



# O Sistema dois-híbridos de *Saccharomyces cerevisiae*

Variações e potencial para a análise funcional e montagem de redes de interações proteína-proteína

## Jenifer Saffi

Farmacêutica, Doutora em Ciências  
saffi@dna.cbiot.ufrgs.br

## Luís Fernando Revers

Biólogo, Mestre em Genética e Biologia Molecular  
Doutorando em Ciências  
revers@dna.cbiot.ufrgs.br

## João Antonio Pêgas Henriques

Doutor em Ciências Naturais  
Professor Titular de Biofísica  
pegas@dna.cbiot.ufrgs.br

Laboratório de Radiobiologia Molecular  
Centro de Biotecnologia da UFRGS  
Porto Alegre, RS

O sistema dois-híbridos (“two-hybrid system”) é um ensaio artificial da transcrição, que se baseia no princípio de que muitas proteínas, incluindo ativadores da transcrição, consistem de múltiplos domínios que funcionam independentemente. Quando domínios individuais são expressos separadamente e aproximados por meio de interações não covalentes, tais domínios podem funcionar coletivamente para reconstituir a atividade da proteína intacta. Ativadores da transcrição são, via de regra, proteínas bipartidas compostas de um domínio de ligação ao DNA (DLD) e de um domínio de ativação da transcrição (DA), que são funcionalmente independentes. A natureza modular dos ativadores da transcrição é explorada no sistema dois-

linhagem apropriada de *Saccharomyces cerevisiae* reconstitui a atividade do ativador da transcrição pela aproximação dos seus domínios funcionais mediados por interações entre X e Y. A interação é facilmente identificada pela expressão de um gene repórter regulado em *cis* por uma região promotora que contém sítios de ligação reconhecidos por DLD. A aplicação mais poderosa e difundida atualmente do sistema dois-híbridos para identificar proteínas interatoras é gerar híbridos DA-Y, onde Y corresponde ao espectro de proteínas codificado por uma biblioteca de cDNA ou fragmentos genômicos. A seleção positiva para colônia que expressem uma presa capaz de interagir com a isca é conseguida pelo uso de genes repórteres prototróficos (por exemplo, *HIS3* e *LEU2*), os quais são transcritos em resposta à interação *isca-presa*, complementando as marcas auxotróficas da linhagem de *S. cerevisiae* específica e permitindo o seu crescimento em meio seletivo. A ativação transcricional paralela de genes repórteres colorimétricos, tais como *lacZ*, serve para confirmar a especificidade da interação e também quantificá-la.

Uma vez que as interações proteína-proteína são críticas a muitos processos biológicos, desde a formação de estruturas celulares macromoleculares e complexos enzimáticos até a regulação de vias de transdução de sinal, diversas abordagens e modificações do sistema dois-híbridos, originalmente descrito por S. Fields e O. Song, em 1989, surgiram para a análise funcional (identificação) de interações potenciais proteína-proteína em diversos sistemas biológicos. Essa revisão tem o objetivo de apresentar

**Tabela 1.** Representantes comerciais de sistemas dois-híbridos

Representante	Website
Clontech	www.clontech.com
OriGene	www.origene.com
Invitrogen	www.invitrogen.com
Stratagene	www.stratagene.com

híbridos através da construção de proteínas de fusão que ligam uma proteína de interesse ao DLD para gerar um domínio de ligação ao DNA híbrido (DLD-X), comumente referido como *isca* (“bait”). Domínios de ativação da transcrição híbridos (DA-Y), comumente referidos como *presas* (“preys”), são geradas através da construção de proteínas de fusão ligando domínios de ativação da transcrição a uma proteína interatora conhecida. A co-expressão de DLD-X e DA-Y em uma

