadas aleatoriamente colônias de cada placa e inoculadas meio líquido nas mesmas condições. No dia seguinte, os clones foram armazenados a -80°C e denominados clones de 1ª geração.

Os clones de 1ª geração foram submetidos a um ensaio de expressão em meio mínimo com histidina e Zeocina, em diferentes concentrações, para observar o efeito da concentração de antibiótico na expressão da proteína L1. Neste experimento, foram analisadas as expressões de sete clones: três selecionados e cultivados em meio com 100µg/ml de Zeocina e quatro selecionados e cultivados em meio com 600µg/ml de Zeocina. Como controle negativo foi usado um clone de *P. pastoris* KM71 sem nenhum vetor e como controle positivo foi usado o clone atual do laboratório para a expressão da proteína L1: uma *P. pastoris* KM71 transformada com o vetor pPICHOLI-1 selecionada e cultivada em meio com 100µg/ml de Zeocina.

Cada colônia foi inoculada em 5ml de meio MDH suplementado com Zeocina, respeitando as concentrações de antibiótico utilizadas nas placas de seleção. O inóculo foi cultivado por 24 horas a 30°C e a 250rpm. Em seguida, 0,5 ml de cultura foi inoculado em 4,5 ml de meio BMMH (também com Zeocina) para induzir a expressão da proteína L1 e este foi

cultivado a 30°C e a 250rpm. Após o período de 48 horas de indução, os clones foram submetidos a uma análise de expressão da proteína L1 por SDS-PAGE e Western Blot.

Em um segundo experimento, realizou-se uma análise de estabilidade com os dois clones que mais expressaram a proteína L1. Eles foram cultivados por 60 gerações em dois tipos de meio líquido: MD (sem histidina) e MDH (com histidina). Após esse período, as culturas foram plaqueadas em seus respectivos meios sólidos e uma colônia de cada placa foi selecionada e novamente cultivada em meio líquido. Os quatro clones resultantes foram denominados clones de 60ª geração. Parte das culturas foi utilizada para fazer um estoque do clone a -80°C e o restante foi utilizado no experimento a seguir.

Por último, foi feito um ensaio de expressão da L1 com os clones da 1ª geração e da 60ª geração para verificar se a marcação auxotrófica era suficiente para manter a expressão dos clones inicialmente selecionados com altas doses de antibiótico. Foram utilizados os mesmos controles do experimento anterior. Os experimentos de expressão foram feitos nas seguintes condições: clone de 1ª geração com expressão em meio BMMH suplementado com 600µg/ml de Zeocina; clone de 1ª geração com expressão em meio BMMH; clone de 1ª geração com expressão em meio BMM; clone de 60ª geração cultivado em meio MDH com expressão em meio BMMH; clone de 60ª geração cultivado em meio MD e com expressão em meio mínimo BMM. Os resultados foram analisados por SDS-PAGE e Western Blot.

Ensaio com o sistema integrativo pINT-AL1

Células competentes de *P. pastoris* KM71 foram transformadas com 1µg do vetor pINT-AL1 linearizado pela enzima XhoI. As células foram plaqueadas em meio YPD sólido com 1000µg/ml de Zeocina e colocadas em uma estufa a 30°C. Após dois a três dias, 50 das colônias formadas foram repicadas em uma nova placa com o mesmo meio seletivo e cultivadas em uma estufa a 30°C. As colônias dessa nova placa foram denominadas clones de 1ª geração.

No dia seguinte, as 50 colônias foram inoculadas em meio líquido seletivo para confirmar a resistência ao antibiótico. Cada uma foi inoculada em 5ml de meio YPD suplementado com Zeocina 1000µg/ml e cultivada por 24 horas a 30°C e a 250rpm. Em

seguida, alíquotas de cada cultura foram utilizadas para fazer um estoque em glicerol dos clones da 1ª. Geração. Esses estoques foram armazenados a -80°C e o restante das culturas foi utilizado no experimento descrito a seguir.

Os 20 clones com as maiores DO₆₀₀ foram selecionados para um ensaio de expressão. As culturas foram centrifugadas por 5 minutos a 4°C e 2000G. Em seguida, os sobrenadantes foram descartados, os precipitados de células foram ressuspensos em 5ml de meio BMMY com Zeocina 1000µg/ml e metanol para induzir a expressão e as culturas foram mantidas por 72 horas a 30°C e a 250rpm. Após esse período, as culturas foram centrifugadas por 5 minutos a 4°C e 2000G, o sobrenadante foi descartado e as células congeladas a -80°C.

Por fim, foi feita uma análise da expressão da proteína L1 por SDS-PAGE e Western Blot. Como controle negativo, foi usado um clone de *P. pastoris* KM71 sem nenhum vetor e como controle positivo, foi usado o clone atual do laboratório para a expressão da proteína L1: uma *P. pastoris* KM71 transformada com o vetor pPICHOLI-1 selecionada e cultivada em meio com 100µg/ml de zeocina.

Na etapa seguinte, realizou-se uma análise de estabilidade com os 50 clones da 1ª geração. Eles foram cultivados por 60 gerações em meio YPD sólido sem Zeocina e, depois desse período, as 50 colônias foram plaqueadas em YPD com 1000µg/ml de Zeocina. Após 24 horas, as colônias formadas foram inoculadas em meio YPD líquido com 1000µg/ml de Zeocina para verificar quais os clones que permaneciam resistentes ao antibiótico após 60 gerações. Depois de 48 horas, os clones que cresceram foram estocados em glicerol a -80°C e receberam o nome de clones de 60ª geração.

Na última etapa, os clones de 60ª geração foram submetidos a um ensaio de expressão da proteína L1 em meio de cultura sem antibiótico. Foram usados os mesmos controles que o experimento anterior. Os clones de 60ª geração foram inoculados em meio YPD líquido sem Zeocina e cultivados por 24 horas a 30°C e a 250rpm. Em seguida, os sobrenadantes foram descartados, os precipitados de células foram ressuspensos em 5 ml de meio BMMY com metanol para induzir a expressão e as culturas foram mantidas por 72 horas a 30°C e a 250rpm. Após esse período, as culturas foram centrifugadas por 5 minutos a 4°C e 2000G, o sobrenadante foi descartado e as células congeladas a -80°C. Em seguida, os resultados foram analisados por SDS-PAGE e Western Blot.

Depois dos ensaios de expressão, alguns clones da 60ª geração foram selecionados para uma quantificação do número de cópias do gene da L1 integradas no genoma. O

experimento foi feito através do método de qPCR. Paralelamente, o ensaio de expressão foi repetido para uma posterior análise comparativa entre expressão da L1 e número de cópias do gene L1.

Por último, dois clones tiveram sua expressão quantificada a partir de um de Western Blot comparativo. A quantificação foi feita a partir de algumas amostras de concentrações conhecidas de L1 purificada.

Ensaio com o vetor integrativo pINT-BL1

Em um primeiro experimento, células competentes de *P. pastoris* KM71 foram transformadas com o vetor pINT-BL1 e plaqueadas em meio cultura sólido com Geneticina. A dose mínima de antibiótico para selecionar clones resistentes é de 250µg/ml, mas as células foram plaqueadas em meios de cultura YPD contendo 500µg/ml e 4000µg/ml de Geneticina. As placas foram mantidas em estufa a 30°C por 72 horas.

Após esse período, todas as colônias formadas foram inoculadas em meio YPD líquido em duas concentrações de antibiótico (independentemente da concentração de antibiótico usada para a seleção): 1000µg/ml e 4000µg/ml de Geneticina. As culturas foram mantidas em estufa a 30°C e 250rpm por 48 horas. Depois, os clones que cresceram em pelo menos uma das condições foram estocados a -80°C e denominados clones de 1ª geração.

Os clones de 1ª geração foram submetidos a um ensaio de expressão na presença Geneticina. Os clones foram inoculados em 5 ml de meio YPD suplementado com Geneticina na mesma concentração em que foram cultivados posteriormente. As culturas foram mantidas por 48 horas a 30°C e a 250rpm até atingir uma OD₆₀₀ de aproximadamente 1 e depois foram centrifugadas por 5 minutos a 4°C e 2000G. Em seguida, os sobrenadantes foram descartados e os precipitados de células ressuspensos em 5 ml de meio BMMY com metanol para induzir a expressão. O meio de cultura BMMY também possuía Geneticina. As culturas foram mantidas por 72 horas a 30°C e a 250rpm.

Após esse período, os clones foram submetidos a uma análise da expressão da proteína L1 por SDS-PAGE e Western Blot. Foram aplicadas amostras de aproximadamente

10µg de proteínas totais no sistema de SDS-PAGE e, como controle negativo, foi usado um clone de *P. pastoris* KM71 sem nenhum vetor.

Em um segundo experimento realizou-se uma análise de estabilidade dos clones analisados no primeiro ensaio de expressão. Os clones de 1ª geração foram inoculados em meio YPD com Geneticina nas mesmas concentrações em que eles foram selecionados. Depois de 48 horas, os clones foram transferidos para meio YPD sólido sem Geneticina e cultivados a 30°C por 60ª gerações.

Após todas as passagens, as colônias foram inoculadas em meio YPD líquido suplementado com 1000µg/ml e 4000µg/ml de Geneticina para identificar quais clones permaneceram estáveis. Novamente, os clones que cresceram em pelo menos uma das condições foram estocados a -80°C e denominados clones de 60ª geração.

Depois, os clones de 60ª geração foram submetidos a um ensaio de expressão da L1. O ensaio foi feito em meio complexo sem antibiótico para analisar a expressão da L1 em células cultivadas por longas gerações em meio de cultura sem pressão seletiva. Como controle negativo, foi usado um clone de *P. pastoris* KM71 sem nenhum vetor.

Após o ensaio de expressão, os extratos celulares foram analisados por SDS-PAGE e Western Blot. Foram aplicadas amostras de aproximadamente 20µg de proteínas totais no sistema de SDS-PAGE.

RESULTADOS

Sistema Epissomal pEPI-AL1

Construção do vetor pEPI-AL1

Os vetores pPICHOLI-1 e pPIC3,5K foram utilizados como base para a construção do vetor epissomal para *P. pastoris* pEPI-A. O vetor pPICHOLI-1 forneceu a seqüência de replicação autônoma PARS1, o gene de resistência à Zeocina e a origem de replicação bacteriana ColE1. O vetor pPIC3,5K forneceu o cassete de expressão do gene AOX1, o gene HIS4 e o gene de resistência à Kanamicina. O vetor pGEM-Teasy-L1 foi utilizado para fornecer o gene da proteína L1 do HPV16 com códons otimizados para *Pichia pastoris* para o vetor pEPI-A. O vetor epissomal contendo o gene da proteína L1 foi denominado **pEPI-AL1**.

As **Ilustrações 7 e 8** mostram, respectivamente, o fluxograma de construção e o mapa do vetor pEPI-AL1.

Células de *P. pastoris* transformadas com o vetor pEPI-AL1 demonstraram resistência aos antibióticos Geneticina e Zeocina e foram capazes de crescer em meio de cultura sem histidina. Os transformantes foram plaqueados nos meios seletivos e, 48 horas depois, todas as placas apresentaram colônias enquanto as placas de controle negativo permaneceram sem colônias.

Ilustração 7. Fluxograma de construção do vetor episomal pEPI-AL1. 1.- Promotor AOX1 de *P. pastoris*. 2- Promotor T7 do bacteriófago T7. 3- Cauda de 6xHis. 4- Sequência de biotilação. 5- T.T. AOX1 de *P. pastoris*. 6- L1 HPV16 codons otimizados. 7- Sequência 3' do AOX1 de *P. pastoris*. 8- rDNA de *H. polymorpha*. 9- HARS36-JR. 10- PARS1 de *P. pastoris*. 11- Gene HIS4 de *P. pastoris*. 12- Promotor TEF1 de *S. cerevisiae*. 13- Promotor sintético EM7. 14- Gene Sh ble: resistência à Zeocina. 15- T.T. CYC1 de *S. cerevisiae*. 16- Promotor GAPDH61. 17- Gene APH: resistência à Kanamicina e Geneticina. 18- Cole1 ori. 19- pMB1 ori. 20- Gene de resistência à Ampicilina. 21- Gene de resistência à Kanamicina.

Ilustração 8. Mapa do vetor pEPI-AL1. **P. AOX1**- Promotor AOX1 de *P. pastoris*. **L1**- gene L1 HPV16 com codons otimizados para *P. pastoris*. **T.T. AOX1**- Terminador da transcrição AOX1 de *P. pastoris*. **HIS4**- Gene HIS4 de *P. pastoris*. **APH**- Gene APH: resistência à Kanamicina e Geneticina. **PARS1**- PARS1 de *P. pastoris*. **P. TEF1**- Promotor TEF1 de *S. cerevisiae*. **P. EM7**- Promotor sintético EM7. **SH BLE**- Gene Sh ble: resistência à Zeocina. **T.T. CYC1**- Terminador da transcrição CYC1 de *S. cerevisiae*. **ColE1 ori**- origem de replicação ColE1.

Influência da dose de Zeocina na expressão da L1

Células competentes de *P. pastoris* KM71 foram transformadas com o vetor pEPI-AL1 e selecionadas em meio mínimo com histidina e Zeocina. Foram utilizadas duas concentrações diferentes de Zeocina para analisar o efeito da dose do antibiótico na expressão da proteína L1. Os clones selecionados foram armazenados como clones de 1ª geração. Em seguida, esses clones foram submetidos a um ensaio de expressão nas mesmas condições de pressão seletiva em que foram selecionados.

A **Ilustração 9** mostra o SDS-Page e o *Western blot* do primeiro experimento de expressão utilizando o vetor pEPI-AL1. Dos sete clones analisados, seis expressaram a proteína L1.

Ilustração 9. Primeira análise de expressão da proteína L1 no sistema pEPI-AL1. **A.** SDS-PAGE. **B.** Western Blot. **1** Padrão de massa molecular LMW. **2** Controle Positivo – Clone de *P. pastoris* KM71 transformado com o vetor pPICHOLI-1L1, seleção, cultivo e expressão meio de cultura com 100µg/ml de Zeocina, expressão em BMMH. **3** Controle Negativo - clone de *P. pastoris* KM71 sem nenhum vetor e expressão em BMMH. **4-7** Clones de 1ª geração, seleção, cultivo e expressão em meio de cultura com 600µg/ml de Zeocina, expressão em BMMH. **8-10** Clones de 1ª geração, seleção, cultivo e expressão em meio de cultura com 100µg/ml de Zeocina, expressão em BMMH.

Os clones selecionados e cultivados com maiores concentrações de Zeocina apresentaram, de maneira geral e qualitativamente, uma expressão da proteína L1 superior aos clones selecionados apenas com 100µg/ml de Zeocina. Os clones das colunas 4 e 6, que obtiveram os maiores níveis de expressão, foram selecionados para a segunda etapa de experimentos e passaram a se chamar clone A e clone B respectivamente.

Expressão da L1 sob pressões seletivas diferentes

Primeiramente realizou-se um ensaio de estabilidade com os clones A e B. Eles foram cultivados em dois tipos de meio mínimo: MD (sem histidina) e MDH (com histidina). Após isso, os clones foram armazenados e denominados clones de 60ª geração.

Em seguida, foi feito um experimento de expressão da L1 com clones da 1ª geração e da 60ª geração para verificar se a marcação auxotrófica era suficiente para manter a expressão dos clones inicialmente selecionados com altas doses de antibiótico. A **Tabela 5** resume o histórico experimental de cada condição analisada nesse ensaio de expressão.

Seleção Inicial	Clone	Geração	Análise de Estabilidade	Ensaio de Expressão
Meio MDH + 600µg/ml de Zeocina	A	1ª	-	Meio BMMH + 600µg/ml de Zeocina
				Meio BMM
				Meio BMMH
		60ª	Meio MD	Meio BMM
			Meio MDH	Meio BMMH
	B	1ª	-	Meio BMMH + 600µg/ml de Zeocina
				Meio BMM
				Meio BMMH
		60ª	Meio MD	Meio BMM
			Meio MDH	Meio BMMH

Tabela 5. Histórico experimental e condições dos ensaios de expressão dos clones da 60ª geração transformados com o vetor pEPI-AL1.

A **Ilustração 10** mostra o SDS-PAGE e o Western Blot resultantes do experimento de expressão. Pode-se observar que as expressões dos clones de 1ª geração cultivados em meio com 600µg/ml de Zeocina ou meio sem histidina foram similares, mas a expressão da L1 nos clones de 1ª geração cultivados em meio não seletivo diminuiu bastante. Também se pode observar que os clones de 60ª geração expressaram a proteína L1 em quantidades muito inferiores às dos clones de 1ª geração.

Ilustração 10. Ensaio de expressão dos clones A e B sob pressões seletivas diferentes. **A.** SDS-PAGE. **B.** Western Blot. **1** Padrão de massa molecular MW. **2** Controle Positivo – Clone de *P. pastoris* KM71 transformado com o vetor pPICHOLI-1L1, seleção, cultivo e expressão em meio de cultura com 100µg/ml de Zeocina, expressão em BMMH. **3** Clone A, 1ª geração, expressão em BMMH com 600µg/ml de Zeocina. **4** Clone A, 1ª geração, expressão em BMM. **5** Clone A, 1ª geração, expressão em BMMH. **6** Clone A, 60ª geração, cultivado em MD e expressão em BMM. **7** Clone A, 60ª geração, cultivado em MDH e expressão em BMMH. **8** Clone B, 1ª geração, expressão em BMMH com 600µg/ml de Zeocina. **9** Clone B, 1ª geração, expressão em BMM. **10** Clone B, 1ª geração, expressão em BMMH. **11** Clone B, 60ª geração, cultivado em MD e expressão em BMM. **12** Clone B, 60ª geração, cultivado em MDH e expressão em BMMH. **14** Controle Negativo - clone de *P. pastoris* KM71 sem nenhum vetor e expressão em BMMH.

Esses resultados mostraram que os clones de 1ª geração expressam níveis semelhantes da L1 quando cultivados em altas doses de antibiótico ou em meio auxotrófico (coluna 3, 4, 8 e 9). Mas quando cultivados em meio não seletivo (5 e 10), os clones perdem quase toda capacidade de expressar a proteína L1. Já os clones da 60ª geração perdem quase toda a capacidade de expressão da L1 quando cultivados por 60 gerações na ausência de Zeocina, independentemente da presença da pressão seletiva auxotrófica (colunas 6, 7, 11 e 12).

Sistema integrativo pINT-AL1

Construção do vetor pINT-AL1

Os vetores pPICHOLI-1, pPIC3,5K e pGEM-Teasy-rDNA foram utilizados como base para a construção do vetor pINT-A para *P. pastoris*. O vetor pPICHOLI-1 forneceu a origem de replicação bacteriana e o gene de resistência a Zeocina, o vetor pPIC3,5K forneceu o cassete de expressão e o vetor pGEM-Teasy-rDNA forneceu a sequência do rDNA.

O vetor pGEM-Teasy-L1 foi utilizado para fornecer o gene da proteína L1 do HPV16 com códons otimizados para *Pichia pastoris*. O vetor integrativo contendo o gene da proteína L1 foi denominado pINT-AL1. O vetor não possui nenhuma sequência de replicação autônoma para que permaneçam nas células apenas os vetores integrados no genoma.

As **Ilustrações 11 e 12** mostram, respectivamente, o fluxograma de construção e o mapa do vetor pEPI-AL1.

Ilustração 11. Fluxograma de construção do vetor integrativo pINT-AL1. 1- Promotor AOX1 de *P. pastoris*. 2- Promotor T7 Phago. 3- Cauda de 6xHis. 4- Sequência de biotilação. 5- T.T. AOX1 de *P. pastoris*. 6- L1 HPV16 codons otimizados. 7- Sequência 3' do AOX1 de *P. pastoris*. 8- rDNA de *H. polymorpha*. 9- HARS36-JR. 10- PARS1 de *P. pastoris*. 11- Gene HIS4 de *P. pastoris*. 12- Promotor TEF1 de *S. cerevisiae*. 13- Promotor sintético EM7. 14- Gene Sh ble; resistência à Zeocina. 15- T.T. CYC1 de *S. cerevisiae*. 16- Promotor GAPDH61. 17- Gene APH; resistência à Kanamicina e Geneticina. 18- ColE1 ori. 19- pMB1 ori. 20- Gene de resistência à Ampicilina. 21- Gene de resistência à Kanamicina.

Ilustração 12. Mapa do vetor integrativo pINT-AL1. **P. AOX1**- Promotor AOX1 de *P. pastoris*. **L1**- gene L1 HPV16 com codons otimizados para *P. pastoris*. **T.T AOX1**- Terminador da transcrição AOX1 de *P. pastoris*. **rDNA** - fragmento do DNA ribossomal de *H. polymorpha*. **P. TEF1**- Promotor TEF1 de *S. cerevisiae*. **P. EM7**- Promotor sintético EM7. **SH BLE**- Gene Sh ble: resistência à Zeocina. **T.T CYC1**- Terminador da transcrição CYC1 de *S. cerevisiae*. **ColE1 ori** - origem de replicação ColE1. **XhoI**- Sítio de linearização do vetor.

Expressão da L1 a partir de vetores integrativos

Células competentes de *P. pastoris* foram transformadas com o vetor pINT-AL1 e selecionadas em meio suplementado com 1000µg/ml de Zeocina. Após três dias surgiram algumas dezenas de colônias por placa. Dessas colônias, 50 foram selecionados aleatoriamente e repicadas em uma nova placa com 1000µg/ml de Zeocina. As colônias foram denominadas clones de 1ª geração.

A partir desses clones, foi realizado um ensaio de expressão da proteína L1. Todos os clones de 1ª geração foram inoculados em meio YPD com 1000µg/ml de Zeocina, mas somente os 20 clones que mais cresceram em 24 horas foram utilizados no ensaio de expressão. Durante a etapa de indução também foi utilizado meio de cultura com Zeocina. Esse primeiro experimento de expressão foi realizado apenas para verificar se era possível expressar a proteína L1 utilizando um vetor integrativo com uma sequência do DNA ribossomal de *H. polymorpha*.

A **Ilustração 13** mostra o SDS-Page e o Western Blot do primeiro experimento de expressão utilizando o vetor pINT-AL1.

Ilustração 13. Ensaio de expressão com os clones de 1ª geração contendo o vetor integrativo pINT-AL1. Foram analisados 18 clones selecionados, cultivados e induzidos em meio de cultura com 1000µg/ml de Zeocina. **A.** SDS-PAGE. **B.** Western Blot. **1** Padrão de massa molecular MW. **2** Controle Negativo 1 - Clone de *P. pastoris* KM71 sem vetor cultivado em BMMY. **3.** Controle Negativo 2 - Clone de *P. pastoris* KM71 sem vetor cultivado em YPD. **4** Nada. **5.** Controle Positivo - Clone de *P. pastoris* KM71 transformado com o vetor pPICHOLI-1L1, seleção, cultivo e expressão em meio de cultura com 100µg/ml de Zeocina, expressão em BMMY. **6** Nada. **7-15** Clones de 1ª geração, seleção, cultivo e expressão em meio de cultura com 1000µg/ml de Zeocina, expressão em BMMY.