**Tabela 2.** Endereços da Internet com informações sobre dois-híbridos

Endereço da Internet	Informações
<a href="http://www.uib.no/aasland/two-hybrid.html">http://www.uib.no/aasland/two-hybrid.html</a>	informa diversos links para sistemas dois-híbridos
<a href="http://xanadu.mgh.harvard.edu:/brentlabweb/andyweb/m&amp;b.html">http://xanadu.mgh.harvard.edu:/brentlabweb/andyweb/m&amp;b.html</a>	artigo de R. Brent e A. Mendelsohn sobre as aplicações do sistema dois-híbridos na biotecnologia
<a href="http://www.fccc.edu/research/labs/golemis/com_sources1.html">http://www.fccc.edu/research/labs/golemis/com_sources1.html</a>	coleção de plasmídeos de fusão com LexA
<a href="http://www.fccc.edu/research/labs/golemis/main_false.html">http://www.fccc.edu/research/labs/golemis/main_false.html</a>	coleção de falsos positivos

uma síntese sobre algumas das variações do sistema dois-híbridos original, suas aplicações na análise funcional de interações proteína-proteína (e genômica/ proteômica), bem como divulgar o potencial dessa tecnologia na era pós-genômica.

### O sistema dois-híbridos original e seus derivados

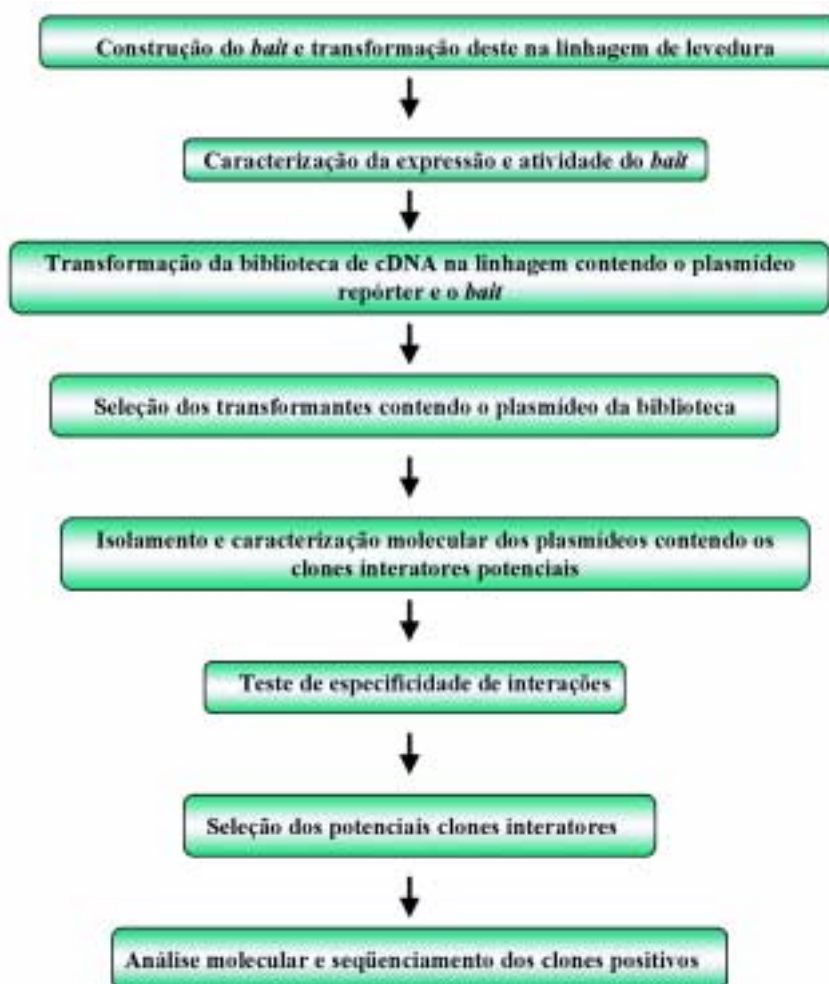
O sistema dois-híbridos básico evoluiu, na última década, para facilitar o seu uso em uma gama muito diversa de aplicações e, conseqüentemente, para facilitar ou prover a obtenção de dados para compreensão de interações protéicas mais complexas e variáveis. A figura 1A mostra um esquema do funcionamento básico do sistema dois-híbridos e algumas das variações surgidas. Todos esses sistemas apresentam três componentes básicos essenciais: 1º) vetores de levedura para a expressão de uma proteína conhecida fusionada ao DLD; 2º) vetores de levedura que direcionam a expressão de proteínas codificadas por um cDNA fusionado ao DA; e 3º) genes repórteres de levedura que contêm sítios de ligação para o DLD.