Todos os 18 clones de 1ª geração expressaram a proteína L1. Não foi detectada a expressão da L1 nos controles negativos ou no controle positivo. Provavelmente, o controle positivo perdeu a capacidade de expressão da L1. Os géis de SDS-PAGE corados com Comassie indicam que foram aplicadas quantidades muito grandes de extratos celulares e que ocorreu bastante degradação das proteínas. Isso explica a imagem obtida no experimento de Western Blot. Entretanto, os resultados obtidos não invalidam a ensaio. Os sinais de Western Blot, apesar de distorcidos, são específicos da L1, porque o controle negativo não apresentou nenhum sinal, e porque o padrão dos sinais é similar ao padrão de degradação da L1.

Expressão da L1 na ausência de pressão seletiva

No próximo experimento, os 50 clones de 1ª geração foram submetidos a um ensaio de estabilidade. Depois do cultivo em meio não seletivo por 60 gerações, os clones foram transferidos para um meio de cultura com 1000µg/ml de Zeocina para verificar os clones que permaneciam resistentes ao antibiótico. Todos os clones cresceram em meio sólido com 1000µg/ml de Zeocina. Entretanto, somente 24 clones cresceram quando as colônias foram inoculadas em meio líquido contendo a mesma concentração de antibiótico. Esses 24 clones possuíam colônias bem maiores do que as dos clones que não cresceram. Esses clones resistentes foram armazenados e denominados clones de 60ª geração.

Por último, os clones de 60ª geração foram submetidos a um ensaio de expressão da L1. Todo o ensaio foi feito em meio complexo sem antibiótico para analisar a expressão de células estáveis na ausência de pressão seletiva.

A **Ilustração 14** mostra o SDS-Page e o Western blot do segundo ensaio de expressão, utilizando os clones de 60ª geração.

Ilustração 14. Ensaio de expressão com os clones de 60ª geração contendo o vetor integrativo pINT-AL1. Foram analisados 24 clones selecionados em meio de cultura com 1000µg/ml de Zeocina, cultivados e induzidos na ausência do antibiótico. **A.** SDS-PAGE. **B.** Western Blot. **(+)** Controle Positivo - Clone de *P. pastoris* KM71 transformado com o vetor pPICHOLI-1L1, seleção, cultivo e expressão em meio de cultura com 100µg/ml de Zeocina, expressão em BMMY. **(-)** Controle Negativo - Clone de *P. pastoris* KM71 sem vetor em BMMY. **(60ª geração)** Clones de 60ª geração, expressão em BMMY.

Dos 24 clones avaliados, somente três não expressaram a proteína L1 (Clones: coluna 2 - gel 1; coluna 10 - gel 2; coluna 4 - gel 3). Não foi possível identificar com precisão quais clones possuíam os maiores níveis de expressão devido à degradação parcial da L1. Em função disso, foi contabilizada a soma das intensidades de todas as bandas para selecionar

os clones com o maior nível de expressão. Seis dos clones foram selecionados para um novo ensaio de expressão e para uma quantificação do número de cópias (Clones: coluna 9 – gel 1; colunas 2, 4, 8 e 9 – gel 2; coluna 2 – gel 3).

Quantificação do número de cópias

Os seis clones citados acima tiveram seus DNAs genômicos extraídos para quantificar o número de cópias do gene da L1 integradas no genoma através do método de qPCR. Paralelamente o ensaio de expressão foi repetido.

A **Tabela 6** mostra os resultados do experimento de quantificação do número de cópias do gene L1 integradas no genoma dos clones. Nenhum dos amplicons foi amplificado na amostra de água e somente o amplicon do gene HIS4 foi amplificado na amostra de DNA genômico da levedura selvagem (KM71). A diluição 2^{12} da amostra do vetor pEPI-A foi utilizada como Calibrador, porque obteve o valor de CP do gene HIS4 mais próximo aos dos valores obtidos nas amostras dos clones. Nesse experimento, a eficiência média de amplificação do gene L1 foi de 1,986 e a do gene HIS4 foi de 1,947. Considerando que a eficiência de amplificação dos dois genes foram muito semelhante entre si e próxima de 2, pode-se assumir, segundo a PE Applied Biosystems, um valor de eficiência igual a 2,0 para todas as reações do experimento. Nesse caso o modelo matemático pode ser simplificado para:

$$\text{ratio} = \frac{(E_{\text{ref}})^{CP_{\text{Sample}}}}{(E_{\text{target}})^{CP_{\text{Sample}}}} \div \frac{(E_{\text{ref}})^{CP_{\text{Calibrator}}}}{(E_{\text{target}})^{CP_{\text{Calibrator}}}} = \frac{(2)^{CP_{\text{Sample}}}}{(2)^{CP_{\text{Sample}}}} \div \frac{(2)^{CP_{\text{Calibrator}}}}{(2)^{CP_{\text{Calibrator}}}}$$

$$\text{ratio} = 2^{-[\Delta CP_{\text{sample}} - \Delta CP_{\text{Calibrator}}]} = 2^{-\Delta \Delta CP}$$

Utilizando essa equação, foram obtidos os números de cópias da proteína L1 integradas no genoma de cada clone.

Quantificação do Número de Cópias - qPCR								
Amostra		CP _s - L1			CP _s - HIS4			ΔCP _a (HIS4/L1)
Clone	gDNA	MÉDIA	DESPAD	CV (%)	MÉDIA	DESPAD	CV (%)	
KM71	25 ng	24.93	0.42	1.70	17.23	0.02	0.11	-7.69
1	25 ng	12.93	0.14	1.08	17.78	0.18	1.04	4.85
2	25 ng	11.83	0.07	0.62	17.60	0.03	0.18	5.77
3	25 ng	12.08	0.29	2.36	17.78	0.07	0.39	5.70
4	25 ng	10.61	0.09	0.81	17.21	0.04	0.22	6.60
5	25 ng	13.05	0.14	1.04	18.28	0.10	0.53	5.23
6	25 ng	13.26	0.10	0.76	18.14	0.11	0.59	4.88
Amostra de pEPI-A		CP _c - L1			CP _c - HIS4			ΔCP _c (HIS4/L1)
Diluição		MÉDIA	DESPAD	CV (%)	MÉDIA	DESPAD	CV (%)	
2 ⁴		10.76	0.00	0.01	11.06	0.02	0.18	0.29
2 ⁶		12.74	0.05	0.42	13.05	0.01	0.04	0.31
2 ⁸		13.34	0.02	0.14	13.74	0.11	0.77	0.40
2 ¹⁰		15.36	0.05	0.35	15.84	0.13	0.84	0.48
2 ¹² (Calibrador)		17.54	0.11	0.60	17.83	0.15	0.85	0.29
2 ¹⁴		19.64	0.03	0.16	19.94	0.12	0.62	0.30
2 ¹⁶		21.55	0.06	0.28	21.97	0.18	0.82	0.41
2 ¹⁸		23.75	0.18	0.75	23.67	0.10	0.42	-0.08
2 ²⁰		25.79	0.10	0.37	25.52	0.00	0.01	-0.27
Clones		ΔΔCP (ΔCP _s /ΔCP _c)		Número de Cópias				
KM71		-7.99		0				
1		4.56		24				
2		5.48		45				
3		5.41		43				
4		6.31		79				
5		4.94		31				
6		4.59		24				

Tabela 6. Resultado do experimento de qPCR para quantificação do número de cópias do gene L1 em clones da 60ª geração transformados com o vetor pINT-AL1. Foram analisados os 6 clones de 60ª geração e um clone de KM71 selvagem. Como controle interno da reação, foi utilizado o gene HIS4, presente em apenas uma cópia no genoma da *P. pastoris*. Como controle externo (Calibrador), foi utilizada uma amostra diluída de vetor pEPI-A, que possui em sua estrutura uma cópia do gene L1 e do gene HIS4. O clone de KM71 selvagem foi utilizado como controle positivo da amplificação do gene HIS4 e controle negativo da amplificação do gene L1. Os resultados foram obtidos a partir de um experimento em triplicata.

Em seguida foi feita uma comparação entre os níveis de expressão dos seis clones e o número de cópias do gene L1 integradas ao genoma. A **Ilustração 15** mostra a comparação entre os níveis de expressão dos seis clones, analisados por Western Blot, e o número de cópias do gene da L1, obtido através do experimento anterior. Não foi possível estabelecer uma relação de proporcionalidade entre a dosagem gênica e o nível de expressão nesses seis clones.

Ilustração 15. Comparação entre expressão da L1 e dosagem gênica. Neste experimento foram analisados seis clones de 60ª geração transformados com o vetor pINT-AL1. A figura mostra o Western Blot do ensaio de expressão e o número de cópias do gene L1 integradas no genoma de cada clone. Um clone de KM71 selvagem foi utilizado como controle negativo.

Localização do sítio de integração

Neste experimento, os seis clones da 60ª geração tiveram seus DNAs genômicos analisados por PCR. Foram realizados dois PCRs para amplificar as regiões do DNA ribossomal de *P. pastoris* onde possivelmente ocorre a integração dos vetores. Caso tivesse ocorrido a integração dos vetores no interior das regiões amplificadas, os produtos de PCR teriam um tamanho igual ou superior ao tamanho da região amplificada mais o tamanho do vetor de expressão.

A **Ilustração 16** mostra as eletroforeses em géis de agarose dos produtos de PCR. Pode-se observar que os clones apresentaram um padrão de produtos diferente do padrão do clone KM71 selvagem. Os clones das colunas 7 e 10 apresentaram as maiores diferenças em ambos os ensaios de PCR (A e B). O tamanho da região intergênica 18S-25S é de 3360bp, e o tamanho da região 5S é de 100 bp.

Ilustração 16. Resultados do experimento de amplificação das regiões do rDNA de *P. pastoris*. Foram analisados seis clones da 60ª geração transformados com o vetor pINT-AL1. **A** Eletroforese dos produtos de PCR da amplificação da região intergênica entre o DNA ribossomal 18S e 25S (tamanho padrão= 3360bp); gel de ágarose 0,5%. **B** Eletroforese dos produtos de PCR da região 5S do DNA ribossomal (tamanho padrão = 100bp); gel de ágarose 3%. **C** Ilustração do padrão de tamanho molecular 1kb Plus, que foi usado nas duas eletroforeses. **Coluna 1** Padrão de tamanho molecular **Coluna 3** Controle Negativo - *Pichia* KM71. **Colunas 5-10** Seis clones da 60ª geração transformados com o pINT-AL1.

Quantificação da Expressão da L1

Por último, dois dos seis clones da 60ª geração foram submetidos a um ensaio de expressão da proteína L1. Além deles, dois clones de KM71 transformados com o vetor pPICHOLI-1L1 foram submetidos a um ensaio de expressão na presença de 100µg/ml de Zeocina. As culturas foram induzidas por 48 horas e após isso, seus extratos celulares foram analisados por Western Blot. Foram aplicadas amostras de 2 µl e 5 µl dos extratos celulares diluídos e também 10 ng e 20 ng de proteína L1 purificada.

A **Ilustração 17** mostra o resultado do experimento de Western Blot. As leituras das intensidades das bandas formadas foram realizadas pelo instrumento Eagle Eye II da Stratagene. Os resultados indicam uma expressão de aproximadamente 18,6 mg/L para o clone A-PICHOLI (coluna 1 e 2) e 17,8 mg/L para o clone B-PICHOLI (coluna 3 e 4). Os clones A-PICHOLI e B-PICHOLI possuem o plasmídeo episomal pPICHOLI-1L1 e foram selecionados e induzidos na presença de Zeocina. Os dois clones da 60ª geração do experimento anterior

(Ilustração 15, colunas 2 e 3) tiveram uma expressão de aproximadamente 20,4 mg/L e 29,7 mg/L respectivamente.