### Sistema três híbridos: co-expressão de enzimas modificadoras

Um exemplo da necessidade de introduzir variações na técnica originalmente descrita veio do fato de que muitas proteínas de organismos eucarióticos superiores sofrem extensivas modificações pós-traducionais as quais são essenciais para o correto desempenho funcional. Esse fato é particularmente problemático quando o gene

de interesse (*isca*) é, por exemplo, originário de vias de transdução de sinal de um organismo eucarioto superior. Nesses casos, modificações pós-traducionais criam sítios de reconhecimento críticos e necessários para as interações protéicas dessas cascatas.

A solução para esse problema foi co-expressar a (enzima fosforiladora) proteína quinase adequada para efetuar as modificações pós-traducionais juntamente com a *isca* e a *presa*, facilitando as modificações pós-traducionais ausentes em *S. cerevisiae*. De



Representação esquemática das etapas de um screening de dois-híbridos.

forma semelhante, outras modificações foram introduzidas para facilitar o estudo de interações protéicas dependentes de complexos multiméricos, nos quais uma determinada proteína possui mais afinidade de ligação/interação com outro complexo ou consequência de interações múltiplas de efeito cumulativo com outros complexos ou componentes distintos. Nesses casos, a co-expressão de um “parceiro molecular” previamente identificado pode prover uma interface de interação adequada para identificar outros componentes do complexo multimérico. A co-expressão de uma isca auxiliar para permitir a identificação de interações protéicas em complexos multiméricos foi então denominada de sistema três-híbridos (figuras 1B e 1C).

### Sistema dois-híbridos reverso

No sistema dois-híbridos original, as interações proteína-proteína são selecionadas positivamente pelo crescimento das colônias mediado pela transcrição dos genes repórteres. No sistema dois-híbridos reverso, a ativação da transcrição de genes repórteres de seleção negativa foi implementada para indicar perda de interação *isca-presa* (figura 1D). Entre os diferentes genes marcadores tóxicos para *S. cerevisiae*, os mais utilizados são *URA3* e *CYH2*. O gene repórter *URA3* é tradicionalmente o mais utilizado, porque permite ambas as seleções-positiva e negativa - em meio que contém ácido 5-fluoracético - 5-FOA, um análogo não-tóxico à enzima orotidina-5'- fosfato descarboxilase, codificada pelo gene *URA3* (cujas meta-

bolização leva à produção de 5-fluoracil, um produto tóxico) e em meio mínimo sem uracil, respectivamente. No caso do gene repórter *CYH2*, interações protéicas do tipo selvagem conferem sensibilidade à cicloheximida, permitindo, portanto, a seleção da perda de interação entre isca-presa.

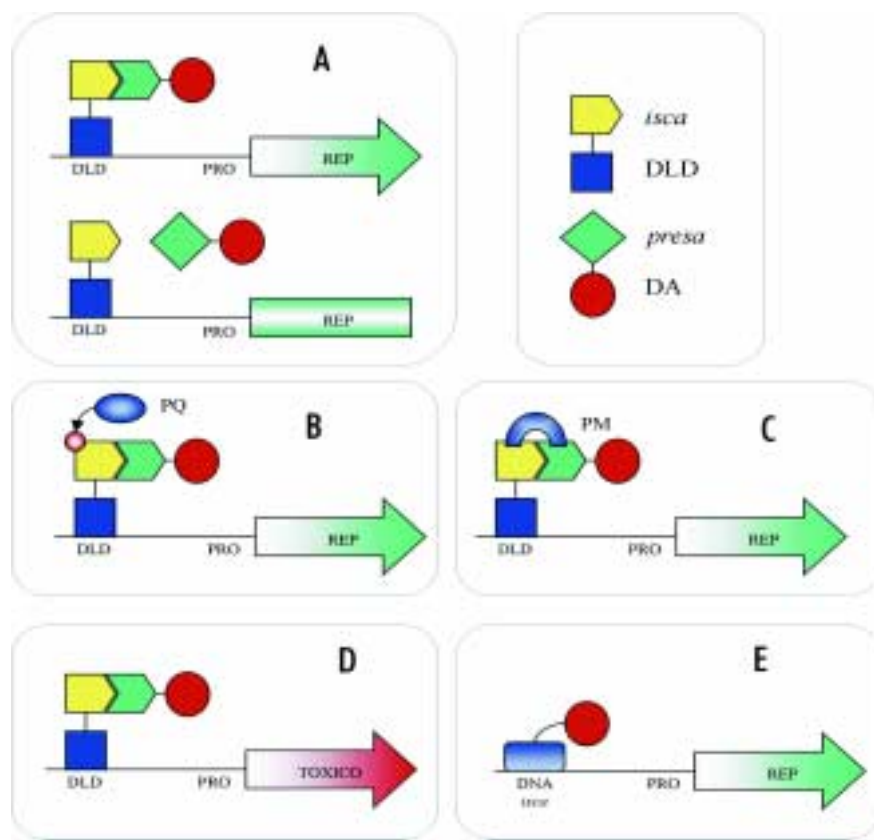
Essa variação do sistema dois-híbridos é especialmente aplicada no estudo específico de interações previamente determinadas, que permitem avaliar o grau de interação *isca-presa* e identificar os domínios responsáveis pela interação. Através desse sistema, é possível também selecionar mutações ou moléculas que dissociam ou previnem interações proteicas. Outra potencialidade dessa técnica é o seu emprego na terapêutica. Muitas doenças são causadas por disfunções mediadas por interações proteína-proteína, proteína-DNA ou receptor-ligante. A determinação dessas interações, seguida de um desenvolvimento de drogas e triagem de produtos capazes de atuar sobre essas interações representam uma área de pesquisa de ampla aplicação e significância farmacêutica.

### O sistema mono-híbrido

Os sistemas mono-híbridos foram desenvolvidos para identificar proteínas que interagem com uma seqüência específica de DNA (figura 1E). A estratégia inclui a utilização de genes repórteres similares aos do sistema dois-híbridos, que são regulados *cis* por uma cópia única ou multicópias em *tandem* de uma seqüência de DNA, que contém um domínio de ligação específico ou a região promotora de um determinado gene, por exemplo. Uma biblioteca de presas-DA, essencialmente equivalente à biblioteca do sistema dois-híbridos, é utilizada para identificar proteínas capazes de se ligarem ao domínio regulador e ativarem a transcrição do gene repórter.

### O sistema dois-híbridos na era pós-genômica

As recentes publicações de seqüências genômicas completas de diversos organismos têm produzido quantidades vastas de seqüências de