Ilustração 17. Quantificação da expressão da proteína L1 por Western Blot. **1 e 2** Clone A-PICHOLI, transformado com pPICHOLI-1L1, selecionado e induzido na presença de 100 µg/ml de Zeocina. Amostras de 2µl e 5µl de extrato diluído, expressão de 18,6 mg/L da L1. **3 e 4** Clone B-PICHOLI, transformado com pPICHOLI-1L1, selecionado e induzido na presença de 100 µg/ml de Zeocina. Amostras de 2µl e 5µl de extrato diluído, expressão de 17,8 mg/L da L1. **5 e 6** Aliquotas da L1 purificada. Amostras de 20ng e 10ng. **7 e 8** Clone 1 (coluna 2 da ilustração 15), expressão na ausência de Zeocina, amostras de 2µl e 5µl de extrato diluído, expressão de 20,4 mg/L da L1. **9 e 10** Clone 2 (coluna 3 da ilustração 15), expressão na ausência de Zeocina, amostras de 2µl e 5µl de extrato diluído, expressão de 29,7 mg/L da L1.

Sistema Integrativo pINT-BL1

Construção do vetor pINT-BL1

Os vetores pINT-A e pGEM-Teasy-G61APH foram utilizados como base para a construção do vetor pINT-B para *P. pastoris*. O vetor pGEM-Teasy-G61APH forneceu o gene de resistência a Geneticina/Kanamicina controlado pelo promotor fraco GAPDH61 e o vetor pINT-A forneceu o cassete de expressão, a sequência do rDNA ribossomal e a origem de replicação ColE1. O vetor pGEM-Teasy-L1 foi utilizado para fornecer o gene da proteína L1 do HPV16 com códons otimizados para *Pichia pastoris*. O vetor integrativo contendo o gene da proteína L1 foi denominado pINT-BL1. O vetor não possui nenhuma sequência de replicação autônoma para que permaneçam nas células apenas os vetores integrados no genoma.

As **Ilustrações 18 e 19** mostram, respectivamente, o fluxograma de construção e o mapa do vetor pINT-BL1.

Ilustração 18. Fluxograma de construção do vetor integrativo pINT-BL1. 1- Promotor AOX1 de *P. pastoris*. 2- Promotor do bacteriófago T7. 3- Cauda de 6xHis. 4- Sequência de biotilação. 5- T.T. AOX1 de *P. pastoris*. 6- L1 HPV16 codons otimizados. 7- Sequência 3' do AOX1 de *P. pastoris*. 8- rDNA de *H. polymorpha*. 9- HARS36-JR. 10- PARS1 de *P. pastoris*. 11- Gene HIS4 de *P. pastoris*. 12- Promotor TEF1 de *S. cerevisiae*. 13- Promotor sintético EM7. 14- Gene Sh ble: resistência à Zeocina. 15- T.T. CYC1 de *S. cerevisiae*. 16- Promotor GAPDH61. 17- Gene APH: resistência à Kanamicina e Geneticona. 18- ColE1 ori. 19- pMB1 ori. 20- Gene de resistência à Ampicilina. 21- Gene de resistência à Kanamicina.

Ilustração 19. Mapa do vetor integrativo pINT-BL1. **P. AOX1**- Promotor AOX1 de *P. pastoris*. **L1**- gene L1 HPV16 com codons otimizados para *P. pastoris*. **T.T. AOX1**- Terminador da transcrição AOX1 de *P. pastoris*. **rDNA** – DNA ribossomal de *H. polymorpha*. **APH**- Gene APH: resistência à Geneticina/Kanamicina. **GAPDH61**- Fragmento do promotor GAPDH de *H. polymorpha*. **ColE1 ori** – origem de replicação ColE1. **NaeI** – Sítio de linearização do vetor.

Expressão da L1 nos clones de 1ª geração

Células competentes de *P. pastoris* foram transformadas com o vetor pINT-BL1 e selecionadas em meio suplementado com 500µg/ml e 4000µg/ml de Geneticina. Após três dias, surgiram colônias apenas na placa com 500µg/ml de Geneticina. As 25 colônias geradas foram inoculadas em meio YPD líquido com 500µg/ml e 4000µg/ml de Geneticina. Dos 25 clones, 17 cresceram: 14 resistentes a 4000µg/ml e três resistentes a 1000µg/ml. Os clones foram armazenados e denominados clones de 1ª geração.

A partir desses clones foi realizado um ensaio de expressão da proteína L1. A **Ilustração 20** mostra o SDS-Page e o Western Blot do primeiro experimento de expressão utilizando o vetor pINT-BL1. Neste experimento foram analisadas as expressões dos 17 clones de 1ª geração. O ensaio foi feito em meio complexo suplementado com as mesmas doses de antibiótico utilizadas na seleção.

Ilustração 20. Ensaio de expressão com os clones de 1ª geração contendo o vetor integrativo pINT-BL1. Foram analisados 17 clones selecionados, cultivados e induzidos em meio de cultura com altas doses de Geneticina. **A.** SDS-PAGE. **B.** Western Blot. **1.** Padrão de massa molecular MW. **2.** Controle Negativo - Clone de *P. pastoris* KM71 sem vetor cultivado em BMMY. **3-16.** Clones de 1ª geração, seleção com 500 µg/ml de Geneticina, cultivo e expressão com 4000 µg/ml de Geneticina. **17-19.** Clones de 1ª geração, seleção com 500 µg/ml de Geneticina, cultivo e expressão com 1000 µg/ml de Geneticina.

Todos os clones, menos um (coluna 11), expressaram a proteína L1. Não existe diferença visual no Western Blot entre a expressão dos clones resistentes a 1000µ/ml de Geneticina (colunas 17, 18 e 19) e a dos clones resistentes a 4000µ/ml de Geneticina.

Expressão da L1 nos clones de 60ª geração

Primeiramente foi realizada uma análise de estabilidade dos 17 clones analisados no primeiro ensaio de expressão. Os clones foram cultivados em meio YPD sólido sem

Geneticina por 60 gerações. Após todas as passagens, as colônias foram inoculadas em meio YPD líquido suplementado com 1000µg/ml e 4000µg/ml de Geneticina e cultivadas por 48 horas.

Nenhum clone manteve a resistência às concentrações anteriores de antibiótico. Onze dos clones cresceram no meio de cultura com 1000µg/ml Geneticina e os demais perderam a resistência à Geneticina (entre eles, os três clones selecionados em 1000µg/ml Geneticina). Os clones resistentes foram estocados e denominados clones de 60ª geração.

Por último, os clones de 60ª geração foram submetidos a um ensaio de expressão da L1. O ensaio foi feito em meio complexo sem antibiótico para analisar a expressão da L1 em células cultivadas por longas gerações em meio de cultura sem pressão seletiva.

A **Ilustração 21** mostra o SDS-Page e o Western blot do segundo ensaio de expressão. Utilizando os clones de 60ª geração. Os clones foram cultivados em meio complexo sem Geneticina. Dos 11 clones avaliados, 9 expressaram a proteína L1.

Ilustração 21. Ensaio de expressão com os clones de 60ª geração contendo o vetor integrativo pINT-BL1. Foram analisados 11 clones, selecionados em altas doses de Geneticina, mas cultivados e induzidos em meio de cultura sem Geneticina. **A.** SDS-PAGE. **B.** Western Blot. **1.** Padrão de massa molecular MW. **2 e 15.** Controle Negativo - Clone de *P. pastoris* KM71 sem vetor cultivado em BMMY. **3-13.** Clones da 60ª geração, cultivo e expressão sem Geneticina.

Vetores pINT-CL1 e pINT-DL1

Os vetores foram construídos para aumentar as opções de desenvolvimento de um sistema de expressão em larga escala. Futuramente os vetores serão testados em *P. pastoris* ou em *H. polymorpha*.

Construção dos vetores pINT-CL1 e pINT-DL1

Os vetores pINT-A e pINT-B foram utilizados como base para a construção dos vetores pINT-C e pINT-D, respectivamente. A única diferença desses dois vetores em relação aos seus vetores base reside na sequência de integração ao genoma. Os vetores pINT-C e pINT-D possuem o fragmento da sequência HARS36 no lugar da sequência do DNA ribossomal. A sequência HARS36 foi fornecida pelo vetor pHARS36JR. A diferença entre os vetores pINT-C e pINT-D está no marcador de seleção, o vetor pINT-C confere resistência à Zeocina, e o vetor pINT-D à Geneticina.

O vetor pGEM-Teasy-L1 foi utilizado para fornecer o gene da proteína L1 do HPV16 com códons otimizados para *Pichia pastoris*. Os vetores integrativos contendo a proteína L1 foram denominados pINT-CL1 e pINT-DL1.

As **Ilustrações 22, 23 e 24** mostram o fluxograma de construção e o mapa dos vetores pINT-CL1 e pINT-DL1.

Ilustração 22. Fluxograma de construção dos vetores pINT-CL1 e pINT-DL1. Primeira etapa. **1.** Promotor AOX1 de *P. pastoris*. **2.** Promotor do bacteriófago T7. **3.** Cauda de 6xHis. **4.** Sequência de biotilação. **5.** T.T. AOX1 de *P. pastoris*. **6.** L1 HPV16 codons otimizados. **7.** Sequência 3' do AOX1 de *P. pastoris*. **8.** rDNA de *H. polymorpha*. **9.** HARS36-JR. **10.** PARS1 de *P. pastoris*. **11.** Gene HIS4 de *P. pastoris*. **12.** Promotor TEF1 de *S. cerevisiae*. **13.** Promotor sintético EM7. **14.** Gene Sh ble: resistência à Zeocina. **15.** T.T. CYC1 de *S. cerevisiae*. **16.** Promotor GAPDH61. **17.** Gene APH: resistência à Kanamicina e Geneticina. **18.** ColE1 ori. **19.** pMB1 ori. **20.** Gene de resistência à Ampicilina. **21.** Gene de resistência à Kanamicina.

Ilustração 23 Fluxograma de construção dos vetores pINT-CL1 e pINT-DL1. Segunda etapa. 1- Promotor AOX1 de *P. pastoris*. 2- Promotor do bacteriófago T7. 3- Cauda de 6xHis. 4- Sequência de biotilação. 5- T.T. AOX1 de *P. pastoris*. 6- L1 HPV16 codons otimizados. 7- Sequência 3' do AOX1 de *P. pastoris*. 8- rDNA de *H. polymorpha*. 9- HARS36-JR. 10- PARS1 de *P. pastoris*. 11- Gene HIS4 de *P. pastoris*. 12- Promotor TEF1 de *S. cerevisiae*. 13- Promotor sintético EM7. 14- Gene Sh ble: resistência à Zeocina. 15- T.T. CYC1 de *S. cerevisiae*. 16- Promotor GAPDH61. 17- Gene APH: resistência à Kanamicina e Geneticina. 18- ColE1 ori. 19- pMB1 ori. 20- Gene de resistência à Ampicilina. 21- Gene de resistência à Kanamicina.

Ilustração 24. Mapa dos vetores pINT-CL1 e pINT-DL1. **P. AOX1**- Promotor AOX1 de *P. pastoris*. **L1**- gene L1 HPV16 com codons otimizados para *P. pastoris*. **T.T. AOX1**- Terminador da transcrição AOX1 de *P. pastoris*. **HARS36JR** – Fragmento da sequência HARS36 de *H. polymorpha*. **P. TEF1**- Promotor TEF1 de *S. cerevisiae*. **P. EM7**- Promotor sintético EM7. **SH BLE**- Gene Sh ble: resistência à Zeocina. **T.T. CYC1**- Terminador da transcrição CYC1 de *S. cerevisiae*. **APH**- Gene APH: resistência à Geneticina/Kanamicina. **GAPDH61**- Fragmento do promotor GAPDH de *H. polymorpha*. **ColE1 ori** – origem de replicação ColE1. **SalI**. Sítio de linearização dos vetores.