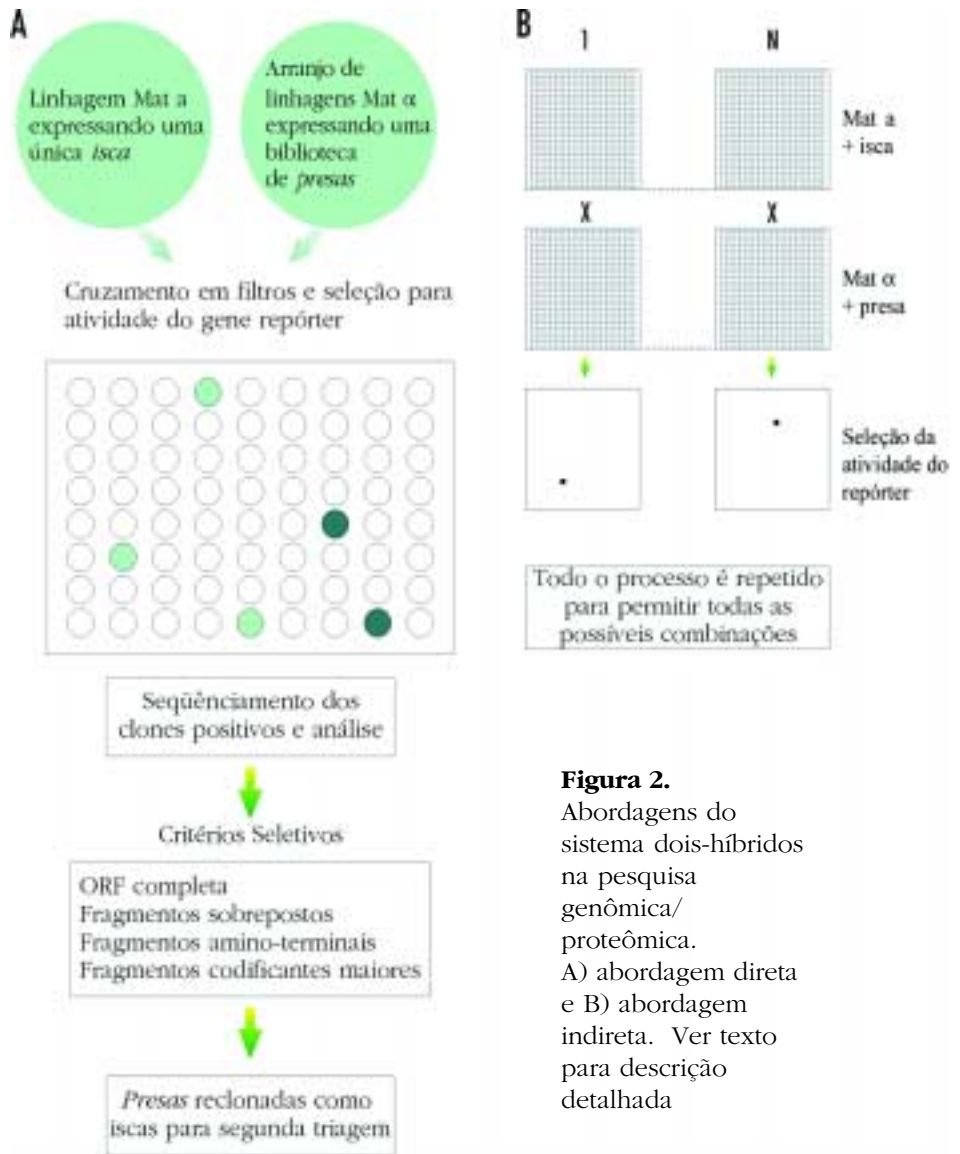
**Figura 1.** O sistema dois-híbridos e suas variações. A) Sistema dois-híbridos original mostrando a situação de interação e ausência de interação; B e C) Sistema três-híbridos; D) Sistema dois-híbridos reverso e E) Sistema mono-híbrido. REP = gene repórter; PRO = região promotora do gene repórter contendo os sítios de ligação da maquinaria de transcrição DLD = domínio de ligação ao DNA; DA = domínio de ativação da transcrição; PQ = proteína quinase; PM = proteína mediadora de interação

genes e das proteínas predictas codificadas por eles. Porém, a maioria das novas seqüências não compartilha homologia a proteínas já caracterizadas funcionalmente ou domínios proteicos conhecidos. Conseqüentemente, esses enormes bancos de dados ainda não possuem informações funcionais correspondentes às proteínas potencialmente codificadas por essas seqüências genômicas. Decifrar o papel funcional de cada gene seqüenciado representa o maior desafio da biotecnologia moderna na era pós-genômica. Sob essa perspectiva, o sistema dois-híbridos coloca-se como uma tecnologia de grande potencial a ser explorado.

O sistema dois-híbridos apresenta três grandes vantagens sobre outros ensaios alternativos para a identificação funcional de genes: baseado em um sistema genético de seleção poderoso em um organismo modelo, de biologia conhecida e fácil manipulação, como é o de *S. cerevisiae*, permite testar seqüências potencialmente codificadoras em grande escala, em experimentos relativamente simples e de baixo custo; além disso, a metodologia baseia-se em um ensaio conduzido *in vivo*, portanto, não é limitado por condições artificiais de ensaios realizados *in vitro*.

O isolamento de interações físicas com proteínas previamente caracterizadas funcionalmente pode ser indicativo de uma função potencial para a nova proteína e, conseqüentemente, servir para direcionar futuros experimentos. Essas abordagens, em combinação com análises bioinformáticas de conservação de seqüências e funções em diferentes organismos, podem efetivamente se complementar e contribuir para a elaboração de mapas de interações proteína-proteína em diferentes organismos.