DISCUSSÃO

Em experimentos anteriores, o laboratório tentou expressar a proteína L1 selvagem do HPV16 em *Hansenula polymorpha* e em *Pichia pastoris*, mas nenhum sucesso foi obtido. A expressão só ocorreu quando foi utilizado um vetor de expressão episomal (pPICHOLI-1) e um gene L1 sintético com códons otimizados para a *P. pastoris* (Bazan et al., 2009). Esses resultados garantiram que a expressão da proteína L1 em *P. pastoris* era viável, mas ainda eram necessárias melhorias no sistema para a produção em larga escala.

Sistema de expressão episomal

Neste projeto, foi construído um vetor episomal contendo um marcador auxotrófico com o objetivo de eliminar o uso de antibióticos na expressão em larga escala, entretanto a pressão seletiva exercida pelo meio de cultura sem histidina é baixa e uma única cópia do gene HIS4 é suficiente para restabelecer as funções normais da célula. Isso significa que é necessário fazer uma triagem de centenas de clones para achar alguns com muitas cópias. Para contornar esse problema, primeiro foi feita uma seleção em altas doses de antibióticos.

Nos experimentos, células de *P. pastoris* KM71 foram transformadas com o vetor pEPI-A-L1 e selecionadas em meio mínimo com histidina e Zeocina. Para observar a influência da concentração de antibiótico na expressão da proteína L1, as células foram selecionadas e submetidas a ensaios de expressão em baixa e alta dose de antibiótico. Os resultados mostraram que a expressão da L1 é dependente da pressão seletiva e, conseqüentemente, dependente do número de cópias do plasmídeo. Esse fenômeno confirma a hipótese inicial de que a expressão intracelular da proteína L1, assim como a de muitas das proteínas heterólogas intracelulares, é dependente do número de cópias do gene.

Em um segundo momento, realizou-se uma análise de estabilidade com os dois clones que mais expressaram a proteína L1. Os clones foram cultivados em dois tipos de meio mínimo por 60 gerações: MD (com histidina; sem pressão seletiva) e MDH (sem

histidina; com pressão seletiva). Após isso, foi feito um ensaio de expressão da L1 com clones da 1ª geração em meio MD, MDH/Zeocina e MDH, e clones da 60ª geração em meio MD e MDH. O objetivo foi verificar se a marcação auxotrófica era suficiente para manter a expressão dos clones inicialmente selecionados com altas doses de antibiótico.

O cultivo por diversas gerações em meio de seleção auxotrófico ou em meio não seletivo fez com que os clones selecionados inicialmente com altas doses de antibiótico perdessem quase toda a expressão da proteína L1. Isso ocorre porque a pressão seletiva do marcador auxotrófico é muito menor do que a pressão de grandes doses de antibiótico, resultando na perda dos plasmídeos a um nível basal necessário à síntese de histidina, reduzindo, assim, a expressão da proteína heteróloga. Esses resultados tornam esse sistema de expressão inviável para a produção em larga escala.

Vetores integrativos com o rDNA de *H. polymorpha*

Os vetores integrativo pINT-A e pINT-B utilizam o fragmento do rDNA de *Hansenula polymorpha* (Klabunde et al., 2002). Essa sequência de DNA permite a integração de múltiplas cópias de vetores no genoma de diversas leveduras (Klabunde et al., 2003; Steinborn et al 2005, 2006). Por isso, esse projeto se propôs a avaliar o uso desta sequência em *P. pastoris*.

A ausência de uma sequência de replicação autônoma, o período prolongado de cultivo em meio não seletivo e a seleção em concentrações de antibióticos bem maiores do que as utilizadas para a seleção convencional sugerem que os vetores contendo o rDNA de *Hansenula polymorpha* se integra de maneira estável ao genoma da *Pichia pastoris*.

Esse fenômeno foi confirmado pelo experimento de PCR em tempo real, que quantificou o número de cópias do gene L1 integradas no genoma dos clones. Mesmo depois do cultivo prolongado na ausência de pressão seletiva, foi possível identificar clones resistentes a altas doses de antibiótico e que possuíam em média mais de 30 cópias do gene L1.

Apesar dos fatos comprovarem a integração, eles não confirmam o local em que ela ocorreu. Por esse motivo foi feito um ensaio de PCR para tentar verificar se as integrações tinham acontecido nas regiões do DNA ribossomal.

Foram realizados dois experimentos de PCR para amplificar as regiões do DNA ribossomal de *P. pastoris* onde possivelmente ocorreria a integração dos vetores. Os resultados dos experimentos mostraram que a maioria dos clones possuía um padrão de produtos diferente do padrão de produtos do organismo selvagem. Dois clones (Ilustração 16; gel A; colunas 7 e 10) apresentaram um produto de aproximadamente 11 kb, equivalente ao tamanho de um amplicon contendo uma cópia do vetor integrativo. Esses resultados indicam que ocorreu alguma alteração nas sequências do DNA ribossomal dos clones analisados. Entretanto, devido às limitações da metodologia, ainda será necessário fazer um experimento de Southern Blot com os DNAs genômicos dos clones para confirmar a integração no locus do rDNA.

Sistema pINT-AL1

O objetivo central do primeiro ensaio de expressão com o vetor pINT-AL1 foi constatar se o vetor, uma vez integrado no genoma de *P. pastoris*, era capaz de induzir a expressão da proteína L1. Por isso, a pressão seletiva foi mantida para que os clones não perdessem as cópias integradas evitando a possibilidade de um resultado negativo causado pela perda dos vetores. Todos os clones analisados expressaram a proteína L1.

Os ensaios de estabilidade mostraram que 24 dos 50 clones mantiveram a resistência a 1000µg/ml de Zeocina mesmo após o cultivo por 60 gerações em meio não seletivo.

Em um segundo experimento, os clones da 60ª geração resistentes foram submetidos a ensaios de expressão na ausência de antibiótico. Dos 24 clones avaliados, 20 expressaram a proteína L1. Esses resultados sugerem que o sistema criado com o vetor pINT-AL1 é estável.

Não foi possível comparar com precisão o nível de expressão dos clones porque ocorreu muita degradação da proteína L1, mas há um indicativo visual de que os seis clones selecionados possuem um nível de expressão semelhante ou superior ao clone utilizado atualmente pelo laboratório. O experimento de quantificação da expressão realizado com amostras de dois deles sugere uma produtividade da ordem de 20 a 30mg/L de proteína.

Não foi observada uma relação de proporcionalidade entre o nível de expressão da proteína L1 e o número de cópias do gene. Uma possibilidade para explicar o observado é que nem todas as cópias integradas permanecem necessariamente funcionais, e quanto

maior o número de cópias integradas, maior a variação da proporção de cópias funcionais e não funcionais entre os diferentes clones.

Outro aspecto relevante é que o número de cópias pode atingir um limite de eficiência, e após esse limite o aumento do número de cópias não causa mais mudanças no nível de expressão, podendo até causar um prejuízo na expressão da proteína recombinante. Isso ocorre devido a um fenômeno denominado fardo metabólico (Macauley-Patrick S, et al., 2005, Hohenblum H et al., 2004, Mattanovich et. al., 2004, Gatzke R et al., 1995). Nessas condições o gene recombinante já recrutou toda a maquinaria de transcrição disponível ou o nível de tradução é tão alto que ocorre um decréscimo na expressão das outras proteínas celulares, fundamentais para a sobrevivência da célula.

Como todos os seis clones possuíam um número muito grande de cópias a influência da dosagem gênica pode ter sido mascarada pelos fatores descritos acima.

Sistema pINT-BL1

O gene APH de resistência à Geneticina/Kanamicina controlado por um fragmento do promotor do gene gliceraldeído-3-fosfato desidrogenase (GAPDH61) de *H. polymorpha* foi usado como marcador de seleção do vetor pINT-BL1. O fragmento GAPDH61 funciona como um promotor fraco e faz com que seja necessário um maior número de cópias do gene para conferir o mesmo nível de resistência ao antibiótico. Até então, o uso desse marcador só tinha sido descrito em *H. polymorpha*.

No experimento de seleção de células transformadas com o vetor pINT-BL1 cresceram somente 25 colônias nas placas com 500 µg/ml de Geneticina. Essas colônias foram inoculadas em meio líquido com 1000 µg/ml e 4000 µg/ml de Geneticina. 17 clones cresceram: 14 resistentes a 4000µg/ml e três resistentes a 1000µg/ml. Esse resultado comprova a eficiência desse marcador de seleção em *P. pastoris*. Dos 17 clones, 16 expressaram a proteína L1.

Após o ensaio de estabilidade nenhum dos clones manteve-se resistente às mesmas concentrações de antibiótico. Restaram 11 clones que mantiveram uma resistência a 1mg/ml de Geneticina. Dos 11 clones, nove expressaram a proteína L1.

A Geneticina é um antibiótico aminoglicosídeo e age diretamente na subunidade 80S dos ribossomos e bloqueia a elongação das proteínas e não causa nenhum dano direto ao material genético.

Neste projeto, os clones que possuíam o vetor pINT-BL1 apresentaram uma instabilidade mitótica maior do que os que possuíam o vetor pINT-AL1. Poucos clones transformantes foram gerados, um número ainda menor de clones resistiu aos ensaios de estabilidade e os clones que restaram apresentaram uma resistência inferior em relação aos clones de 1ª geração.

Considerando que a única diferença estrutural desses dois vetores é o próprio marcador de seleção, e que o vetor pINT-BL1 é menor do que o pINT-AL1, existem duas hipóteses para explicar o fenômeno observado: o promotor GAPDH61 é pouco eficiente em *P. pastoris* ou os sítios de linearização dos vetores (XhoI no pINT-AL1, e NaeI no pINT-BL1) alteram a probabilidade de integração e a estabilidade mitótica dos vetores no genoma de *P. pastoris*.

Adicionalmente, o próprio mecanismo de ação da Geneticina pode representar uma desvantagem no processo inicial de seleção. Como a Geneticina inibe a síntese de proteínas, as células recém transformadas não têm tempo suficiente para sintetizar a enzima que degradará o antibiótico. Deve ser por este motivo que a Geneticina, na maioria das vezes, só é utilizada como marcador de seleção secundário (Higgins, Cregg, 1998).

CONCLUSÕES

O cultivo por diversas gerações em meio de seleção auxotrófico ou em meio não seletivo fez com que os clones transformados com o vetor pEPI-AL1 e selecionados com altas doses de antibiótico perdessem quase toda a expressão da proteína L1. Esse fenômeno torna esse sistema de expressão inviável para a produção em larga escala.

Os ensaios realizados com o vetor p-INT-AL1 garantem que é possível obter clones estáveis produtores da proteína L1 pela integração dos vetores no genoma. O protocolo de obtenção dos clones se mostrou eficiente para selecionar células contendo múltiplas cópias do gene L1 integradas ao genoma. Os clones verificados possuíam em média mais de 25 cópias do gene L1.

O sistema pINT-AL1 desenvolvido nesse projeto é superior ao sistema anterior do laboratório, que utilizava o vetor pPICHOLI-1. Com o rastreamento de poucas dezenas de clones, foi possível selecionar células com mais de 50 cópias do gene L1 e com uma produção de aproximadamente 30mg/L da proteína L1.

O sistema pINT-BL1 gerou poucos clones e apresentou menor estabilidade mitótica em relação ao clone pINT-AL1, mas nível de expressão e o número de cópias dos clones precisam ser quantificados para poder determinar qual dos dois sistemas é mais eficiente.