Mundialmente, um número grande de grupos de pesquisa tem explorado a utilidade do sistema dois-híbridos para análise genômica/proteômica de genomas inteiros ou o mapeamento de todas as interações protéicas envolvidas em um processo celular específico como, por exemplo, o ciclo celular. Nesses projetos, duas abordagens têm sido basicamente utilizadas. A abordagem direta, conhecida como "mating assay", consiste em



**Figura 2.** Abordagens do sistema dois-híbridos na pesquisa genômica/proteômica. A) abordagem direta e B) abordagem indireta. Ver texto para descrição detalhada

efetuar o cruzamento da linhagem MAT $\alpha$  de *S. cerevisiae* expressando uma única isca DLD-X com a linhagem MAT $\alpha$ , expressando uma biblioteca genômica de presas DA-Y. As interações *isca-presa* são, então, identificadas pela ativação dos genes repórteres. Clones positivos são seqüenciados e analisados criteriosamente para eliminar falsas interações. Os clones prioritários são subclonados como novas iscas e utilizados em novas triagens (figura 2A). Um exemplo representativo de abordagem indireta foi a metodologia empregada para análise do genoma de *S. cerevisiae*, onde arranjos de clones em linhagens MAT $\alpha$  e MAT $\alpha$  de *S. cerevisiae* expressando os cerca de 6.000 genes genômicos como fusões DA-Y (*presas*) e DLD-X (*iscas*), respectivamente, foram cruzados e as interações selecionadas de acordo com a expres-

são dos genes repórteres. O processo inteiro é então repetido para obter todas as combinações possíveis entre DA-Y e DLD-X (figura 2B). Só em levedura, já foram identificadas em torno de 1.200 pares de interações e outros genomas já estão sendo estudados por esse sistema, como os de *Schizosaccharomyces pombe*, *C. elegans* e *Drosophila*.

### Limitações do sistema dois-híbridos

O sistema dois-híbridos exige expressão estável e, principalmente, o dobramento correto das proteínas híbridas. Embora foram introduzidas modificações na técnica para permitir modificações pós-traducionais, a expressão das proteínas híbridas pode não ser estável, não ter o dobramento correto ou até mesmo ser tóxica.

Uma metodologia muito utilizada e confiável para confirmar a interação de duas proteínas híbridas é testar a sua habilidade de manter a interação quando clonadas, reciprocamente, como *isca* e *presa*. A verificação da interação entre híbridos recíprocos (DLD-X x DA-Y e DLD-Y x DA-X) é um forte indicativo de uma interação com relevância biológica. Entretanto, outras evidências que suportem a interação verificada no sistema dois-híbridos são necessárias para mostrar que a interação é real. Técnicas bioquímicas e de imunohistoquímica são as mais utilizadas. A co-immunoprecipitação dos candidatos, através da expressão e purificação de proteínas etiquetadas com HA (hemaglutinina) ou GST (glutathione S transferase), são as técnicas de maior uso atualmente. A co-transfecção de linhagens de células de mamíferos contendo os cDNAs das duas proteínas de interesse para obter altos níveis de expressão, seguidas de coimmunoprecipitação, e/ou imunolocalização, também são empregadas para verificar e confirmar interações observadas pelo sistema dois-híbridos.

Algumas das limitações do sistema dois-híbridos nem sempre são relatadas na literatura. Entre elas, inclui-se um número muito alto de clones sem atividade biológica. Esses chamados “falsos positivos” podem ser reduzidos com a utilização de baixos níveis de expressão de proteínas híbridas e o uso de genes repórteres múltiplos, que contenham diferentes promotores. Por ser um ensaio artificial de transcrição, as interações encontradas por esse método devem ser inicialmente consideradas apenas como uma “hipótese de interação”, até que sejam validadas através de algum outro método de significância biológica. Quando uma mesma proteína é encontrada sendo interatora de várias iscas diferentes, é bastante provável tratar-se de uma “falsa interação” e, portanto, deve ser utilizada com bastante cautela. Exemplo comum de falsas interações são as proteínas de citoesqueleto, assim como proteínas envolvidas em transcrição.