Os vetores pINT-CL1 e pINT-DL1 foram construídos e devem ser avaliados na leveduras *P. pastoris*.

Existe ainda a possibilidade de se criar um processo de cotransformação de vetores contendo sequências de integração e marcadores de seleção diferentes na tentativa de gerar clones com um número ainda maior de cópias do gene L1.

Com o sucesso deste projeto, o sistema de expressão desenvolvido poderá ser utilizado não só para a produção industrial da proteína L1 do HPV16 como também para expressar proteínas de capsídeos de outros vírus, inclusive proteínas L1 de outros tipos de HPV, além de outras proteínas de interesse biotecnológico.

Por fim, a capacidade de se auto-agregar em partículas virais da proteína L1 abre a possibilidade de utilizar essas VLPs como *Delivery Systems* de genes e antígenos, que podem ser empregados na construção de vacinas polivalentes.

REFERÊNCIAS

Bazan BS, Chaves AAM, Aires LAJ, Cianciarullo AM, Garcea RL, Ho PL. Expression and characterization of HPV-16 L1 capsid protein in *Pichia pastoris*. Archives of Virology, 2009; vol. 154, p 1609–1617.

Brown DR, Schroeder JM, Bryan JT, Stoler MH, Fife KH. Detection of multiple human papillomavirus types in Condylomata acuminata lesions from otherwise healthy and immunosuppressed patients. Journal of Clinical Microbiology. 1999; 37(10):3316–3322.

Buck CB, Cheng N, Thompson CD, Lowy DR, Steven AC, Schiller JT, Trus BL. Arrangement of L2 within the Papillomavirus Capsid. Journal of Virology, June 2008, p. 5190–5197.

Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiology Review; 2000.vol. 24, p. 45–66.

Cereghino GPL, Cereghino JL, Ilgen C, Cregg JM. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Current Opinion in Biotechnology; 2002. vol. 13, p. 329–332.

Chen XS, Garcea RL, Goldberg I, Casini G, Harrison SC. Structure of Small Virus-like Particles Assembled from the L1 Protein of Human Papillomavirus 16. Molecular Cell, 2000. Vol. 5, 557–567.

Chen XS, Casini G, Harrison SC, Garcea RL. Papillomavirus capsid protein expression in *Escherichia coli*: purification and assembly of HPV11 and HPV16 L1. Journal of Molecular Biology; 2001.v. 307, p. 173–182.

Clare JJ, Rayment FB, Ballantine SP, Sreekrishna K, Romanos MA. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. Biotechnology (N Y). 1991 May;9(5):455–60.

Clifford GM, Smith JS, Plummer M, et al. Human papillomavirus types in invasive cervical cancer worldwide: a metaanalysis. Br J Cancer 2003; 88: 63–73.

Clifford GM, Smith JS, Aguado T, et al. Comparison of HPV type distribution in high-grade cervical lesions and cervical cancer: a meta-analysis. *Br J Cancer* 2003; 89: 101–105.

Cregg, JM, Barringer, KJ, Hessler, AY, e Madden, KR. *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol* 1985. 5, 3376-3385.

Mattanovich D, Gasser B, Hohenblum H, Sauer M. Stress in recombinant protein producing yeasts. *Journal of Biotechnology*, 2004, vol. 113, pp. 121–135.

Doorbar J. Molecular biology of human papillomavirus infection and cervical cancer *Clinical Science*; 2006. v. 110, p. 525–541.

Gatzke R, Weydemann U, Janowicz ZA, Hollenberg CP: Stable multicopy integration of vector sequences in *Hansenula polymorpha*. *Appl Microbial Biotechnol* 1995, 43:844-849.

Gellissen G, Hollenberg CP. Application of yeasts in gene expression studies: A comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* - a review. *Gene*; 1997. v.190 (1): p. 87-97.

Gellissen, G. Heterologous protein production in methylotrophic yeasts. *Applied Microbiology and Biotechnology*. 2000. Vol. 54 (6), p. 741-750.

Gellissen G, Kunze G, Gaillardin C, Cregg JM, Berardi E, Veenhuis M, van der Klei I. New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica* - A comparison. *Fems Yeast Research*; 2005. v. 5 (11): p. 1079-1096.

Gellissen G. *Production of Recombinant Proteins – Novel Microbial and Eukaryotic Expression Systems*. Wiley- VCH, 2005.

Gietz RD, Woods RA. Genetic transformation of yeast. *Biotechniques*; 2001. v. 30(4): p. 816-20, 822-6, 828.

Harford N, Cabezon T, Colau B, Delisse AM, Rutgers T, de Wilde M. Construction and characterization of a *Saccharomyces cerevisiae* strain (RIT4376) expressing hepatitis B surface antigen. *Postgrad Med Journal*; 1987.vol. 63: p 65-70.

Harper DM, for the GlaxoSmithKline HPV Vaccine Study Group. Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. *Lancet*; 2004. v. 364: p. 1757–1765.

Higgins DR., Cregg JM. *Methods in molecular biology*; 1998. Vol. 103: *Pichia* Protocols. Humana Press; Totowa, New Jersey.

Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol Bioeng*, 2004, vol 85(4): 367–375.

Hollenberg CP, Gellissen G. Production of recombinant proteins by methylotrophic yeasts. *Current opinion in biotechnology*; 1997.vol. 8 (5): 554-560.

Human Papillomaviruses/IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. IARC Monographs on the evaluation of carcinogenic risk to humans. Vol. 90: Lyon, France, 2005.

ICTVdB Management (2006). 00.099.0.02. Alphapapillomavirus. In: ICTVdB - The Universal Virus Database, version 4. Büchen-Osmond, C. (Ed), Columbia University, New York, USA.

Klabunde J, Diesel A, Waschk D, Gellissen G, Hollenberg CP, Suckow M. Single-step co-integration of multiple expressible heterologous genes into the ribosomal DNA of the methylotrophic yeast *Hansenula polymorpha*. *Appl Microbiol Biotechnol.*; 2002. Vol. 58: p. 797–805.

Klabunde, J; Kunze, G; Gellissen, G, et al. Integration of heterologous genes in several yeast species using vectors containing a *Hansenula polymorpha*-derived rDNA-targeting element. *FEMS Yeast Research*; 2003. Vol.4, p. 185-193.

Liu Y, Li Y, Liu L, Hu X, Qiu B. Design of vectors for efficient integration and transformation in *Hansenula polymorpha*. *Biotechnol Lett.*; 2005. vol. 27(19):1529-34.

Lowy DR, Kirnbauer R, Schiller JT. Genital human papillomavirus infection. Proceedings of the National Academy of Sciences of the United States of America; 1994. Vol. 91, p.2436-2440.

Lowy DR, Schiller JT. Prophylactic human papillomavirus vaccines. J Clin Invest.; 2006. May; vol. 116(5): p. 1167-73.

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. Yeast, 2005, vol. 22, pp. 249–270.

Mach H, Volkin DB, Troutman RD, Wang B, Luo Z, Jansen KU, et al. Disassembly and reassembly of yeast-derived recombinant human papillomavirus virus-like particles (HPV VLPs). J Pharm Sci.; 2006. Oct; vol. 95(10): 2195-206.

Mansur M, Cabello C, Hernandez L, et al. Multiple gene copy number enhances insulin precursor secretion in the yeast *Pichia pastoris* Biotechnology letters; 2005. Vol. 27 (5): 339-345.

Marx H, Mecklenbrauker A, Gasser B, et al. Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. FEMS YEAST RESEARCH; 2009. Dec., v.9, p. 1260-1270.

Melmer G. Biopharmaceuticals and the industrial environment. In Production of recombinant proteins – Novel microbial and eukaryotic expression systems Edited by: Gellissen G. Weinheim: Wiley- VCH; 361-383, 2005.

Munoz N, for the International Agency for Research on Cancer Multicenter Cervical Cancer Study Group. Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer. N Eng. J Méd 2003; 348(6): 518-527.

MOBITEC. pPICHOLI Shuttle Vector System. Product Information and Instructions. August, 2003.

Narisawa-saito M, kiyono T. Basic mechanisms of high-risk human papillomavirus-induced carcinogenesis: roles of e6 and e7 proteins. cancer sci; 2007. vol. 98 (10): p 1505-1511.

Pfaffl MW. A new mathematical model for relative quantification of real-time RT-PCR. Nucleic Acids Research, 2001, Vol. 29, p. 2003-2007.

Ramackers, C.; Ruijter, J.M.; Deprez, R.H.L.; Moorman, A.F.M. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data, *Neurosci Lett*, 2003. Vol.339, p.62-66, 2003.

Rebrikov DV, Trofimov DY. Real-Time PCR: A Review of Approaches to Data Analysis. *Applied Biochemistry and Microbiology*, 2005. vol. 42, n.5, p. 455-461.

Roden RBS, Hubbert NL, Kirnbauer R, Christensen NS, Lowy DR, Schillher JT. Assessment of the Serological Relatedness of Genital Human Papillomaviruses by Hemagglutination Inhibition. *Journal of Virology*; 1996. v. 70, n. 5, p. 3298-3301.

Russel C, Mawson J, Yu PL. Production of recombinant products in yeasts: a review. *Aust J Biotechnol*; 1991. Vol. 5: p. 48–55.

Sambrook J, Russell DW. *Molecular Cloning: A Laboratory Manual*, 2002 third ed., Cold Spring Harbor Laboratory Press, New York.

Scheurer ME, Tortolero-Luna G, Adler-Storthz K. Human papillomavirus infection: biology, epidemiology, and prevention. *Int J Gynecol Cancer*; 2005. Vol. 15: p. 727–746.

Schmiedeskamp MR, Kockler DR. Human papillomavirus vaccines. *Annals of Pharmacotherapy*, 2006. Vol.40 (7-8), p. 1344-1352.

Scorer CA, Clare JJ, McCombie WR, Romanos MA, Sreekrishna K. Rapid selection using G418 of high copy number transformants of *Pichia pastoris* for high-level foreign gene expression. *Biotechnology (N Y)*. 1994 Feb;12(2):181-4.

Schmidt FR. Recombinant expression systems in the pharmaceutical industry. *Appl Microbiol Biotechnol* (2004) 65: 363–372.

Smith JS, Lindsay L, Hoots B, Keys J, Franceschi S, Winer R. Human papillomavirus type distribution in invasive cervical cancer and high-grade cervical lesions: A meta-analysis update. *Int. J. Cancer*; 2007. vol.121,p. 621–632.

Sohn, JH; Choi, ES; Kim, CH, et al. Title: A novel autonomously replicating sequence (

ARS) for multiple integration in the yeast *Hansenula polymorpha* DL-1. *Journal of bacteriology*, 1996. vol. 178, no.15, p. 4420–4428.

Sohn JH, Choi ES, Kang HA, Rhee JS, Agaphonov MO. A dominant selection system designed for copy-number-controlled gene integration in *Hansenula polymorpha* DL-1. *Appl Microbiol Biotechnol*; 1999. Vol. 51(6): p. 800-807.

Stanley M, Lowy DR, Frazer I. Chapter 12: Prophylactic HPV vaccines: Underlying mechanisms. *Vaccine* 2006 Aug 21;24 Suppl 3:S106-13.

Steinborn G, Gellissen G, Kunze G. Assessment of *Hansenula polymorpha* and *Arxula adeninivorans* derived rDNA-targeting elements for the design of *Arxula adeninivorans* expression vectors. *FEMS Yeast Res*; 2005. Vol. 5(11): p.1047-54.

Steinborn G, Boer E, Scholz A, Tag K, Kunze G, Gellissen G. Application of a wide-range yeast vector (CoMedtrade mark) system to recombinant protein production in dimorphic *Arxula adeninivorans*, methylotrophic *Hansenula polymorpha* and other yeasts. *Microb Cell Fact*; 2006. Vol. 14: p. 5:33.

Thierry F. Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology*; 2009. p.375–379.

Tomita YE, Ramos CRR, Nascimento ALO, HO PL. Isolation of genomic DNA from *Pichia pastoris* without hydrolases. *Biotechnologia Aplicada* 2002, vol. 19, p. 167-168.

Villiers ET, Fauquet C, Broker TR, Bernard HU, Hausen H. Classification of papillomaviruses. *Virology*; 2004. vol.324: p. 17– 27.

WHO/ICO Information Centre on HPV and Cervical Cancer (HPV Information Centre). Human Papillomavirus and Related Cancers in Brazil. Summary Report 2010. [30/03/2010]. Available at www.who.int/hpvcentre

Woodman CBJ, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. *Nature reviews*; 2007. Vol 7: p. 11-22.

World Health Organization, Department of Immunization, Vaccines and Biologicals, Initiative for Vaccine Research Team. Report of the Consultation on Human Papillomavirus

vaccines. World Health Organization; April 2005; Geneva, Switzerland. Geneva: Switzerland; October 2005; WHO/IVB/05.16

World Health Organization. Preparing for the introduction of HPV vaccines: policy and programme guidance for countries. Technical Consultation on Sexual and Reproductive Health Programmes and HPV Vaccines, March 14-16 2006; Montreux, Switzerland. Geneva : Switzerland; 2006.

Zhang Wei, Carmichael J, Ferguson J, Inglis S, Ashrafian H, Stanley M. Expression of Human Papillomavirus Type 16 L1 Protein in *Escherichia coli*: Denaturation, Renaturation, and Self-Assembly of Virus-like Particles in Vitro. *Virology*; 1998. Vol. 243: p. 423–431.

Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci.*; 2006. vol. 11: p.2286–2302.

Zhu TC, Gu MJ, Sun C, et al. A systematical investigation on the genetic stability of multi-copy *Pichia pastoris* strains. *Biotechnology Letters*; 2009. Vol. 31, p. 679–684.

Xu YF, Zhang YQ, Xu XM, Song GX. Papillomavirus virus-like particles as vehicles for the delivery of epitopes or genes. *Archives of Virology*, 2006, vol. 151, no11, pp. 2133-2148.

APÊNDICE A – BIOLOGIA MOLECULAR DA **INFECÇÃO DO CÓLO UTERINO POR HPV.**

As características de cada infecção variam drasticamente em função do sítio de infecção, do tipo de HPV e da interação com o hospedeiro. A maioria das infecções é assintomática e auto-limitada. Em geral um episódio de infecção (com ou sem alteração citológica) se resolve dentro do primeiro ou segundo ano após a exposição. Estudos de detecção por PCR indicam que menos da metade das mulheres continuam infectadas pelo mesmo HPV após um ano. (WHO, 2005, IARC 2007).

As infecções por HPV de alto risco tendem a durar mais que as de baixo risco e, conforme o tempo de persistência aumenta, a probabilidade de resolução da infecção diminui enquanto a chance de um diagnóstico de lesão pré-cancerosa aumenta. O tempo de evolução da doença é longo e são necessários cerca de dez anos para o surgimento de um câncer.

Nos próximos parágrafos serão descritas as principais características moleculares da infecção. Os dados apresentados foram retirados de quatro revisões importantes sobre o tema (Zheng, Baker, 2006, Doorbar, 2006, Narisawa-Saito, Kyono, 2007, Woodman et al., 2007).

O ciclo de vida do HPV esta intimamente ligado à biologia das células do hospedeiro. Esse vírus infecta queratinócitos da camada basal do epitélio escamoso estratificado. Em condições normais, as células escamosas epiteliais crescem de forma estratificada, com aquelas na camada basal se dividindo e funcionando como células tronco. Após a divisão, uma das células filha migra pra uma camada superior e inicia a fase terminal de diferenciação enquanto a outra permanece na camada basal em um ciclo lento de autorenovação da população. A replicação e a expressão das proteínas virais são finamente reguladas e dependentes desse processo de diferenciação que ocorre durante a replicação e externalização dos queratinócitos.

Os vírions do HPV acessam as células da camada basal através de traumas e micro-lesões do epitélio. A internalização dos vírion parece ser facilitada pela ligação à

proteoglicanos de heparan sulfato e a alguns receptores secundários como a Integrina $\alpha 6$. Os vírions penetram por endocitose e a transferência do DNA viral é facilitada pela proteína L2.

Após a entrada no núcleo, ocorre a ativação do promotor que direciona a expressão das proteínas E1 e E2 e o genoma se estabelece em uma forma epissomal se replicando junto ao genoma hospedeiro. As proteínas E1 e E2 são responsáveis por recrutar a maquinaria celular de replicação e segregação. A proteína E2 ainda atua como fator de transcrição dos genes da região EARLY através da ligação à região regulatória LCR. Em baixas concentrações a proteína E2 funciona como ativador da transcrição e em altas como inibidor da transcrição dos oncogenes E6 e E7. Esse comportamento paradoxal faz com que as proteínas virais se mantenham em concentrações baixas, evitando assim, a detecção pelo sistema imune. O HPV pode permanecer nesse estágio por períodos muito longos, apenas replicando o genoma epissomal lentamente.

Na maioria das lesões de baixo grau, o vírus esta predominantemente na forma epissomal, entretanto em lesões de alto grau e cânceres, o vírus esta integrado no genoma do hospedeiro.

A desrepressão dos oncogenes virais pode ocorrer após a integração do genoma viral no genoma do hospedeiro. Frequentemente a integração pode levar a disruptura da ORF de E2 e eliminar a repressão transcricional das proteínas E6 e E7. Acredita-se que esse é o evento responsável pelos altos níveis de expressão dos oncogenes nas células basais, processo imprescindível para a imortalização, transformação e carcinogênese das células (Thierry, 2008).

Nas lesões cervicais causadas pelo HPV, a proliferação acentuada das células epiteliais suprabasais é atribuída à expressão dos oncogenes E6 e E7, que não estão mais reprimidos como nas células basais. Em condições normais as células suprabasais escapam do ciclo celular e sofrem o processo de diferenciação terminal para produzir a barreira protetora da pele. Em células infectadas, os oncogenes impedem a saída do ciclo celular e a diferenciação normal das células. Essa habilidade de direcionar as células pra Fase S do ciclo é necessária, em conjunto com a expressão das proteínas E1 e E2, para a replicação do vírus.

As proteínas E6 e E7 foram extensamente estudadas e hoje sabemos que elas são os principais fatores envolvidos no processo de carcinogênese. Ambas as proteínas inibem ou degradam as proteínas p53 e pRb (*retinoblastoma protein*), duas das principais proteínas

supressoras de tumor. Além disso, essas proteínas multifuncionais possuem diversas outras propriedades que colaboram com o desenvolvimento do câncer. A **Ilustração 25** (Narisawa-Saito, Kyono, 2007) mostra a interação com alguns dos principais alvos celulares ligados à transformação maligna. Estudos mostram que as proteínas E6 e E7 dos HPV de baixo risco ligam-se aos seus alvos com menor afinidade do que as de alto risco.

Ilustração 25. Função cooperativa das proteínas E6 e E7 no desenvolvimento do câncer cervical. O principal gatilho do processo é a superexpressão de E6 e E7, que geralmente ocorre devido à integração do genoma viral nos cromossomos do hospedeiro. Uma vez superexpressas, as proteínas agem de forma cooperativa em diversas vias de controle celular e induzem instabilidade genômica e cromossomal.

No estágio intermediário, mudanças no sistema de sinalização celular ativam preferencialmente o promotor P670 dependente de diferenciação. A ativação promove o aumento da expressão das proteínas relacionadas à amplificação do genoma viral (E1, E2, E4 e E5). Nesse momento ocorre um aumento acentuado no número de cópias do DNA viral. A proteína E5 é um oncogene que inibe a proteína supressora de tumor p21 e também diminui a degradação do receptor de EGF (fator de crescimento epidermal). Essas duas condições contribuem para a manutenção de um ambiente propício à amplificação genômica viral.

O estágio final do ciclo de vida do HPV requer o empacotamento do DNA genômico. As proteínas do capsídeo viral (L1 e L2) começam a ser expressas no estágio final, logo após a amplificação genômica, com a expressão da L2 precedendo a da L1. A expressão preferencial das proteínas estruturais nesse estágio ocorre devido a mecanismos transcricionais e pré-traducionais modulados pelo estágio de diferenciação dos queratinócitos.

A formação do capsídeo viral ocorre exclusivamente no núcleo celular. Primeiramente a proteína L2 é translocada para o núcleo e, com a ajuda de outros fatores, se associa ao DNA viral. A proteína L1 se associa em pentâmeros no citoplasma e em seguida é translocada para o núcleo. A associação da proteína L2 ao pentâmero de L1 ocorre por interações hidrofóbicas e ela se posiciona no centro do pentâmero. As associações entre os pentâmeros também ocorre através de interações hidrofóbicas entre as porções C-terminal da L1 e através de pontes de dissulfeto.

Por fim, os vírions são expulsos dos queratinócitos pela ajuda da proteína E4. Essa proteína é capaz de destruir a malha de queratina que protege os queratinócitos, e dessa forma auxilia a evasão dos vírions através do envelope cornificado na superfície celular. Os novos vírions liberados pelas células reinfectam as células da camada basal e reiniciam o ciclo. A **Ilustração 26** (Doorbar, 2006) apresenta um resumo esquemático do ciclo de vida do vírus HPV.

Adicionalmente é importante mencionar que, apesar de pequeno, o genoma do HPV possui características bastante complexas. Em primeiro lugar, o vírus não possui mRNAs monocistrônicos. Os mRNAs são bicistrônicos, tricistrônicos ou até policistrônicos. Uma mesma ORF pode estar presente em diferentes transcritos, os quais também podem ser controlados por diferentes promotores. Esses transcritos também não traduzem necessariamente todas suas ORFs e os mecanismos de seleção preferencial de início da tradução ainda são desconhecidos. Isso torna o estudo de expressão gênica do HPV extremamente difícil. (Zheng, Baker, 2006, Doorbar, 2006).

Ilustração 26. Ciclo de vida do vírus HPV. As fases da infecção estão diagramadas à direita. As diferentes camadas celulares do epitélio estão diagramadas à esquerda. Os períodos de expressão das proteínas virais e da replicação do DNA estão representados por setas coloridas. Porções mais claras das setas indicam baixa expressão ou replicação, e porções mais escuras indicam alta expressão ou replicação. A coloração das células epiteliais também indica a presença das proteínas virais de mesma coloração. Os promotores responsáveis pelas expressões de cada elemento estão indicados abaixo das setas

APÊNDICE B – RESPOSTA IMUNE À INFECÇÃO POR HPV DO COLO UTERINO.

O sistema imune desempenha um papel fundamental na defesa contra a progressão da infecção por HPV. Pacientes com o sistema imune defectivo podem desenvolver lesões disseminadas e refratárias ao tratamento.

O tempo necessário para a liberação dos vírions desde o início do ciclo do HPV é de aproximadamente três semanas, o mesmo tempo necessário para as células basais sofrerem a diferenciação completa. Entretanto, o período entre a infecção inicial e o surgimento das lesões é imprevisível, podendo variar de semanas até meses, sugerindo que o vírus possui a capacidade de escapar da detecção pelo sistema imune (Stanley, 2006, IARC, 2007).