Além da detecção de falsos positivos, também pode acontecer o opo-

to: interações já detectadas por outros métodos podem não aparecer em um ensaio de dois-híbridos; são os chamados “falsos negativos”. São vários também os motivos que levam ao aparecimento desses falsos negativos. Primeiro, DLD-X e DA-Y podem não se encontrar no núcleo. Segundo, as proteínas X e/ou Y podem não interagir em um contexto de fusão entre DLD e DA. Terceiro, a interação entre X e Y pode depender de modificação pós-translacional, que é ausente em uma célula de levedura. Por último, já foi detectado no sistema dois-híbridos, que um número de interações proteína-proteína só pode ser visualizada se X ou Y estiverem com domínios truncados. Estima-se que um percentual de até 45% de falsos negativos pode acontecer num sistema dois-híbridos. Isso sugere que as análises em larga escala sejam importantes para obtenção rápida e parcial de interações proteicas de um determinado proteoma de interesse. Entretanto, ainda precisam ser desenvolvidas, em larga escala, outras metodologias para a obtenção completa e segura de interações.

De qualquer maneira, esse é, sem dúvida nenhuma, um dos ensaios mais difundidos entre a comunidade científica na última década, gerando não só uma série de parceiros moleculares, mas, também, unindo pesquisadores de diversas áreas e de diversos lugares do mundo, com interesses comuns.

O Laboratório de Radiobiologia Molecular do Centro de Biotecnologia da UFRGS começou, em 1996, a implementar a técnica original do sistema dois-híbridos para análise de interações entre proteínas codificadas pelos genes *PSO* e as proteínas de reparo de DNA de *Saccharomyces cerevisiae*. Vários *screenings* já foram realizados com sucesso no laboratório, revelando interações importantes, como entre o gene de reparo *PSO5/RAD16* e o *SGS1*, envolvido com envelhecimento precoce na levedura. Atualmente, já se dispõe de dois sistemas para triagem de interações proteicas, podendo-se ainda implementar com facilidade outras variações da técnica, conforme a abordagem experimental.

## Bibliografia Recomendada

Bartel P.L. e Fields S. (1997) The yeast two-hybrid system. Oxford University Press, 1st Ed., New York.

Colas P. e Brent R. (1998) The impact of two-hybrid and related methods on biotechnology. *Tibtech* 16: 355-363.

Drees B. L. (1999) Progress and variations in two-hybrid and three hybrid technologies. *Curr Opin Chem Biol* 3: 64-70.

Fashena S.J., Serebriiskii I., Golemis E. A. (2000) the continued evolution of two-hybrid screening approaches in yeast: how to outwit different preys with different baits. *Gene* 250: 1-14.

Fields S. E. e Song O. (1989). A novel genetic system to detect protein:protein interactions. *Nature* 340: 245-246.

Goffeau A. (2000) Four years of post-genomic life with 6000 yeast genes. *FEBS Letters* 480: 37-41.

Ito T., Tashiro K., Muta S. *et al* (2000) Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci* 97: 1143-1147.

Legrain P. e Selig L. (2000) Genome-wide protein interaction maps using two-hybrid systems. *FEBS Letters* 480: 32-36.

Saffi, J. Feldmann H. Winnacker E. L. E Henriques J. A. P. (2001) Interaction of the yeast Pso5/Rad16 and Sgs1 proteins: influences on DNA repair and aging. *Mutat Res* (in press).

Topcu Z. e Borden K. L. B. (2000) The yeast two-hybrid system and its pharmaceutical significance. *Pharm Res* 17: 1049-1055.

Uetz P. e Hughes R. E. (2000) Systematic and large scale two-hybrid screens. *Curr Opin in Microbiol* 3: 303-308.

Uetz P., Glot L., Cagney G., Mansfield T., Judson R. S. *et al* (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403: 623-627.

Vidal M. e Endoh H. (1999) Prospects for drug screening using the reverse two-hybrid system. *Tibtech* 17: 374-381.