Parte dessa capacidade resulta do próprio ciclo de vida do vírus. Primeiramente porque não ocorre viremia - a infecção é sítio-específica e se limita ao tecido epitelial. Em segundo lugar, a infecção não causa citólise ou morte citopática, porque a etapa de construção e liberação dos vírions acontece somente nas células mais superficiais do epitélio, onde os queratinócitos terminais já estão em um processo natural de morte celular e descamação. Sem esses fenômenos quase não há a liberação de citocinas pro-inflamatórias, essenciais para a ativação das células dendríticas e do sistema imune como um todo. Além do mais, a própria superficialidade do processo de formação viral dificulta o acesso pelas células imunes (Stanley, 2006, IARC, 2007).

Adicionalmente ao controle da expressão das proteínas do capsídeo (L1 e L2), a linha temporal de expressão das proteínas EARLY também beneficia o escape do sistema imune. Nas células mais inferiores do estrato, as proteínas EARLY são expressas em concentrações inferiores àsquelas necessárias para a detecção pelo sistema imune. Nesse estágio a replicação tem a função apenas de manter o número de cópias viriais, e o estágio de amplificação ocorre somente em estratos celulares superiores (Doorbar, 2006, Stanley, 2006, Scheurer et al., 2005).

Mesmo dentro dessas condições, os queratinócitos infectados ainda deveriam ativar uma importante via de defesa, a secreção de Interferon tipo 1, INF- α e INF- β . Eles têm

propriedades antivirais, antiproliferativas, antiangiogênicas e ativam as células dendríticas imaturas, agindo como uma ponte entre a imunidade inata e adaptativa. Os HPV de alto risco regulam negativamente a expressão de IFN e as oncoproteínas E6 e E7 do HPV16 interagem diretamente com componentes da via de sinalização dos INF, bloqueando a cascata de sinalização (Doorbar, 2006, Stanley, 2006, Scheurer et al., 2005).

Frequentemente, a penetração dos capsídeos virais representa um sinal ativador das células dendríticas, mas existem evidências de que as Células de Langerhans não são ativadas pela captura dos capsídeos do HPV, um fenômeno que inibe a migração e maturação dessas células e, assim, o início da resposta imune celular contra os capsídeos virais (Doorbar, 2006, Stanley, 2006, Scheurer et al., 2005).

Em resumo, o cenário descrito mostra que o HPV escapa eficientemente do sistema imune inato e atrasa a ativação do sistema imune adaptativo.

Apesar dessas habilidades do HPV, na maioria dos casos, o sistema imune desenvolve uma resposta eficaz contra as infecções genitais desse vírus. A maior parte das informações sobre a natureza da resposta imune celular vem de estudos imunohistológicos de lesões com regressão espontânea. As lesões persistentes são caracterizadas pela ausência de células imunes, existindo apenas poucos linfócitos CD8 e células mononucleares. Estudos histológicos de lesões em regressão revelam grandes infiltrados de células T (CD4+ e CD8+) e macrófagos (Stanley, 2006, Scheurer et al., 2005).

Adicionalmente, ocorre a produção de anticorpos neutralizantes. Esses anticorpos são direcionados contra epitopos conformacionais da proteína L1 expostos na superfície do capsídeo. Enquanto a resposta imune inata é imediata, a produção de anticorpos contra HPV de alto risco pode demorar de seis meses a um ano para acontecer, porém a produção de anticorpos contra HPV de baixo risco parece ter início logo que a infecção acontece. Os níveis de anticorpos produzidos nas infecções são muito baixos mesmo nos picos de titulação e parecem não aumentar durante episódios de reinfecção. Provavelmente o principal mecanismo imunológico responsável pela resolução e controle da infecção por HPV é uma resposta imune celular eficaz (Stanley, 2006, Scheurer et al., 2005).

APÊNDICE C – SEQUÊNCIA DE DNA DO GENE DA L1 COM CÓDONS OTIMIZADOS PARA *P. PASTORIS*

Sequência de DNA:

ACC ATG TCT TTG TGG TTG CCA TCT GAG GCT ACT GTT TAC TTG CCA CCA GTT CCA GTT TCC AAA GTT GTT
 TCC ACT GAC GAG TAC GTT GCT AGA ACT AAC ATC TAC TAC CAC GCT GGT ACT TCC AGA TTG TTG GCT GTT
 GGT CAC CCA TAC TTC CCA ATC AAG AAG CCA AAC AAC AAC AAG ATC TTG GTT CCA AAG GTT TCC GGA TTG
 CAG TAC AGA GTT TTC AGA ATC CAC TTG CCA GAC CCA AAC AAG TTT GGT TTC CCA GAC ACT TCC TTC TAC
 AAC CCA GAC ACT CAG AGA TTG GTT TGG GCT TGT GTT GGT GTT GAA GTT GGT AGA GGA CAG CCA TTG GGT
 GTT GGT ATT TCC GGT CAC CCA TTG TTG AAC AAG TTG GAC GAC ACT GAA AAC GCT TCT GCT TAC GCT GCT
 AAC GCT GGT GTT GAC AAC AGA GAA TGT ATC TCC ATG GAC TAC AAG CAG ACT CAG TTG TGT TTG ATC GGT
 TGT AAG CCA CCA ATT GGT GAA CAT TGG GGA AAG GGT TCC CCA TGT ACA AAC GTT GCT GTT AAC CCT GGT
 GAT TGT CCA CCA TTG GAG TTG ATC AAC ACT GTT ATC CAG GAC GGT GAC ATG GTT GAT ACT GGT TTC GGT
 GCT ATG GAC TTC ACT ACT TTG CAG GCT AAC AAG TCC GAA GTT CCA TTG GAC ATC TGT ACT TCC ATC TGT
 AAG TAC CCA GAC TAC ATC AAG ATG GTT TCC GAA CCA TAC GGT GAC TCT TTG TTC TTC TAC TTG AGA AGA
 GAA CAG ATG TTT GTT AGA CAC TTG TTC AAC AGA GCT GGT GCT GTT GGT GAA AAC GTT CCA GAC GAC TTG
 TAC ATT AAG GGT TCC GGT TCC ACT GCT AAC TTG GCT TCC TCC AAC TAC TTT CCA ACT CCA TCC GGT TCT
 ATG GTT ACT TCC GAC GCT CAG ATC TTC AAT AAG CCA TAC TGG TTG CAA AGA GCA CAA GGT CAC AAC AAC
 GGT ATC TGT TGG GGT AAC CAG TTG TTC GTT ACT GTT GTT GAC ACT ACT AGA TCC ACT AAC ATG TCC TTG
 TGT GCT GCT ATT TCC ACT TCC GAG ACT ACT TAC AAG AAC ACT AAC TTC AAG GAG TAC TTG AGA CAC GGT
 GAA GAG TAC GAC TTG CAG TTC ATC TTC CAG TTG TGT AAG ATC ACT TTG ACT GCT GAC GTT ATG ACT TAC
 ATC CAC TCC ATG AAC TCC ACT ATT TTG GAG GAC TGG AAC TTC GGA TTG CAA CCA CCA CCA GGT GGT ACT
 TTG GAG GAC ACT TAC AGA TTC GTT ACT TCC CAG GCT ATT GCT TGT CAA AAG CAC ACT CCA CCT GCT CCA
 AAG GAA GAT CCA TTG AAG AAG TAC ACT TTC TGG GAG GTT AAC TTG AAG GAG AAG TTC TCT GCT GAC TTG
 GAC CAA TTC CCA TTG GGT AGA AAG TTC TTG TTG CAG GCT GGA TTG AAG GCT AAG CCA AAG TTC ACT TTG
 GGA AAG AGA AAG GCT ACT CCT ACT ACT TCT TCC ACT TCC ACT ACT GCT AAG AGA AAG AAG AGA AAA TTG
 TAA TAG TAG

Sequência de aminoácidos:

Met SLWLPSEATVYLPPVPVSKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNK
 ILVPKVSGLQYRVFRIHLPDPNKF GFDTSFYNPDTQRLVWACVGVGVGRGQPLGVGIS
 GHPLLNLKDDTENASAYAANAGVDNRECISM DYKQTQLCLIGCKPPIGEHWGKGS PCT
 NVAVNPGDCPPLELINTVIQDGDMDVTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIK
 MVSEPYGDSLFFYLRRREQMFVRHLFN RAGAVGENVPDDL IYKSGSGSTANLASSNYFPTP
 SGSMVTSDAQIFNKPYWLQRAQGHNN GICWGNQLFVTVDTTTRSTNMSLCAAISTSET
 TYKNTNFKEYLRHGEEYDLQFIFQLCKITLTADVMTYIHS MNSTILEDWNFGLQPPPGGT
 LEDTYRFVTSQAIACQKHTPPAPKEDPLKKYTFWEVNLKEKFSADLDQFPLGRKFLLQAG
 LKAKPKFTLGKRKATPTTSSTSTTAKRKRKL Stop Stop Stop

APÊNDICE D – SEQUÊNCIAS DE DNA RIBOSSOMAL DE LEVEDURAS

A **Ilustração 27** mostra a estrutura da unidade do DNA ribossomal de *H. polymorpha* e o fragmento da sequência que foi utilizado nos vetores pINT-A e pINT-B (Código de acesso NCBI: gi|18652300|gb|AF467695.1).

Ilustração 27. Estrutura da unidade do DNA ribossomal de *H. polymorpha*. O fragmento de DNA ribossomal utilizado nos vetores está indicado na figura. Ele possui parte da sequência intergênica NTS2, a sequência do gene do RNA ribossomal 5S, a sequência intergênica NTS1 e parte da sequência do gene do RNA ribossomal 25S.

O genoma da *Pichia pastoris* foi publicado em junho de 2009. A *Pichia* possui apenas cerca de 21 cópias da unidade do DNA ribossomal e o gene do RNA ribossomal 5S não faz parte da unidade. Ele está distribuído pelo genoma da *Pichia* e também possui cerca de 21 cópias. As sequências dos DNAs ribossomais da *P. pastoris* não possuem quase nenhuma similaridade com a sequência do fragmento do rDNA de *H. polymorpha*.

Sequências do DNA ribossomal de P. pastoris.

Sequência 1 - Gene do RNA ribossomal 5S.

TCGGTTGCGGCCATATCTAGCAGAAAAGCACCGTTTCCCGTCCGATCAACCGTAGTTAAGCTGCTAAGAGCCTGAC
CGAGTAGTGTAGTGGGAGACCATACGCGAAACTCAGGTGCTGCAATCTC

Sequência 2 – Unidade do DNA Ribossomal contendo os genes dos RNAs ribossomais 5.8S, 18S e 25S (aproximadamente 7450 bp).

[illegible]

APÊNDICE D – MECANISMOS DE INTEGRAÇÃO

HOMÓLOGA

A **Ilustração 28** mostra os dois mecanismos clássicos de recombinação homóloga entre a sequências de DNA de um vetor e a sequência de DNA genômico. Na recombinação por substituição, as sequências das extremidades do vetor linearizado são similares a duas sequências genômicas não adjacentes. Nesse tipo de recombinação, a sequência genômica contida entre os dois sítios de recombinação é eliminada. Na recombinação por inserção, as sequências das extremidades do vetor linearizado são similares a duas sequências genômicas adjacentes. Nesse tipo de recombinação, o vetor é inserido no cromossomo sem causar deleções no material genético.

Neste projeto, os vetores integrativos pINT-AL1 e pINT-BL1 contendo um fragmento do rDNA de *H. polymorpha* tinham como objetivo se integrar por inserção no DNA ribossomal de *P. pastoris*.

Ilustração 28. Mecanismos de integração por recombinação homóloga. A ilustração utiliza como exemplo a sequência do gene AOX1. Na integração por inserção, não ocorre a deleção do gene AOX1, mas na integração por substituição sim. A região vermelha (GOI) representa o gene que se deseja integrar no genoma.